การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ชัยมงคล ปินะสา

วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้
ปีการศึกษา 2566
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ชัยมงคล ปินะสา

วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้
ปีการศึกษา 2566
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

วิทยานิพนธ์ เรื่อง "การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"

ของ ชัยมงคล ปินะสา

ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการสอบวิทยานิพนธ์

(รองศาสตราจารย์ ดร.พัชรี จันทร์เพ็ง)

..... ประธานที่ปรึกษาวิทยานิพนธ์

(รองศาสตราจารย์ ดร.สำราญ มีแจ้ง)

..... กรรมการที่ปรึกษาวิทยานิพนธ์

(รองศาสตราจารย์ ดร.น้ำทิพย์ องอาจวาณิชย์)

..... กรรมการที่ปรึกษาวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร.อ้อมธจิต แป้นศรี)

..... กรรมการผู้ทรงคุณวุฒิภายใน

(รองศาสตราจารย์ ดร.กฤธยากาญจน์ โตพิทักษ์)

อนุมัติ

.....

(รองศาสตราจารย์ ดร.กรองกาญจน์ ชูทิพย์)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ผู้วิจัย ชัยมงคล ปินะสา

ประธานที่ปรึกษา รองศาสตราจารย์ ดร.สำราญ มีแจ้ง

กรรมการที่ปรึกษา รองศาสตราจารย์ ดร.น้ำทิพย์ องอาจวาณิชย์

ผู้ช่วยศาสตราจารย์ ดร.อ้อมธจิต แป้นศรี

ประเภทสารนิพนธ์ วิทยานิพนธ์ ปร.ด. นวัตกรรมทางการวัดผลการเรียนรู้, มหาวิทยาลัยนเรศวร, 2566

คำสำคัญ การคัดเลือกข้อสอบข้อถัดไป, โครงข่ายประสาทเทียม, การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ, เว็บแอพพลิเคชั่น

บทคัดย่อ

การวิจัยครั้งนี้มีจุดมุ่งหมาย 1) เพื่อพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 2) เพื่อประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ซึ่งผู้วิจัยประยุกต์การใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาผสมผสานเข้ากับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) การดำเนินการวิจัยแบ่งออกเป็น 4 ขั้นตอน ดังนี้ 1) การพัฒนาวิธีการจัดคลังข้อสอบ 2) พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม 3) พัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ และ 4) การหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้แบ่งเป็น 2 ส่วน ได้แก่ กลุ่มตัวอย่างที่ใช้ในการหาคุณภาพข้อสอบคัดเข้าคลังข้อสอบ จำนวน 500 คน และกลุ่มตัวอย่างทดลองใช้เว็บแอพพลิเคชั่น จำนวน 200 คน ซึ่งเป็นนักศึกษาหลักสูตรครุศาสตร์บัณฑิต ชั้นปีที่ 3-4 ที่กำลังศึกษาหลักสูตรทางการศึกษาของกลุ่มมหาวิทยาลัยราชภัฏ สังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม เครื่องมือที่ใช้ในการวิจัย ประกอบด้วยแบบทดสอบวัดสมรรถนะความรู้วิชาชีพครูพหุมิติ และแบบประเมินประสิทธิภาพของเว็บแอพพลิเคชั่นต่อการทดสอบ สถิติที่ใช้วิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ตรวจสอบคุณภาพของข้อสอบรายข้อด้วยค่า MDISC MDIFF และตรวจสอบความ

สอดคล้องของโมเดลด้วยค่า SSR RMSR และ GFI ตรวจสอบประสิทธิภาพของโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม อยู่ 4 ค่า คือ 1) ค่าความแม่นยำ (Precision) 2) ค่าความระลึก (Recall) 3) ค่าความถ่วงดุล (F-measure) และ 4) ค่าความถูกต้อง (Accuracy)

ผลการวิจัยพบว่า

1) คลังข้อสอบที่ผ่านการคัดเลือกเข้าคลังข้อสอบ มีจำนวน 242 ข้อ ได้ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (Multidimensional Discrimination: MDISC) ในคลังข้อสอบอยู่ระหว่าง 0.430 ถึง 1.000 และค่าความยากแบบพหุมิติของข้อสอบ (Multidimensional Difficulties: MDIFF) ในคลังข้อสอบ อยู่ระหว่าง -2.432 ถึง 2.199 และค่า SSR RMSR และ GFI มีความสอดคล้องของโมเดลอยู่ในเกณฑ์ที่กำหนด

2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม พบว่าได้โมเดลโครงข่ายประสาทเทียมที่มีชั้นข้อมูลเข้าจำนวน 25 โหนด ชั้นข้อมูลซ่อน^{1} 10 โหนด ชั้นข้อมูลซ่อน^{2} 10 โหนด ชั้นข้อมูลซ่อน^{3} 10 โหนด และชั้นข้อมูลออก 2 โหนด (โมเดล 25-10-10-10-2) ใช้อัลกอริธึม Neural Net ประเมินประสิทธิภาพความแม่นยำในการคัดเลือกข้อสอบข้อถัดไป ด้วยวิธี 10 Fold-Cross-Validation พบว่า มีค่าความถูกต้อง (Accuracy) และมีความแม่นยำ (Precision) 90% มีค่าความถ่วงดุล (F-measure) 94.44% และวัดความถูกต้องของโมเดล มีค่าความระลึก (Recall) 100%

3) พัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นสามารถใช้งานบนระบบปฏิบัติการได้ทุกระบบที่มี Web Browser ทั้ง Notebook Teblet และ Mobile สามารถวิเคราะห์ความสามารถของผู้สอบได้ถูกต้องแม่นยำทั้งความสามารถแยกตามมิติและความสามารถโดยรวม และยังสามารถจำแนกระดับความสามารถของผู้สอบออกเป็น 7 ระดับย่อย คือ สูงมาก สูง ค่อนข้างสูง ปกติ ค่อนข้างต่ำ ต่ำ และต่ำมาก

4) การหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ พบว่าผลการประเมินความเหมาะสมของโปรแกรมทั้งผู้เชี่ยวชาญและผู้ใช้งานมีความเหมาะสมที่สุดทุกรายด้าน แสดงให้เห็นว่า ของเว็บแอพพลิเคชั่นมีความเหมาะสมเพียงพอที่จะนำไปใช้ทดสอบจริงได้

Title WEB APPLICATION DEVELOPMENT FOR MULTIDIMENSIONAL OPTIMIZED TESTING BY APPLYING THE NEURAL NETWORK TO SELECT THE NEXT EXAM

Author Chaimongkhon Pinasa

Advisor Associate Professor Samran Mejang

Co-Advisor Associate Professor Namthip Ongrardwanich

Assistant Professor Omthajit Pansri

Academic Paper Ph.D. Dissertation in Innovation of Learning Measurement - (Type 2.1), Naresuan University, 2023

Keywords Next exam selection, Artificial Intelligence, Artificial Neural Network, Multidimensional Computerized Adaptive Testing

ABSTRACT

The purpose of this study was 1) creating a web application for many tests and performing test compliance. Neural networks are used to choose the next test. 2) Evaluate the quality of web applications and perform different, specific tests. Neural networks are used to choose the next test. Researchers combined artificial intelligence (AI) studies with multidimensional computerized adaptive testing to use neural networks to choose the next test. The research is divided into four steps:

1) Designing the test system 2) designing the method for selecting the next test using the neural network technique. 3) Create a web application to test multiple requirements ; and 4) Find the performance of web applications to test multiple requirements. The models used in this study are divided into two groups. This section contains the criteria used to determine the quality of test papers. Model number. Try web apps 400 and 200 for B.Ed. students. Students from Year 3 to Year 4 are taking courses at the Royal Roads University Group in the Ministry of Higher Education, Science, Research, and Innovation. Research tools include tests that measure intelligence. Multidimensional teacher and satisfaction survey for testing web applications. Statistics used to analyze data include mean and standard

deviation. Use the value to check the quality of each test. Select MDISC MDIFF and check if the pattern matches the value. SSR RMSR and GFI examine the performance of the subsequent test selection model using a neural network approach. There are four values, as follows: 1) Precision 2) Recall 3) F-measure and 4) Accuracy value.

The research results are as follows:

1) Valid test data It contains a large number of 242 license test Multidimensional Classification (Multidimensional Discrimination: MDISC) data checks. 0.430 to 1.000, and Multidimensional Difficulties (MDIFF) are in the middle of the test. The total price was between -2.432 and 2.199 SSR. RMSR and GFI model consistency is within specified standards.

2) Develop a method to select the next test using a neural network method. A neural network model with 25 input methods is shown. Hidden layer: 10 nodes Hidden layer 2: 10 nodes Hidden layer: 10 nodes and question 2 nodes (model 25-10-10-10-2). Using an algorithmic neural network to evaluate the effectiveness of the evaluation Then select the gender and correct it. We found the right value by measuring the performance with the 10-fold cross-validation method. (Accuracy and precision) (Precision) Percentage 90 There is a counterbalance. (F-measure) Percentage: 94.44 and measure the accuracy of the model. There is a commemorative value. (Recall) Percentage 100

3) Develop a web application for multidimensional adaptive testing by applying a neural network to select the next exam developed that can be used on any operating system available. Web Browser for All Notebooks, Tablets, and Mobile Able to analyze test takers' abilities accurately both by dimension and overall ability, and can also classify test takers' proficiency levels into 7 sub-levels: very high, high, relatively high, normal, and relatively low. And very low.

4) Finding Web Application Performance for Multidimensional Tailored Testing It was found that the results of the assessment of the suitability of

the program by both experts and users were suitable in all aspects. Shows that the web application is suitable enough to be used for practical testing.

ประกาศคุณูปการ

วิทยานิพนธ์ฉบับนี้สำเร็จจุล่วงด้วยดีด้วยความกรุณาอย่างดียิ่งจาก รองศาสตราจารย์ ดร. สำราญ มีแจ้ง ประธานที่ปรึกษาวิทยานิพนธ์ รองศาสตราจารย์ ดร.น้ำทิพย์ องอาจวาณิชย์ และ

ผู้ช่วยศาสตราจารย์ ดร.อ้อมธจิต แป้นศรี กรรมการที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้ความรู้ และคำแนะนำที่เป็นประโยชน์ตลอดจนให้กำลังใจกับผู้วิจัยด้วยดีตลอดมา ผู้วิจัยขอขอบพระคุณเป็น อย่างสูงมา ณ โอกาสนี้

ขอขอบพระคุณ รองศาสตราจารย์ ดร.พัชรี จันทร์เพ็ง ประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.กฤธยากาญจน์ โตพิทักษ์ ผู้ทรงคุณวุฒิสอบวิทยานิพนธ์ ที่ให้คำแนะนำ เพื่อให้ วิทยานิพนธ์มีความถูกต้องตามหลักวิชาการ จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จด้วยดีขอขอบพระคุณ ผู้เชี่ยวชาญที่ให้ความอนุเคราะห์ให้คำแนะนำและตรวจสอบเครื่องมือทุกฉบับให้มีความสมบูรณ์

มากยิ่งขึ้น

ขอขอบคุณมหาวิทยาลัยทุกมหาวิทยาลัยที่ให้ความอนุเคราะห์ทำให้การเก็บข้อมูลสำเร็จลุล่วง ไปด้วยดี ขอขอบคุณมหาวิทยาลัยราชภัฏเลย ที่ให้โอกาสได้ศึกษาต่อในครั้งนี้เพื่อนำความรู้

ไปพัฒนาวิชาชีพ รวมทั้งคณาจารย์ในสาขาทุกคน และเพื่อนๆ ปริญญาเอก สาขานวัตกรรม ทาง

การวัดผลการเรียนรู้ รุ่นที่ 1 อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.กฤธยากาญจน์ โตพิทักษ์

ที่ให้คำแนะนำ ให้ความรู้ ความเมตตา และแรงกระตุ้นแก่ผู้วิจัยในการทำวิทยานิพนธ์และ ขอบคุณ

ทุกกำลังใจที่มีให้กับผู้วิจัยเสมอมา

ผู้วิจัยขอขอบคุณครอบครัวที่ให้คำแนะนำให้ความช่วยเหลือสนับสนุนทุนทรัพย์และ

เป็นกำลังใจในการทำวิจัยในครั้งนี้ด้วยดีตลอดมา คุณประโยชน์ที่ได้พึงได้รับจากวิทยานิพนธ์ เล่มนี้

ขอมอบให้กับบิดา มารดา ผู้บ่มเพาะความรู้ให้แก่ผู้วิจัย

ชัยมงคล ปินะสา

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ค

บทคัดย่อภาษาอังกฤษ ..... ฉ

ประกาศคุณูปการ ..... ฌ

สารบัญ ..... ญ

สารบัญตาราง ..... ฑ

สารบัญภาพ ..... ด

บทที่ 1 บทนำ 1

ความเป็นมาและความสำคัญของปัญหา 1

คำถามการวิจัย 5

จุดมุ่งหมายของการวิจัย 6

ขอบเขตการวิจัย 6

นิยามศัพท์เฉพาะ 8

ประโยชน์ที่ได้รับจากการวิจัย 10

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 12

แนวคิดทฤษฎีเกี่ยวกับการพัฒนาเว็บแอพพลิเคชั่น 13

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) 14

โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional IRT Models) 35

การทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing) 52

โครงข่ายประสาทเทียม (Artificial Neuron Network) 67
งานวิจัยที่เกี่ยวข้อง.....84
กรอบแนวคิดในการวิจัย.....92
บทที่ 3 วิธีดำเนินการวิจัย 93
ระยะที่ 1 การพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู 95
ระยะที่ 2 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม 112
ระยะที่ 4 การหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 126
บทที่ 4 ผลการวิเคราะห์ข้อมูล..... 128
ตอนที่ 1 ผลการพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู 129
ตอนที่ 2 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม 143
ตอนที่ 3 ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 152
ตอนที่ 4 ผลการหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะ แบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป..... 162
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ 166
สรุปผล..... 167
อภิปรายผล 173

ข้อเสนอแนะ 178
บรรณานุกรม 180
ภาคผนวก 187
ภาคผนวก ก ผลการวิเคราะห์คุณภาพเครื่องมือ 188
ภาคผนวก ข คำสั่งและผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยการวิเคราะห์องค์ประกอบ เชิงยืนยันเชิงซ้อนตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ โดยโปรแกรม NOHARM 201
ภาคผนวก ค การวิเคราะห์โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยใช้โปรแกรม RapidMiner Studio Educational 10.1) 249
ภาคผนวก ง Code PHP สำหรับโครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 277
ภาคผนวก จ คู่มือการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดย ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 287
ภาคผนวก ฉ แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานของแอพพลิเคชั่น 315
สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ 315
ที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 315
(สำหรับผู้เชี่ยวชาญ) 315
ภาคผนวก ช แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานของแอพพลิเคชั่น 319
สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ 319
ที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป 319
(สำหรับผู้ใช้งาน) 319

ภาคผนวก ซ หนังสือขอความร่วมมือในการวิจัย 324

ประวัติผู้วิจัย 330

สารบัญตาราง

หน้า

ตารางที่ 1 ลักษณะของข้อมูลโมเดลที่ใช้วิเคราะห์ และผู้นำเสนอโมเดลลักษณะของข้อมูลโมเดลที่ใช้วิเคราะห์ ผู้นำเสนอโมเดล 19

ตารางที่ 2 ผลการตอบและการประมาณค่าความสามารถ 34

ตารางที่ 3 สรุปความแตกต่างของการวิเคราะห์องค์ประกอบ (Factor Analysis) และ MIRT 40

ตารางที่ 4 สรุปความแตกต่างของการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ 43

ตารางที่ 5 โครงสร้างของการจัดแบบทดสอบเฟล็กซิเลเวล 62

ตารางที่ 6 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหาเรื่อง การเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ 97

ตารางที่ 7 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหาเรื่อง จิตวิทยาการศึกษา จำนวน 50 ข้อ 98

ตารางที่ 8 การกำหนดโครงสร้างพฤติกรรม(มิติ) ที่วัดเนื้อหาเรื่อง การประกันคุณภาพการศึกษา จำนวน 50 ข้อ 99

ตารางที่ 9 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหา เรื่อง วัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ..... 100

ตารางที่ 10 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่เนื้อหาเรื่อง หลักสูตรและการสอน จำนวน 50 ข้อ..... 101

ตารางที่ 11 โครงสร้างของข้อสอบจำแนกตามจุดประสงค์ทางการศึกษาด้านพุทธิปัญญาและสมรรถนะความรู้ทางวิชาชีพครูที่ออกข้อสอบ จำแนกตามมิติของข้อสอบ..... 102

ตารางที่ 12 วิธีการสุ่มแบบตัวอย่างแบบหลายขั้นตอน (Multi-stage random Sampling) 104

ตารางที่ 13 เกณฑ์พิจารณาค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC)..... 108

ตารางที่ 14 เกณฑ์พิจารณาค่าความยากแบบพหุมิติของข้อสอบ (MDIFF)..... 108

ตารางที่ 15 จำนวนข้อสอบที่ผ่านเกณฑ์จากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติที่ได้รับคัดเลือกเข้าคลังข้อสอบ จำแนกตามมิติของข้อสอบ 109

ตารางที่ 16 จำนวนข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูที่ได้รับการคัดเลือกเข้าคลังข้อสอบ จำแนกตามมิติของข้อสอบ 130

ตารางที่ 17 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 1 การเปลี่ยนแปลงบริบทโลก (a_1)..... 131

ตารางที่ 18 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 2 จิตวิทยาการศึกษา (a_2)..... 132

ตารางที่ 19 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 3 การประกันคุณภาพการศึกษา (a_3) 134

ตารางที่ 20 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 4 วัดผลและวิจัยการเรียนรู้ (a_4) 136

ตารางที่ 21 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 5 หลักสูตรและการสอน (a_5) 138

ตารางที่ 22 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู ที่ผ่านการคัดเลือกเข้าคลังข้อสอบ 140

ตารางที่ 23 รายละเอียดข้อมูลสำหรับทำเหมืองข้อมูล ข้อสอบ 5 ข้อ..... 144

ตารางที่ 24 ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบ 147

ตารางที่ 25 คะแนนความสามารถTheta ที่ปรับเป็นคะแนนมาตรฐาน (scale score) 159
ตารางที่ 26 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้เชี่ยวชาญ..... 162
ตารางที่ 27 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้ใช้งาน..... 164
ตารางที่ 28 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู ที่ผ่านการคัดเลือกเข้าคลังข้อสอบ 190

สารบัญภาพ

หน้า

ภาพที่ 1 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 3 พารามิเตอร์.....20

ภาพที่ 2 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 1 พารามิเตอร์.....21

ภาพที่ 3 เส้นโค้งลักษณะข้อสอบของโมเดลโลจิสแบบ 4 พารามิเตอร์ 22

ภาพที่ 4 ช่วงความเชื่อมั่น (confidence interval ; $\underline{\theta}$, $\bar{\theta}$) ของความสามารถที่ได้จากการประมาณค่า ($\hat{\theta}$).....24

ภาพที่ 5 ช่วงแห่งความเชื่อมั่นแอซิมโทติก 95% ($\underline{\theta}$, $\bar{\theta}$) สำหรับความสามารถ θ 25

ภาพที่ 6 พหุมิติระหว่างข้อคำถามและพหุมิติภายในข้อคำถาม, โดย (Wu et al., 2007)...37

ภาพที่ 7 รูปแบบโมเดล MIRT ของ Embretson & Reise (2000).....47

ภาพที่ 8 รูปแบบโมเดล MIRT ของ Reckase (2009).....48

ภาพที่ 9 โครงสร้างของการทดสอบเฉพาะบุคคลแบบสองขั้นตอน 57

ภาพที่ 10 โครงสร้างของการทดสอบเฉพาะบุคคลรูปแบบพีระมิดขนาดขั้นคงที่.....58

ภาพที่ 11 โครงสร้างของแบบทดสอบรูปพีระมิดขนาดขั้นแปรผัน.....59

ภาพที่ 12 โครงสร้างของแบบทดสอบรูปพีระมิดแบบโรบินส์มอนโร 59

ภาพที่ 13 โครงสร้างของแบบทดสอบรูปพีระมิดแบบสกัดการสะท้อนกลับ.....60

ภาพที่ 14 โครงสร้างของแบบทดสอบรูปพีระมิดแบบรักษาการสะท้อนกลับ.....60

ภาพที่ 15 โครงสร้างของแบบทดสอบรูปพีระมิดแบบมีหลายข้อในแต่ละขั้น.....61

ภาพที่ 16 โครงสร้างของแบบทดสอบรูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือกของข้อสอบเพื่อแยกทาง.....61

ภาพที่ 17 โครงสร้างของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบบแบ่งชั้น.....63

ภาพที่ 18 ตัวอย่างขั้นตอนและกฎเส้นทางของ MST 65
ภาพที่ 19 ตัวอย่างขั้นตอนของวิธี MST.....66
าพที่ 20 ระบบเซลล์ประสาทของมนุษย์.....69
ภาพที่ 21 ส่วนประกอบของโครงข่ายประสาทเทียม 70
ภาพที่ 22 โครงสร้างการทำงานของโครงข่ายประสาทเทียม 71
ภาพที่ 23 ฟังก์ชันลีคกี้เรคติไฟด์ลินเนียยูนิต 73
ภาพที่ 24 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว 74
ภาพที่ 25 โครงข่ายประสาทเทียมแบบหลายชั้น 74
ภาพที่ 26 ความแตกต่างระหว่างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก 77
ภาพที่ 27 โมเดล Neural Net อัลกอริทึม Deep Learning 77
ภาพที่ 28 โครงสร้างของโครงข่ายประสาทเทียมแบบวนกลับ.....79
ภาพที่ 29 รูปโปรแกรม Rapid miner studio 10.1 ประเภทสำหรับการศึกษา (Educational Program).....84
ภาพที่ 30 Decision Item Tree Network (Yu Hsieh & Chen Kuo, 2010, p. 1161).....88
ภาพที่ 31 CAT Based on Decision Item Subtree (Yu Hsieh & Chen Kuo, 2010, p. 1164) 88
ภาพที่ 32 กรอบแนวคิดที่ใช้ในการวิจัย 92
ภาพที่ 33 ขั้นตอนการดำเนินการวิจัย.....94
ภาพที่ 34 การวิเคราะห์เนื้อหาการออกข้อสอบ พร้อมระบุขอบข่ายประเด็น และระดับพฤติกรรม.....95
ภาพที่ 35 ขั้นตอนการพัฒนาวิธีการจัดคลังข้อสอบประเมินมิติความรู้ทางวิชาชีพครู.... 111
ภาพที่ 36 โมเดลโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไป..... 115

ภาพที่ 37 การทดสอบประสิทธิภาพโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไป 117
ภาพที่ 38 ขั้นตอนการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม (ann) 120
ภาพที่ 39 ขั้นตอนพัฒนาเว็บแอพพลิเคชั่น 121
ภาพที่ 40 แผนผังบริบท (Context Diagram) ของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ 122
ภาพที่ 41 โครงสร้างหน้าจอทั่วไปของเว็บแอพพลิเคชั่นการทดสอบ 123
ภาพที่ 42 โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ 123
ภาพที่ 43 โครงสร้างหน้าจอการทดสอบ 123
ภาพที่ 44 โครงสร้างหน้าจอการรายงานผลการทดสอบ 124
ภาพที่ 45 โครงสร้างหน้าจอ Login 124
ภาพที่ 46 โครงสร้างหน้าจอการจัดการข้อสอบ 124
ภาพที่ 47 โครงสร้างหน้าจอรายงานผลการจัดการข้อสอบ 125
ภาพที่ 48 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก 131
ภาพที่ 49 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก 132
ภาพที่ 50 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา 133
ภาพที่ 51 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา 134
ภาพที่ 52 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา 135

ภาพที่ 53 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา..... 136

ภาพที่ 54 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้..... 137

ภาพที่ 55 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้..... 138

ภาพที่ 56 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้..... 139

ภาพที่ 57 ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC)..... 141

ภาพที่ 58 ค่าความยากแบบพหุมิติของข้อสอบ (MDIFF)..... 142

ภาพที่ 59 แสดงการกระจายระหว่างค่าความยากแบบพหุมิติของข้อสอบ กับค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ..... 143

ภาพที่ 60 สร้างโมเดลการคัดเลือกข้อสอบข้อถัดไป จากโปรแกรม RapidMiner Studio Educational 10.1..... 145

ภาพที่ 61 โมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) กรณีตอบถูก และตอบผิด..... 146

ภาพที่ 62 เปรียบเทียบค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึกและค่าความถ่วงดุล ของการคัดเลือกข้อสอบข้อถัดไป..... 147

ภาพที่ 63 สมการโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) (25-10-10-10-2)..... 150

ภาพที่ 64 Code PHP สำหรับคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN)..... 151

ภาพที่ 65 หน้าจอแรกของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป เมื่อเข้าใช้งานโปรแกรมที่เว็บไซต์ http://examadaptive.com/index.php 152

ภาพที่ 66 หน้าจอแสดงความหมายและขั้นตอนการทำงานของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยสรุป 153

ภาพที่ 67 หน้าจอการเข้าระบบและลงทะเบียนสมัครสมาชิก 153

ภาพที่ 68 หน้าจอแสดงแบบฟอร์มให้ผู้สอบพิมพ์ข้อมูลและเลือกข้อมูลของตนเองตามความเป็นจริง เมื่อกดปุ่ม ลงทะเบียน 154

ภาพที่ 69 หน้าจอแรกของการเข้าใช้งานโปรแกรม 154

ภาพที่ 70 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้ 155

ภาพที่ 71 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้ ผู้สอบตรวจสอบข้อมูลเบื้องต้น ของผู้สอบ แล้วกดปุ่ม บันทึก 156

ภาพที่ 72 หน้าจอแสดงข้อมูลรายวิชา และรายละเอียดแบบทดสอบ ผู้สอบอ่านคำชี้แจงให้ละเอียด ก่อนกด เริ่มทำแบบทดสอบ 157

ภาพที่ 73 หน้าจอแสดงตัวเลือกตอบ 5 ตัวเลือก 157

ภาพที่ 74 หน้าจอแสดงข้อมูลตามที่ผู้สอบลงทะเบียนไว้ และประวัติผลการทำแบบทดสอบ 158

ภาพที่ 75 หน้าจอแสดงผลรายงานผลการใช้แบบทดสอบ 158

ภาพที่ 76 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบ Theta (θ), ค่า Scale Score และผลการประเมินแยกตามมิติ 159

ภาพที่ 77 หน้าจอแสดงผลคะแนนที่ปรับในสเกลมาตรฐาน (scale score) 160

ภาพที่ 78 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบและผลการประเมิน .. 160

ภาพที่ 79 หน้าจอเกณฑ์การประเมินความสามารถของผู้สอบแยกตามมิติของข้อสอบ. 161

ภาพที่ 80 หน้าจอทำแบบประเมิน..... 161
ภาพที่ 81 กราฟแสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้เชี่ยวชาญ
..... 163
ภาพที่ 82 เก็บข้อมูลประวัติผลการทำแบบทดสอบเพื่อเก็บสถิติและดูพัฒนาการ..... 164
ภาพที่ 83 กราฟแสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้ใช้งาน165

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การทดสอบและประเมินสมรรถนะทางวิชาชีพครูเพื่อขอรับใบอนุญาตประกอบวิชาชีพครู โดยผู้เข้าศึกษาตามหลักสูตรปริญญาและประกาศนียบัตรบัณฑิตที่คุรุสภารับรอง เมื่อสำเร็จการศึกษาแล้วต้องผ่านการทดสอบจึงจะรับใบอนุญาตประกอบวิชาชีพครูได้ ที่ผ่านมามีผู้สำเร็จและเข้าสอบทุกปีมีจำนวนมากทั่วประเทศ ลักษณะการสอบ และขั้นตอนในการทดสอบใช้เวลานานรวมถึงการสร้างข้อสอบให้เป็นมาตรฐานเดียวกัน การวิเคราะห์และหาคุณภาพของข้อสอบและการจัดคลังข้อสอบยังไม่มีระบบ ต่อมาได้มีการปรับเป็นการทดสอบด้วยระบบอิเล็กทรอนิกส์เป็นการอำนวยความสะดวกให้กับผู้สอบและเป็นการปรับรูปแบบให้ทันสมัย รวมถึงมีความรวดเร็วในกระบวนการตรวจข้อสอบถูกต้อง และแม่นยำ สำหรับข้อสอบที่ใช้จะเป็นข้อสอบการบูรณาการความรู้ ทักษะ และคุณลักษณะที่แสดงสมรรถนะตามขอบเขตการทำงานของวิชาชีพครู นอกจากนี้ได้จัดสิ่งอำนวยความสะดวกแก่ผู้เข้าสอบทุกคน รวมถึงมีมาตรการป้องกันการแพร่ระบาดของเชื้อไวรัสโคโรนา 2019

การนำเทคโนโลยีทางคอมพิวเตอร์มาใช้ในการทดสอบในปัจจุบันนี้ ซึ่งเทคโนโลยีทางคอมพิวเตอร์ได้เข้ามาช่วยเสริมสร้างประสิทธิภาพในการตรวจข้อสอบ การวิเคราะห์ข้อสอบ การตัดเกรด และรายงานผล ซึ่งมีการใช้กันอย่างกว้างขวางตั้งแต่ช่วงปี ค.ศ. 1970-1980 รวมทั้งยังมีการนำมาใช้สร้างคลังข้อสอบ (item bank) ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing) ซึ่งการทดสอบแบบปรับเหมาะนั้นมีความสามารถมากกว่าการทดสอบแบบดั้งเดิมที่เขียนลงบนกระดาษคำตอบ (paper-and-pencil tests) ซึ่งเป็นการเพิ่มความแม่นยำในการวัด (Segall, 1996, pp. 331-354) ดังนั้นจึงมีการพัฒนาแนวคิดเกี่ยวกับทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (MIRT) ซึ่งมีแนวคิดมาจาก 2 กลุ่ม คือ Spearman (1927) and Thurstone (1947) และ Lazarsfeld (1950) (พัชรี จันทร์เพ็ง, 2550, น. 22) (Segall, 2010, pp. 57-75) และเมื่อนำคอมพิวเตอร์มาใช้ในกระบวนการทดสอบจึงเรียกการทดสอบนี้ว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) การทดสอบดังกล่าวนี้เป็นการทดสอบที่มีประสิทธิภาพในการวัดสูง ลดจำนวนข้อสอบลงได้มากกว่าการทดสอบด้วยการทดสอบแบบ CAT ประมาณ 30 -50% และลดจำนวนข้อสอบได้มากกว่าการทดสอบแบบดั้งเดิมประมาณ 70% และไม่สูญเสียความแม่นยำในการทดสอบ (Frey and Seitz, 2009, p. 93)

โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory : MIRT) เป็นแนวคิดที่ขยายมาจากโมเดลการตอบสนองข้อสอบ (Unidimensional Item Response

Theory Model; UIRT Model) เพื่อแก้ไขจุดอ่อนของข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติโดยมีพื้นฐานมาจากการวิเคราะห์องค์ประกอบและทฤษฎีการตอบสนองข้อสอบ (Reckase, 2009, pp. 63 -70) เป็นโมเดลการตอบสนองข้อสอบแบบพหุมิติ (ศิริชัย กาญจนวาสี, 2550, น. 84) ซึ่งแบบทดสอบที่นำมาวัดจะต้องเหมาะสมกับผู้สอบ ทำให้มีการพัฒนาการทดสอบแบบปรับเหมาะ (Adaptive testing) ขึ้นมา ทฤษฎีการตอบสนองข้อสอบมีประสิทธิภาพและมีประโยชน์ในการพัฒนาการทดสอบ การวิเคราะห์ข้อสอบและคะแนนของผู้สอบ (Nathan, 2011, p. 1) ดังนั้น การทดสอบแบบปรับเหมาะจึงควรเป็นไปในลักษณะที่จัดการทดสอบให้เป็นรายบุคคล โดยทำการคัดเลือกข้อสอบที่เหมาะสมกับผู้สอบ (Tailored) ในการประมาณค่าความสามารถของผู้สอบได้อย่างถูกต้องแม่นยำ เมื่อประมาณค่าความสามารถได้อย่างมีความคลาดเคลื่อนของผู้สอบต่ำแล้วจึงยุติการสอบ (Hambleton, Swaminathan and Rojers, 1991, p. 145)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีกระบวนการในการทดสอบที่เป็นพื้นฐานอยู่ 6 กระบวนการ (Weiss and Kingsbury, 1984, pp. 361-375) ได้แก่ 1) รูปแบบของการตอบข้อสอบ 2) คลังข้อสอบ 3) จุดเริ่มต้นของการทดสอบ 4) วิธีการเลือกข้อสอบ 5) การประมาณค่าความสามารถของผู้สอบ 6) เกณฑ์ยุติการทดสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computer Adaptive Testing: MCAT) (Reckase, 2009, p. 313) นำเสนอ 4 กระบวนการ ได้แก่ 1) พัฒนาคลังข้อสอบ 2) การเลือกข้อสอบ 3) การประมาณค่าความสามารถ 4) เกณฑ์ยุติการทดสอบ รายละเอียดในแต่ละกระบวนการเป็นดังนี้ 1) รูปแบบของการตอบข้อสอบ (Hambleton, Swaminathan and Rojers, 1991, pp. 17-18) แนะนำว่าการใช้รูปแบบ 3 พารามิเตอร์ คือ มีความยาก อำนาจจำแนก และการเดา มีความเหมาะสมกับการทดสอบมากกว่า 1 พารามิเตอร์ และ 2 พารามิเตอร์ แต่ในปัจจุบันทฤษฎีการตอบสนองข้อสอบมีโมเดลที่ใช้กันอย่างแพร่หลายในเช่น โมเดล 1PL 2PL และ 3PL ซึ่งรู้จักกันแพร่หลายต่อมา Barton & Lord (1981, p. 2) ได้เสนอแนวคิดเกี่ยวกับ 4PLคือ ความสะเพร่า 2) คลังข้อสอบ ซึ่งคลังข้อสอบที่ดีจะทำให้การทดสอบมีประสิทธิภาพข้อสอบที่มีจำนวนมาก คลังข้อสอบที่ดีควรจะมีดังนี้ คือ ค่าอำนาจจำแนกสูง ความยากควรจะกระจายครอบคลุมความสามารถและค่าการเดาเข้าใกล้ศูนย์ (Fen, 2000, p. 8) 3) จุดเริ่มต้นของการทดสอบจะมีผลต่อความถูกต้องแม่นยำของการประมาณค่าจำนวนข้อสอบที่ผู้สอบทำจะมีผลต่อเกณฑ์ยุติการทดสอบ ดังนั้น จุดเริ่มต้นของการทดสอบควรทดสอบที่ระดับความเหมาะสมของความสามารถของผู้สอบ ถ้าไม่ทราบระดับความสามารถของผู้สอบที่เหมาะสมโดยทั่วไปก็จะใช้ภูมิหลังของผู้สอบเป็นสารสนเทศในการตัดสินใจเป็นจุดเริ่มต้นของการทดสอบ ถ้าไม่ทราบภูมิหลังของผู้สอบจุดเริ่มต้นที่ดีคือใช้ความยากระดับปานกลาง แต่ถ้าความยากของข้อสอบในคลังข้อสอบมีการกระจายเป็นโค้งปกติคือ มีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนเท่ากับ 1 จุดเริ่มต้นของการทดสอบจะใช้ที่ความยากเท่ากับศูนย์ 4) วิธีการเลือกข้อสอบและการประมาณค่า

ความสามารถของผู้สอบ นิยมใช้วิธีที่สอดคล้องกัน (ศิริชัย กาญจนวาสี, 2550, น. 202) คือ กรณีที่ใช้การประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุด นิยมใช้วิธีการคัดเลือกข้อสอบที่มีความยาก (bi) สอดคล้องกับระดับความสามารถที่ประมาณได้ (θ) หรือนิยมใช้กับวิธีการคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศสูงสุด สอดคล้องกับระดับความสามารถที่ประมาณได้ (θ) ส่วนการประมาณค่าความสามารถด้วยวิธีเบส์ นิยมใช้วิธีการคัดเลือกข้อสอบข้อที่ยังไม่ได้นำมาใช้ในการทดสอบนั้น ที่จะให้ค่าความแปรปรวนของค่าความสามารถที่คาดหวังมีค่าต่ำสุดตลอดระยะเวลาที่ผ่านมา นักวิจัยมีความพยายามที่จะพัฒนาการทดสอบในแต่ละกระบวนการให้ประสิทธิภาพมากที่สุด เพราะในแต่ละกระบวนการมีผลต่อการทดสอบ เช่น Diao และ Reckase (2009, pp. 1-13) พบว่าการประมาณค่าด้วยวิธีความน่าจะเป็นสูงสุดด้วยข้อสอบจำนวน 50 ข้อ ต้องใช้เวลานานมากกว่าจะสามารถประมาณค่าได้ แต่วิธีแบบเบส์สามารถประมาณค่าความสามารถได้อย่างรวดเร็ว และเมื่อข้อสอบมีความยาวมากขึ้นวิธีแบบเบส์จะมีความลำเอียงน้อยกว่าวิธีความน่าจะเป็นสูงสุด ทัศน์ศิรินทร์ สว่างบุญ (2554, น. 148-152) พบว่า วิธีการเลือกข้อสอบขั้นแรกด้วยวิธี Kullback-Leibler information มีคุณภาพการทดสอบสูงกว่าวิธี Fisher information รังสรรค์ มณีเล็ก (2540, น. 171-174) พบว่าวิธีประมาณค่าความสามารถของผู้สอบ เกณฑ์การยุติการสอบ และความสามารถของผู้สอบมีผลต่อความเที่ยงตรงตามสภาพของแบบทดสอบ และความยากของข้อสอบข้อแรกเกณฑ์การคัดเลือกข้อสอบ วิธีประมาณค่าความสามารถของผู้สอบ เกณฑ์การยุติการสอบและความสามารถของผู้สอบมีผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบคลังข้อสอบ (Item pool) มีความสำคัญในการทดสอบ (Flaugher, 2000) และมีผลต่อประสิทธิภาพของการทดสอบ (Ben and David, 2012, pp. 1-20) สำหรับคลังข้อสอบที่มีขนาดพอเหมาะนั้นควรมีประมาณ 100-200 ข้อจะให้ผลดีที่สุด (Weiss, 1988, p. 372) และคลังข้อสอบที่ดีควรมีการกระจายของค่าพารามิเตอร์และการกระจายของเนื้อหา (Weiss, 1985; 2011, pp. 1-23) นอกจากนี้ พบว่าการประมาณค่าความสามารถของผู้สอบที่ใช้คลังข้อสอบจากโมเดล 4PL มีความถูกต้องมากกว่าคลังข้อสอบโมเดล 3PL (Liao, Ho & Yen, 2012, p. 1686) และเกณฑ์ยุติการสอบส่งผลต่อประสิทธิภาพของการทดสอบ (Nathan and Thomson, 2007, pp 5-6)

วิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้กันเป็นมาตรฐานในการเลือกข้อสอบอยู่หลายวิธี ได้แก่ 1) การคัดเลือกข้อสอบข้อถัดไปที่ค่าความยากของข้อสอบ (b-Parameter) ใกล้เคียงกับความสามารถของผู้สอบ 2) การคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์สารสนเทศสูงสุด (Maximum Information Criterion : MI) มีงานวิจัยที่เกี่ยวข้องกับการคัดเลือกข้อสอบสามารถจำแนกกลุ่มได้ 2 กลุ่ม คือ กลุ่มที่ 1 ใช้ปัญญาประดิษฐ์ระดับการเรียนรู้ของเครื่อง (Machine learning) ประเภทต้นไม้ตัดสินใจ (Decision Tree) เพื่อที่จะคาดการณ์หรือตัดสินใจได้อย่างชัดเจน ได้แก่ Bayesian Networks (Hsieh and Kuo, 2010) และ Genetic Algorithms (Phankokkruad and Woraratpanya, 2012)

การคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Sukhanonsawat, 2013) หรือวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีอาณานิคมมด (Thanapattanadol, 2018) การคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ส่วนใหญ่คัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงมาใช้ในการทดสอบ ทำให้ข้อสอบบางข้อในคลังข้อสอบถูกเลือกใช้ในการทดสอบบ่อยครั้งเกินไป จึงมีการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีกระบวนการควบคุมความถี่ในการใช้ข้อสอบให้มีความเหมาะสม เช่น การคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคข์และการควบคุมการใช้ข้อสอบสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Premthongsuk, 2017) ซึ่งนำวิธีการสุ่มตัวอย่างแบบมีระบบมาควบคุมการใช้ข้อสอบ และการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม และกลุ่มที่ 2 ใช้ปัญญาประดิษฐ์ระดับการเรียนรู้เชิงลึก (Deep Learning) ประเภทโครงข่ายประสาทเทียม (Neural Network) โดยเป็นปัญญาประดิษฐ์ที่มีระดับการเรียนรู้แบบ Supervised learning ซึ่งเป็นการเรียนรู้ที่ต้องมีชุดข้อมูลมาฝึกฝน เช่น พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบแบบ CAT โดยใช้วิธีการจำลองข้อมูล (บูรพา วิถีปัญญา, 2563) การนำวิธีการคัดเลือกข้อสอบข้อถัดไปที่นิยมใช้กันเป็นมาตรฐานในการเลือกข้อสอบอยู่หลายวิธี ได้แก่ 1) การคัดเลือกข้อสอบข้อถัดไปที่ค่าความยากของข้อสอบ (b-Parameter) ใกล้เคียงกับความสามารถของผู้สอบ 2) การคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์สารสนเทศสูงสุด (Maximum Information Criterion) (Birnbaum, 1968)

ปัจจุบันมีการประยุกต์วิธีวิเคราะห์ข้อมูลจากศาสตร์ทางวิศวกรรมคอมพิวเตอร์มาใช้ในการวิจัยทางสังคมศาสตร์ ได้แก่ การวิเคราะห์เครือข่ายใยประสาท (neural network analysis) (Ader และ Bramsen, 1998; Ragothaman, Davies and Moen, 1998; Sinha and McKim, 2000; Gonzalez and DesJardins, 2002) ซึ่งเป็นการเปรียบเทียบสถิติแบบดั้งเดิม ได้แก่ การวิเคราะห์ถดถอยโลจิสติค การวิเคราะห์จำแนกกับการวิเคราะห์เครือข่ายใยประสาทแบบต่างๆ เพื่อศึกษาประสิทธิภาพการพยากรณ์และการจำแนกกลุ่มผลลัพธ์ที่ได้มีค่าใกล้เคียงกัน ทั้งนี้การวิเคราะห์ด้วยเครือข่ายใยประสาทไม่มีข้อตกลงเบื้องต้นใดๆ และเป็นการวิเคราะห์ที่มีความแม่นยำสูงเพราะหลักการคำนวณด้วยเครือข่ายใยประสาทใช้การวิเคราะห์ผ่านการทำงานของข้อมูลแอบแฝง (hidden node) หลายตัวโดยผ่านฟังก์ชั่นแปลงค่า ผลค่าน้ำหนักที่ได้ จึงเป็นการประมาณค่าของโปรแกรมที่ไม่สามารถหาความสัมพันธ์หรือเปรียบเทียบค่าใดๆ ได้และมีการวิเคราะห์เครือข่ายใยประสาทมาใช้ศึกษาความสัมพันธ์เชิงสาเหตุ เช่น งานวิจัยของ Ader และ Bramsen (1998), ศจีมาจ ขวัญเมือง (2548) ที่ทดลองนำข้อมูลการวิจัยที่วิเคราะห์ด้วยสมการเชิงโครงสร้างที่ใช้โปรแกรมลิสเรลมาวิเคราะห์ด้วยเทคนิคเครือข่ายใยประสาท ผลการทดลองพบว่า ค่าน้ำหนักของการคำนวณด้วยเครือข่ายใยประสาท มีความสอดคล้องกับค่าอิทธิพลของการวิเคราะห์ด้วยโปรแกรมลิสเรล โดย

พารามิเตอร์ในโมเดลลิสเรลที่มีค่าอิทธิพลสูงจะมีค่าน้ำหนักในการวิเคราะห์ด้วยเครือข่ายใยประสาทสูง ให้ผลการวิเคราะห์สอดคล้องกันเช่นกัน

โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) เป็นศาสตร์แขนงหนึ่งของทางด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) เป็นโปรแกรมทางคณิตศาสตร์ที่พัฒนาขึ้นด้วยวัตถุประสงค์ที่จะสร้างโปรแกรมการวิเคราะห์ให้มีความสามารถในการเรียนรู้ การจำ การจัดการกับข้อมูลที่คลุมเครือ ซับซ้อน กำกวม และไม่สมบูรณ์ได้ ทำให้การประมวลผลมีความแม่นยำ รวดเร็วและมีประสิทธิภาพสูง โดยอาศัยพื้นฐานหลักการทำงานของระบบประสาทของมนุษย์เป็นเทคนิคที่เน้นการจัดกระทำกับข้อมูล (data driven) มาสร้างเป็นข้อสรุป หลักเครือข่ายที่นิยมใช้ในการวิเคราะห์เปรียบเทียบค่าน้ำหนักในคือ Back propagation (แพร่กระจายย้อนกลับ) ซึ่งเป็นการปัจจุบันกับค่าน้ำหนักในอดีต โดยพิจารณาจากค่าความคลาดเคลื่อน ผลการวิเคราะห์จะสิ้นสุดเมื่อผลรวมของค่าเฉลี่ยความคลาดเคลื่อนมีค่าน้อยกว่าระดับที่ผู้วิจัยกำหนด

จากการทดสอบและประเมินสมรรถนะทางวิชาชีพครูเพื่อขอรับใบอนุญาตประกอบวิชาชีพครูในปัจจุบัน และงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ทำให้ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป เพื่อใช้สำหรับทดสอบและประเมินสมรรถนะทางวิชาชีพครูให้มีความทันสมัยยิ่งขึ้น โดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artifitial Intelligence : AI) มาประยุกต์ใช้กับเทคนิคการวิเคราะห์ข้อมูลทางการศึกษาด้วยเทคนิควิเคราะห์แนวใหม่ คือ โครงข่ายประสาทเทียม (Artificial Neuron Network : ANN) การวิเคราะห์ข้อมูลด้วยวิธีดังกล่าวเน้นการจัดกระทำกับข้อมูล (data driven) โดยตรงจึงให้สารสนเทศที่สอดคล้องกับข้อมูลจริงมากที่สุดและเป็นวิธีการวิเคราะห์ที่ผ่อนคลายข้อตกลงเบื้องต้น สามารถนำไปประยุกต์ใช้กับข้อมูลที่มีจำนวนน้อยได้ ผลการวิจัยนี้จึงเป็นอีกแนวทางหนึ่งในการการทดสอบแบบปรับเหมาะแบบพหุมิติที่เป็นประโยชน์และมีคุณค่าทางวิชาการ

คำถามการวิจัย

1. เว็บแอพพลิเคชั่นสำหรับการทดสอบสมรรถนะทางวิชาชีพครูแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้น มีรูปแบบและความสามารถในการทำงานอย่างไร

2. เว็บแอพพลิเคชั่นสำหรับการทดสอบสมรรถนะทางวิชาชีพครูแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป มีคุณภาพอยู่ในระดับใด

จุดมุ่งหมายของการวิจัย

1. เพื่อพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบสมรรถนะทางวิชาชีพครูแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

2. เพื่อประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบสมรรถนะทางวิชาชีพครูแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ขอบเขตการวิจัย

งานวิจัยนี้มุ่งเน้นที่การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป และหาประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยประยุกต์ใช้โครงข่ายประสาทเทียม มีรายละเอียดดังนี้

1. ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักศึกษาหลักสูตรครุศาสตรบัณฑิต ชั้นปีที่ 3-4 ที่กำลังศึกษาหลักสูตรทางการศึกษาของกลุ่มมหาวิทยาลัยราชภัฏ สังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม สำหรับประชากรที่ใช้ในครั้งนี้เป็นนักศึกษาในระดับชั้นปีที่ 3-4 เท่านั้น เนื่องจากในการทดสอบครั้งนี้เน้นการนำไปใช้เพื่อการประเมินผลสรุปรวม (summative evaluation) เพื่อเป็นข้อมูลย้อนกลับให้กับหลักสูตร/ สาขาวิชา/ คณะ ใช้ประโยชน์ในการพัฒนาปรับปรุงแก้ไขการจัดการเรียนการสอนต่อไป ดังนั้นจึงเหมาะที่จะใช้นักศึกษาในระดับชั้นปีที่ 3 ขึ้นไป ซึ่งเป็นระดับชั้นที่เรียนรายวิชาทางด้านวิชาชีพครูและวิชาเฉพาะครบถ้วนแล้ว และกำลังจะออกไปฝึกประสบการณ์วิชาชีพครูในสถานศึกษาจริงอีกทั้งเป็นการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครู

2. การทดสอบและประเมินสมรรถนะทางวิชาชีพครูในการวิจัยครั้งนี้ เป็นการทดสอบและประเมินสมรรถนะทางวิชาชีพครูตามประกาศคณะกรรมการคุรุสภา เรื่องหลักเกณฑ์และวิธีการทดสอบ และประเมินสมรรถนะทางวิชาชีพครู พ.ศ. 2563 ซึ่งกำหนดองค์ประกอบของสมรรถนะทางวิชาชีพที่จะทดสอบและประเมินไว้ ดังนี้ กลุ่มวิชาชีพครู ได้แก่ หลักสูตรและการสอน จิตวิทยาการประกันคุณภาพการศึกษา วัดผลและวิจัยการเรียนรู้ และการเปลี่ยนแปลงบริบทโลกโดยเฉพาะกลุ่มวิชาชีพครู นับว่าเป็นกลุ่มที่มีความสำคัญสำหรับบัณฑิตครูของของมหาวิทยาลัยราชภัฏ และบัณฑิตทุกคนของมหาวิทยาลัยจะต้องสอบผ่านในวิชาเหล่านี้และจะต้องผ่านการทดสอบและการประเมินจนได้รับใบประกอบวิชาชีพครู ทุกคน

3. เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปในครั้งนี้ มีองค์ประกอบของการทดสอบแบบปรับเหมาะแบบพหุมิติ ประกอบด้วยขั้นตอน 5 ขั้นตอน (Veldkamp & Matteucci, 2013; Thompson & Weiss, 2011) ได้แก่

1) การสร้างคลังข้อสอบ (Item pools)

การสร้างคลังข้อสอบ (Item Pools) สร้างตามหลักเกณฑ์ที่ถูกต้องและถูกเก็บไว้อย่างเป็นระบบ วิธีจัดคลังข้อสอบโดยประยุกต์วิธีการของ Chang and Ying (1996), Baker (2001), Hasmy (2014), ศักดิ์ชัย จันทะแสง (2561) ซึ่งแบ่งคลังข้อสอบออกเป็น 4 ชั้น ตามช่วงของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) ชั้นที่ 1 คือ 0.35 - 0.64 ชั้นที่ 2 คือ 0.65 - 1.34 ชั้นที่ 3 คือ 1.35 - 1.70 และชั้นที่ 4 คือ 1.71 ขึ้นไป ร่วมกับการจัดสมดุลเนื้อหา (Content Balancing) ภายในชั้นของค่าอำนาจจำแนก ได้แบ่งเป็น 5 กลุ่ม ตามมิติความรู้ทางวิชาชีพครู ได้แก่ 1) มิติหลักสูตรและการสอน 2) มิติจิตวิทยา 3) มิติการประกันคุณภาพการศึกษา 4) มิติวัดผลและวิจัยการเรียนรู้ และ 5) มิติการเปลี่ยนแปลงบริบทโลก

2) การเริ่มต้นการทดสอบ (Initiation/Starting point)

การคัดเลือกข้อสอบข้อแรก (Initial Item) ใช้ข้อสอบที่มีความยากปานกลาง กำหนดให้ Initial ability = 0 เป็นจุดเริ่มต้นของการทดสอบ

3) การคัดเลือกข้อสอบข้อถัดไป (Item selection algorithm/Item selection)

การคัดเลือกข้อสอบ (Item Selection) กำหนดการคัดเลือกข้อสอบด้วยวิธีการปัญญาประดิษฐ์ (Artifitial Intelligence: AI) มาประยุกต์ใช้กับเทคนิคการวิเคราะห์โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN)

4) การประมาณความสามารถ (Ability Estimation/ Scoring algorithm)

วิธีประมาณค่าความสามารถของผู้สอบ (Ability Estimate) ผู้วิจัยใช้การประมาณค่าความสามารถผู้สอบด้วยวิธีเบส์ปรับใหม่ (Bayesian Updating) เนื่องจากการศึกษาของ รังสรรค์ มณีเล็ก (2540) พบว่า การประมาณค่าความสามารถด้วยวิธี Bayesian Estimation มีความตรงเชิงสภาพสูงกว่าการประมาณค่าแบบวิธี Maximum Likelihood Estimation

5) เกณฑ์การยุติการทดสอบ (Termination criterion/ Stopping rules) (Veldkamp and Matteucci, 2013; Thompson and Weiss, 2011)

เกณฑ์ยุติการทดสอบ (Termination Criteria) ใช้การจำกัดจำนวนข้อสอบ (Fixed Length) เนื่องจากมีการเปรียบเทียบวิธีวิทยาของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ ซึ่งงานวิจัยที่ผ่านมาไม่ว่าจะเป็นการจำลองข้อมูลหรือการเก็บข้อมูลจริง ใช้เกณฑ์การจำกัดจำนวนข้อสอบ (Fixed Length) นี้ สำหรับงานวิจัยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) ที่ผ่านมานั้น ส่วนใหญ่จะใช้การจำกัดจำนวนข้อสอบที่ 10, 20 ,30 ,40 ,50 และ 60 ข้อ (Chen, 2009; Diao & Reckase, 2009; Finch, 2010) ซึ่งงานวิจัยของ Diao and Reckase (2009) ปรากฏว่า การใช้ข้อสอบจำนวน 50 ข้อ ความลำเอียงจะน้อยกว่าการใช้ข้อสอบจำนวน 20 ข้อ แต่ยังระบุไม่ได้ว่าจำนวนข้อสอบเท่าใดจึงจะเหมาะสมสำหรับการทดสอบปรับเหมาะแบบพหุมิติ และเกณฑ์

ยุติการทดสอบที่นิยมใช้กันอยู่ในขณะนี้มี 2 ลักษณะ ได้แก่ จำกัดจำนวนข้อสอบ และกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (ศิริชัย กาญจนวาสี, 2555, น. 204) หรืออาจจะใช้ทั้ง 2 เกณฑ์ร่วมกัน ในการวิจัยครั้งนี้ ใช้เกณฑ์การยุติการสอบแต่ละมิติ ด้วยการจำกัดจำนวนข้อสอบ (Fixed length) จำนวน 20 ข้อ หรือกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ที่ SE ≤ 0.30

นิยามศัพท์เฉพาะ

ในการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปเพื่อให้มีความเข้าใจตรงกันผู้วิจัยจึงได้นิยามศัพท์เฉพาะ ดังนี้

1. การทดสอบปรับเหมาะแบบพหุมิติ หมายถึง การผสมผสานแนวคิดของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response theory: MIRT) กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerized adaptive testing: CAT) ซึ่งเป็นการทดสอบที่ลดความยาวของข้อสอบ และเพิ่มความแม่นยำในการประมาณค่า ความสามารถของผู้สอบซึ่งหลักการคล้ายกับการทดสอบปรับเหมาะแบบเอกมิติ โดยมีความแตกต่างเพียงค่า พารามิเตอร์อำนาจจำแนกจะมีหลายค่าตามมิติที่มุ่งวัด ดังนั้นการคำนวณค่าต่างๆ จะใช้ข้อมูล ในลักษณะของเวคเตอร์

2. การทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing: MCAT) หมายถึง การทดสอบที่คัดเลือกข้อสอบจากคลังข้อสอบมาให้มีความเหมาะสมกับความสามารถของผู้สอบ เป็นการผสมผสานแนวคิดระหว่างทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ซึ่งเป็นการทดสอบที่มีประสิทธิภาพสูง มีความแม่นยำ และลดจำนวนข้อสอบ แบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) มี 5 ขั้นตอน ได้แก่ 1) การสร้างคลังข้อสอบ (Item pools) 2) การเริ่มต้นการทดสอบ (Starting point) 3) การคัดเลือกข้อสอบข้อถัดไป (Item selection) 4) การประมาณความสามารถ (Scoring algorithm) และ 5) เกณฑ์การยุติการทดสอบ (Stopping rules)

3. มิติของข้อสอบ หมายถึง การแบ่งข้อสอบในแบบทดสอบฉบับนี้ออกเป็น 5 มิติ โดยให้สอดคล้องกับการทดสอบและประเมินสมรรถนะทางวิชาชีพครูตามประกาศคณะกรรมการคุรุสภาเรื่องหลักเกณฑ์และวิธีการทดสอบ และประเมินสมรรถนะทางวิชาชีพครู พ.ศ. 2563 ซึ่งกำหนดองค์ประกอบของสมรรถนะทางวิชาชีพที่จะทดสอบและประเมินไว้ดังนี้ คือ 1) มิติหลักสูตรและการสอน 2) มิติจิตวิทยา 3) มิติการประกันคุณภาพการศึกษา 4) มิติวัดผลและวิจัยการเรียนรู้ และ 5) มิติการเปลี่ยนแปลงบริบทโลก

4. คลังข้อสอบ (Item bank) หมายถึง ระบบการจัดเก็บข้อสอบที่มีคุณภาพ ซึ่งได้จากการสร้างตามหลักเกณฑ์ที่ถูกต้อง ข้อสอบที่อยู่ในคลังเพื่อเก็บข้อมูลวิจัยในครั้งนี้เป็นข้อสอบที่ผู้วิจัยพัฒนาขึ้นจากการรวบรวมข้อสอบประเมินสมรรถนะทางวิชาชีพที่ทดสอบและประเมินในกลุ่มวิชาชีพครู เป็นข้อสอบเลือกตอบ 5 ตัวเลือกมาวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ด้วยโมเดลปกติสะสมแบบพหุมิติกับโอกาสการเดาข้อสอบ (Multidimensional normal ogive model guessing) โดยใช้โปรแกรม NOHARM 4.0 เมื่อกำหนดค่าพารามิเตอร์โอกาส (c) กำหนดให้ทุกข้อมีค่าคงที่เท่ากับ 0.20 (ถ้าข้อสอบเป็นชนิด 5 ตัวเลือก) แล้วคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ จัดทำเป็นคลังข้อสอบ (Item bank) ได้ข้อสอบทั้งหมด 242 ข้อ จำแนกเป็น 5 มิติ คือ 1) มิติการเปลี่ยนแปลงบริบทโลก จำนวน 50 ข้อ 2) มิติจิตวิทยาการศึกษา จำนวน 40 ข้อ 3) มิติการประกันคุณภาพการศึกษา จำนวน 45 ข้อ 4) มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ และ 5) มิติหลักสูตรและการสอน จำนวน 47 ข้อ

5. วิธีการคัดเลือกข้อสอบข้อถัดไป (Next Item Selection) หมายถึง วิธีการคัดเลือกข้อสอบในคลังข้อสอบที่ยังไม่ถูกเลือกใช้ในการทดสอบ ซึ่งเป็นข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งเป็นขั้นตอนหนึ่งในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

6. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม (Artificial Neural Network Selection: ANN) หมายถึง การคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติด้วยเว็บแอพพลิเคชั่นโดยประยุกต์ใช้หลักการของโครงข่ายประสาทเทียม ซึ่งเป็นศาสตร์แขนงหนึ่งของทางด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) ใช้เทคนิคที่เลียนแบบการทำงานของสมองมนุษย์ ซึ่งประกอบด้วยเซลล์ประสาท (Neuron) และแต่ละเซลล์จะถูกเชื่องโยงกันเป็นโครงข่าย ซึ่งในซอฟต์แวร์ Neuron จะเรียกว่า “โหนด (Node)” และแต่ละโหนดจะถูกแบ่งออกเป็นชั้น (Layer) โดยหลักการของ Deep Learning ก็จะเป็น ANN ที่มีโหนดหลายๆ ชั้น ทำให้สามารถประมวลผลได้ครั้งละจำนวนมาก ช่วยให้การเรียนรู้ของเครื่องสามารถให้ผลลัพธ์ในการตัดสินใจและคาดการณ์ได้แม่นยำมากยิ่งขึ้น

7. ประสิทธิภาพของโมเดลการคัดเลือกข้อสอบข้อถัดไป ด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) หมายถึง การนำโมเดลไปใช้งานจริงได้นั้น จำเป็นจะต้องทราบประสิทธิภาพของโมเดลก่อน โดยทั่วไปแล้วจะมีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่างๆ อยู่ 4 ค่า คือ

7.1 ค่าความแม่นยำ (Precision) หมายถึง ค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์

7.2 ค่าความระลึก (Recall) หมายถึง จำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล

7.3 ค่าความถ่วงดุล (F-measure) หมายถึง ค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก

7.4 ค่าความถูกต้อง (Accuracy) หมายถึง จำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดลโดยพิจารณารวมทุกคลาส

8. เว็บแอพพลิเคชั่น หมายถึง โปรแกรมประยุกต์ที่ใช้การทดสอบสมรรถนะทางวิชาชีพครูแบบปรับเหมาะแบบพหุมิติในแต่ละสมรรถนะความรู้ทางวิชาชีพครู ผู้ทดสอบสามารถเข้าทดสอบได้ทุกที่ทุกเวลา ผ่านคอมพิวเตอร์ โน๊ตบุ๊ก สมาร์ทโฟน แท็บเล็ต มีความครบสมบูรณ์ในตัวเอง สื่อเนื้อหา แบบทดสอบ มีความสัมพันธ์และสอดคล้องกับเนื้อหาการใช้งาน และมีการรายงานผลแบบเรียลไทม์ (Real-Time System) และเผยแพร่ภายใต้สัญญาอนุญาตสิทธิ์แบบ Creative Commons ด้วยเงื่อนไข “CC BY NC SA”

9. การประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ หมายถึง การสอบถามข้อมูลความคิดเห็นของผู้ใช้งานเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติเพื่อประเมินความเหมาะสมของเว็บแอพพลิเคชั่นในด้านต่างๆ ดังนี้ 1) ด้านความสามารถของแอพพลิเคชั่น (Functional Requirement) 2) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) 3) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability) 4) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual)

ประโยชน์ที่ได้รับจากการวิจัย

ประโยชน์ด้านวิชาการ

1. ได้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ที่สามารถลดจำนวนข้อสอบและเวลาที่ใช้ทดสอบลงแต่ยังคงประสิทธิภาพของการทดสอบที่สามารถประเมินระดับความสามารถของผู้สอบได้อย่างถูกต้อง

2. ได้วิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมที่สามารถลดจำนวนข้อสอบและเวลาที่ใช้ทดสอบลงแต่ยังคงประสิทธิภาพของการทดสอบที่สามารถประเมินระดับความสามารถของผู้สอบได้อย่างถูกต้อง ดังนั้นนักวิจัยที่ทำวิจัยเกี่ยวข้องกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถนำวิธีการคัดเลือกข้อสอบข้อถัดไปตามวิธีที่นำเสนอไปประยุกต์ใช้กับการทดสอบเรื่องต่างๆ ได้

3. เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป สามารถนำไปใช้ทดสอบความรู้ทางการศึกษาในรายวิชาใดก็ได้ เพียงแต่ข้อสอบที่ใช้ทดสอบต้องมีลักษณะการตรวจให้คะแนน 2 ค่า คือ ถูก–ผิด หรือ 0–1 และมีรายการคำตอบอย่างน้อย 4 รายการ

4. ได้ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์เข้ากับการทดสอบแบบปรับเหมาะ

ด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) เป็นเทคนิคการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แนวใหม่ที่ทันสมัยทำให้มีประสิทธิภาพยิ่งขึ้นในการคัดเลือกข้อสอบและเป็นวิธีการวิเคราะห์ที่ผ่อนคลายข้อตกลงเบื้องต้น สามารถนำไปประยุกต์ใช้กับข้อมูลที่มีจำนวนน้อยได้ ผลการวิจัยนี้จึงเป็นอีกแนวทางหนึ่งในการวิเคราะห์ข้อมูลที่เป็นประโยชน์และมีคุณค่าทางวิชาการ

ประโยชน์ด้านการนำไปใช้

1. เป็นแนวทางในการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปในระดับชั้นอื่นๆ ต่อไป

2. ได้แนวทางการนำเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป และข้อสอบใน Item bank ผ่านการวิเคราะห์คุณภาพข้อสอบตามแนวคิดทฤษฎี MIRT สามารถนำไปใช้เป็นเครื่องมือในการทำวิจัยต่อไป

3. ได้สารสนเทศเกี่ยวกับการทดสอบและประเมินสมรรถนะทางวิชาชีพครูด้านความรู้ และประสบการณ์วิชาชีพตามมาตรฐานวิชาชีพครู และเป็นการประเมินผลสรุปรวม (summative evaluation) เพื่อเป็นข้อมูลย้อนกลับให้กับหลักสูตร/ สาขาวิชา/ คณะ ใช้ประโยชน์ในการพัฒนาปรับปรุงแก้ไขการจัดการเรียนการสอน อีกทั้งเป็นการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครู

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องตามลำดับดังนี้

1. แนวคิดทฤษฎีเกี่ยวกับการพัฒนาเว็บแอพพลิเคชั่น
2. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)
3. โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional IRT Models)
4. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing)
5. โครงข่ายประสาทเทียม (Artificial Neuron Network)
6. การศึกษาในสถานการณ์จำลอง (Simulation Study)
7. งานวิจัยที่เกี่ยวข้อง
8. กรอบแนวคิดในการวิจัย

แนวคิดทฤษฎีเกี่ยวกับการพัฒนาเว็บแอพพลิเคชั่น

1. ความหมายของเว็บแอปพลิเคชัน

เว็บแอปพลิเคชั่น (Web Application) คือ โปรแกรมประยุกต์ที่ถูกเขียนขึ้นสามารถเข้าใช้งานผ่านทางเว็บบราวเซอร์ (Web Browser) โดยอาศัยโปรโตคอล http(s) ซึ่งทำให้เหมาะสำหรับงานที่ต้องการข้อมูลแบบเรียลไทม์หรือแบบเวลาจริง ข้อมูลต่างๆ ที่อยู่ในระบบหรือบนเว็บแอปพลิเคชัน (Web Application) จึงมีการไหลเวียนแบบออนไลน์ส่งผลให้เกิดความพึงพอใจกับผู้ใช้งาน เพราะสามารถโต้ตอบได้ทันทีกับผู้ใช้บริการ เว็บแอปพลิเคชั่นสามารถเชื่อมต่ออินเทอร์เน็ตที่มีความเร็วต่ำกว่าทำให้ผู้ใช้บริการสามารถใช้โปรแกรมได้จากทุกแห่งทั่วโลก ผ่านคอมพิวเตอร์ สมาร์ทโฟนแท็บเล็ต ผ่านเครือข่ายอินเทอร์เน็ต ตัวอย่างระบบออนไลน์ที่เหมาะกับเว็บแอปพลิเคชัน เช่น ระบบการสั่งซื้อสินค้าออนไลน์ ระบบการจองสินค้าและบริการต่างๆ ระบบงานบุคลากร และระบบการทดสอบ เป็นต้น

2. กระบวนการพัฒนาเว็บแอพพลิเคชัน

ผู้วิจัยได้ประยุกต์ใช้กระบวนการพัฒนาเว็บไซต์แบ่งออกเป็นขั้นตอนต่างๆ จำนวน 4 ขั้นตอน (ธวัชชัย ศรีสุเทพ,2544; ดวงพร เกี๋ยงคำ, 2551) เพื่อให้เหมาะสำหรับใช้เป็นแนวทางในการพัฒนาระบบ โดยมีรายละเอียดดังนี้

2.1 วิเคราะห์และกำหนดความต้องการ (Analysis and Requirement Definition)

การกำหนดเป้าหมายของเว็บไซต์ที่ชัดเจน ศึกษาผู้ใช้เพื่อให้สามารถระบุกลุ่มผู้ใช้ และความต้องการของผู้ใช้ และศึกษาคู่แข่งเพื่อกำหนดกลยุทธ์ในการแข่งขัน

2.2 ออกแบบและพัฒนา (Design and Development)

การจัดทำแผนผังโครงสร้างข้อมูลออกแบบระบบนำทาง ส่วนประกอบต่างๆ ในหน้าเว็บเพจ เพื่อออกแบบการใช้ข้อมูลและเป็นแนวทางในการพัฒนารวมถึงการออกแบบด้านกราฟิกต่างๆ เช่น สี รูปแบบอักษร ภาพโลโก้ ปุ่มเมนู แบนเนอร์โฆษณา เป็นต้น

2.3 สร้างและทดสอบ (Construction and Testing)

การสร้างเว็บเพจโดยพัฒนามาจากขั้นตอนที่ 2 คือ ออกแบบและพัฒนาโดยเนื้อหาหรือข้อมูลต่างๆ จะถูกนำมาใส่และจัดรูปแบบ ระบบนำทางถูกสร้าง การเชื่อมโยงต่างๆ องค์ประกอบเสริมต่างๆ ฯลฯ อย่างไรก็ตาม เมื่อมีการสร้างเว็บเพจจริงอาจมีการปรับแต่งตามความเหมาะสมเพื่อตอบสนองกับความต้องการของผู้ใช้งาน เพื่อให้เว็บเพจที่ออกแบบกลายเป็นเว็บไซต์ที่สมบูรณ์ รวมถึงการออกแบบฐานข้อมูล จดโดเมนและการเตรียมเว็บเซิฟเวอร์ติดตั้งระบบเว็บไซต์ทดสอบระบบกับสภาพแวดล้อมที่ใช้งานจริงเพื่อเผยแพร่บนอินเทอร์เน็ต

2.4 ดูแลและปรับปรุงต่อเนื่อง (Maintenance and Innovation)

นำปัญหาจากการทดสอบระบบกับสภาพแวดล้อมที่ใช้งานจริงมาทำการปรับปรุง แก้ไขข้อผิดพลาดต่างๆ รวมถึงการประเมินผลความพึงพอใจของผู้ใช้งานการกำหนดแนวทางการบำรุงรักษาเว็บไซต์ทั้งเนื้อหาและโครงสร้างข้อมูลให้ทันสมัยอยู่เสมอ

3. เทคโนโลยีที่ใช้ในการสร้างและพัฒนาเว็บแอปพลิชัน (Web Application)

เว็บแอพพลิเคชั่น (Web Application) ในวิศวกรรมซอฟต์แวร์ คือ โปรแกรมประยุกต์ที่เข้าถึงด้วยโปรแกรมค้นดูเว็บผ่านเครือข่ายคอมพิวเตอร์อย่างอินเทอร์เน็ตหรืออินทราเน็ต เว็บแอปพลิเคชันเป็นที่นิยมเนื่องจากความสามารถในการอัปเดต และดูแลโดยไม่ต้องแจกจ่ายและติดตั้งบนเครื่องผู้ใช้ ตัวอย่างเว็บแอปพลิเคชัน ได้แก่ เว็บเมล การพาณิชย์อิเล็กทรอนิกส์ การประมูลออนไลน์ กระดานสนทนา บล็อก วิกิ เป็นต้น Technology ที่ wg-soft ใช้ทำ Web Application ได้แก่ .Net, Java และ PHP การใช้งานคอมพิวเตอร์ส่วนบุคคลที่ติดตั้งโปรแกรม Microsoft office ที่ประกอบด้วย Word สำหรับพิมพ์เอกสาร Excel สำหรับสร้างตารางคำนวณพวกนี้เรียกกันว่า Desktop Application ซึ่งจะติดตั้งบนเครื่องคอมพิวเตอร์ส่วนบุคคลโปรแกรมสำหรับงานบัญชีที่บางหน่วยงานติดตั้งที่เครื่องคอมพิวเตอร์เป็นลักษณะ Client-Server Application เก็บฐานข้อมูลไว้ที่เซิร์ฟเวอร์ (ธีรพล ด่านวิริยะกุล, 2559)

ในการพัฒนาเว็บแอพพลิเคชั่นภาษาที่ถูกเรียกใช้มีหลายภาษา เช่น HTML5,ASP/ASP.Net, PHP และ Java Script เป็นต้น รวมทั้งยังมีโปรแกรมมากมายที่สามารถจะนำไปใช้เป็นเครื่องมือในการพัฒนา เช่น Macromedia Dreamweaver Macromedia Dreamweaver MX และ Microsoft Visual Studio.NET เป็นต้น ผู้พัฒนาระบบเลือกใช้เป็น Microsoft Visual Studio.NET มาใช้เพื่อเป็นเครื่องมือในการพัฒนาและเลือกใช้ PHP เนื่องจากเรียนรู้ง่ายและ PHP ฝังเข้าไปใน HTML และใช้โครงสร้างและไวยากรณ์ภาษาง่ายๆ และใช้ HTML5 ซึ่ง HTML5 นี้เป็นภาษาที่ถูกพัฒนาต่อมาจากภาษา HTML และยังพัฒนาขึ้นมาโดย The Web Hypertext Application Technology Working. Group) โดยได้มีการปรับเพิ่ม Feature หลายๆ อย่างเข้ามาเพื่อให้ผู้พัฒนาสามารถใช้งานได้ง่ายมากยิ่งขึ้น สามารถทำงานควบคู่กับ CSS3 ได้ดี ช่วยให้สามารถเพิ่มลูกเล่นต่างๆ บนเว็บไซต์ได้สวยงามมากยิ่งขึ้น

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีการวัดที่ถูกคิดขึ้นโดยมีวัตถุประสงค์เพื่อเสนอหลักการพื้นฐานที่สำคัญสำหรับการทำนาย ประมาณค่า และสรุปอ้างอิงเกี่ยวกับความสามารถหรือคุณลักษณะที่สามารถวัดได้ด้วยแบบทดสอบ (Hambleton and Swaminathan, 1985, p. 53)

เพื่อให้เกิดความชัดเจนเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยจึงขอเสนอรายละเอียดเป็นด้านๆ ดังนี้

1. ประวัติความเป็นมา

ทฤษฎีการตอบสนองข้อสอบมีประวัติความเป็นมาเรียงตามลำดับดังนี้

ค.ศ. 1916 บิเนท์และไซมอน (Binet and Simon) ได้กำหนดระดับผลการปฏิบัติ (performance levels) ของผู้สอบตามตัวแปรอิสระต่างๆ และใช้ผลการปฏิบัติดังกล่าวไปพัฒนาแบบทดสอบ

ค.ศ. 1936 ริชาร์ดสัน (Richadson) ได้หาความสัมพันธ์ระหว่างค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและทฤษฎีการทดสอบแบบเดิม

ค.ศ. 1943-1944 ลอว์เลย์ (Lowley) ได้คิดวิธีการใหม่ๆ เพื่อใช้สำหรับการประมาณค่าพารามิเตอร์

ค.ศ. 1952 ลอร์ด (Lord) ได้นำเสนอโมเดลนอร์มอลโอไจฟ์ (normal ogive model) แบบ 2 พารามิเตอร์โดยนำวิธีการประมาณค่าพารามิเตอร์และการนำโมเดลไปประยุกต์ใช้

ค.ศ. 1957-1958 เบอร์นบอม (Birnbaum) ได้นำเสนอโมเดลโลจิสติก (Logistic model) เพื่อใช้แทนโมเดลนอร์มอลโอไจฟ์

ค.ศ. ราสช์ (Rasch) ได้พัฒนาโมเดลการตอบสนองข้อสอบ 3 โมเดล โดยพรรณนาไว้ในหนังสือ Probabilistic Model for some intelligence and Attainment Tests ซึ่งผลงานชุดนี้มีอิทธิพลต่อไรท์ (Wright) ซึ่งอยู่ในอเมริกาและนักจิตวิทยาบางคน เช่น แอนเดอร์เซน และฟิชเชอร์ (Andersen and Fisher) ในยุโรป

ค.ศ. 1967 ไรท์ (Wright) ได้เป็นผู้นำและกระตุ้นให้มีการวิจัยในอเมริกาเกี่ยวกับโมเดลของราส์ช (Rasch) ตลอดช่วงทศวรรษ 1970

ค.ศ. 1968 ลอร์ดและโนวิค (Lord and Novick) ได้นำเสนอทฤษฎีคุณลักษณะแฝง (theory of latent trait)

ค.ศ. 1969 ไรท์และปันซาปากีสาน (Wright and Panchapakesan) ได้พรรณนาวิธีการประมาณค่าพารามิเตอร์ในโมเดลของราส์ช โดยใช้โปรแกรมไบคาล (BICAL) และในปีเดียวกันเซมีจีมา (Samejima) ได้พรรณนาโมเดลการตอบสนองข้อสอบโมเดลใหม่พร้อมทั้งการประยุกต์ใช้โมเดลนี้ใช้ได้กับข้อมูลที่ได้จากการตอบหลายแบบและแบบต่อเนื่อง (polychotomous and continuous) และปรับปรุงโมเดลมิติเดียว (unidimensional) ไปสู่โมเดลแบบหลายมิติ (multidimension)

ค.ศ. 1972 บอกซ์ (Bock) ได้สนับสนุนแนวคิดใหม่ๆ หลายแนวคิดเกี่ยวกับการประมาณค่าพารามิเตอร์

ค.ศ. 1974 ลอร์ด (Lord) ได้พรรณนาวิธีการประมาณค่าพารามิเตอร์วิธีใหม่ๆ โดยใช้โปรแกรมโลจิส (LOGIST) และฟิชเชอร์ (Fisher) ได้นำเสนอโปรแกรมการวิจัยเกี่ยวกับโมเดลโลจิสติกเชิงเส้นตรง (linear logistic models)

ค.ศ. 1976 ลอร์ด (Lord) ได้พัฒนาโปรแกรมโลจิสเพื่อให้สามารถประมาณค่าพารามิเตอร์ในโมเดลโลจิสติกได้ดียิ่งขึ้น

ค.ศ. 1977 เบเกอร์ (Baker) นำเสนอผลการทบทวนวิธีการประมาณค่าพารามิเตอร์และได้ตีพิมพ์ประเด็นต่างๆ เป็นกรณีพิเศษเกี่ยวกับการใช้ทฤษฎีการตอบสนองข้อสอบในการวัดผลในวารสาร Journal of Educational Measurement ซึ่งเขียนโดยนักวิจัยหลายคน เช่น บาซอว์ (Bashaw), ลอร์ด (Lord), มาร์โค (Marco), เรนซ์ (Rentz), ยูริ (Urry) และไรท์ (Wright) เป็นต้น

ค.ศ. 1979 ไรท์และสโตน (Wright and Stone) ได้เขียนหนังสือเกี่ยวกับแบบแผนของแบบทดสอบที่ดีที่สุด (best test design) โดยใช้โมเดลของราส์ช (Rasch)

ค.ศ. 1980 ลอร์ด (Lord) ได้ทบทวนพัฒนาการของทฤษฎีการตอบสนองข้อสอบและการประยุกต์ใช้โมเดล 3 พารามิเตอร์ โดยเขียนในหนังสือ Application of Item Response Theory to Practical Testing Problems

ค.ศ. 1980 ไวส์ (Weiss) ได้บรรณาธิกรณ์รายงานการประชุมเกี่ยวกับการดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยคอมพิวเตอร์หรือการทดสอบแบบ ซี เอ ที (Computerized Adaptive Testing: CAT) เมื่อปี 1976

ค.ศ. 1982 ลอร์ดและคณะ (Lord and Staff) ได้ปรับปรุงโปรแกรมโลจิสครั้งที่ 2 เพื่อให้ใช้ได้สะดวกกว่าเดิม

2. คุณสมบัติของรูปแบบการตอบสนองข้อสอบ

แฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan,1985, p. 2) ได้สรุปคุณสมบัติของรูปแบบการตอบสนองข้อสอบไว้ ดังนี้

1. เป็นรูปแบบที่อธิบายเกี่ยวกับปฏิบัติการของผู้สอบในการทำแบบทดสอบ ว่าสามารถอธิบายได้ในคุณลักษณะหนึ่งหรือหลายคุณลักษณะ ซึ่งเรียกว่าคุณลักษณะแฝง

2. รูปแบบการตอบข้อสอบเป็นความสัมพันธ์ระหว่างปฏิบัติการของผู้สอบในการตอบข้อสอบทั้งที่สังเกตได้และคุณลักษณะแฝงหรือความสามารถที่ถือว่าเป็นปฏิบัติการหลักของพฤติกรรมการทำแบบทดสอบ

3. รูปแบบการตอบข้อสอบที่ดีจะต้องให้วิธีทางในการประมาณคะแนนของผู้สอบได้ตรงกับคุณลักษณะแฝงนั้น

4. คุณลักษณะแฝง อธิบายได้จากปฏิบัติการที่สังเกตได้ของผู้สอบในการตอบข้อสอบ ความสัมพันธ์ระหว่างปริมาณของการทำข้อสอบหรือคะแนนของผู้สอบ (Test Performance) กับปริมาณความสามารถของผู้สอบ (Ability) สามารถเขียนในรูปของความสัมพันธ์ทั่วๆ ไปดังนี้ (ตรึงใจ พูลผลอำนวย, 2534, อ้างถึงใน Lord, 1980)

$$P = f(Ui_{1, 2, 3, \ldots, k}; _{j})$$

เมื่อ P แทน ผลการสอบ

f แทน ฟังก์ชัน

Ui แทน ผลการตอบข้อสอบ ตอบถูก Ui = 1 ตอบผิด Ui = 0

$_{1 ,2 ,3, \ldots,k}$ แทน ความสามารถ (Ability หรือ Trait) ที่ 1 ,2 ,3, ... ,k

$_{j}$ แทน ค่าพารามิเตอร์ของข้อสอบข้อที่ j

จากความสัมพันธ์ดังกล่าวนี้เป็นฟังก์ชันความสัมพันธ์ในลักษณะทั่วๆ ไป นักวัดผลการศึกษาจะต้องหาโมเดลทางคณิตศาสตร์ (Mathematical Model) ที่เหมาะสมเพื่อใช้แทนฟังก์ชันความสัมพันธ์ดังกล่าว โดยอาศัยข้อตกลงเบื้องต้นต่างๆ ของทฤษฎีการตอบสนองข้อสอบ

3. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1. ความเป็นมิติเดียว (Unidimensional) หมายถึง ข้อสอบแต่ละข้อในแบบทดสอบจะต้องวัดความสามารถหรือคุณลักษณะเดียวกัน (Ui dimensionality) หรือข้อสอบเหล่านี้มีความเป็นเอกพันธ์ (Homogeneous) การกำหนดเช่นนี้ก็เพื่อให้รูปแบบของทฤษฎีนี้มีความหมายซับซ้อนน้อยลง และง่ายแก่การแปลความหมายของคะแนนที่ได้จากแบบทดสอบ

วิธีการตรวจสอบว่าแบบทดสอบนั้นวัดในมิติเดียวหรือไม่นั้นทำได้หลายวิธี เช่น โดยการวิเคราะห์องค์ประกอบ (Factor Analysis) แล้วสังเกตค่าไอเก้น (Eigen value) ค่าสูงสุดว่าแตกต่างจากค่าอื่นๆ อย่างชัดเจนหรือไม่ (Hambleton and Cook, 1977) ได้เสนอแนะว่าการที่จะทำให้เกิดความเป็นมิติเดียว (Unidimensional) ได้ง่ายวิธีหนึ่งคือ การเลือกข้อสอบที่มีค่าอำนาจจำแนกสูง เนื่องจากข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีสหพันธ์กับคะแนนแบบทดสอบทั้งหมดสูง ยูริ (Urry, 1977) เสนอไว้ว่าควรเลือกข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ 0.8 ขึ้นไป

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง โอกาสในการตอบข้อสอบแต่ละข้อได้ถูกต้องเป็นอิสระจากกัน นั่นคือการตอบข้อสอบข้อใดข้อหนึ่งได้ถูกหรือผิดจะไม่มีผลต่อการตอบข้ออื่นๆ ด้วย (Lord, 1980; Hambleton and Swaminathan,1985) และแฮมเบิลตันกับสวามินาธาน (Hambleton and Swamianthan, 1985) มีความเห็นตรงกันว่าถ้ามีความเป็นมิติเดียวกันแล้ว ความเป็นอิสระในการตอบข้อสอบก็จะมีเอง

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve) เป็นฟังก์ชั่นทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบนั้นได้ถูกต้องกับระดับความสามารถที่วัดได้โดยชุดของข้อสอบ หรือแบบทดสอบนั้นn(Hambleton and Swaminathan, 1985) ซึ่งจะเห็นว่าโอกาสผู้สอบตอบข้อสอบถูก (Probability) จะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) ในแต่ละข้อมีคุณสมบัติไม่แปรเปลี่ยน (Invariant) ไปตามกลุ่มตัวอย่างหรือตัวผู้สอบ (Hambleton and Cook, 1977) ฉะนั้นโอกาสในการตอบข้อสอบถูกในแต่ละข้อจึงไม่แปรเปลี่ยนด้วย

เนื่องจากมีความเชื่อหลายความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบได้ถูกต้องกับระดับความสามารถที่วัดได้โดยใช้ชุดของข้อสอบหรือแบบทดสอบนั้น จึงทำให้เกิดมีรูปแบบของโค้งลักษณะข้อสอบ (Item Response Model) ขึ้นมาหลายรูปแบบ เช่น Guttman Perfect Scale, Latent Distance Model, Linear model, Normal or give Model, Nominal Response Model และ Grade Response Model (จักรกฤษณ์ สำราญใจ, 2531) แต่รูปแบบที่นิยมใช้กันแพร่หลายคือ รูปแบบโลจิสติก (Logistic Model) รูปแบบนี้แบ่งออกเป็น 4 รูปแบบตามจำนวนพารามิเตอร์ของโค้งลักษณะข้อสอบ (Hambleton and Swaminathan, 1985)

4. โมเดลตามทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับปริมาณความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการ (สำเริง บุญเรืองรัตน์, 2529)

$$P = f(\theta)$$

เมื่อ P แทน ผลการสอบ (Performance)

θ แทน ความสามารถหรือคุณลักษณะ (Ability หรือ Trait)

f แทน ฟังก์ชัน (Function)

เนื่องจากข้อมูลที่ได้รับจากผู้สอบนั้นมีหลายลักษณะ ดังนั้นจึงได้มีการพัฒนาโมเดลที่แสดงความสัมพันธ์นี้ในลักษณะต่างกัน เพื่อให้แต่ละโมเดลมีความเหมาะสมกับข้อมูลที่ได้ โมเดลของทฤษฎีการตอบสนองข้อสอบสามารถจำแนกออกได้เป็น 3 กลุ่มใหญ่ๆ ตามลักษณะของข้อมูลที่ได้จากการสอบ 3 ลักษณะคือ คำตอบที่เป็นข้อมูลแบบแบ่งสอง (Dichotomous) ข้อมูลแบบพหุ (Multichotomous) และข้อมูลแบบต่อเนื่อง (Continuous) โมเดลที่มีความเหมาะสมกับลักษณะข้อมูลดังกล่าวมีดังนี้ (รังสรรค์ มณีเล็ก, 2540, อ้างอิงใน Hambleton and Swaminathan, 1985, pp.33 – 52)

ตารางที่ 1 ลักษณะของข้อมูลโมเดลที่ใช้วิเคราะห์ และผู้นำเสนอโมเดลลักษณะของข้อมูลโมเดลที่ใช้วิเคราะห์ ผู้นำเสนอโมเดล

ลักษณะของข้อมูล	โมเดลที่ใช้วิเคราะห์	ผู้นำเสนอโมเดล
Dichotomous	Latent Linear,	Lazarsfeld & Henry (1968)
	Latent Distance	Guttman (1944)
	Perfect scale	Lord (1952)
	One, Two, Three Parameter Normal Ogive	Birnbaum (195, 1958a, 1958b, 1968), Lord & Novick (1968) ,Wright & Stone (1979)
	One, Two, Three Parameter Logistic	McDonald (1957) Barton & Lord (1981)
	Four Parameter Logistic	
Multicategory Scoring	Nominal Response	Bock (1972)
	Graded Response	Samejima (1969)
	Partial Credit Model	Master (1982)
Continuous	Continuous Response	Samejima (1972)

จากตารางที่ 1 ในแต่ละโมเดลก็มีฟังก์ชันทางคณิตศาสตร์แตกต่างกันออกไป ซึ่งสามารถแสดงได้ตามลำดับของการคิดค้นโมเดล ดังนี้

3.1 โมเดลนอร์มอลโอไจฟ์แบบ 2 พารามิเตอร์ (Two Parameter Normal Ogive Model) ลอร์ด (Lord) เป็นผู้เสนอโมเดลนี้ โค้งลักษณะข้อสอบจะได้จากสมการ

$$P_i(\theta) = \int_{a}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{Z^2/2} dz$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถจะตอบข้อที่ i ได้ถูกต้อง

b แทน ค่าความยากของข้อสอบข้อที่ i

a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

Z แทน ความเบี่ยงเบนปกติ จากการแจกแจงที่ค่าเฉลี่ย b

และความเบี่ยงเบนมาตรฐาน 1/a

3.2 โมเดลโลจิสติกแบบ 2 พารามิเตอร์ (Two Parameter Logistic Model) เบอร์นบอม (Birnbaum) เป็นผู้เสนอโมเดลนี้ ซึ่งโค้งลักษณะข้อสอบจะได้จากฟังก์ชันการแจกแจง ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i = 1, 2, 3, ..., n)$$

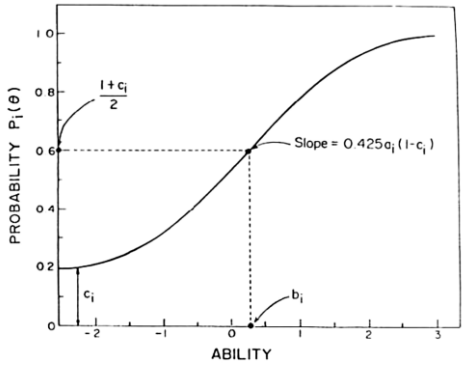
ค่าของ $P_i(\theta)$, a, b และมีความหมายเหมือนเดิม ส่วน D เป็นค่าที่ใช้ในการปรับแก้ (Scaling factor) ซึ่งถ้ากำหนดให้ D = 1.7 แล้วค่าของ $P_i(\theta)$ จากโมเดลนอร์มอลโอไจฟ์และโมเดลโลจิสติกแบบ 2 พารามิเตอร์จะต่างกันน้อยกว่า .01 ในทุกๆ ค่าของที่ประมาณได้ และโมเดลโลจิสติกนั้นนำไปใช้ได้สะดวกกว่าโมเดลนอร์มอลโอไจฟ์ในการประมาณค่า $P(\theta)$

3.3 โมเดลโลจิสติกแบบ 3 พารามิเตอร์ (Three Parameter Logistic Model) โมเดลนี้พัฒนามาจากโมเดลโลจิสติกแบบ 2 พารามิเตอร์ โดยเพิ่มค่าพารามิเตอร์การเดา (c) เข้าไปในฟังก์ชันอีก 1 ค่า ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i = 1, 2, 3, ..., n)$$

พารามิเตอร์ c เป็นค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมากๆ แต่สามารถตอบข้อสอบได้ถูกต้อง ค่าความยากง่ายของข้อสอบ (b) จะอยู่บนสเกลความสามารถ (θ) ตรงจุดที่ค่าความชันของโค้งมีค่าสูงสุด ซึ่งค่าความชันของโค้งที่สูงสุดมีค่าเท่ากับ 0.425(1-c) โค้งลักษณะข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีความชันมากกว่าโค้งลักษณะข้อสอบที่มีค่าอำนาจจำแนกต่ำความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบที่มีความสามารถเท่ากับค่าความยากง่ายของข้อสอบจะเท่ากับ (1 + c)/2 ซึ่งโค้งแสดงลักษณะของข้อสอบตามโมเดลนี้มีลักษณะดังภาพ 1

โอกาสตอบถูก

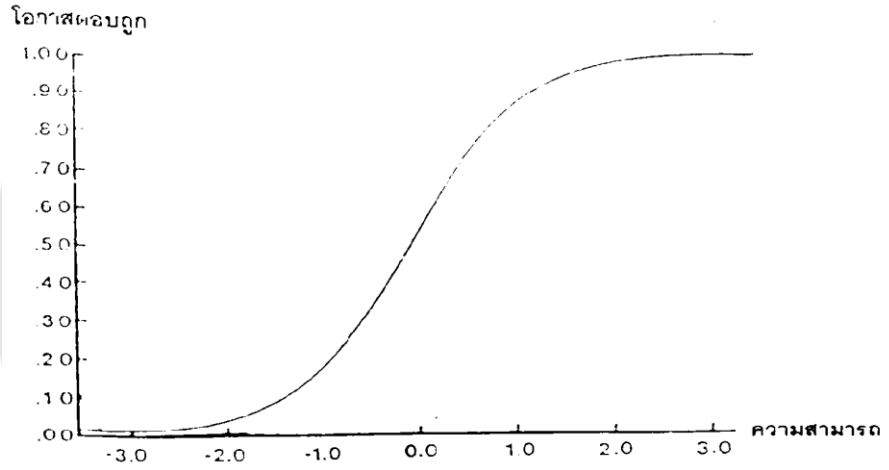


ภาพที่ 1 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 3 พารามิเตอร์

3.4 โมเดลโลจิสติกแบบ 1 พารามิเตอร์หรือราส์ชโมเดล (one parameter Logistic Model, Rasch Model) ราสช์ (Rasch) ซึ่งเป็นนักคณิตศาสตร์ชาวเดนมาร์กเป็นผู้เสนอโมเดลนี้ ซึ่งเป็นโมเดลที่เป็นกรณีเฉพาะของโมเดลโลจิสติก 3 พารามิเตอร์ของเบอร์นบอม (Birnbaum) ที่มีข้อตกลงว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันและมีค่าการเดาน้อยมากๆ ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i=1, 2, 3, ..., n)$$

เมื่อ a แทนค่าอำนาจจำแนกของข้อสอบ ซึ่งกำหนดให้มีค่าเท่ากันทุกข้อ ซึ่งมีโค้งแสดงลักษณะของข้อสอบ ดังแสดงในภาพ 2



ภาพที่ 2 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 1 พารามิเตอร์

3.5 โมเดลโลจิสติกแบบ 4 พารามิเตอร์ (Four Parameter Logistic Model) แมคโดนัล (McDonald) บาร์ตันและลอร์ด (Barton and Lord) เป็นผู้นำเสนอโมเดลนี้โดยมีความเชื่อว่าผู้สอบที่มีความสามารถสูงๆ นั้นบางครั้งอาจตอบข้อสอบผิดได้ด้วยความสะเพร่าหรือการเฉลยข้อสอบที่ผิดของผู้สร้างข้อสอบ ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

$$P_{i(\theta)} = c_i + (\gamma_i - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i = 1, 2, 3, ..., n)$$

เมื่อ γ_i เป็นค่าพารามิเตอร์ความสะเพร่าซึ่งมีค่าต่ำกว่า 1 เล็กน้อย โมเดลนี้เป็นเพียงโมเดลที่ศึกษาตามทฤษฎีเท่านั้น ส่วนการปฏิบัติจริงนั้นยังไม่ค่อยได้นำมาใช้

โมเดลโลจิสแบบ 4 พารามิเตอร์ (Four-Parameter Logistic Model) และโค้งลักษณะข้อสอบ (ICC) สามารถแสดงได้ดังนี้

$$P(X_i = 1 | \theta) = c_i + \frac{d_i - c_i}{1 + e^{-a_i(\theta - b)}}$$

โดย θ = ความสามารถของผู้สอบ

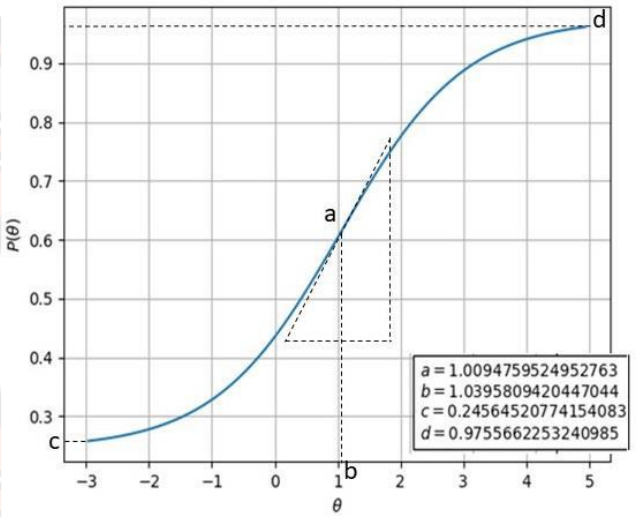
ai = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i (Discrimination Parameter)

bi = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i (Daffily Parameter)

ci = ค่าพารามิเตอร์โอกาสในการเดาข้อสอบถูก (Guessing Parameter)

di = ค่าพารามิเตอร์ความสะเพร่าและอื่นๆ (Upper Asymptote)

e = 2.71828 ค่าคงที่ของลอการิทึมธรรมชาติ (Natural Log)



ภาพที่ 3 เส้นโค้งลักษณะข้อสอบของโมเดลโลจิสแบบ 4 พารามิเตอร์

3.6 โมเดลนอร์มอลโอไจฟ์แบบ 1, 3 และ 4 พารามิเตอร์ (One Three and Four Parameter Normal Ogive Model) ลอร์ด (Lord) เป็นผู้นำเสนอโมเดลนี้ ซึ่งมีฟังก์ชันของการแจกแจงดังนี้

$$P_i(\theta) = \int_{a}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{Z^2/2} dz$$

$$P_i(\theta) = c + (1 - c) \int_{a}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{Z^2/2} dz$$

$$P_i(\theta) = c + (\gamma - c) \int_{a}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{Z^2/2} dz$$

เมื่อเปรียบเทียบกับโมเดลโลจิสติกแล้วพบว่า สมการจะมีลักษณะคล้ายๆ กัน แต่สมการของโมเดลโลจิสติกจะง่ายกว่า ซึ่งในทางปฏิบัติแล้วไม่ค่อยได้ใช้โมเดลนอร์มอลโอไจฝ์เท่าใดนัก เนื่องจากมีข้อจำกัดในเรื่องความยุ่งยากในการประมาณค่า และจะใช้โมเดลโลจิสติกแทนสำหรับโมเดลนอร์มอลโอไจฝ์นั้นเป็นการศึกษาตามทฤษฎีเท่านั้น

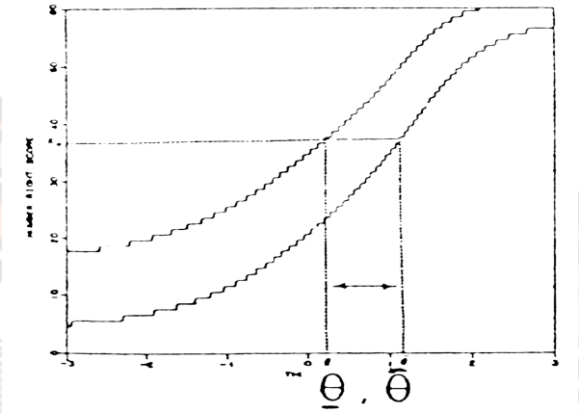
5. ฟังก์ชันสารสนเทศ (Information Function)

ในสถานการณ์ทั่วไป ถ้ามั่นใจค่อนข้างมากว่าเหตุการณ์อย่างหนึ่งจะเกิดขึ้นแสดงว่ามีข่าวสารข้อมูลหรือสารสนเทศเกี่ยวกับเหตุการณ์นั้นมากพอสมควร ในทางกลับกันถ้าไม่มีข่าวสารข้อมูลเกี่ยวกับเหตุการณ์นั้นหรือมีน้อยความมั่นใจก็จะมีน้อยตามไปด้วย ในการอ้างอิงเชิงสถิติความแม่นยำของการประมาณค่าพารามิเตอร์ของกลุ่มประชากรอาจดูได้จากช่วงกว้างของค่าประมาณ ถ้าไม่มีสารสนเทศใดๆ เกี่ยวกับประชากรเลยก็อาจจะต้องประมาณค่าเป็นค่าใดๆ ในช่วง – ถึง + แต่ถ้ามีสารสนเทศเกี่ยวกับประชากรบ้าง ช่วงของค่าประมาณจะแคบเข้านั่นคือความแม่นยำในการประมาณเริ่มมีมากขึ้น ตามปกติความแม่นยำของการประมาณค่าพารามิเตอร์จะแสดงได้ด้วยค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า กล่าวคือ ถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีมากความแม่นยำของการประมาณค่าก็จะมีน้อยเพราะช่วงของค่าประมาณจะกว้าง ในทางกลับกันความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีน้อย ความแม่นยำของการประมาณค่าก็จะมีมาก ช่วงของค่าประมาณจะแคบแสดงว่าค่าสารสนเทศมีความสัมพันธ์กับความคลาดเคลื่อนมาตรฐานของการประมาณค่า โดยมีความสัมพันธ์ในทิศทางกลับกันคือ ความคลาดเคลื่อนมาตรฐานของการประมาณค่าเท่ากับ $1\sqrt{\text{สารสนเทศ}}$ ในทฤษฎีการตอบสนองข้อสอบจะใช้ผลการตอบแบบทดสอบประมาณค่าความสามารถของผู้สอบ การประเมินคุณภาพของแบบทดสอบดูได้จากความถูกต้องแม่นยำในการประมาณค่าความสามารถโดยใช้คะแนนจากแบบทดสอบ ค่าสารสนเทศจากแบบทดสอบจะเป็นดัชนีบ่งชี้ถึงความถูกต้องแม่นยำของการประมาณค่า (จักรกฤษณ์ สำราญใจ, 2531, น. 40, อ้างอิงใน Birnbaum, 1968, p. 418) เนื่องจากฟังก์ชันลักษณะของข้อสอบ $P_i(\theta)$ ขึ้นอยู่กับพารามิเตอร์ของข้อสอบข้อที่ i คือ ai bi และ ci ดังนั้นค่าพารามิเตอร์ของข้อสอบ จึงเป็นตัวกำหนดค่าสารสนเทศของข้อสอบและแบบทดสอบ จึงอาจกล่าวได้ว่าค่าสารสนเทศของข้อสอบหรือแบบทดสอบเป็นดัชนีผสม (Composite Index) ที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลายลักษณะรวมเป็นดัชนีเพียงตัวเดียวเพื่อชี้ถึงคุณภาพของข้อสอบหรือแบบทดสอบ อีกทั้งคุณสมบัติความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบค่าสารสนเทศจึงเหมาะที่จะใช้เป็นดัชนีบอกคุณภาพของข้อสอบหรือแบบทดสอบได้ดีกว่าค่าสถิติหรือค่าดัชนีอื่นๆ ตามแนวคิดของทฤษฎีการวัดแบบมาตรฐานเดิม(จักรกฤษณ์ สำราญใจ, 2531, น. 42)

ทฤษฎีการตอบสนองข้อสอบได้กล่าวถึงฟังก์ชันสารสนเทศไว้ โดยแบ่งออกเป็นฟังก์ชันสารสนเทศของคะแนน (Score Information Function) ฟังก์ชันสารสนเทศของข้อสอบ (Item

Information Function) และฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function) ซึ่งแต่ละเรื่องมีรายละเอียดดังนี้ (รังสรรค์ มณีเล็ก, 2540, อ้างอิงใน Hambleton and Swaminathan, 1985, p. 124)

5.1 ฟังก์ชันสารสนเทศของคะแนน (Score Information Function) เมื่อเรานำคะแนนผลการสอบของผู้สอบแต่ละคนที่มีความสามารถต่างๆ กันมาแจกแจงความถี่และสร้างเป็นแผนภูมิก็จะพบว่า คะแนนในแต่ละระดับความสามารถส่วนใหญ่จะตกอยู่ในช่วง 95% ของการแจกแจงในแต่ละระดับความสามารถ ดังภาพที่ 4



ภาพที่ 4 ช่วงความเชื่อมั่น (confidence interval ; $\underline{\theta}$, $\bar{\theta}$) ของความสามารถที่ได้จากการ ประมาณค่า ($\hat{\theta}$)

เนื่องจากคะแนนนั้นเป็นจำนวนนับที่ไม่ต่อเนื่อง จึงทำให้ภาพที่เห็นเป็นรูปขั้นบันได จากภาพ กล่าวได้ว่าผู้สอบที่ได้คะแนน x_0 จะมีความสามารถอยู่ในช่วง ($\underline{\theta}$, $\bar{\theta}$) ด้วยความเชื่อมั่น 95%

ถ้าให้ z = x/n ดังนั้น การถดถอยของ z ลงบน θ จะได้

$$\mu_{z|\theta} = \frac{1}{n}\sum_{i-1}^{n} \mathrm{P}(\theta) = \xi$$

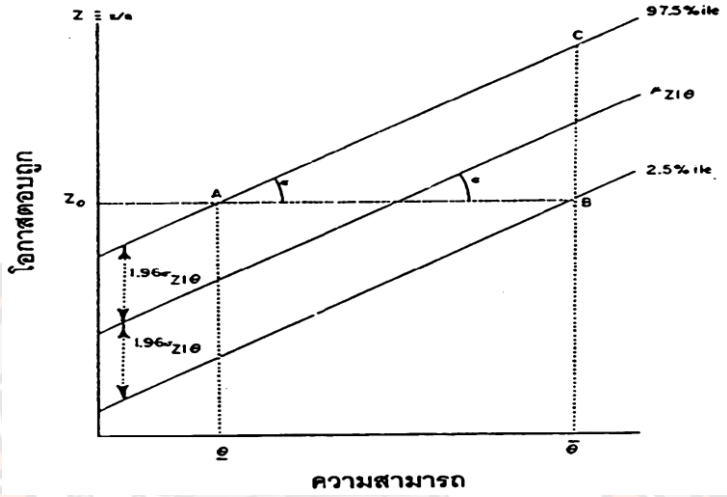
ถึงแม้ว่า n จะมากขึ้นแต่ก็ไม่ทำให้การถดถอยนี้เปลี่ยนไป ความแปรปรวนของ z สำหรับที่คงที่ ก็คือ

$$\sigma^2{}_{z|\theta} = \frac{1}{n^2}\sum_{i-1}^{n} \mathrm{P_i}\, Q_i = (\frac{1}{n})(\overline{PQ} - \sigma^2{}_{p|\theta})$$

ซึ่งเมื่อ n มีค่ามากขึ้น σ^2 ก็จะมีค่าเข้าใกล้ 0

เมื่อนำค่า z มาเขียนเป็นแผนภูมิของการแจกแจง ลักษณะของโค้งการแจกแจงก็จะมีลักษณะเหมือนกับ ภาพประกอบ 13 เมื่อลากเส้นถดถอยของ z ลงบนเส้นถดถอยนี้จะอยู่ระหว่าง

กลางฟังก์ชันแบบขั้นบันได 2 ฟังก์ชัน ซึ่งอยู่ที่ตำแหน่งเปอร์เซ็นไทล์ที่ 2.5 และ 97.5 ของการแจกแจง z เมื่อกำหนดให้และเมื่อ n มีค่าสูงขึ้นการถดถอยก็จะไม่มีการเปลี่ยนแปลงแต่ค่าความแปรปรวนจะลดลงเข้าหา 0 เมื่อฟังก์ชันขั้นบันไดลู่เข้าหากัน จนกระทั่งการแจกแจงของ z เป็นโค้งปกติ ตำแหน่งเปอร์เซนไทล์ที่ 2.5 และ 97.5 จะอยู่ห่างจากเส้นถดถอยประมาณ 1.96 $\sigma_{z|\theta}$ หรือ $\frac{1.96}{n}\sqrt{\sum_{i=1}^{n} P_i Q_i}$ ดังภาพที่ 5



ภาพที่ 5 ช่วงแห่งความเชื่อมั่นแอซิมโทติก 95% ($\underline{\theta}$, $\bar{\theta}$) สำหรับความสามารถ θ

$$\tan-\alpha = \frac{\overline{CB}}{\overline{AB}} = \frac{2(1.96\sigma_{z|\theta})}{\overline{AB}}$$

เมื่อ $\overline{AB}$ แทน ช่วงของความเชื่อมั่นสำหรับความสามารถ θ

(asymptotic confidence interval ; $\underline{\theta}$, $\bar{\theta}$)

$\overline{CB}$ แทน ค่าเบี่ยงเบนมาตรฐานของ Z

ดังนั้น

$$\frac{1}{(\overline{AB})^2} = \frac{\frac{(d\mu_{z|\theta})^2}{d\theta}}{(3.92)^2 Var(Z|\theta)}$$

เมื่อ $\frac{1}{(\overline{AB})^2}$ แทน ฟังก์ชันสารสนเทศของคะแนนใช้สัญลักษณ์ $I\{\theta, y\}$

ซึ่งเบอร์นบอม (ผจงจิต อินทสุวรรณ, 2528, น. 100, อ้างอิงใน Birnbaum, 1968) ได้นิยามฟังก์ชันสารสนเทศของคะแนน y ใดๆ ดังนี้

$$\left|\{\theta, y\} = \frac{\frac{(d\mu_{y|\theta})^2}{d\theta}}{Var(y|\theta)}\right.$$

ค่าสารสนเทศของคะแนนนี้จะแปรเปลี่ยนไปตามระดับความสามารถซึ่งสรุปได้ว่า

(1) ที่ความสามารถระดับเดียวกัน เส้นถดถอยที่มีความชันมากกว่าจะให้สารสนเทศของคะแนนสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า

(2) ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ต่ำจะให้ค่าสารสนเทศของคะแนนสูง

5.2 ฟังก์ชันสารสนเทศของแบบทดสอบ (test information function) ค่าความสามารถโดยประมาณที่เป็นไปได้สูงสุด (maximum likelihood) ในสภาวะปกติจะมีการแจกแจงในลักษณะเอซิมโทติก มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ (รังสรรค์ มณีเล็ก, 2540, อ้างอิงใน Lord, Frederic M., 1980)

$$Var(\hat{\theta}|\theta_0) = \frac{1}{\zeta[(\frac{d\,\mathrm{In}\,L}{d\theta})^2\theta_0]}$$

$$\frac{1}{Var(\hat{\theta}|\theta_0)} = \zeta\{[\sum_{i-1}^{n}(u_i - P_i)\,P_i'/P_i\,Q_i]^2|\theta_0\}$$

$$= \zeta\{[\sum_{i-1}^{n}(u_i - P_i)\,P_i'|P_i\,Q_i][\sum_{j-1}^{n}(u_j - P_j)\,P_j'/P_j\,Q_j]|\theta_0\}$$

$$= \sum_{i-1}^{n}\sum_{j-1}^{n}\frac{P_{i0}'P_{j0}'}{P_{i0}P_{j0}\,Q_{i0}Q_{j0}}\zeta[(u_i - P_i)(u_j - P_j)|\theta_0]$$

$$\frac{1}{Var(\hat{\theta}|\theta_0)} = \sum_{i-1}^{n}\frac{P_i'^2}{P_i^2\,Q_i^2}Var(u_{i0}/\theta_0) = \sum_{i-1}^{n}\frac{P_{i0}'^2}{P_{i0}^2\,Q_{i0}^2}P_{i0}\,Q_{i0}$$

$$Var(\hat{\theta}|\theta_0) = \frac{1}{\sum_{i-1}^{n}\frac{P_i'^2}{P_i\,Q_i}}$$

$$I\{\theta\} = I\{\theta, \hat{\theta}\} = \sum_{i-1}^{n}\frac{P_i'^2}{P_i\,Q_i}$$

เมื่อ L แทน ฟังก์ชันความเป็นไปได้ (likelihood function)

I{θ} แทน ฟังก์ชันสารสนเทศของแบบทดสอบ

$P_i Q_i$ แทน ความน่าจะเป็นในการตอบถูก และตอบผิดตามลำดับ

P_i' แทน ค่าความลาดชันของเส้นโค้ง

หรือเขียนในรูปอื่นได้อีกเช่น

$$I(\theta) = \sum_{i-1}^{n} \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

ลักษณะของฟังก์ชันสารสนเทศของแบบทดสอบ สรุปได้ดังนี้

(1) ฟังก์ชันสารสนเทศของแบบทดสอบได้ถูกกำหนดขึ้นสำหรับชุดของข้อสอบที่แต่ละจุดของสเกลความสามารถ

(2) ค่าของสารสนเทศของแบบทดสอบเป็นผลมาจากคุณภาพและจำนวนของข้อสอบ

(3) ที่ความสามารถระดับเดียวกัน เส้นถดถอยที่มีความชันมากกว่าจะให้ค่าสารสนเทศของแบบทดสอบสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า

(4) ข้อสอบที่มีค่าความแปรปรวนต่ำจะส่งผลให้ค่าสารสนเทศของแบบทดสอบมีค่าสูง

(5) ค่าสารสนเทศของแบบทดสอบจะไม่ขึ้นกับการจัดหมู่เฉพาะของข้อสอบ ข้อสอบแต่ละข้อเป็นอิสระจากกัน

(6) ค่าสารสนเทศของแบบทดสอบมีความสัมพันธ์แบบผกผันกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

5.3 ฟังก์ชันสารสนเทศของข้อสอบ (item information function) ข้อสอบแต่ละข้อจะมีฟังก์ชันแสดงสารสนเทศ ได้แก่ (Lord, 1980; Baker, 1992)

$$I_i\{\theta\} = \frac{P_i'^2}{P_i Q_i}$$

หรือ เขียนในรูปอื่นได้อีก เช่น

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

$$I_i\{\theta\} = D^2 a_i^2 \frac{Q_i}{P_i} \frac{P_i - c_i^2}{1 - c_i}$$

$$I_i\{\theta\} = \frac{D^2 a_i^2 (1 - c_i)}{(c_i + e^{DL_i})(1 - e^{DL_i})^2}$$

เมื่อ $L_i = a_i(\theta - b_i)$

ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อจะรวมกันเป็นฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบจะขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบข้อสอบและค่าความแปรปรวนที่มีเงื่อนไขที่แต่ละระดับของความสามารถ ค่าความชันสูงๆ ค่าความแปรปรวนต่ำๆ จะทำให้ค่าสารสนเทศของข้อสอบมีค่าสูง ซึ่งจะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ การแจกแจงของฟังก์ชันสารสนเทศของข้อสอบมีลักษณะเป็นรูประฆังคว่ำ ค่าสารสนเทศที่สูงสุดจะอยู่ที่จุด b บนสเกลความสามารถสำหรับโมเดลการตอบโลจิสติกแบบ 1 และ 2 พารามิเตอร์ ส่วนโมเดล 3 พารามิเตอร์นั้น ค่าสารสนเทศของข้อสอบข้อที่ I จะสูงสุดที่จุด max

$$\max = b_i + \frac{1}{Da_i} \ln[\frac{1}{2} + \frac{\sqrt{1+8c_i}}{2}]$$

สำหรับโมเดล 1 พารามิเตอร์นั้นค่าสูงสุดของค่าสารสนเทศของข้อสอบจะคงที่ สำหรับโมเดล 2 พารามิเตอร์ ค่าสูงสุดของค่าสารสนเทศของข้อสอบจะเป็นสัดส่วนโดยตรงกับกำลังสองของค่าอำนาจจำแนก ถ้าค่าอำนาจจำแนกของข้อสอบสูงก็จะทำให้ค่าสารสนเทศของข้อสอบมีค่ามาก ส่วนโมเดล 3 พารามิเตอร์นั้น ค่าสารสนเทศของข้อสอบสูงสุดจะมีค่าเท่ากับ (Hambleton and Swaminathan , 1985, p. 105; Lord, 1980, p. 152)

$$I_i\{\theta\}_{max} = \frac{D^2 a_i^2}{8(1-c_i)^2}[1 - 20c_i - 8c_i^2 + (1+8c)^{3/2}]$$

ลักษณะของสารสนเทศของข้อสอบ

(1) ผลรวมค่าสารสนเทศของข้อสอบทุกข้อ คือ ค่าสารสนเทศของแบบทดสอบ

(2) ค่าฟังก์ชันสารสนเทศของข้อสอบจะขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบสนองข้อสอบและค่าความแปรปรวนที่มีเงื่อนไขที่แต่ละระดับของความสามารถ ถ้าค่าความชันมากๆ และค่าความแปรปรวนต่ำๆ จะทำให้ค่าสารสนเทศของข้อสอบมีค่าสูงซึ่งจะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ

6. ประสิทธิภาพสัมพัทธ์ (Relative Efficiency)

หากต้องการเปรียบเทียบแบบทดสอบ 2 ฉบับ ที่วัดคุณลักษณะเดียวกันว่าฉบับใดจะมีประสิทธิภาพสูงกว่ากันก็สามารถทำได้โดยการใช้ประสิทธิภาพสัมพัทธ์ ซึ่งค่าประสิทธิภาพสัมพัทธ์ของคะแนนจากแบบทดสอบ y กับคะแนนจากแบบทดสอบ x เป็นอัตราส่วนของฟังก์ชันสารสนเทศของคะแนน ดังนี้ (Lord, 1980)

$$RE\{y,x\} = \frac{I\{\theta,y\}}{I\{\theta,x\}}$$

ถ้า $RE\{y, x\} = 1$ แสดงว่า แบบทดสอบทั้งสองฉบับมีความแม่นยำในการประมาณค่าความสามารถเท่าๆ กัน

$RE_{y,x} > 1$ แสดงว่า แบบทดสอบ Y มีความแม่นยำในการประมาณค่าความสามารถมากกว่าแบบทดสอบ X

$RE\{y, x\} < 1$ แสดงว่า แบบทดสอบ Y มีความแม่นยำในการประมาณค่าความสามารถน้อยกว่าแบบทดสอบ X

คะแนน x และ y อาจเป็นคะแนนจากแบบทดสอบสองฉบับที่ต่างกันแต่วัดความสามารถเดียวกันหรืออาจเป็นคะแนนจากแบบทดสอบฉบับเดียวกัน แต่ให้คะแนนด้วยวิธีที่ต่างกัน ค่าประสิทธิภาพสัมพัทธ์ของคะแนนจะแปรเปลี่ยนไปในแต่ละระดับความสามารถ ซึ่งสามารถเขียนให้อยู่ในรูปของการแจกแจงคะแนนจริงได้ดังนี้

$$RE\{y,x\} = \frac{(\sigma^2_{x|\xi})P^2(\xi)}{\sigma^2_{y|\eta(\xi)}q^2[\eta(\xi)]}$$

ถ้าใช้คะแนนจากการสอบแทนคะแนนจริง จะได้

$$RE\{y,x\} = \frac{n_y x(n_x - x) f_x^2}{n_x y(n_y - y) f_y^2}$$

เมื่อ p2(ξ), q2[η (ξ)] แทน การแจกแจงความถี่ของคะแนนจริง และ η ตามลำดับ

$\sigma^2_{x|\xi}$ แทน ค่าความแปรปรวนของคะแนนจากแบบทดสอบ X และ y ตามลำดับ

x , y แทน คะแนนที่ตอบถูกจากแบบทดสอบ x และ y ที่ตำแหน่งเปอร์เซนไทล์เดียวกัน

f_x , f_y แทน ความถี่สัมพัทธ์ระหว่างกลุ่ม x และ y ตามลำดับ

n_x , n_y แทน จำนวนข้อสอบในแบบทดสอบ x และ y ตามลำดับ

7. การประมาณค่าความสามารถของผู้สอบ

ทฤษฎีการตอบสนองข้อสอบมีความเชื่อว่าสิ่งที่อยู่เบื้องหลังและส่งผลต่อการทดสอบก็คือความสามารถของผู้สอบ (Ability) ซึ่งอาจเป็นความสามารถเพียงอย่างเดียวหรือหลายๆ อย่างต่อผลการสอบในครั้งหนึ่งๆ ก็ได้ ความสามารถดังกล่าวถูกกำหนดขึ้นเพื่อใช้ในการระบุคุณลักษณะภายใน (Trait หรือ Characteristic) ที่ต้องการวัดจากผู้สอบความสามารถนั้นเป็นสิ่งที่ถูกพรรณนาว่า

เป็นสิ่งที่วัดได้ด้วยชุดของข้อสอบ อาจนิยามความสามารถอย่างกว้างๆ ว่าเป็นตัวแปรด้านความถนัดหรือผลสัมฤทธิ์ทางการเรียน เช่น ความเข้าใจในการอ่าน ความสามารถทางจำนวน หรืออาจนิยามในวงแคบ เช่น ความสามารถในการคูณจำนวน หรือตัวแปรที่เกี่ยวกับบุคลิกภาพ เช่น สังกัปแห่งตน แรงจูงใจ ความสามารถเหล่านี้มิใช่เป็นสิ่งที่มีอยู่ตายตัว แต่สามารถเปลี่ยนแปลงได้ตลอดเวลา โดยเฉพาะเมื่อได้รับการเรียนการสอนเพิ่มขึ้นในการแปลความหมายของความสามารถนั้นก็จะใช้ผลที่ได้จากการศึกษาความเที่ยงตรงเชิงโครงสร้าง (Hambleton and Swaminathan, 1985, p. 55) ความสามารถ (Ability) และคะแนนจริง (True Score) นั้นเป็นสิ่งเดียวกันแต่ใช้มาตราวัดต่างกัน มาตราวัดคะแนนจริงขึ้นอยู่กับข้อสอบทั้งหลายในแบบทดสอบ ส่วนมาตราวัดความสามารถนั้นจะเป็นอิสระจากข้อสอบทั้งหลายในแบบทดสอบ ดังนั้นค่าของความสามารถจึงมีประโยชน์มากกว่าคะแนนจริงเมื่อต้องการเปรียบเทียบแบบทดสอบต่างฉบับที่วัดความสามารถเดียวกัน ลอร์ดและโนวิค (Lord and Novick, 1980, pp. 45–46) กล่าวว่า คะแนนจริงเป็นค่าที่คาดหวังของคะแนนที่วัดได้ ผู้สอบทุกคนที่มีความสามารถเท่ากับ จะมีคะแนนจริงของจำนวนคำตอบถูกเท่ากัน คือ

$$\sum_{i-1}^{n} P_i(\theta)$$

ขั้นตอนในการประมาณค่าความสามารถของผู้สอบนั้นมี ดังนี้

(1) รวบรวมข้อมูลการตอบของผู้สอบ

(2) เปรียบเทียบความเหมาะ (fit) ระหว่างโมเดลการวิเคราะห์ข้อมูลกับข้อมูลที่ได้มาและเลือกใช้โมเดลสำหรับวิเคราะห์ข้อมูลที่เหมาะสม

(3) ประมาณค่าพารามิเตอร์ต่างๆ

(4) กำหนดสเกลและแปลงค่าความสามารถให้อยู่ในรูปของสเกลง่ายๆ

ในการประมาณค่าความสามารถของผู้สอบนั้นมีอยู่หลายวิธีด้วยกัน แต่พอจะจำแนกออกเป็น 2 กลุ่มใหญ่ๆ ได้แก่ วิธีความเป็นไปได้สูงสุด (maximum likelihood) และวิธีของเบส์ (Bayesian)

การประมาณค่าความสามารถของผู้สอบ ถ้าผู้สอบตอบข้อสอบถูกทุกข้อหรือผิดหมดทุกข้อ การประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุดก็จะไม่เหมาะสม เว้นแต่ว่าจะนำเอาผลการตอบของผู้สอบเหล่านั้นออกไปจากกลุ่มผู้สอบก่อน หากทราบค่าสารสนเทศภายแรก (Prior Information) การประมาณค่าความสามารถด้วยวิธีของเบส์ก็จะเกิดประสิทธิภาพมากกว่าค่าสารสนเทศภายแรกที่มักนิยมใช้กันก็คือความสามารถของผู้สอบที่มีการแจกแจงเป็นรูปโค้งปกติมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 (Hambleton and Swaminathan, 1985, p. 92)

การกำหนดรูปร่างของการแจกแจงเป็นรูปโค้งปกตินับว่าค่อนข้างสะดวก เบอร์นบวม (Birnbaum) เชื่อว่าการแจกแจงภายแรกของความสามารถเป็นฟังก์ชันโลจิสติก (logistic density function) ซึ่งมีสมการ ดังนี้

$$f(\theta) = \frac{\exp.(\theta)}{[1+\exp.(\theta)]^2}$$

หัวใจของวิธีการประมาณค่าความสามารถตามวิธีการของเบส์ คือ ทฤษฎีของเบส์ (Bayes theorem) ซึ่งมีความสัมพันธ์กับ condition probability และ marginal probability

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)}$$

ในการประมาณค่าความสามารถของผู้สอบ A อาจถูกกำหนดให้เป็น a และ B เป็นชุดของคำตอบของข้อสอบ n ข้อ ได้คะแนน u แล้วสมการ $P(B|A)$ เขียนใหม่ได้เป็น

$$P(\theta_a|u) = \frac{P(u|\theta_a)\,P(\theta_a)}{P(u)}$$

เนื่องจาก θa เป็นตัวแปรต่อเนื่อง ดังนั้นจึงถูกแปลความว่าเป็นฟังก์ชันความหนาแน่น (density function) เพื่อเป็นการป้องกันการจำสับสนระหว่าง P(θa) กับ P(θ) ซึ่งเป็นฟังก์ชันการตอบสนองข้อสอบ (item response function) จึงใช้ f(θa) แทน P(θ) ดังนั้น จึงเขียนสมการใหม่ได้ ดังนี้

$$f(\theta_a|u) = \frac{f(u|\theta_a)\,f(\theta_a)}{f(u)}$$

เมื่อ f(u) แทน ค่าคงที่

f (θ_a u)| แทน ความหนาแน่นภายหลัง (posterior density) ของ θ_a

f(a) แทน ความหนาแน่นภายแรก (prior density) ของ θ_a

แต่เนื่องจาก f (lu θ_a) เป็นฟังก์ชันความเป็นไปได้ (likelihood function) ของคะแนน ดังนั้นจึงสามารถเขียนสมการ f (u θ_a) ให้อยู่ในรูปของการแปรผันได้ ดังนี้

f (θ_a ul) $_{L(ua)f()}$ หรือ Posterior Likelihood x Prior

วิธีการประมาณค่าความสามารถของผู้สอบตามวิธีการของเบส์นี้ ได้ถูกนำไปประยุกต์เป็นวิธีต่างๆ อีกหลายวิธี แต่วิธีที่ได้รับความนิยมมากคือ วิธีของเบส์ที่ปรับใหม่ (Bayesian Updating) ซึ่งเสนอโดยโอเวน (อำนาจ เกษศรีไพร, 2549, อ้างอิงใน Owen, 1975) ซึ่งมีวิธีการประมาณค่า ดังนี้

ในกรณีที่ตอบข้อสอบถูก

$$\theta_{m+1} = \theta_m + (1-c)\left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2}+\sigma_m^2}}\right)\left(\frac{O(D)}{B}\right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 + \left[1 - \left(\frac{1-c}{1+\frac{1}{a^2\sigma_m^2}}\right)\left(\frac{O(D)}{B}\right)\left(\frac{(1-c)O(D)}{B}\right) - D\right]$$

$$D = \frac{(b-\theta)}{\sqrt{\frac{1}{a^2+\sigma^2}}}$$

เมื่อ และ $B = c + (1-c)A(-D)$

เมื่อ θ_m แทน ค่าความสามารถของผู้สอบที่ประมาณได้ก่อนตอบข้อสอบข้อที่ $m+1$ ตามปกติแล้วถ้าไม่ทราบความสามารถเบื้องต้นของผู้สอบจะกำหนดให้ θ_m = 0.000

θ_{m+1} แทน ความสามารถของผู้สอบที่ประมาณค่าได้หลังจากตอบข้อสอบข้อที่ $m+1$ แล้ว

σ_m^2 แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อสอบข้อที่ $m+1$ ตามปกติแล้วถ้าไม่ทราบความแปรปรวนดังกล่าวมาก่อน จะกำหนดให้ σ_m^2 = 1.000

σ_{m+1}^2 แทน ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเมื่อตอบข้อสอบข้อที่ $m+1$ แล้ว

a_i แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ $m+1$

b_i แทน ค่าความยากของข้อสอบข้อที่ $m+1$

c_i แทน ค่าการเดาของข้อสอบข้อที่ $m+1$

D แทน จุดบนแกน x

$O(D)$ แทน ค่าออร์ดิเนท (Ordinate) ของโค้งปกติที่จุด D

$A(D)$ แทน พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุดถึงจุด D

ในกรณีที่ตอบผิด

$$\theta_{m+1} = \theta_m - \left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2} + \sigma_m^2}} \right) \left(\frac{O(D)}{A(D)} \right)$$

$$\sigma_{m+1}^2 = \sigma_m^2 + \left[1 - \frac{\left(\frac{O(D)}{1 + \frac{1}{a^2 \sigma_m^2}} \right) \left(\frac{O(D)}{A(D)} + D \right)}{A(D)} \right]$$

ตัวอย่างการประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ เมื่อกำหนดให้ $\theta_m = 0.0000$ และ $\sigma_m^2 = 1.0000$ ข้อสอบทั้งหมดมีค่าอำนาจจำแนกเท่ากับ 1.00 ค่าการเดาเท่ากับ .20 (ในแต่ละขั้นของการทดสอบ การคัดเลือกข้อความยากของข้อสอบจะถูกต้องโดยแท้จริงกับระดับความสามารถของผู้สอบที่ประมาณค่าแรกเริ่มนั่น คือ $b_i - \theta_m = 0$) จากตัวอย่างการตอบของข้อสอบจำนวน 6 ข้อ มีผลการตอบเป็น ถูก, ผิด, ถูก, ถูก, ผิด , ผิด, สามารถนำมาประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถได้จากการแทนค่าในสูตร ดังนี้

$$= 0.376$$

$$\sigma_1^2 = 1.0 \left\{ 1.0 - \left(\frac{1.0 - 0.20}{1.0 + \frac{1.0}{1.0}} \right) \left(\frac{0.3989}{0.20 + (0.80)(0.50)} \right) \left(\frac{(0.80)(0.3989)}{0.20 + (0.8)(0.5)} \right) \right\}$$

$$= 0.859$$

จากผลการคำนวณที่ผู้สอบตอบถูกในข้อแรก สามารถประมาณค่าความสามารถของผู้สอบได้เท่ากับ 0.376 ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถเท่ากับ $\sqrt{0.859}$ นั่นคือเท่ากับ 0.927 และเมื่อผู้สอบตอบข้อสอบข้อที่ 2 ผิด สามารถนำมาประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ ได้ดังนี้

$$\theta_1 = 0.376 - \left(\frac{0.859}{\sqrt{1.0+0.859}}\right)\left(\frac{0.3989}{0.50}\right)$$

$$= -0.127$$

$$\sigma_1^2 = 0.859\left\{1.0 - \left(\frac{0.3989}{1.0+\frac{1.0}{0.859}}\right)\left(\frac{0.3989}{0.25}\right)\right\}$$

$$= 0.606$$

และจากผลการประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถของผู้สอบ จำนวน 6 ข้อ แสดงผลการตอบและการประมาณค่าได้ ดังตารางที่ 2

ตารางที่ 2 ผลการตอบและการประมาณค่าความสามารถ

ข้อสอบ	ผลการตอบ	ค่าประมาณความสามารถ (θ)	ค่าความแปรปรวน (σ^2)	ค่าความคลาดเคลื่อนมาตรฐาน (σ)
1	ถูก	0.375	0.859	0.927
2	ผิด	-0.127	0.606	0.778
3	ถูก	0.127	0.541	0.736
4	ถูก	0.359	0.487	0.698
5	ผิด	0.040	0.385	0.620
6	ผิด	-0.221	0.317	0.563

จาก *Applied Psychometrics Boston: Houghton Mifflin Company*, Thorndike, R.L., 1982, p. 305.

จากตารางที่ 2 แสดงให้เห็นว่ากระบวนการในการคัดเลือกข้อสอบข้อใหม่ให้มีค่าความยากใกล้เคียงที่สุดกับความสามารถที่ประมาณค่าได้ใหม่ มีผลทำให้ค่าความสามารถที่ประมาณค่าได้ใหม่มีค่าเปลี่ยนแปลงไป กล่าวคือ ถ้าตอบข้อสอบถูกค่าความสามารถที่ประมาณได้ใหม่จะมีค่าสูงขึ้นและถ้าตอบผิดค่าความสามารถที่ประมาณได้ใหม่จะมีค่าลดลง ในขณะที่ความคลาดเคลื่อน มาตรฐานของการประมาณค่าจะมีค่าลดลงเรื่อยๆ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถลดลงถึงจุดที่กำหนดไว้ คือ .30 กระบวนการทดสอบจะยุติ

การประมาณค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ (เกณฑ์ยุติการทดสอบ) เนื่องจากการวิจัยครั้งนี้ใช้วิธีการประมาณค่าความสามารถของผู้สอบโดย ใช้วิธีของเบส์

ฉะนั้น การประมาณค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ จึงใช้ค่า σ_m^2 ที่คำนวณได้ในการตอบข้อสอบแต่ละข้อมาถอดรากที่สองแล้วเทียบกับเกณฑ์ การวิจัยครั้งนี้กำหนดให้ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.3 หมายความว่า ถ้าค่าการถอดรากที่สองของความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.3 โปรแกรมจะยุติการทดสอบครั้งนั้นๆ

โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional IRT Models)

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายในหรือความสามารถที่มีอยู่ในตัวบุคคล (Latent Trait Theory) กับผลการตอบข้อสอบหรือข้อคำถามโดยใช้โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าความน่าจะเป็นในการเดาข้อสอบได้ถูกต้อง (c) โดยอยู่บนฐานความคิดที่สำคัญ 2 ประการ คือ 1) ผลการตอบข้อสอบหรือข้อคาถามของผู้สอบ สามารถอธิบายได้ด้วยความสามารถที่มีอยู่ภายในของผู้สอบmและ 2) ความสัมพันธ์ระหว่างผลการตอบข้อสอบหรือข้อคำถามกับความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วยฟังก์ชันลักษณะข้อสอบหรือโค้งลักษณะข้อสอบอันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์เรียกว่าฟังก์ชันโลจิส (Logistic Function) หรือใกล้เคียงกับฟังก์ชันปกติสะสม (Normal Ogive Function) (ศิริชัย กาญจนวาสี, 2555, น. 53)

ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้นที่สำคัญ ได้แก่ 1) แบบทดสอบมีลักษณะเป็นมิติเดียว (Unidimensional Test) กล่าวคือ แบบทดสอบที่ใช้ในการวัดความสามารถจะต้องเป็นแบบทดสอบที่ใช้วัดความสามารถด้านเดียว หรือมีความเป็นเอกพันธ์ (Homogeneous Items) ทั้งนี้เพื่อให้ง่ายต่อการรวมคะแนนและการแปลความหมายของคะแนน 2) ความเป็นอิสระข้อสอบ (Local Independence) กล่าวคือ ข้อสอบแต่ละข้อเป็นอิสระไม่เกี่ยวข้องกัน ซึ่งความน่าจะเป็นในการที่ผู้สอบจะทาข้อสอบข้อหนึ่งถูกนั้น ไม่เกี่ยวข้องกับความน่าจะเป็นในการตอบข้อสอบข้ออื่นถูกและข้อสอบแต่ละข้อจะปรากฏอยู่ในตำแหน่งใดของแบบทดสอบก็ได้ ซึ่งจะไม่มีผลต่อการตอบของผู้สอบ 3) การรู้คำตอบที่ถูกต้อง (Know-Correct Assumption) กล่าวคือ ในการตอบข้อสอบข้อใดนั้นผู้สอบต้องรู้คำตอบที่ถูกต้องจึงจะตอบข้อสอบข้อนั้นถูกแต่ถ้าผู้สอบตอบข้อสอบใดผิด แสดงว่าผู้สอบไม่ทราบคำตอบที่ถูกต้องของข้อสอบนั้น และ 4) โค้งคุณลักษณะของ หมายถึง โอกาสที่ผู้สอบผู้หนึ่งจะตอบข้อสอบข้อหนึ่งๆ ถูกนั้น ขึ้นอยู่กับโค้งคุณลักษณะข้อสอบของแต่ละรูปแบบที่ใช้ ไม่ได้ขึ้นอยู่กับการแจกแจงของกลุ่มประชากร หมายความว่ารูปร่างของเส้นโค้งคุณลักษณะข้อสอบไม่ได้ขึ้นอยู่กับการกระจายความสามารถของกลุ่มผู้สอบ จึงเห็นได้ว่าลักษณะของข้อคำถามจะไม่ขึ้นอยู่กับลักษณะของกลุ่มผู้สอบ จึงเป็นข้อคาถามที่มีคุณสมบัติคงที่ตลอดเวลา ดังนั้นจากข้อตกลงเบื้องต้นดังกล่าว

แบบทดสอบจะต้องมีมิติเดียว ซึ่งเป็นโมเดลการตอบสนองข้อสอบแบบเอกมิติ หากใช้กระบวนการวัดและการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ โดยมีความเชื่อว่าข้อสอบที่สร้างขึ้นนั้น วัดมิติเชิงเดี่ยว (Single Dimension) จะทำให้ฝ่าฝืนข้อตกลงเบื้องต้นของทฤษฎีได้ง่าย และก่อให้เกิดประเด็นปัญหา คือ 1) ข้อสอบวัดความสามารถที่ต้องการจริงหรือไม่ 2) ค่าสถิติที่ได้จากการวัดไม่สามารถบ่งบอกความแตกต่างของมิติความสามารถที่ต้องการวัดได้ (Miller & Hirsch, 1992) ดังนั้นนักวัดผลจึงได้พยายามคิดทฤษฎีการทดสอบขึ้นใหม่เพื่อแก้ปัญหาที่เกิดขึ้น เรียกว่า “ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ” (Multidimensional Item Response Theory: MIRT)

ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ เป็นการพัฒนาแนวคิดมาจากหลายสาขาวิชา เช่น การศึกษา จิตวิทยา สถิติศาสตร์ และการแพทย์ เป็นต้น เพื่อสนองต่อกระบวนการทางการศึกษาของแต่ละสาขาวิชาในปัจจุบันที่มีความซับซ้อนมากขึ้น และเนื่องจากข้อจากัดของโมเดลการตอบสนองข้อสอบแบบเอกมิติ (UIRT) ในกรณีที่ผู้สอบใช้ความสารถมากกว่า 1 ในการตอบข้อคำถามและในกรณีที่ข้อคำถามต้องการสอบวัดความสามารถจำนวนมาก เพื่อกำหนดคำตอบหรือแนวทางแก้ไขที่ถูกต้อง แนวคิดของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติเกิดมาจากแนวคิดพื้นฐานสำคัญจาก 2 กลุ่มแนวคิด โดยกลุ่มแรกเป็นการศึกษาโดย Spearman ในปี ค.ศ. 1927 และ Thurstone ในปี ค.ศ.1947 ที่พบว่าโมเดลการตอบสนองข้อสอบแบบพหุมิติมีที่มาจากการวิเคราะห์องค์ประกอบ (Factor Analysis) เนื่องจากลักษณะของการวิเคราะห์องค์ประกอบมีความสัมพันธ์โดยตรงกับการวิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติ ส่วนอีกกลุ่มหนึ่งเป็นการศึกษาของ Lazarsfeld ในปี ค.ศ. 1950 กล่าวว่าโมเดลการตอบสนองข้อสอบแบบพหุมิติได้รับอิทธิพลมาจากทฤษฎีการตอบสนองข้อสอบแบบเอกมิติเช่นเดียวกับการวิเคราะห์องค์ประกอบ โดยทฤษฎีการตอบสนองข้อสอบแบบพหุมิติมีสองประเด็นที่เป็นอิทธิพลของแนวคิดนี้คือ 1) ความเข้าใจในการเพิ่มเนื้อหาทำให้มีความชัดเจนยิ่งขึ้นว่าสิ่งที่วัดมีความซับซ้อนมากกว่าเดิม 2) ความซับซ้อนสามารถแทนได้ด้วยโมเดลหรือทฤษฎี (Reckase, 2009, pp. 63-64)

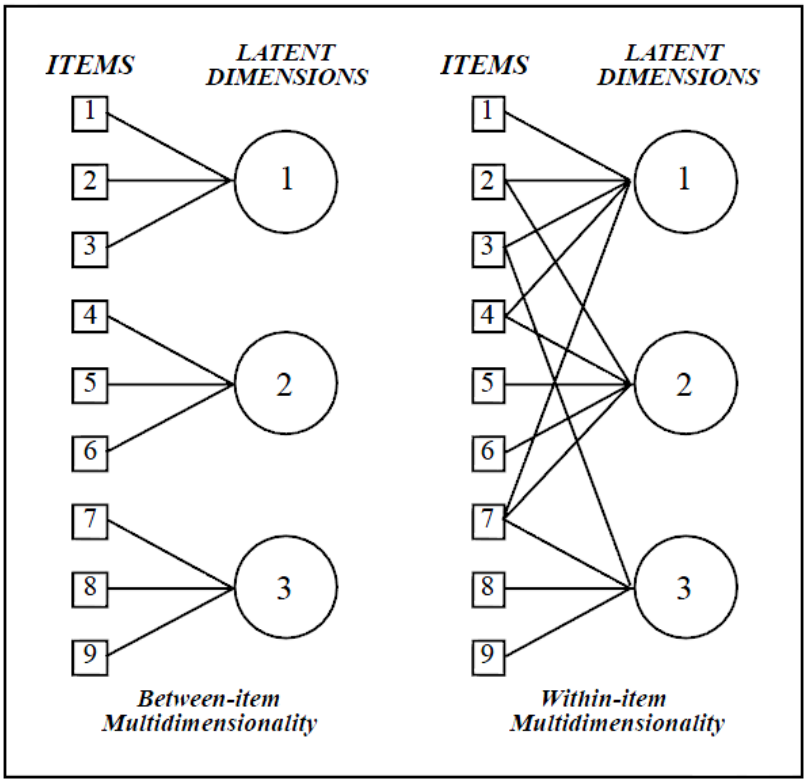
ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติมีข้อตกลงเบื้องต้นที่สอดคล้องกับโมเดลการตอบสนองข้อสอบแบบพหุมิติที่เกี่ยวกับคุณลักษณะของผู้สอบ และคุณลักษณะของข้อสอบ ดังนี้

1. การเพิ่มขึ้นทางเดียว (Monotonicity Assumption) เป็นข้อตกลงเบื้องต้นที่กำหนดว่าเมื่อโครงสร้างตามสมมติฐานที่ต้องการวัดเพิ่มขึ้น ความน่าจะเป็นในการตอบข้อสอบถูกต้องก็น่าจะมีเพิ่มขึ้นด้วย กล่าวคือ ความน่าจะเป็นของการเลือกตอบข้อสอบที่ถูกต้องจะให้คะแนนเป็นถูกหรือไม่ถูกเพิ่มขึ้นหากสมาชิกของ θ-vector เพิ่มขึ้น

2. การนิยามฟังก์ชันความน่าจะเป็นในการตอบข้อสอบ ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบ โดยใช้โครงสร้างตามสมมติฐานสอดคล้องกับฟังก์ชันที่ได้นิยามขึ้น

3. ความเป็นอิสระในการตอบข้อสอบ (Local Independence Assumption) เป็นข้อตกลงเบื้องต้นที่กาหนดว่าความน่าจะเป็นในการตอบข้อสอบถูกหลายข้อได้ถูกต้อง พิจารณาจากผลคูณของความน่าจะเป็นในการตอบข้อสอบถูกหลายข้อ เมื่อความน่าจะเป็นคำนวณอย่างมีเงื่อนไขบนจุดในสเปซที่นิยามโดยสมมติฐาน คือการตอบข้อสอบของบุคคลหนึ่งไม่มีผลมาจากอีกบุคคลหนึ่ง

4. คุณลักษณะแฝงที่ต้องการวัดมากกว่า 1 องค์ประกอบ และแต่ละคุณลักษณะแฝงมีความสัมพันธ์กันชนิดของแบบสอบพหุมิติที่มีความแตกต่างกัน 2 ชนิด (Wilson & Hoskens, 2005; Wang et al., 2011; Wu et al., 2007; Baghaei, 2012) ได้แก่ 1) ความเป็นพหุมิติระหว่างข้อคำถาม (Between-items Multidimensionality) มีลักษณะของข้อคำถามหนึ่งข้อจะวัดคุณลักษณะแฝงเดียว ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน และ 2) ความเป็นพหุมิติภายในข้อคำถาม (Within-items Multidimensionality) ลักษณะของข้อคำถามหนึ่งข้อจะวัดได้หลายคุณลักษณะแฝง (มากกว่าหนึ่ง) ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน ดังภาพที่ 6



ภาพที่ 6 พหุมิติระหว่างข้อคำถามและพหุมิติภายในข้อคำถาม, โดย (Wu et al., 2007)

จากการศึกษาแนวคิดของ IRT ในระยะแรกอยู่บนข้อตกลงเบื้องต้นของพารามิเตอร์ที่อธิบายของผู้สอบในมิติความสามารถเพียงมิติเดียว ลอร์ด โนวิค และ ราสช์ (Reckase, 2009, อ้างอิงใน Lord and Novick, 1968; Rasch, 1960) ซึ่งง่ายต่อการฝ่าฝืนข้อตกลงเบื้องต้นและส่งผลต่อการ

ประมาณค่าพารามิเตอร์ (Camilli, G., Wang, M.-m., & Fesq, J., 1995) ดังนั้น การศึกษาในช่วงแรกจึงพิจารณาลักษณะการวัดแบบพหุมิติในมุมมองของ IRT ดังที่จะนำเสนอต่อไปนี้

Rasch (Reckase, 2009, อ้างอิงใน Rasch, 1960) ในช่วงแรก Rasch ได้ศึกษาเฉพาะโมเดล UIRT ต่อมาในปี ค.ศ. 1962 ได้นำเสนอโมเดลที่ยอมรับโดยทั่วไปรวมถึงความเป็นไปได้ของความสามารถของผู้สอบ โดยนำเสนอในรูปของเวคเตอร์มากกว่าค่าสเกลล่า ซึ่งสามารถแสดงได้ดังสมการที่ (1)

$$P(\mu_{ij}|\theta_j,\eta_j)=\frac{1}{\gamma(\theta_j,\eta_i)}e^{f(u_{ij})'\theta_j+g(u_{ij})'\eta_j+\theta_j'h(u_{ij})\eta_j+l(u_{ij})} \quad \text{.....(1)}$$

เมื่อ f, g, h และ l แทน ฟังก์ชันของคะแนนในแต่ละข้อ

μ_{ij} แทน ค่าสถิติที่ได้จากการสังเกตได้จากค่าพารามิเตอร์ของผู้สอบและแบบสอบ

η แทน เวคเตอร์ค่าพารามิเตอร์ของข้อสอบ

γ แทน ฟังก์ชัน Normalizing ที่จะสามารถรับประกันได้ว่าค่าที่อยู่ในฟังก์ชันของ IRT อยู่ในช่วง 0 และ 1

จากการพิจารณาโมเดลข้างต้นจะเห็นว่า มีการระบุระดับความสามารถของผู้สอบในรูปของเวคเตอร์ค่าพารามิเตอร์ โดยการประมาณค่าพารามิเตอร์ของผู้สอบและข้อสอบมีความเป็นอิสระจากกัน เมื่อพิจารณาส่วนประกอบอื่นๆ เช่น f, g, h และ l ซึ่งเป็นเวคเตอร์ที่เพียงพอสำหรับประมาณค่าได้จากข้อมูลในการตอบของผู้สอบ สำหรับในกรณีของข้อมูลที่มีลักษณะเป็นเอกมิติ การให้คะแนนจะมี 2 ค่า คือ 0 หรือ 1 โดยไม่ได้คำนึงถึงลักษณะข้อสอบดังนั้นการพิจารณาเฉพาะจำนวนที่ตอบข้อสอบได้ถูกต้อง ถือเป็นค่าสถิติที่เพียงพอสำหรับการพิจารณาค่า

ถ้าหากมีการตอบสนองข้อสอบในมิติที่ต่างกัน 2 ด้าน โดยคะแนนการตอบข้อสอบแต่ละข้อของมิติที่ 1 และ 2 คือ และ ดังนั้นค่าสถิติที่เพียงพอสำหรับประมาณค่า และ คือ และ เมื่อ n คือ จำนวนข้อที่ตอบข้อสอบได้ถูกต้อง อย่างไรก็ตามเนื่องจาก และเป็นค่าคงที่ข้ามข้อสอบ ดังนั้นคะแนนที่ได้แต่ละมิติเป็นฟังก์ชันของ n เท่านั้น ดังนั้นค่า และ ที่ประมาณได้ จึงมีความสัมพันธ์กันอย่างสมบูรณ์ นั่นคือโมเดลที่ได้ยังคงเป็นโมเดลแบบเอกมิตินั่นเอง จากปัญหาดังกล่าวนี้มีนักวิจัยหลายท่านพยายามที่จะแบ่งคะแนนให้มากกว่า 2 ค่า เช่น จัดกระทำกับข้อสอบจากลักษณะ “Dichotomous Items” แปลงให้เป็น “Single Polytomous Item” (Reckase, 2009, as cited in Rackase, 1972) เพื่อจะนำไปสู่การวิเคราะห์แบบ “Polytomous Items” Kelderman (Reckase, 2009, as cited in Kelderman, 1994) ดังนั้นฟังก์ชันการให้คะแนนจึงมีความซับซ้อนมากขึ้น หรือกำหนดให้ฟังก์ชันการให้คะแนนแตกต่างในแต่ละข้อก่อนที่จะวิเคราะห์ด้วยโมเดลราสช์ เพื่อทำให้การวิเคราะห์มี

ความเป็นเหตุเป็นผลมากขึ้นสำหรับการวิเคราะห์คุณลักษณะของข้อสอบ (Reckase, 2009, as cited in Glass, 1992) จะเห็นได้ว่าไม่นิยมนำโมเดลราสช์มาใช้ในการวิเคราะห์แบบพหุมิติ เนื่องจากมีความซับซ้อนของกระบวนการและความถูกต้องเกี่ยวกับฟังก์ชันการให้คะแนนที่มีความเฉพาะเจาะจงต่อมา Lord และ Novick (Reckase, 2009, อ้างอิงใน Lord and Novick, 1968) ได้เสนอโมเดลพื้นฐานของ MIRT ที่บ่งชี้นิยามของมิติคุณลักษณะแฝงได้อย่างสมบูรณ์และข้อตกลงเบื้องต้นความเป็นอิสระ ในที่นี้หมายถึง คุณลักษณะของกลุ่มผู้สอบต่างๆที่มีค่าเหมือนกันของมิติความสามารถ , , ..., เป็นการกระจายของคะแนนผลการตอบข้อนั้นของผู้สอบแต่ละคนไม่สัมพันธ์กัน เมื่อ k คือจำนวนของมิติ ส่วนในที่นี้เป็นเวคเตอร์มิติคุณลักษณะที่สมบูรณ์ นอกจากนี้ Lord และ Novick แสดงให้เห็นถึงความสัมพันธ์ระหว่า.โมเดลปกติสะสมตามทฤษฎีตอบสนองข้อสอบแบบเอกมิติและโมเดลองค์ประกอบร่วมอย่างชัดเจนมากยิ่งขึ้น

นอกจากนี้ Samejima (Reckase, 2009, as cited in Samejima, 1974) ได้เสนอสูตรที่ใช้โดยทั่วไป โดยมีข้อตกลงเบื้องต้นว่าการตอบที่มีลักษณะการตรวจให้คะแนน 2 ค่าหรือมากกว่า 2 ค่าเป็นผลมาจากการแบ่งเป็นส่วนย่อยๆ จากการตอบของตัวแปรต่อเนื่อง ซึ่ง Samejima ได้พัฒนาโมเดล MIRT สำหรับข้อสอบซึ่งมีลักษณะเป็นการตอบแบบต่อเนื่อง แสดงได้ดังสมการ

$$P_{zi(\theta_j)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_i(\theta_j - b_j)} e^{-\frac{t^2}{2}dt}$$

แม้ว่าโมเดลของ Samejima เป็นโมเดลแรกที่ทำให้แนวคิดเกี่ยวกับโมเดล MIRT มีความชัดเจน แต่จากการศึกษาของ Bejar (Reckase, 2009, อ้างอิงใน Bejar, 1977) พบว่า โมเดลดังกล่าวไม่สามารถนำไปประยุกต์ใช้ได้จริงในทางปฏิบัติ ซึ่งสาเหตุอาจเกิดจากการตอบข้อสอบมีการพิจารณาเป็นตัวแปรต่อเนื่องซึ่งไม่สอดคล้องกับบริบทของแบบสอบทั้งทางด้านการศึกษาและจิตวิทยา แต่อย่างไรก็ตามถ้ามีการประเมินเกี่ยวกับทักษะการปฏิบัติ โมเดลดังกล่าวก็จะเหมาะในการประยุกต์ใช้ได้มากขึ้น (Reckase, 1997)

1.1 เปรียบเทียบวิธีการของการวิเคราะห์องค์ประกอบ และ MIRT (Comparison of the Factor Analysis and MIRT Approach)

จากการศึกษาถึงวิธีการวิเคราะห์องค์ประกอบและ MIRT ของ Bock และ Aitkin; Samejima; McDonad (Reckase, 2009, อ้างอิงใน Bock and Aitkin, 1981; Samejima, 1974; McDonad. 1967) พบว่า สูตรทางสถิติของทั้งสองวิธีจะมีลักษณะเป็นเมตริกซ์ของการตอบข้อสอบที่ตรวจให้คะแนนแบบ 2 ค่า ซึ่งความเป็นจริงโปรแกรมคอมพิวเตอร์ที่พัฒนาโดย Bock และคณะ ในปี ค.ศ. 1988 สามารถนำมาประยุกต์ใช้ทั้งการวิเคราะห์องค์ประกอบและ MIRT ดังแสดงในตารางที่ 3

ตารางที่ 3 สรุปความแตกต่างของการวิเคราะห์องค์ประกอบ (Factor Analysis) และ MIRT

Methodology	Factor Analysis	MIRT
1. Focus	มุ่งไปที่เทคนิคการสกัดข้อมูลและสังเคราะห์ข้อมูลให้มีขนาดเล็กลง (Data Reduction Technique) เพื่อให้ได้มิติที่น้อยที่สุดของข้อมูลและง่ายต่อความเข้าใจสำหรับการบ่งชี้คุณลักษณะต่างๆ	ไม่ได้มุ่งไปที่เทคนิคการสกัดข้อมูลและสังเคราะห์ข้อมูลให้มีขนาดเล็กลงแต่มุ่งไปที่ปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบ
2. Input Variable	สนใจ Correlation หรือ Variance/Covariances Matrices มากกว่าค่าเฉลี่ยและความสัมพันธ์ระหว่างตัวแปรซึ่ง ค่าเฉลี่ย และ S.D. มีผลน้อยมากหรืออาจจะไม่มีผลต่อการวิเคราะห์ที่ได้	ค่าเฉลี่ย และ S.D. ของการตอบข้อสอบแทนด้วยค่าพารามิเตอร์ความยากและอำนาจจำแนกของข้อสอบ ซึ่งการประมาณค่าและการตีความหมายของค่าพารามิเตอร์เหล่านี้ถือเป็นส่วนประกอบสำคัญในการวิเคราะห์
3. Data Analysis	เมตริกซ์ที่ใช้เป็นข้อมูลในการวิเคราะห์ คือ Z-Score Matrix นั่นคือเป็นเมตริกซ์ความสัมพันธ์ (Correlation Matrices) ที่มีการแปลงคะแนนแบบเส้นตรงของคะแนนที่สังเกตได้ไปสู่ Z-Score ส่วนค่าความแปรปรวนร่วม (Covariance Matrices) คือเมตริกซ์ของการแปลงคะแนนเชิงเส้นตรงของจำนวนคะแนนที่ตอบข้อสอบได้ถูกต้องโดยมีค่าเฉลี่ยเท่ากับ 0	เป็นเมตริกซ์ของฟังก์ชันทางคณิตศาสตร์ในการตอบสนองข้อสอบที่จะระบุถึงคุณลักษณะในมิติของ (θ)
4. Goodness of Fit Under Hypothesis Model	โมเดลสมมติฐานมุ่งไปที่เมตริกซ์ความแปรปรวนและเมตริกซ์ความแปรปรวนร่วม โดยภาพรวมของกลุ่มผู้สอบ มากกว่าตัวแปรใดตัวหนึ่งหรือการเลือกเฉพาะ	โมเดลที่แสดงปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบ มีเป้าหมายของการระบุถึงความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องแม่นยำของแต่ละบุคคลในแต่ละมิติของ (θ)โดยให้ความสำคัญกับข้อสอบ

Methodology	Factor Analysis	MIRT
	กลุ่มย่อยกลุ่มใดกลุ่มหนึ่ง มุ่งไปที่การวัดโดยภาพรวมมากกว่าเงื่อนไขการวัดความสอดคล้องของข้อมูล	ที่ไม่สอดคล้องกับโมเดลหรือไม่สอดคล้องกับการทำนายความน่าจะเป็นโดยเฉพาะในช่วงพิสัยของความสามารถ ซึ่งเงื่อนไขของการวัดที่สอดคล้องกับแบบสอบชุดหนึ่งจะต้องอยู่ภายใต้เงื่อนไขความสอดคล้องของผู้สอบ (Liou & Chang. 1992) และความเหมาะสมของการวัด (Drasgow, Levine & McLaughlin, 1991)
5. Analysis & Application	แม้ว่าวิธีการวิเคราะห์จะมุ่งไปที่การหมุนแกนตัวประกอบ (Procrustes Rotation) และสัมประสิทธิ์ความสอดคล้องเพื่อที่จะแสวงหาองค์ประกอบร่วมกันแต่ยังมีลักษณะที่เน้นย้ำน้อยกว่า MIRT	วิธีการวิเคราะห์จะมีการศึกษาข้ามแบบทดสอบและกลุ่มผู้สอบ จุดมุ่งหมายของการวิเคราะห์เพื่อให้อยู่ในสเกลร่วมกันโดยมีประโยชน์ในการจัดเก็บเป็นคลังข้อสอบ หรือการจัดชุดของแบบทดสอบเพื่อวัดโครงสร้างหรือคุณลักษณะที่มุ่งวัด (Fixed-Form Test Construction) หรือการสร้างแบบสอบให้เหมาะสมกับความสามารถของผู้สอบ (Adaptive Treating)

1.2 ประเภทของโมเดลการตอบสนองข้อสอบแบบพหุมิติ

โมเดลการตอบสนองข้อสอบแบบพหุมิติจะประกอบด้วย ค่าพารามิเตอร์ (ความสามารถ) ของผู้สอบตั้งแต่ 2 พารามิเตอร์ขึ้นไป ซึ่งจะทำให้โมเดลมีความสอดคล้องกับข้อมูลเมื่อความแตกต่างกันของผู้สอบมีความแตกต่างกันอย่างเป็นระบบซึ่งข้อสอบมีทั้งข้อง่ายและข้อยากในโมเดลที่เป็นแบบพหุมิติ ค่าอำนาจจำแนกของข้อสอบจะมีผลต่อมิติในข้อสอบ แต่ละชุดสัญลักษณ์ที่แสดงในโมเดลมี ดังนี้

X_{ip} = ผลการตอบข้อสอบข้อที่ i ของบุคคล p (0=ผิด, 1=ถูก)

θ_{pm} = ค่าระดับความสามารถหรือคุณลักษณะแฝงของผู้สอบ p (m≥2)

b_i = ค่าความยากของข้อสอบ ข้อที่ i

δ_i = ค่าจุดตัดความง่าย (easiness intercept) ของข้อสอบ ข้อที่ i

a_{im} = ค่าอำนาจจำแนกของข้อสอบข้อที่ i บนมิติ m

c_i = ค่าโอกาสการเดาข้อสอบได้ถูก (guessing) ของข้อสอบข้อที่ i

Embretson and Reise (2000) นำเสนอโมเดลการตอบสนองข้อสอบแบบพหุมิติเป็น 2 ประเภท คือ โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ (Exploratory MIRT Model) และโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยัน (Confirmatory MIRT Model) ซึ่งมีลักษณะคล้ายคลึงกับการวิเคราะห์องค์ประกอบ โดย Exploratory MIRT Model เป็นโมเดลที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบมากกว่า 1 มิติ เพื่อให้โมเดลมีความสอดคล้องกับข้อมูลโดยไม่ได้กำหนดจำนวนขององค์ประกอบไว้ล่วงหน้า ซึ่งตรงข้ามกับ Confirmatory MIRT Model เป็นโมเดลที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในมิติที่มีความเฉพาะเจาะจงซึ่งคล้ายคลึงกับการวิเคราะห์องค์ประกอบ โดยการวิเคราะห์จะเกี่ยวข้องกับการระบุความสัมพันธ์ของข้อสอบไปยังมิติต่างๆ รายละเอียดมีดังต่อไปนี้

(1) โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ (Exploratory MIRT Model) ในการวิเคราะห์องค์ประกอบซึ่งมีลักษณะคล้ายกับโมเดล MIRT ดังที่กล่าวมาแล้วข้างต้น ความจริงแล้วเมื่อพิจารณาตามข้อตกลงเบื้องต้นสามารถพิสูจน์ได้ว่าเป็นโมเดลเดียวกัน Takane and de Leeuw (Reckase, 2009, อ้างอิงใน Takane and de Leeuw, 1988) ซึ่ง McDonald (1967) ได้เสนอโมเดลการวิเคราะห์องค์ประกอบแบบไม่ใช้เส้นตรง โดยสร้างจากการรวมพื้นฐานของการวิเคราะห์องค์ประกอบ CTT และ IRT เข้าด้วยกัน ซึ่งปัจจุบันจากการศึกษาของ Reckase (2009) ได้ระบุว่าขอบเขตของ Exploratory Factor Analysis จะนิยมใช้กับข้อมูลแบบแบ่ง 2 คล้ายกับโมเดลของ IRT จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับ Exploratory MIRT Model พบว่า มีความสัมพันธ์กันในระดับสูงกับการวิเคราะห์องค์ประกอบโดยศักยภาพของผู้สอบพิจารณาจากการตอบข้อสอบคือ การให้น้ำหนักรวมกันภายใต้มิติของคุณลักษณะที่ต้องการศึกษา

1.3 เปรียบเทียบวิธีการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ

การวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติมีสูตรทางสถิติเหมือนกันเมื่อถูกนาไปใช้กับเมทริกซ์ของการตอบสนองข้อสอบ และสามารถเห็นได้จากการเปรียบเทียบโมเดลที่นาเสนอโดย Bock and Aitken ในปี ค.ศ. 1981 Samejima ในปี ค.ศ. 1974 และ McDonald ในปี ค.ศ. 1967 ซึ่งในความเป็นจริงซอฟต์แวร์สำหรับการวิเคราะห์องค์ประกอบที่ให้สารสนเทศสมบูรณ์ที่นำเสนอโดย Bock et al. ในปี ค.ศ. 1988 ที่สามารถนำมาใช้ในการวิเคราะห์ได้ทั้งการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Reckase, 2009, p. 70) เมื่อพิจารณาถึงแนวความคิดของทั้งการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ พบว่ามีทั้งส่วนที่คล้ายคลึงกันและส่วนที่ต่างกันในประเด็นระเบียบวิธีการวิเคราะห์ (Methodology) โดยในส่วนของความเหมือนกัน คือ

1. กระบวนการทางคณิตศาสตร์ของการวิเคราะห์องค์ประกอบกับโมเดลการตอบสนองข้อสอบแบบพหุมิติ มีลักษณะระเบียบวิธีทางด้านตัวเลขที่คล้ายกัน ซึ่งทั้งสองวิธีพยายามที่จะระบุสเกลสมมุติฐานเพื่อถอดแบบข้อมูลที่จะนาไปสู่การวิเคราะห์ข้อมูล โดยสเกลของการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ จะต้องกำหนดจุดกำเนิดและมีหน่วยของการวัดเหมือนกัน

2. การวิเคราะห์องค์ประกอบในปัจจุบัน มีลักษณะการหมุนแกนที่ปรับปรุงขยายมาจากวิธีการวิเคราะห์ของโมเดลการตอบสนองข้อสอบแบบเอกมิติ ดังนั้นวิธีการหมุนแกนของการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติจึงเหมือนกัน ส่วนในประเด็นแตกต่างกันของการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติดังตารางที่ 4 (Reckase, 2009, pp. 70-71)

ตารางที่ 4 สรุปความแตกต่างของการวิเคราะห์องค์ประกอบและโมเดลการตอบสนองข้อสอบแบบพหุมิติ

การวิเคราะห์องค์ประกอบ	**โมเดลการตอบสนองข้อสอบแบบพหุมิติ**
1. เป็นเทคนิคสำหรับลดข้อมูล (Data reduction technique) โดยมีเป้าหมายหลักเพื่อให้ได้องค์ประกอบจำนวนน้อยที่สุด	1. เป็นเทคนิคสำหรับสร้างโมเดลปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบ
2. โดยทั่วไปไม่สนใจคุณสมบัติของตัวแปรสังเกตได้ เห็นได้จากในการวิเคราะห์เมทริกซ์ความสัมพันธ์ ความแตกต่างของค่าเฉลี่ยและความแปรปรวนของตัวแปรมีผลต่อการวิเคราะห์น้อยมากหรือไม่มีผลเลย และในทางตรงข้ามการวิเคราะห์องค์ประกอบวิธีใหม่ เช่น โมเดลสมการโครงสร้างจะพิจารณาค่าเฉลี่ย ความแปรปรวน และความแปรปรวนร่วมของตัวแปร แต่ก็ไม่ได้ทำเพื่อวัตถุประสงค์ที่จะเข้าใจคุณสมบัติของตัวแปรสังเกตได้	2. สนใจความแตกต่างของค่าเฉลี่ยและความแปรปรวนของคะแนนจากการตอบข้อสอบ เพราะคุณสมบัติของตัวแปรสังเกตได้มีความเกี่ยวข้องโดยตรงกับคุณสมบัติที่สำคัญของข้อสอบ เช่น ความยากและอำนาจจำแนกของข้อสอบ

ปัจจุบันนี้โมเดลของ MIRT ส่วนใหญ่จะช่วยในการพิจารณาด้านโครงสร้างของการพัฒนาแบบสอบนั่นคือ การพิจารณาโครงสร้างที่ต้องการวัดในข้อสอบจำเป็นที่จะต้องอาศัยหลักการของโมเดล IRT เพื่อให้สอดคล้องกับข้อมูลโดยจำนวนของคุณลักษณะที่ต้องการวัดจะต้องมีความสอดคล้องกับข้อมูลเชิงประจักษ์ และเป็นส่วนย่อยของมิติต่างๆ เป็นการสร้างตามหลักการหรือทฤษฎี

ในชุดข้อสอบ ผลที่ได้จากการประมาณค่าพารามิเตอร์ของข้อสอบจะเป็นประโยชน์ในการเลือกข้อสอบเพื่อให้เหมาะสมกับคุณลักษณะแฝง นั่นคือ แม้ว่าลักษณะที่วัดจะมีความเกี่ยวข้องกันบางส่วนก็ตามแต่สามารถเลือกศึกษาเพียงคุณลักษณะใดลักษณะหนึ่งได้

อย่างไรก็ตามเราสามารถประมาณค่าพารามิเตอร์ของผู้สอบที่มีหลายคุณลักษณะแฝงไปพร้อมๆ กันได้ในครั้งเดียวโดยใช้โมเดล MIRT ซึ่งข้อมูลที่ได้จากการตอบของผู้สอบถือว่าเป็นข้อมูลที่สำคัญในการวัดตั้งแต่ 2 คุณลักษณะขึ้นไป โดยในที่นี้ได้จำแนก Exploratory MIRT Model ออกเป็นประเภทย่อย ได้ 2 ประเภท คือ 1) Multidimensional Logistic Model และ 2) Multidimensional Normal Ogive Model ซึ่งมีรายละเอียดดังนี้

โมเดลการตอบสนองรายข้อแบบพหุมิติของ Embretson & Reiseในหัวข้อนี้ผู้วิจัยจะนำเสนอรูปแบบของโมเดลการตอบสนองรายข้อแบบพหุมิติเป็น 2 แนวคิด ได้แก่ รูปแบบโมเดลการตอบสนองรายข้อแบบพหุมิติของ Embretson& Reise (Reckase, 2009, p. 86, as cited in Embretson & Reise, 2000) และรูปแบบโมเดลการตอบสนองรายข้อแบบพหุมิติของ Reckase (2009, p. 86) โดยมีรายละเอียด ดังนี้

1. รูปแบบโมเดลการตอบสนองรายข้อแบบพหุมิติของ Embretson & Reise ได้แบ่งโมเดลการตอบสนองรายข้อแบบพหุมิติออกเป็น 2 ประเภท ได้แก่ โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ (Exploratory MIRT Model) และโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยัน (Confirmatory MIRT Model) ซึ่งมีลักษณะคล้ายคลึงกับการวิเคราะห์องค์ประกอบ ดังนี้

1.1 โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ (Exploratory MIRT Model) โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ เป็นโมเดลที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบมากกว่า 1 มิติเพื่อให้โมเดลมีความสอดคล้องกับข้อมูลโดยไม่ได้กำหนดจำนวนขององค์ประกอบไว้ล่วงหน้า แบ่งได้เป็น 2 ลักษณะ คือ โมเดลโลจิสติกแบบพหุมิติ (Multidimensional Logistic Model) และโมเดลปกติสะสมแบบพหุมิติ (Multidimensional Normal Ogive Model) ดังนี้

1.1.1 โมเดลโลจิสติกแบบพหุมิติ (Multidimensional Logistic Model) ประกอบด้วย 3 โมเดล ได้แก่ โมเดลโลจิสติกแบบพหุมิติ 1 พารามิเตอร์ (Multidimensional Rasch Model) โมเดลโลจิสติกแบบพหุมิติชนิด 2 พารามิเตอร์ (Multidimensional 2-Parameter Logistic Model) และโมเดลโลจิสติกแบบพหุมิติชนิด 3 พารามิเตอร์ (Multidimensional 3-Parameter Logistic Model)

1.1.2 โมเดลปกติสะสมแบบพหุมิติ (Multidimensional Normal Ogive Model) ประกอบด้วย 2 โมเดล ได้แก่ โมเดลปกติสะสมแบบพหุมิติ 2 พารามิเตอร์ (Multidimensional

2-Parameter Normal Ogive Model) และโมเดลปกติสะสมแบบพหุมิติ กับโอกาสการเดาข้อสอบ (Multidimensional Normal Ogive Model with Guessing)

1.2 โมเดลการตอบสนองรายข้อแบบพหุมิติเชิงยืนยัน (Confirmatory MIRT Model) เป็นโมเดลที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในมิติที่มีความเฉพาะเจาะจงตามคุณลักษณะแฝงที่กำหนดไว้ล่วงหน้า ซึ่งคล้ายคลึงกับการวิเคราะห์องค์ประกอบโดยการวิเคราะห์เกี่ยวข้องกับการระบุความสัมพันธ์ของข้อสอบไปยังมิติต่างๆ แบ่งได้เป็น 4 ลักษณะ ได้แก่ โมเดลสำหรับการวัดมิติที่ไม่สามารถทดแทนได้ (Models for No compensatory Dimensional) โมเดลสำหรับการวัดการเรียนรู้และการเปลี่ยนแปลง (Models for Learning and Change) โมเดลในการระบุโครงสร้างของระดับคุณลักษณะ (Models with Specified Trait Level Structures) และโมเดลสำหรับการจำแนกกลุ่มบุคคล (Model for Distinct Classes of Persons) โดยมีรายละเอียด ดังนี้

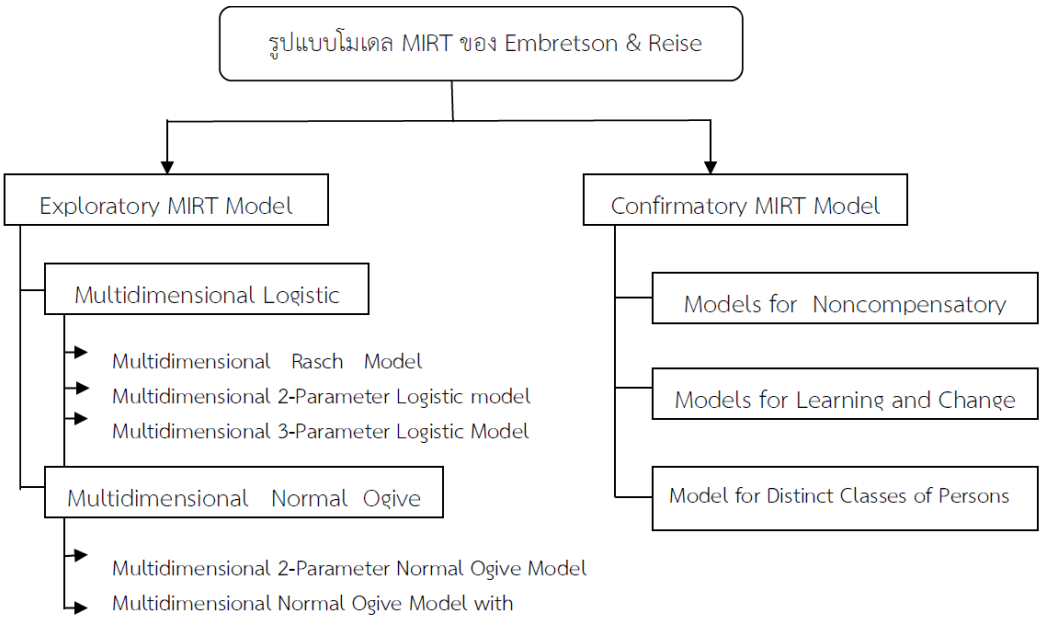
1.2.1 โมเดลสำหรับการวัดมิติที่ไม่สามารถทดแทนได้ (Models for No compensatory Dimensional) ในปี1980 Whitely ได้เสนอโมเดล Multicomponent Latent Trait Model (MLTM) เพื่อวัดองค์ประกอบของกระบวนการที่หลากหลาย ภายใต้ข้อสอบที่มีลักษณะของการพิสูจน์หรือแก้ปัญหา โดยที่ MLTM จะรวมโมเดลคณิตศาสตร์ของกระบวนการในการตอบข้อสอบกับโมเดล IRT ซึ่งจะประมาณค่าทั้งในส่วนของระดับคุณลักษณะ (trait level) และค่าความยากของข้อสอบ โดยมีข้อตกลงเบื้องต้นว่าต้องทำข้อสอบถูกต้องหรือสอดคล้องกับผลสำรวจที่ประมาณค่าจากหลายส่วนประกอบ ถ้าส่วนประกอบใดไม่ประสบความสำเร็จจะส่งผลให้การตอบข้อสอบไม่ถูกต้อง

นอกจากนี้ Embretson (1984) ยังได้เสนอ General Component Latent Trait Model (GLTM) ซึ่งเป็นโมเดลที่มีการรวมเอาโมเดลทางคณิตศาสตร์เกี่ยวกับส่วนประกอบค่าความยากของข้อสอบเช่นเดียวกับ Linear Logistic Test Model (LLTM) ที่ค่าความยากของข้อสอบมีการถ่วงน้ำหนักโดยรวมส่วนประกอบที่เป็นตัวกระตุ้นเข้าด้วยกัน ทั้งนี้โมเดลพื้นฐานทั้ง MLTM และ GLTM ต่างก็มีการกำหนดส่วนประกอบของการตอบเหมือนกับการตอบข้อสอบโดยรวมของแต่ละข้อ เพื่อใช้ในการประมาณค่าพารามิเตอร์ของส่วนประกอบการตอบสนองรายข้อซึ่งเป็นสิ่งที่สามารถสังเกตได้

1.2.2 โมเดลสำหรับการวัดการเรียนรู้และการเปลี่ยนแปลง (Models for Learning and Change) ในปี1991 Embretson ได้พัฒนาโมเดลที่เรียกว่า Multidimensional Rasch Model for Learning and Change (MRMLC) เพื่อใช้ในการแก้ปัญหาเกี่ยวกับการวัดทางจิตวิทยาในการวัดความเปลี่ยนแปลง ซึ่งการออกแบบที่เหมาะสมของโมเดล MRCML คือ ผู้สอบจะถูกสังเกตภายใต้เงื่อนไขตั้งแต่ 2 เงื่อนไขขึ้นไปหรือสองครั้งขึ้นไป ขึ้นอยู่กับความสะดวกหรืออุปสรรคในการปฏิบัติ เช่น การทดสอบ Pretest และ Posttest เป็นต้น ในการใช้งาน MRMLC สามารถปรับปรุงได้และถือเป็นมิติที่แยกออกไป และนำเอากระบวนการทางโครงสร้างของ Wiener มาใช้ในการระบุมิติความสามารถ

1.2.3 โมเดลในการระบุโครงสร้างของระดับคุณลักษณะ (Models with Specified Trait Level Structures) เมื่อแบบสอบที่นำมาใช้อยู่ภายใต้เงื่อนไขที่มีความหลากหลายหรือใช้ในหลายช่วงเวลาและเกี่ยวข้องกับระดับคุณลักษณะที่อาจจะมีหลากหลายระบบด้วย การออกแบบโครงสร้างอาจจะรวมอยู่ในโครงสร้างระดับคุณลักษณะตามโมเดล IRT และระบุคุณลักษณะที่มีเงื่อนไขร่วมกัน ซึ่งจากมุมมองอื่นๆ ในการออกแบบโครงสร้างจะมีการกำหนดคุณลักษณะสำคัญของระดับคุณลักษณะเช่นเดียวกันเนื่องจากระดับของคุณลักษณะเป็นผลมาจากการเปรียบเทียบกับปฏิบัติตามเงื่อนไข และโมเดล IRT ที่นำเสนอแต่ละโมเดลมีการออกแบบโครงสร้างที่ต่างกันไป เช่น ในปี 1995 Embretson ได้นำเสนอGeneral Structured Latent Trait Model (SLTM) โดยมีจุดมุ่งหมายในการออกแบบโครงสร้างของความสามารถหลายๆ ด้านที่มีค่าอำนาจำแนกของข้อสอบไม่เท่ากัน และมีปฏิสัมพันธ์ของข้อสอบภายใต้เงื่อนไขการวัดที่แตกต่างกัน ต่อมาในปี 1997 Wang, Wilson & Adams ได้นำเสนอโมเดลเชิงโครงสร้างในแต่ละมิติที่คล้ายคลึงกันกับโมเดลราสช์ ซึ่งเป็นโมเดลที่สามารถอ้างอิงไปยังสถานการณ์โดยทั่วไปและสามารถนำไปใช้ร่วมกับการออกแบบโครงสร้างอื่นๆ รวมทั้งสามารถใช้ร่วมกับ SLTM บางโมเดลได้

1.2.4 โมเดลสำหรับการจำแนกกลุ่มบุคคล (Model for Distinct Classes of Persons) โมเดล IRT เป็นโมเดลที่นำมากำหนดกลุ่มคุณลักษณะแฝงที่ไม่สามารถสังเกตได้จากตัวแปรภายนอก ซึ่งการแบ่งกลุ่มเกิดขึ้นจากการระบุโมเดลการตอบข้อสอบด้วยระบบที่แตกต่างกันของโมเดลการตอบสนองรายข้อ ซึ่งโมเดลจะประกอบไปด้วยค่าพารามิเตอร์ของกลุ่ม (class) และค่าพารามิเตอร์ของคุณลักษณะ (trait) เพื่อใช้ในการทำนายลักษณะของการตอบสนองรายข้อแม้ว่าเทคนิควิธีของโมเดล IRT จะเป็นการศึกษาเพียง 1 ระดับคุณลักษณะของผู้สอบแต่ละคน แต่สามารถนำมาใช้ในการจำแนกเชิงพหุมิติได้จากการศึกษารูปแบบโมเดลการตอบสนองรายข้อแบบพหุมิติของ Embretson & Reise สรุปได้ว่า มี 2 รูปแบบ ได้แก่ โมเดลการตอบสนองรายข้อแบบพหุมิติเชิงสำรวจ (Exploratory MIRT Model) และโมเดลการตอบสนองรายข้อแบบพหุมิติเชิงยืนยัน (Confirmatory MIRT Model) ซึ่งโมเดลการตอบสนองรายข้อแบบพหุมิติเชิงสำรวจเป็นโมเดลเกี่ยวข้องกับการประมาณค่าพารา- มิเตอร์ของข้อสอบและผู้สอบมากกว่า 1 มิติ เพื่อให้โมเดลมีความสอดคล้องกับข้อมูลโดยไม่ได้กำหนดจำนวนขององค์ประกอบไว้ล่วงหน้า แบ่งได้เป็น 2 ลักษณะ คือ โมเดลโลจิสติกแบบพหุมิติและโมเดลปกติสะสมแบบพหุมิติอีกรูปแบบหนึ่ง คือ โมเดลการตอบสนองรายข้อแบบพหุมิติเชิงยืนยันเป็นโมเดลที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในมิติที่มีความเฉพาะเจาะจงตามคุณลักษณะแฝงที่กำหนดไว้ล่วงหน้า ซึ่งแบ่งได้เป็น 4 ลักษณะ ได้แก่ โมเดลสำหรับการวัดมิติที่ไม่สามารถทดแทนได้โมเดลสำหรับการวัดการเรียนรู้และการเปลี่ยนแปลง โมเดลในการระบุโครงสร้างของระดับคุณลักษณะ และโมเดลสำหรับการจำแนกกลุ่มบุคคล ดังภาพ ต่อไปนี้



ภาพที่ 7 *รูปแบบโมเดล MIRT* ของ Embretson & Reise (2000)

2. รูปแบบโมเดลการตอบสนองข้อสอบแบบพหุมิติของ Reckase (2009, p. 86) ตามแนวคิดของ Reckase (2009) ได้แบ่งโมเดลการตอบสนองรายข้อแบบพหุมิติ ออกเป็น 2 รูปแบบ ได้แก่ โมเดล MIRT ของข้อสอบที่ให้คะแนน 2 ค่า (MIRT Models for Test Item with Two Score Categories) และโมเดล MIRT ของข้อสอบที่ให้คะแนนมากกว่า 2 ค่า (MIRT Models for Test Item with more than Two Score Categories) โดยมีรายละเอียด ดังนี้

2.1 โมเดล MIRT ของข้อสอบที่ให้คะแนน 2 ค่า (MIRT Models for Test Item with Two Score Categories) เป็นโมเดลเชิงสำรวจ ประกอบด้วย โมเดลย่อย 2 โมเดล คือ โมเดลที่สามารถทดแทนกันได้ (Compensatory Extensions of the UIRT Models) และโมเดลที่ไม่สามารถทดแทนกันได้ (Partially Compensatory Extensions of the UIRT Models) ดังนี้

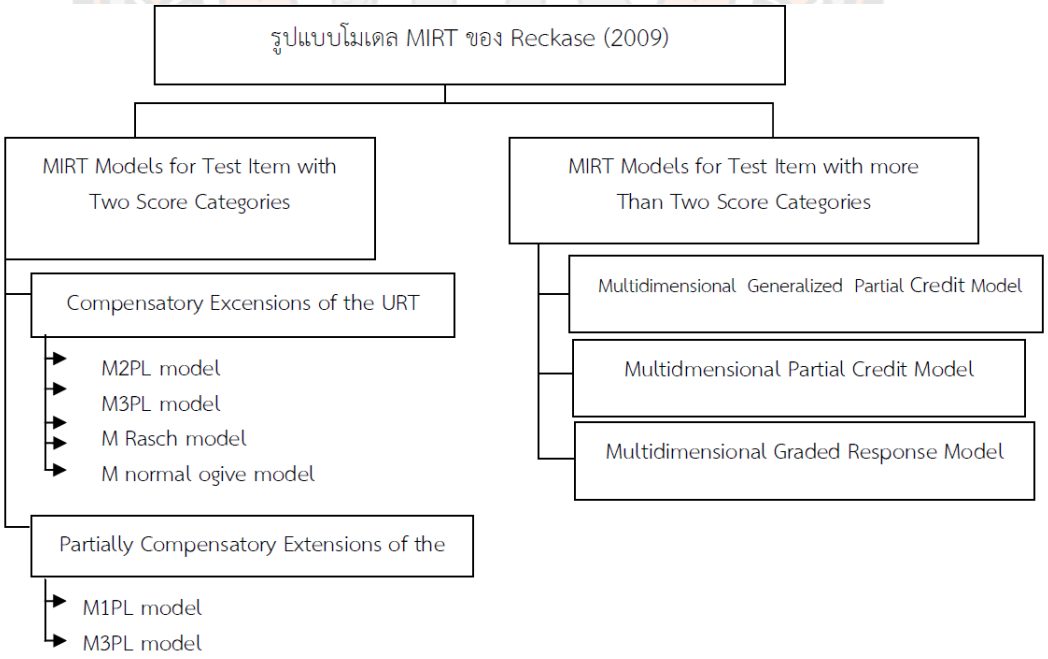
2.1.1 โมเดลที่สามารถทดแทนกันได้ (Compensatory Extensions of the UIRT Models) ประกอบด้วย โมเดลย่อย 4 โมเดล คือ โมเดลพหุมิติที่ขยายจากโมเดลโลจิสติก 2 พารามิเตอร์ (Multidimensional extension of the two-parameter logistic (M2PL) model) โมเดลพหุมิติที่ขยายจากโมเดลโลจิสติก 3 พารามิเตอร์ (Multidimensional extension of the three-parameter logistic (M3PL) model) โมเดลพหุมิติที่ขยายจากราสช์โมเดล (Multidimensional extension of the Rasch model) และ โมเดลพหุมิติที่ขยายจากโมเดลปกติสะสม (Multidimensional extension of the normal ogive model)

2.1.2 โมเดลที่สามารถทดแทนกันได้บางส่วน (Partially Compensatory Extensions of the UIRT Models) ประกอบด้วยโมเดลย่อย 2 โมเดล คือ โมเดลพหุมิติที่ขยายจาก

โมเดลโลจิสติก 1 พารามิเตอร์ (Multidimensional extension of the one-parameter logistic model) และโมเดลพหุมิติที่ขยายจากโมเดลโลจิสติก 3 พารามิเตอร์ (Multidimensional extension of the three-parameter logistic model)

2.2 โมเดล MIRT ของข้อสอบที่ให้คะแนนมากกว่า 2 ค่า (MIRT Models for Test Item with more than Two Score Categories) โมเดลลักษณะนี้มีใช้มานานแล้ว เป็นโมเดลที่อยู่ภายใต้โมเดลแบบทดแทน (compensatory model) เนื่องจากปัจจุบันไม่มีการเสนอโมเดลแบบทดแทนกันได้บางส่วน (Partially Compensatory Model) ที่ให้คะแนนมากกว่า 2 ค่าประกอบด้วยโมเดลย่อย 3 โมเดล ได้แก่ Multidimensional Generalized Partial Credit Model (MGPC) Multidimensional Partial Credit Model (MPC) และ Multidimensional Graded Response Model (MGR)

จากการศึกษารูปแบบโมเดลการตอบสนองข้อสอบแบบพหุมิติตามแนวคิดของ Reckase (2009 : 87) โมเดลการตอบสนองข้อสอบแบบพหุมิติ แบ่งออกเป็น 2 รูปแบบ ได้แก่โมเดล MIRT ของข้อสอบที่ให้คะแนน 2 ค่า (MIRT Models for Test Item with Two Score Categories) และโมเดล MIRT ของข้อสอบที่ให้คะแนนมากกว่า 2 ค่า (MIRT Models for Test Item with more than Two Score Categories) โดยแต่ละโมเดลมีโมเดลย่อย ดังภาพประกอบต่อไปนี้



ภาพที่ 8 รูปแบบโมเดล MIRT ของ Reckase (2009)

3. ข้อตกลงเบื้องต้น

สำหรับข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ Reckase (2009, p. 257) ได้ระบุไว้ 5 ประการ ได้แก่

1) ตำแหน่ง (location) ที่ถูกวัดของบุคคล หรือระดับความสามารถต้องไม่เปลี่ยนแปลงในระหว่างกระบวนการวัด ซึ่งในความเป็นจริงแล้วเป็นไปได้ยากเนื่องจากบุคคลอาจมีการเรียนรู้จากการมีปฏิสัมพันธ์กับข้อสอบ สภาพแวดล้อมการสอบ เป็นต้น

2) คุณลักษณะของข้อสอบมีความคงที่ข้ามสถานการณ์การทดสอบนั่นคือ พารามิเตอร์ของข้อสอบไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบ

3) ข้อตกลงเกี่ยวกับความเป็นอิสระในการตอบข้อสอบ (local independence assumption) กล่าวคือเป็นข้อตกลงเกี่ยวกับการตอบข้อสอบของบุคคลหนึ่งไม่ได้เป็นผลมาจากผู้ตอบคนอื่นๆ นั่นคือ การตอบข้อสอบของบุคคลไม่สัมพันธ์กัน หรือเป็นอิสระต่อกัน

4) ความสัมพันธ์ระหว่างตำแหน่งของมิติหลายๆ มิติกับความสามารถในการตอบข้อสอบได้ถูกต้องของบุคคลจะแสดงในรูปฟังก์ชันทางคณิตศาสตร์แบบต่อเนื่อง ซึ่งเป็นลักษณะสำคัญในการอธิบายปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ

5) ความน่าจะเป็นในการตอบถูกของบุคคลต้องเพิ่มขึ้น หรืออย่างน้อยต้องไม่ลดลง (คงที่) เมื่อตำแหน่งความสามารถของบุคคลเพิ่มขึ้นในมิติอื่น ซึ่งเรียกว่าข้อตกลงของการเพิ่มขึ้นทางเดียว (monotonicity assumption) แต่ก็มีบางโมเดลที่ละเมิดข้อตกลงเบื้องต้นนี้กล่าวได้ว่าข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ประกอบด้วยข้อตกลงเกี่ยวกับระดับความสามารถของบุคคลคุณลักษณะของข้อสอบ ความเป็นอิสระในการตอบข้อสอบ และการใช้ฟังก์ชันทางคณิตศาสตร์แสดงปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบเหมือนโมเดลการตอบสนองข้อสอบแบบเอกมิติแต่ที่แตกต่างคือข้อตกลงของการเพิ่มขึ้นทางเดียว ที่ไม่พบในโมเดลการตอบสนองข้อสอบแบบเอกมิติ

4. การประมาณค่าพารามิเตอร์

Reckase (2009, p. 138) ได้อธิบายถึงการประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบด้วยเหตุผล 3 ประการ กล่าวคือ ประการที่ 1 โมเดลประกอบด้วยพารามิเตอร์ของข้อสอบและพารามิเตอร์ของผู้สอบ จึงเป็นไปไม่ได้ที่การประมาณค่าพารามิเตอร์ทั้งสองจะเป็นอิสระจากกัน เหตุผลประการที่ 2 คือ พารามิเตอร์ที่ต้องประมาณค่ามีจำนวนมากนั่นคือ ถ้า m เป็นจำนวนของแกนประกอบ (จำนวนมิติ) n เป็นจำนวนของข้อสอบ และ N เป็นจำนวนผู้สอบ ดังนั้นจำนวนพารามิเตอร์ที่ต้องประมาณค่าจะเท่ากับ n (m+1) + (m) และเหตุผลประการที่ 3 คือ มีค่าบางค่าที่หาไม่ได้ (indeterminacies) ในโมเดล เช่น ตำแหน่งของ การเกิดมิติ (location of the origin of

the space) หน่วยการวัดของแต่ละแกนร่วม (coordinate axis) และความสัมพันธ์ของแกนร่วมกับความสามารถของผู้สอบ

จากการศึกษาเอกสารข้างต้นสรุปได้ว่า โมเดลการตอบสนองข้อสอบแบบพหุมิติมีลักษณะคล้ายคลึงกับการวิเคราะห์องค์ประกอบ แบ่งออกเป็น 2 ประเภท คือ โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ และโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยัน โดยโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจเกี่ยวข้องกับพารามิเตอร์ข้อสอบและผู้สอบที่มากกว่า 1 มิติ โดยไม่ได้กำหนดจำนวนองค์ประกอบไว้ล่วงหน้า ส่วนโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยันเกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในมิติที่มีความเฉพาะเจาะจง โดยคล้ายกับการวิเคราะห์องค์ประกอบที่เกี่ยวข้องกับการระบุความสัมพันธ์ของข้อสอบไปยังมิติต่างๆ โดยพบว่าข้อสอบจะประกอบด้วยการออกแบบตามโครงสร้างและเป็นโมเดลในรูปฟังก์ชันเชิงคณิตศาสตร์ที่มีการเชื่อมโยงกับคุณลักษณะที่มีความเฉพาะเจาะจงในหลายลักษณะโมเดลนี้เหมาะสำหรับข้อสอบหรือชิ้นงานของทักษะการปฏิบัติที่ได้มีการกำหนดหรือสร้างทฤษฎีไว้ล่วงหน้าทั้งนี้โมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยันเหมาะสา หรับการนำไปประยุกต์ใช้ได้หลากหลาย เช่น แก้ปัญหาทางด้านการวัดทางจิตวิทยาในการวัดความเปลี่ยนแปลง (Measurement of change) ออกแบบโครงสร้างของความสามารถ และในการวิจัยครั้งนี้ผู้วิจัยได้เลือกที่จะใช้วิธีการวิเคราะห์ตามโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงยืนยันจากโมเดลปกติสะสมแบบพหุมิติ 3 พารามิเตอร์

5. โปรแกรมคอมพิวเตอร์สำหรับการประมาณค่าพารามิเตอร์โมเดลการตอบสนองข้อสอบแบบพหุมิติ

สำหรับโปรแกรมการวิเคราะห์การทดสอบแบบพหุมิตินั้นมีคุณสมบัติที่เหมือนและแตกต่างกันไป จึงควรพิจารณาข้อจำกัดของแต่ละโปรแกรมก่อนที่จะนำมาใช้ประมาณค่าพารามิเตอร์ข้อสอบ หรือผู้สอบเพื่อให้เกิดประสิทธิภาพ และมีความแม่นยำมากที่สุด โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นมาเพื่อใช้ในการประมาณค่าพารามิเตอร์ของโมเดลการตอบสนองข้อสอบแบบพหุมิติมีอยู่หลากหลายโปรแกรมและแต่ละโปรแกรมมีคุณสมบัติที่ต่างกัน เช่น วิธีการประมาณค่า ลักษณะโมเดลการตอบสนองข้อสอบแบบพหุมิติที่สามารถประมาณค่าได้ จำนวนโมเดลการตอบสนองข้อสอบแบบพหุมิติที่สามารถประมาณค่าได้ ความยากง่ายในการใช้งานคู่มือการใช้งานโปรแกรม และค่าใช้จ่ายที่เกิดขึ้นจากการใช้โปรแกรมในประเด็นวิธีการประมาณค่าพารามิเตอร์ของข้อสอบนั้น Reckase (2009, p. 148) ได้นำเสนอไว้ 3 วิธี คือ Maximum likelihood, Bayesian maximum a posteriori และ Least squares และโปรแกรมคอมพิวเตอร์ที่ใช้วิธี Maximum likelihood คือ โปรแกรม TESTFACT และโปรแกรม Conquest โปรแกรมที่ใช้วิธี Bayesian maximum a posteriori คือโปรแกรม BMIRT ส่วนโปรแกรมที่ใช้วิธี Least squares คือ NOHARM และ Mplus

ซึ่งผู้วิจัยจะนำเสนอโปรแกรมคอมพิวเตอร์ที่ผู้วิจัยนำมาใช้ในงานวิจัยนี้เท่านั้น คือ โปรแกรม NOHARM และ มีรายละเอียด ดังนี้

1. โปรแกรม NOHARM (Nomal-Ogive Harmonic Analysis Robust method) เป็นโปรแกรมคอมพิวเตอร์ที่ใช้ประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบสำหรับโมเดลการตอบสนองข้อสอบแบบพหุมิติได้ด้วยวิธีการทั้งสองแบบคือ Exploratory Factor Analysis (EFA) และ Confirmatory Factor Analysis (CFA) โดยชื่อโปรแกรมมาจากคำว่า Normal Ogive by Harmonic Analysis Robust Method การประมาณค่าพารามิเตอร์ข้อสอบใช้วิธีที่เรียกว่า Unweighted Least Squares เป็นโปรแกรมที่ใช้ประมาณค่า Multidimensional Extension of the 2-Parameter Normal Ogive Model และใช้ประมาณค่าพารามิเตอร์ Multidimensional Extension of the 3-Parameter Normal Ogive Model โดยผู้ใช้ต้องคำนวณค่าการเดาจากโปรแกรมอื่น ๆ แล้วนำมาป้อนข้อมูลให้กับโปรแกรม NOHARM การประมาณค่าพารามิเตอร์ข้อสอบของโปรแกรม NOHARM ใช้วิธีการที่แตกต่างจากโปรแกรม TESTFACT และผลของการประมาณค่าพารามิเตอร์จะได้ค่าอำนาจจำแนกที่มีหลายค่าต่อข้อสอบหนึ่งข้อ และได้ค่าพารามิเตอร์ความยากหนึ่งค่าต่อข้อสอบหนึ่งข้อ และเป็นโปรแกรมคอมพิวเตอร์ที่มีลักษณะเป็นฟรีแวร์ (Freeware) ไม่มีลิขสิทธิ์การค้าจึงไม่ต้องมีค่าใช้จ่ายในการใช้งาน สามารถใช้งานได้ทั้งบนระบบดอส (Dos System) และบนระบบวินโดว์ (Windows System) เป็นโปรแกรมที่ทำงานได้รวดเร็วมากอีกทั้งไม่มีข้อจำกัดเกี่ยวกับจำนวนพารามิเตอร์ที่จะประมาณค่าแต่ขึ้นอยู่กับขีดความสามารถของเครื่องคอมพิวเตอร์ที่นำมาใช้งาน แต่โปรแกรม NOHARM จะไม่ประมาณค่าองค์ประกอบของเวกเตอร์ความสามารถ (θ - vector) ซึ่งหากผู้ใช้งานต้องการประมาณค่าส่วนประกอบของเวกเตอร์จะต้องใช้โปรแกรมอื่นๆ ช่วยในการประมาณค่า เช่น TESTFACT (สมประสงค์ เสนารัตน์, 2555, น. 71-72) โปรแกรมนี้สามารถวิเคราะห์ข้อสอบแบบพหุมิติได้มากถึง 50 มิติ อย่างไรก็ตามโปรแกรมนี้จะไม่สามารถประมาณค่าพารามิเตอร์ผู้สอบ และวิเคราะห์ได้เฉพาะข้อสอบที่ให้คะแนนแบบ dichotomous

2. โปรแกรม jMetrik ซึ่งเป็นโปรแกรมที่มีลิขสิทธิ์แบบเปิด (Open source) ที่สามารถนำมาใช้งานได้อย่างถูกกฎหมายและไม่ต้องเสียค่าใช้จ่ายสำหรับวิเคราะห์ข้อมูลที่มีการให้คะแนนแบบหลายค่า (3PL, 4PL, GPCM) และใช้กับโมเดล Generalized Partial Credit Model ซึ่งใช้วิธี marginal maximum likelihood estimation (MMLE) ในการประมาณค่าพารามิเตอร์ร่วมกับโปรแกรม NOHARM version 4.0 ในการคัดเลือกค่าพารามิเตอร์ข้อสอบเข้าคลังข้อสอบ (Item bank)

การทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing)

ในปัจจุบันนี้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้รับความนิยมมากทั้งในการทดสอบทางการศึกษาและจิตวิทยา และเป็นวิธีการเฉพาะที่ใช้ในการประเมินความสามารถและคุณลักษณะแฝงอื่นๆ ในการเลือกข้อสอบจากคลังข้อสอบที่ให้ผู้สอบได้รับข้อสอบตามความสามารถที่แท้จริงของตัวเอง (Frey & Seitz, 2009) เป็นการลดจำนวนข้อสอบให้ผู้สอบตอบข้อสอบและเพิ่มความแม่นยำของการวัด เมื่อเปรียบเทียบกับการทดสอบแบบดั้งเดิมที่มีจำนวนข้อสอบคงที่ในแบบทดสอบ (Fixed Item Test: FIT) และเป็นการขยายแนวคิด จากโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model : MIRT) ซึ่งถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบ ส่งผลต่อการตอบข้อสอบ และการทดสอบแบบปรับเหมาะกำลังได้รับความนิยมอย่างมากในการนำไปปฏิบัติ จึงเกิดการรวมกันของทั้งสองวิธีนี้เรียกว่า การทดสอบปรับเหมาะแบบพหุมิติ (Multidimensional Adaptive Testing: MAT) (Segall, 1996, p. 62; Reckase, 2010, p. 95)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติเป็นการขยายมิติความสามารถการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ (CAT) สู่การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) ซึ่งผู้เขียนได้นำเสนอสาระสำคัญเกี่ยวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ คือ แนวคิดพื้นฐานของลักษณะพหุมิติ, การวิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติ, ลักษณะของโมเดลการตอบสนองข้อสอบแบบพหุมิติ และองค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ ดังนี

3.1 แนวคิดพื้นฐานของลักษณะพหุมิติ

ลักษณะความเป็นพหุมิติมีข้อตกลงเบื้องต้นที่ว่า คุณลักษณะที่จะทำการวัดมีความสัมพันธ์กันซึ่งมีอยู่ 2 ลักษณะ คือ 1) ความเป็นพหุมิติระหว่างข้อคำถาม (between-items multidimensionality) เป็นข้อคำถามวัดคุณลักษณะแฝงเดียว มีหลายคุณลักษณะแฝง ทั้งนี้แต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน และ 2) ความเป็นพหุมิติภายในข้อคำถาม (within-items multidimensionality) เป็นของข้อคำถามวัดหลายคุณลักษณะแฝง มีหลายคุณลักษณะแฝง ซึ่งเป็นแนวคิดพื้นฐานของลักษณะพหุมิติ (ชัยวิชิต เชียรชนะ, 2552; เมษา นวลศรี, 2559; สุกัญญา บุญศรี, 2560) โมเดลการตอบสนองข้อสอบแบบพหุมิติถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบ ส่งผลต่อการตอบข้อสอบ ดังนั้นพารามิเตอร์ความสามารถของผู้สอบ จึงมีตั้งแต่ 2 พารามิเตอร์ขึ้นไป ดังนั้น การพิจารณาความสามารถหลายมิติของบุคคลน่าจะช่วยทำให้โมเดลมีความสอดคล้องกับข้อมูลได้ดียิ่งขึ้น ทั้งนี้แต่ละคุณลักษณะแฝงมีความสัมพันธ์กันเมื่อผู้สอบใช้ความสามารถหลายมิติที่แตกต่างกัน อย่างเป็นระบบในการตอบข้อสอบย่อมมีความสัมพันธ์กับระดับ

ความยากและอำนาจจำแนกของข้อสอบรายข้อ และมีผลต่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง โดยมีสัญลักษณ์ของค่าต่าง ๆ ในโมเดล ดังนี้ (ศิริชัย กาญจนวาสี, 2555, น. 82-83)

X_{ip} = ผลการตอบข้อสอบข้อที่ i ของบุคคล p (0=ผิด, 1=ถูก)

θ_{pm} = ค่าระดับความสามารถหรือคุณลักษณะแฝงของผู้สอบ p (m≥2)

b_i = ค่าความยากของข้อสอบ ข้อที่ i

δ_i = ค่าจุดตัดความง่าย (easiness intercept) ของข้อสอบ ข้อที่ i

a_{im} = ค่าอำนาจจำแนกของข้อสอบข้อที่ i บนมิติ m

c_i = ค่าโอกาสการเดาข้อสอบได้ถูก (guessing) ของข้อสอบข้อที่ i

จากแนวคิดและความสำคัญของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model : MIRT) ทำให้นักการศึกษาหลายท่านเริ่มหันมาสนใจใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory) ในงานวิจัยมากขึ้น โดยได้นำวิธีการนี้มาผสมผสานเข้ากับการทดสอบปรับเหมาะแบบพหุมิติ (Multidimensional Adaptive Testing: MAT) และได้พัฒนาต่อเนื่องจนถึงปัจจุบันและเมื่อนำเทคโนโลยีคอมพิวเตอร์มาใช้ในกระบวนการทดสอบ จึงเรียกว่าการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) ซึ่งวิธีการทดสอบดังกล่าวนี้เป็นการทดสอบที่มีประสิทธิภาพในการวัดสูง ไม่ว่าจะเป็นความแม่นยำในการวัด การลดความยาวของข้อสอบ และเป็นการวัดที่ตรงกับสภาพจริงของลักษณะข้อสอบ

3.2 องค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ เป็นการขยายมิติความสามารถของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติสู่การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ ซึ่งการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติด้วยคอมพิวเตอร์ ประกอบด้วยขั้นตอน 5 ขั้นตอน ได้แก่ 1) การสร้างคลังข้อสอบ (Item pools) 2) การเริ่มต้นการทดสอบ (Initiation/Starting point) 3) การคัดเลือกข้อสอบข้อถัดไป(Item selection algorithm/ Item selection) 4) การประมาณความสามารถ (Ability Estimation/Scoring algorithm) และ 5) เกณฑ์การยุติการทดสอบ (Termination criterion/Stopping rules) (Veldkamp & Matteucci, 2013; Thompson & Weiss, 2011) โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

1. คลังข้อสอบ (Item Pools)

คลังข้อสอบ (Item Pools, Item Bank) เป็นแหล่งรวบรวมข้อสอบหรือคำถามแบบทดสอบ การทดสอบด้วยแบบทดสอบดั้งเดิม ผู้สอบทั้งหมดจะได้รับข้อสอบชุดเดียวกันแต่สำหรับการทดสอบแบบปรับเหมาะผู้สอบจะได้รับข้อสอบเหมาะสมกับความสามารถของผู้สอบแต่ละคน

ดังนั้นต้องใช้ข้อสอบจำนวนมากสำหรับสร้างคลังข้อสอบ ทำให้คลังข้อสอบสำหรับการทดสอบแบบปรับเหมาะมีขนาดใหญ่กว่าการทดสอบดั้งเดิม (Thompson & Weiss, 2011; Babcock & Weiss, 2012; ชัยวิชิต เชียรชนะ, 2552)

2. การเริ่มต้นการทดสอบ (Starting the First Item)

การกำหนดจุดเริ่มต้นการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีหลายวิธี โดยทั่วไปจะใช้ข้อสอบข้อแรกทีมีค่าความยากปานกลางจัดให้กับผู้สอบแต่ละคน เมื่อไม่ทราบค่าสารสนเทศเบื้องต้นหรือระดับความสามารถของผู้สอบ จึงจำเป็นต้องพิจารณาว่าจะใช้เกณฑ์ใดในการเริ่มต้นการทดสอบซึ่งแนวทางในการกำหนดเกณฑ์เริ่มต้นการทดสอบมีหลากหลายแนวทาง ดังนี้ (Becker et al., 2008; Thompson & Weiss, 2011; kirisci et al., 2012)

1) ใช้ข้อสอบทีมีค่าความยากปานกลางจัดให้กับผู้สอบแต่ละคนเมื่อไม่ทราบค่าสารสนเทศเบื้องต้นของผู้สอบ

2) ทำการทดสอบเบื้องต้นก่อนการทดสอบจริงและนำผลมากำหนดจุดเริ่มต้นการทดสอบ

3) หากไม่มีข้อมูลอื่นประกอบการพิจารณา นิยมใช้ค่าเฉลี่ยคุณลักษณะแฝงหรือระดับความสามารถของประชากร (Average Theta) ซึ่งมีค่าเท่ากับ 0.00 เป็นจุดเริ่มต้นในการเลือกข้อสอบข้อแรกให้เหมาะสมกับระดับความสามารถของผู้สอบดังกล่าว

4) ใช้วิธีการสุ่มเลือกข้อสอบด้วยโปรแกรมคอมพิวเตอร์

5) หากมีผลการทดสอบของมาตรวัดนั้นๆ ในครั้งก่อนหน้าอาจเริ่มต้นด้วยค่าคุณลักษณะแฝงหรือระดับความสามารถของผู้สอบที่ได้จากครั้งก่อน

6) พิจารณาจากข้อมูลอื่นๆ เช่น ผลการทดสอบจากมาตรวัดอื่น เกรดเฉลี่ย แรงจูงใจในการทดสอบ และฐานะทางสังคม เป็นต้น

3. การคัดเลือกข้อสอบ (Item Selection)

สำหรับการคัดเลือกข้อสอบมีเป้าหมายเพื่อลดความแปรปรวนของเวกเตอร์ความสามารถแบบพหุมิติชั่วคราว (Frey & Seitz, 2009) วิธีการคัดเลือกข้อสอบที่ใช้กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติอาศัยภาวะน่าจะเป็นสูงสุดหรือต่ำสุดจากการประมาณค่า (θ) แต่ละวิธีสามารถใช้ร่วมกันกับวิธีการประมาณค่าทุกแบบและยังสามารถใช้วิธีการคัดเลือกข้อสอบหลายวิธีร่วมกันได้อีกด้วย (Reckase, 2010, p. 327) โดยมีวิธีการคัดเลือกที่สำคัญ ดังนี้

3.1 Maximize the Determinant of the Fisher Information Matrix การคัดเลือกข้อสอบด้วยเมทริกซ์สารสนเทศฟิชเชอร์ ดำเนินการภายใต้เงื่อนไขการประมาณค่าเวกเตอร์ความสามารถที่กำหนดจากเวกเตอร์ค่าเฉลี่ยเท่ากับ 0 และเวกเตอร์ความสามารถจริง (Ture θ) มี

การแจกแจงแบบปกติหลายตัวแปรด้วยเมทริกซ์ความแปรปรวนและความแปรปรวนร่วม (Variance–Covariance Matrix) ที่เกี่ยวข้องกับเมทริกซ์สารสนเทศของฟิชเชอร์ (Fisher Information Matrix) (Segall, 1996; Frey & Seitz, 2009; Reckase, 2010, p. 330)

3.2 Largest Decrement in the Volume of the Bayesian Credibility Ellipsoid วิธีการประมาณค่าแบบเบส์ ที่นำเสนอโดย Segall (2010, pp. 65- 74) นำไปสู่การคัดเลือกข้อสอบที่จะลดปริมาณของข้อสอบได้อย่างมาก (Frey & Seitz, 2009; Reckase, 2009, pp. 331-332)

3.3 Maximize Kullback-Leibler Information วิธีการ Posterior Expected Kullback–Leibler Information ถูกนำเสนอ ให้ใช้ในการคัดเลือกข้อสอบในการทดสอบแบบปรับเหมาะที่เป็นมิติเดียว Chang and Ying (1996) และต่อมา Veldkamp and van der Linden (2002) ได้เสนอทางเลือกให้ใช้ วิธีการ Posterior Expected Kullback–Leibler Information ในการคัดเลือกข้อสอบแบบปรับเหมาะแบบพหุมิติ (MCAT)โดยมีเกณฑ์การคัดเลือกข้อสอบจากระยะห่างของการกระจายของการตอบสนอง จากข้อสอบข้อที่จะถูกเลือกให้ใช้ทดสอบข้อต่อไปเป็นตำแหน่งที่ใช้ประมาณค่าความสามารถของผู้สอบกับความสามารถที่แท้จริงของผู้สอบ (θ) ด้วยการคาดหมายเอาจากการตอบข้อสอบ (Wang, Chang, and Boughton, 2012; Mulder, van der Linden, 2009, p. 85; Reckase, 2009, pp. 334-335)

4. การประมาณค่าความสามารถ (Ability Estimation)

การประมาณค่าความสามารถของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) มีเกณฑ์การประมาณค่าที่แตกต่างกันหลายแนวทาง เช่น Maximum Likelihood Criterion, Maximum a Posteriori Bayesian Criterion, Newton-Raphson Method และ Least Squares Criterion (Reckase, 2009, p. 139) แต่โดยทั่วไปการประมาณค่าความสามารถของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) ที่นิยมมีอยู่ 2 วิธี คือ วิธีการประมาณค่าแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) และวิธีการประมาณค่าแบบเบส์ (Bayesian Estimation) (Reckase, 2009, p. 314)

การประมาณค่าความสามารถของผู้สอบด้วยวิธีเบส์ที่ปรับใหม่ (Bayesian Updating) ซึ่งเสนอโดยโอเวน (Owen) เป็นวิธีการที่ได้รับความสนใจมากจนถึงปัจจุบันนี้ วิธีการนี้ใช้ความรู้เกี่ยวกับโครงสร้างของความแปรปรวนร่วมระหว่างมิติของการวัดสำหรับการเลือกข้อสอบและการประมาณค่าความสามารถ ส่งผลให้จำนวนข้อสอบที่ใช้ลดลง ซึ่งเป็นคุณสมบัติที่จำเป็นของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ และการประมาณค่าความสามารถของผู้สอบที่ให้ค่าเที่ยงตรงเชิงสภาพสูง (รังสรรค์ มณีเล็ก, 2540)

5. การยุติการทดสอบ (Stopping Rule)

กฎการยุติการทดสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยทั่วไปใช้กฎการยุติการทดสอบมี 2 ลักษณะได้แก่ กฎความยาวคงที่ (Fixed-Length) และ จำนวนข้อสอบผันแปร (variable length) (Reckase, 2009, p. 335) และการพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard error of estimation: SE) (Thissen, 1990) แต่การกำหนดเกณฑ์การยุติการทดสอบขึ้นอยู่กับวัตถุประสงค์ในการทดสอบลักษณะคลังข้อสอบและการควบคุมการปฏิบัติงานต่างๆ เกณฑ์ในการพิจารณาการยุติการทดสอบมีหลากกลายเกณฑ์ให้เลือกใช้ (Wainer et al., 2001, p. 113; Thompson & Weiss, 2011; Kuo et al., 2015) ดังนี้

5.1 ยุติการทดสอบเมื่อข้อสอบในคลังข้อสอบหมด

5.2 ยุติการทดสอบเมื่อการประมาณค่าคุณลักษณะแฝงมีการเปลี่ยนแปลงเล็กน้อยหรือไม่เปลี่ยนแปลงเลย

5.3 กฎความยาวคงที่ (Fixed-Length) และ จำนวนข้อสอบผันแปร (variable length) (Reckase, 2009, p. 335) สำหรับงานวิจัยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) ที่ผ่านมานั้น ส่วนใหญ่จะใช้การจำกัดจำนวนข้อสอบที่ 10, 20, 30, 40, 50 และ 60 ข้อ (Chen, 2009; Diao & Reckase, 2009; Finch, 2010) ซึ่งงานวิจัยของ Diao and Reckase (2009) ปรากฏว่าการใช้ข้อสอบจำนวน 50 ข้อ ความลำเอียงจะน้อยกว่าการใช้ข้อสอบ จำนวน 20 ข้อ แต่ยังระบุไม่ได้ว่าจำนวนข้อสอบเท่าใดจึงจะเหมาะสมสำหรับการทดสอบปรับเหมาะแบบพหุมิติ

5.4 การพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of estimation: SE) โดยทั่วไปจะใช้กฎการยุติการทดสอบจากค่าความคลาดเคลื่อนมาตรฐานกำหนดไว้ที่ 0.3 (Thissen, 1990) ส่วนงานวิจัยที่ผ่านมาจะกำหนด SE ≤ .55, .50 และ 0.30 (Seo, 2011; Thissen, 1990; รังสรรค์ มณีเล็ก, 2540; สมประสงค์ เสนารัตน์, 2555; โสพส สุขานนท์สวัสดิ์เสรี ชัดแช้ม และกฤษณะ ชินสาร, 2556; สุกัญญา บุญศรี, 2560)

ดังนั้น จุดสิ้นสุดการทดสอบมีเกณฑ์ในการยุติการทดสอบที่นำเสนอหลากหลายเกณฑ์ เช่น หยุดเมื่อจำนวนข้อสอบที่ทดสอบครบตามจำนวน หรือเมื่อการวัดมีความแม่นยำตามเกณฑ์หรือใช้ทั้งสองเกณฑ์ร่วมกัน ในการศึกษาจากข้อมูลจำลองส่วนใหญ่ การทดสอบแบบปรับเหมาะแบบพหุมิติจะยุติการทดสอบภายหลังการสอบครบตามจำนวนข้อสอบที่กำหนดไว้ล่วงหน้า (Frey and Seitz, 2009, p. 91)

3.3 ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ไวส์ (Weiss, 1974, pp. 78-110) แฮมเบิลตันและสวามินาทาน (Hambleton and Swaminathan, 1985, p. 297) ได้จำแนกการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบออกเป็น 2 ประเภทใหญ่ๆ คือ

1) การทดสอบสองขั้นตอน (two-stage strategies)

2) การทดสอบหลายขั้นตอน (multi-stage strategies)

(1) แบบแยกทางคงที่ (fixed branching model)

(1.1) รูปพีระมิดขนาดขั้นคงที่ (constant step size pyramid)

(1.2) รูปพีระมิดขนาดขั้นแปรผัน (variable step size pyramid)

(1.3) รูปพีระมิดข้างตัด (truncated pyramid)

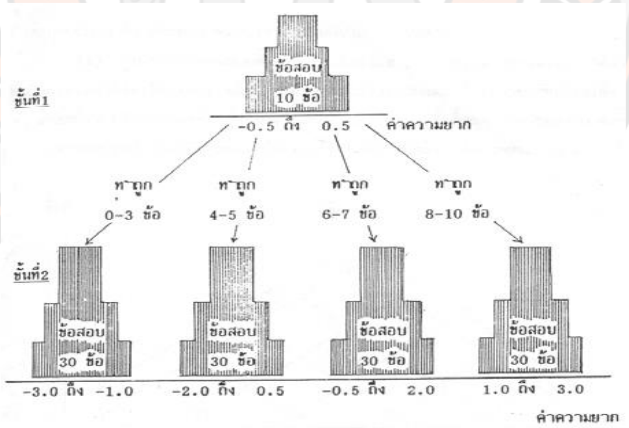
(1.4) รูปพีระมิดที่มีหลายข้อในแต่ละขั้น (multiple item pyramid)

(1.5) รูปพีระมิดที่ให้น้ำหนักแก่ตัวเลือก (differential response option branching)

(1.6) แบบทดสอบเฟล็กซิเลเวล (flexilevel test)

(2) แบบแยกทางแปรผัน (variable branching model)

3.3.1 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบสองขั้นตอนการทดสอบสองขั้นตอนเป็นการทดสอบโดยใช้ข้อสอบ 2 ชุด การดำเนินการสอบเริ่มด้วยให้ ผู้สอบทำแบบทดสอบเพื่อแยกทาง (routing test) ผลจากการทดสอบนี้จะเห็นตัวชี้วัดว่าผู้สอบจะได้รับแบบทดสอบชุดใดในขั้นที่ 2 ซึ่งเรียกแบบทดสอบในขั้นนี้ว่า แบบทดสอบเพื่อวัดผล(Measurement test) โดยหลักการแล้วผู้สอบที่ได้คะแนนน้อยจากการทดสอบแยกทางถือว่าเป็นผู้ที่มีความสามารถต่ำจึงจะได้รับข้อสอบเพื่อวัดผลชุดที่ง่าย ส่วนผู้ที่ได้คะแนนจากการทดสอบเพื่อแยกทางมากหรือมีความสามารถสูง จะได้รับข้อสอบเพื่อวัดผลชุดที่ยากขึ้น ดังภาพที่ 9



ภาพที่ 9 โครงสร้างของการทดสอบเฉพาะบุคคลแบบสองขั้นตอน

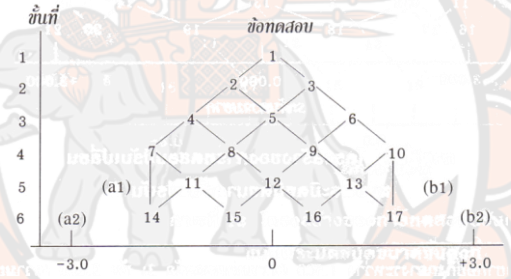
จากภาพที่ 9 จะเห็นว่ามีข้อสอบเพื่อแยกทางในขั้นที่ 1 จำนวน 10 ข้อ ถ้าผู้สอบตอบคำถามถูก 0 – 3 ข้อ จะได้ทำข้อสอบเพื่อวัดผลในขั้นที่ 2 ในชุดที่ง่ายที่สุดจำนวน 30 ข้อ ส่วนผู้ที่มีความสามารถสูงขึ้นก็จะได้รับข้อสอบในชุดที่ยากขึ้นตามลำดับ

3.3.2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบหลายขั้นตอน

การทดสอบหลายขั้นตอน เป็นการทดสอบที่มีโครงสร้างของแบบทดสอบเป็นรูปพีระมิดโดยใช้ค่าความยากง่ายของข้อสอบเป็นเกณฑ์ ในการจัดเรียงข้อสอบที่มีความยาก ง่าย ปานกลาง จะอยู่จุดยอดของพีระมิด การดำเนินการสอบจะเริ่มด้วยให้ผู้สอบตอบข้อสอบข้อแรกที่มีความยากง่าย ปานกลาง ถ้าผู้สอบตอบคำถามถูกข้อสอบข้อต่อไปจะยากขึ้น และถ้าผู้สอบตอบผิดข้อสอบข้อต่อไปก็จะง่ายลงการทดสอบจะดำเนินการเช่นนี้ไปเรื่อยๆ จนกระทั่งถึงเกณฑ์ยุติการสอบตามที่กำหนดไว้การสอบจึงสิ้นสุดลง การทดสอบหลายขั้นตอนสามารถแยกเป็นประเภทย่อยๆ ได้ดังนี้

(1) แบบแยกคงที่

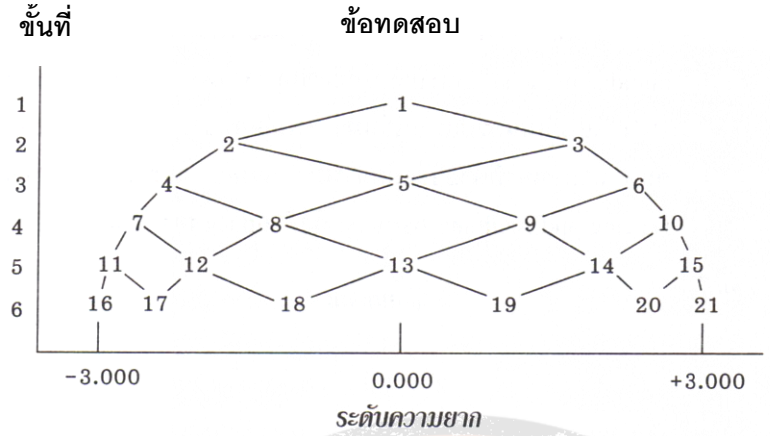
(1.1) รูปพีระมิดขนาดขั้นคงที่เป็นการจัดชุดแบบทดสอบเป็นรูปสามเหลี่ยมโดยมีจำนวนข้อสอบในแต่ละขั้นเท่ากับลำดับที่ของขั้น ผู้สอบจะต้องตอบข้อสอบขั้นละ 1 ข้อ โดยเริ่มตอบข้อ 1 ก่อน ถ้าตอบถูกก็จะไปตอบข้อที่ 3 ซึ่งยากกว่าข้อ 1 แต่ถ้าตอบผิดก็จะไปตอบข้อ 2 ซึ่งง่ายกว่าข้อ 1 ดังภาพที่ 10



ภาพที่ 10 โครงสร้างของการทดสอบเฉพาะบุคคลรูปแบบพีระมิดขนาดขั้นคงที่

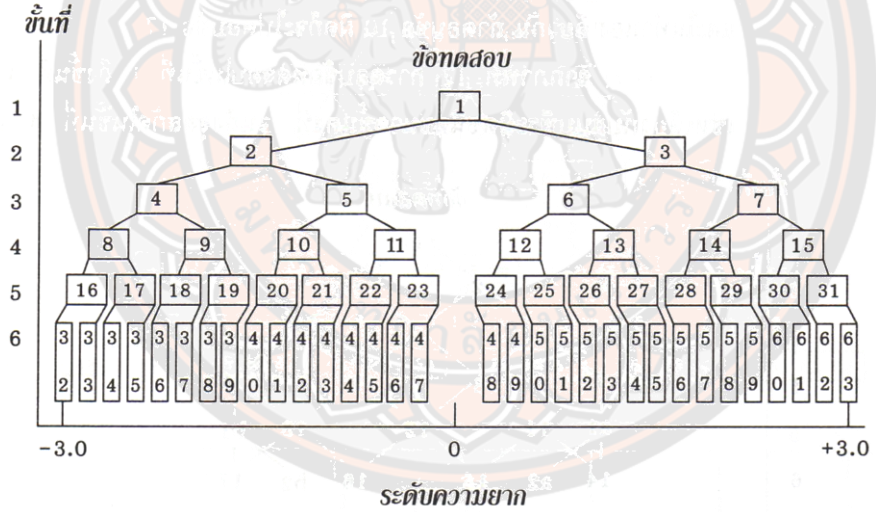
(1.2) รูปพีระมิดขนาดขั้นแปรผัน

การจัดชุดแบบทดสอบลักษณะนี้ ได้กำหนดช่วงห่างของค่าความยากง่ายของข้อสอบในขั้นเดียวกันไม่เท่ากัน โดยข้อสอบที่มีความยากง่ายปานกลางจะมีช่วงห่างของค่าความยากง่ายมากกว่าข้อสอบที่ยากหรือข้อสอบที่ง่าย ดังภาพที่ 11



ภาพที่ 11 โครงสร้างของแบบทดสอบรูปพีระมิดขนาดขั้นแปรผัน

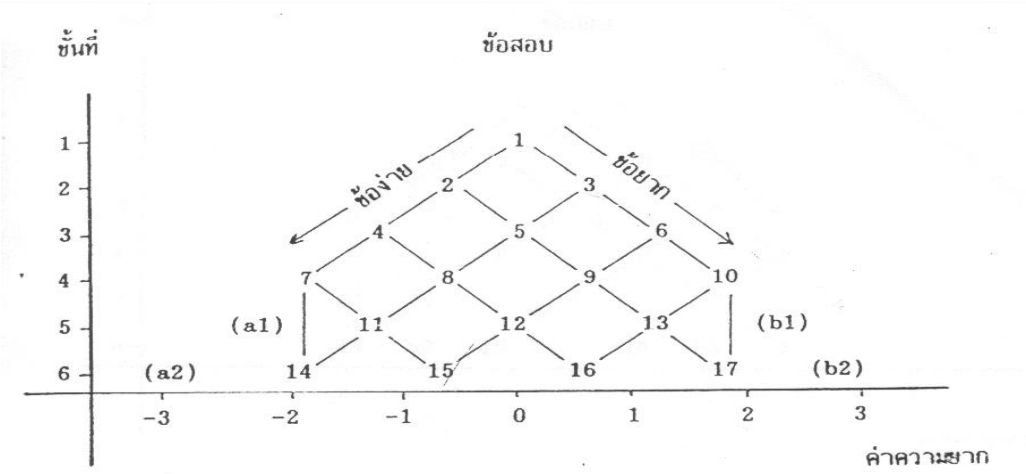
ต่อมาได้มีการพัฒนาโครงสร้างของแบบทดสอบรูปพีระมิดขนาดขั้นแปรผันใหม่ โดยมีการเพิ่มจำนวนข้อสอบในแต่ละขั้นให้มีจำนวนเป็น 2 เท่าของจำนวนข้อสอบ ในขั้นที่ผ่านมารูปแบบใหม่นี้มีชื่อว่าโรบินส์มอนโร (Robbins Monro) ซึ่งมีโครงสร้างดังภาพที่ 12



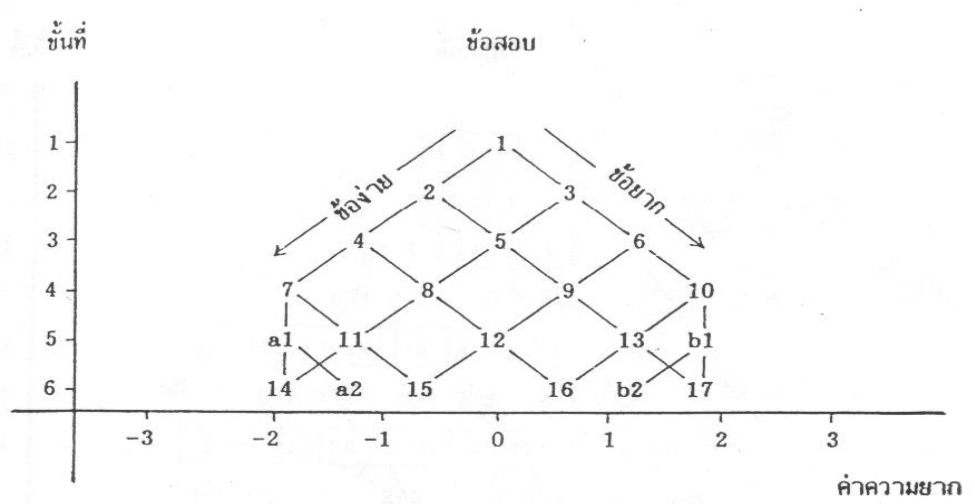
ภาพที่ 12 โครงสร้างของแบบทดสอบรูปพีระมิดแบบโรบินส์มอนโร

(1.3) รูปพีระมิดแบบข้างตัด

เป็นรูปแบบที่มุสสิโอ (Weiss, 1974, as cited in Mussio, 1973) ที่สร้างขึ้นเพื่อลดจำนวนข้อสอบจากโครงสร้างแบบทดสอบรูปพีระมิดแบบขนาดขั้นคงที่ ซึ่งจำแนกเป็นแบบสกัดการสะท้อนกลับ (Reflecting Barrier) และรักษาการสะท้อนกลับ (Retaining Barrier) ดังภาพที่ 13 และ 14 ตามลำดับ



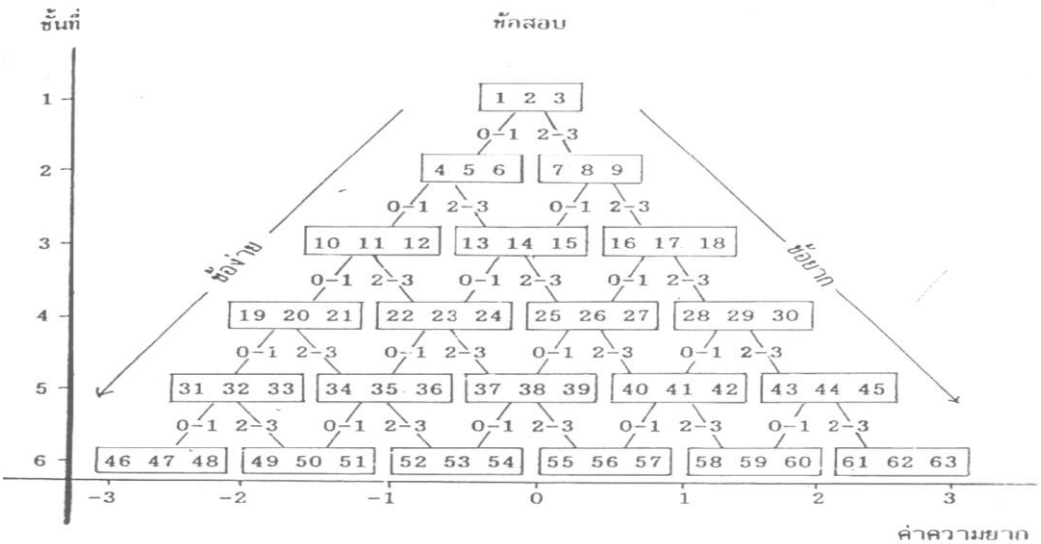
ภาพที่ 13 โครงสร้างของแบบทดสอบรูปพีระมิดแบบสกัดการสะท้อนกลับ



ภาพที่ 14 โครงสร้างของแบบทดสอบรูปพีระมิดแบบรักษาการสะท้อนกลับ

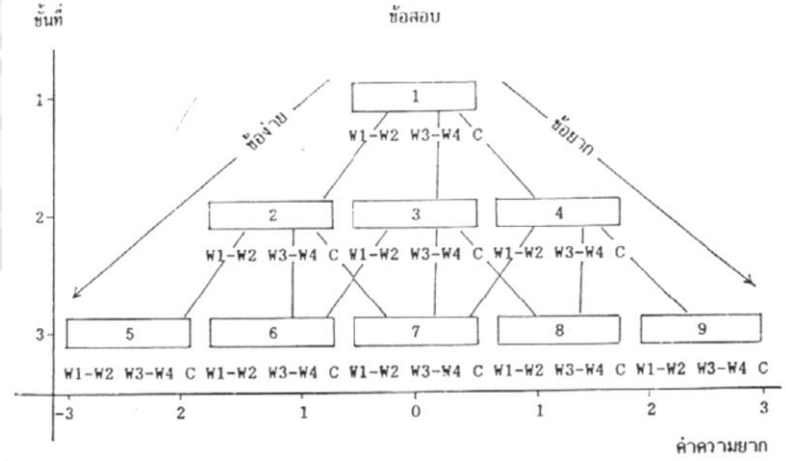
(1.4) รูปพีระมิดแบบมีหลายข้อในแต่ละขั้น

รูปแบบนี้กำหนดให้มีข้อสอบหลายข้อในแต่ละขั้น ทั้งนี้เพื่อให้การแยกทางมีความถูกต้องมากยิ่งขึ้น ซึ่งมีโครงสร้างของการจัดข้อสอบดังภาพที่ 15



ภาพที่ 15 โครงสร้างของแบบทดสอบรูปพีระมิดแบบมีหลายข้อในแต่ละขั้น

(1.5) รูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือกรูปแบบนี้คล้ายกับรูปแบบที่มีหลายข้อในแต่ละขั้น ต่างกันตรงที่จะมีการกำหนดน้ำหนักให้กับตัวเลือกแต่ละข้อดังภาพที่ 16



ภาพที่ 16 โครงสร้างของแบบทดสอบรูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือกของข้อสอบเพื่อแยกทาง

(1.6) การทดสอบเฟล็กซิเลเวล

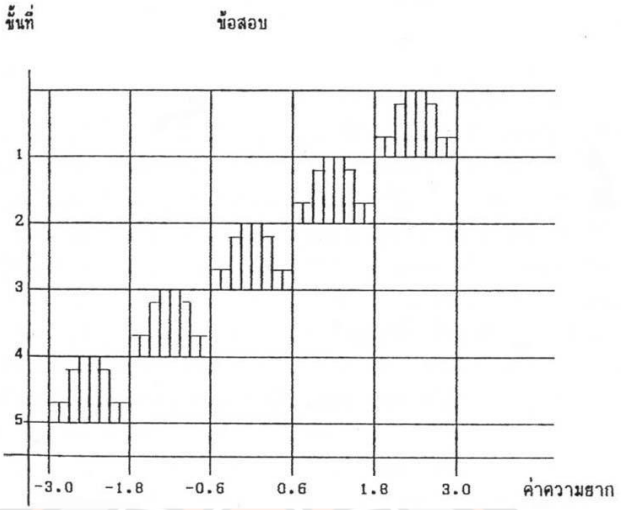
ลอร์ด (Lord, 1980, pp. 114-117) เป็นผู้นำเสนอการทดสอบลักษณะนี้ขึ้นมา การจัดเรียงข้อสอบจะเรียงจากง่ายไปหายาก ข้อที่มีความยากง่ายปานกลางจะถูกนำเสนอแก่ผู้สอบก่อน เช่น ถ้าข้อสอบมีทั้งหมด 75 ข้อ โครงสร้างของการจัดข้อสอบลักษณะนี้สามารถแสดงได้ดังตารางที่ 5

ตารางที่ 5 โครงสร้างของการจัดแบบทดสอบเฟล็กซิเลเวล

ข้อ 0. (มีความยากอยู่กึ่งกลางระหว่างข้อสอบกลุ่มง่ายกับกลุ่มยาก)	
(ข้อสอบกลุ่มง่าย)	(ข้อสอบกลุ่มยาก)
(พิมพ์สีแดง)	(พิมพ์สีน้ำเงิน)
ข้อ 1. (ข้อที่ยากที่สุดในกลุ่มนี้)	ข้อ 1. (ข้อที่ง่ายที่สุดในกลุ่มนี้)
ข้อ 2.	ข้อ 2.
ข้อ 3.	ข้อ 3.
:	:
ข้อ 37. (ข้อที่ง่ายที่สุดในกลุ่มนี้)	ข้อ 37. (ข้อที่ยากที่สุดในกลุ่มนี้)

กระดาษคำตอบของการทดสอบลักษณะนี้จะทำให้ผู้สอบทราบได้ทันทีหลังจากการตอบแต่ละข้อว่า ถ้าตอบผิดจุดแดงจะขึ้นมา (ข้อสอบง่ายกว่าเดิม) ถ้าตอบถูกต้องจุดน้ำเงินจะขึ้นมา (ข้อสอบยากกว่าเดิม) ซึ่งอาจใช้วิธีการเจาะหรือแกะกระดาษที่ทับตัวเลือกแต่ละข้อ ผู้สอบจะต้องตอบข้อสอบครึ่งหนึ่งของข้อสอบทั้งหมด การทดสอบจึงจะยุติวิธีการง่ายๆ เพื่อให้ผู้สอบสามารถยุติการสอบได้ด้วยตนเอง ก็คือ จัดเรียงข้อสอบด้านยากใหม่โดยกลับขั้ว จากง่ายลงไปหายากที่สุด เป็นจากยากที่สุดลงไปหาง่าย (เหมือนการจัดข้อสอบด้านง่าย) การตอบก็จะมีลักษณะเหมือนเดิมคือ ตอบข้อสอบที่ยากง่ายปานกลางก่อน ถ้าตอบผิดก็จะไปตอบข้อที่ง่ายกว่าเดิม ถ้าตอบถูกก็จะไปตอบข้อที่ยากกว่าเดิมตอบไปจนกระทั่งพบว่าในแต่ละแถวของข้อสอบ (มี 2 ข้อ) ถูกตอบครบทุกแถว แถวละ 1 ข้อจึงยุติการสอบ

ต่อมาได้มีการปรับปรุงโครงสร้างของการทดสอบลักษณะนี้ โดยแบ่งความยากง่ายของข้อสอบออกเป็นช่วงๆ ในแต่ละช่วงความยากง่ายจะมีข้อสอบหลายข้อ การจัดแบบทดสอบลักษณะนี้เรียกว่า แบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบบแบ่งชั้น (stradaptive test) ซึ่งมีโครงสร้างของการจัดข้อสอบ ดังภาพที่ 17



ภาพที่ 17 โครงสร้างของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบบแบ่งชั้น

3.4 การทดสอบแบบหลายขั้นตอน (multi-stage strategies) รูปแบบแยกทางแปรผัน (variable branching model)

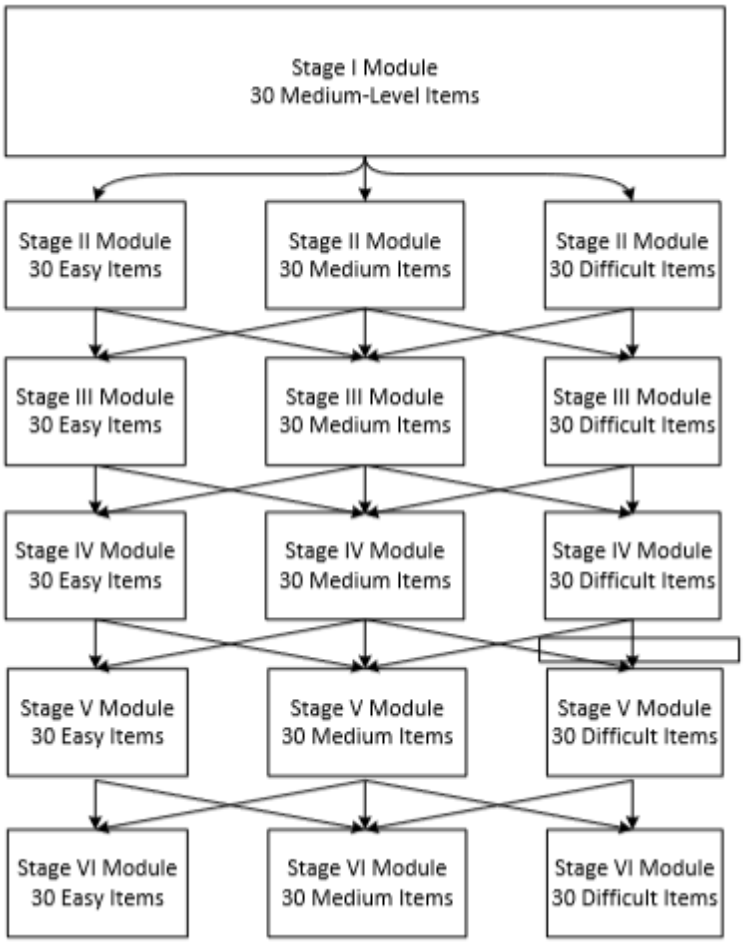
การทดสอบแบบแยกทางแปรผันนี้ ต่างจากการทดสอบแบบแยกทางคงที่ตรงที่ไม่สามารถกำหนดเส้นทางของการตอบข้อสอบได้แน่นอนตายตัว ข้อสอบที่จะตอบข้อต่อไปนั้นขึ้นอยู่กับผลการตอบข้อสอบข้อที่ผ่านมา ซึ่งต้องมีการประมาณค่าความสามารถของผู้สอบทุกครั้งหลังจากตอบข้อสอบแต่ละข้อ รูปแบบของการทดสอบแบบปรับเหมาะแบบหลายขั้นตอน (multi-stage strategies) ซึ่งถูกพัฒนาขึ้นมาตามความต้องการที่หลากหลายขององค์กรการทดสอบที่ต่างกัน รูปแบบนี้ส่วนใหญ่จะแตกต่างกันเกี่ยวกับระดับของการปรับเหมาะที่เกิดขึ้น (Zenisky, 2004, as cited in Brossman & Guille, 2014) สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) ข้อสอบแต่ละข้อจะถูกเลือกขึ้นมาตามความสามารถของผู้สอบก่อนที่จะทำข้อสอบข้อนั้น สำหรับรูปแบบอื่นๆ เช่น การทดสอบแบบหลายขั้นตอน (MST) จะเป็นชุดของข้อสอบที่ถูกเลือกตามความสามารถของผู้สอบก่อนที่จะทาข้อสอบชุดนั้น ซึ่งกลุ่มของรูปแบบหลังๆ จะมีลักษณะคล้ายกับการทดสอบเส้นตรงแบบดั้งเดิมในเรื่องของรูปแบบของแบบสอบในท้ายที่สุดจะถูกบริหารจัดการจากการบริหารการสอบก่อนหน้า (Armstrong & Little, 2003; Guille et al.,2011; Luecht & Sireci, 2011; Zenisky, 2004, as cited in Brossman & Guille, 2014)

ในการบริหารจัดการการทดสอบของวิธี MST นั้นจะเหมือนกับวิธีการแบบเส้นตรงแบบดั้งเดิม แต่แตกต่างกันตรงที่วิธี MST จะอนุญาตให้ในแต่ละรายการทดสอบนั้นสามารถทบทวนรายละเอียดด้านสถิติ เนื้อหา และอื่น ๆ เช่น ส่วนประกอบของกราฟ ภาพและเสียง ก่อนที่จะบริหารจัดการสอบในขั้นต่อไป การอนุญาตนี้เพื่อช่วยควบคุมคุณภาพการติดตามที่สูงกว่าวิธีการที่ใช้ใน CAT

โดยทั่วไป ดังนั้นการใช้วิธี MST จึงเป็นการเพิ่มประโยชน์ขององค์ประกอบด้านการปรับเหมาะที่มากกว่าเมื่อเปรียบเทียบกับวิธีเส้นตรงแบบดั้งเดิม

การดำเนินการของ MST นั้นจะใช้ข้อสอบชุดแรกเป็นกลุ่มๆ หรือเรียกว่า โมดูล (Modules) ก่อนการสอบนั้น เช่น ในแต่ละโมดูลจะมีรายละเอียดของค่าสถิติและเนื้อหา ซึ่งโมดูลของ MST นี้จะถูกสร้างเป็นรูปแบบบล็อก โดยทั่วไปในแต่ละโมดูลจะสร้างตามระดับของเนื้อหาจับคู่กับรายการเนื้อหาในแบบสอบทั้งฉบับ ในส่วนของรายละเอียดด้านสถิตินั้นจะแตกต่างกันระหว่างโมดูล เช่น ความแตกต่างของแต่ละโมดูลจะเกี่ยวข้องกับระดับความยากง่ายของข้อสอบ ยกตัวอย่าง เช่น การทดสอบขององค์กรใช้วิธี MST ซึ่งสร้างโมดูลยาก (โมดูลนี้สำหรับผู้สอบที่มีความสามารถสูง) โมดูลง่าย (โมดูลนี้สำหรับผู้สอบที่ไม่สามารถทำข้อสอบได้ดี) และโมดูลปานกลาง (โมดูลนี้สำหรับผู้สอบที่ทำข้อสอบได้ไม่มากหรือน้อยเกินไป) หลังจากที่แต่ละโมดูลถูกสร้างขึ้น นักพัฒนาแบบสอบจะทบทวนว่าโมดูลเหล่านั้นมีเนื้อหาและค่าสถิติตามนั้นหรือไม่โดยทั่วไประหว่างการบริหารการทดสอบแบบ MST ผู้สอบทั้งหมดจะใช้โมดูลเหมือนกันซึ่งจะใช้โมดูลที่มีความยากของข้อสอบในระดับปานกลาง บนพื้นฐานของการตอบข้อสอบในโมดูลแรกนี้ ผู้สอบแต่ละคนจะถูกประมาณค่าคะแนนความสามารถดังนั้น จะอยู่บนฐานของกฎการตัดสินที่เฉพาะเจาะจงก่อนที่ผู้สอบจะถูกใช้ในโมดูลที่ 2 ต่อไปซึ่งประกอบด้วยข้อสอบที่ยาก ปานกลาง หรือง่าย ซึ่งได้จากการประมาณค่าความสามารถในขั้นตอนที่ 1 ถ้าแบบสอบประกอบด้วยขั้นตอนเพียง 2 ขั้นตอน การทดสอบก็จะยุติหลังจากขั้นตอนที่ 2 อย่างไรก็ตามการทดสอบแบบ MST นั้นโครงสร้างมักจะประกอบด้วยมากกว่า 2 ขั้นตอน (Zenisky, 2004) ในสถานการณ์นี้

หลังจากขั้นตอนที่ 2 แล้ว คะแนนความสามารถของผู้สอบแต่ละคนจะถูกประมาณค่าซ้ำบนพื้นฐานการตอบทั้งหมด ทั้งขั้นตอนที่ 1 และ 2 แล้วผู้สอบก็จะข้ามไปในขั้นตอนที่ 3 ซึ่งมีทั้งข้อสอบยาก ปานกลาง หรือง่าย ขึ้นอยู่กับการประมาณค่าความสามารถของผู้สอบก่อนหน้า กระบวนการนั้นจะทำต่อเนื่องไปจนกระทั่งครบหมดทุกขั้นตอน หลังจากนั้นการทดสอบก็จะสิ้นสุด คะแนนความสามารถของผู้สอบแต่ละคนจะถูกประมาณค่าขึ้นอยู่กับผลการตอบข้อสอบทั้งหมด ดังตัวอย่างในภาพที่ 2 แสดงเส้นทางของรูปแบบ MST ประกอบด้วย 6 ขั้นตอน ขั้นตอนที่ 2–6 ประกอบด้วย 3 โมดูล (ง่าย ปานกลาง ยาก) ซึ่งในแต่ละโมดูลประกอบด้วยข้อสอบ 30 ข้อ โดยจะมีโครงสร้างแบบ 1-3-3-3-3

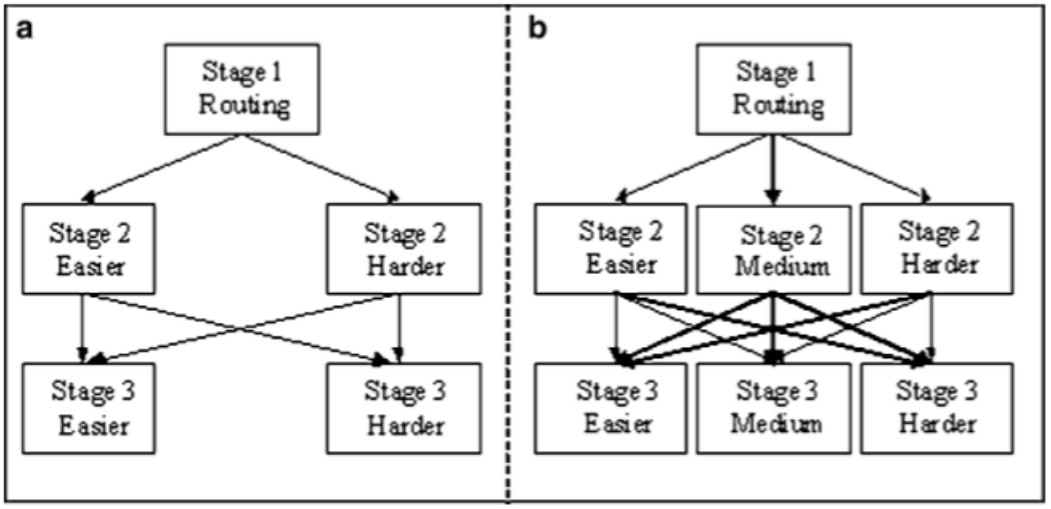


ภาพที่ 18 ตัวอย่างขั้นตอนและกฎเส้นทางของ MST

ที่มา: Brossman and Guille (2014)

นอกจากนี้วิธีการทดสอบแบบหลายขั้นตอน (MST) ยังสามารถกล่าวได้ว่ามีประสิทธิภาพมากสำหรับการทดสอบแบบปรับเหมาะเพื่อการการทดสอบเชิงจำแนก เมื่อเปรียบเทียบกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) รูปแบบ MST มีองค์ประกอบที่เพิ่มความหลากหลาย เช่น กลุ่ม โมดูล ขั้นตอน และเส้นทาง (Luecht, 2000 cited in Ryoungsun Park et al, 2014) ซึ่งโมดูลเป็นการเก็บรวบรวมข้อสอบที่เป็นหน่วยที่เล็กที่สุดของ MST โดยธรรมชาติของการปรับเหมาะของ MST จะเกิดขึ้นหลังจากผู้สอบตอบข้อสอบทั้งหมดในโมดูลนั้น ข้อสอบภายในโมดูลถูกควบคุมด้วยคุณลักษณะค่าสถิติที่เหมือนกัน แม้ว่าในส่วนของเนื้อหาและคุณลักษณะที่ไม่ใช่ค่าสถิติจะแปรเปลี่ยน ขั้นตอนต่อมาของหน่วยใน MST จะยังคงใช้หนึ่งโมดูลหรือมากกว่าได้เนื่องจากผู้สอบจะทำข้อสอบเพียงหนึ่งโมดูลภายในแต่ละขั้นตอน คุณลักษณะของค่าสถิติของโมดูลแปรเปลี่ยนตาม

จุดมุ่งหมายของการทดสอบกลุ่มของหน่วยที่ใหญ่ที่สุดของ MST และยังมีขั้นตอนที่หลากหลายอยู่ ดังภาพที่ 18



ภาพที่ 19 ตัวอย่างขั้นตอนของวิธี MST

จาก: Zenisky, Hambleton, & Luecht, (2010)

ในงานวิจัยเรื่องนี้ได้บูรณาการแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) หลอมรวมกับการทดสอบแบบหลายขั้นตอน (Multi-Stage Test Design) ซึ่งเป็นการผ่อนคลายหรือลดข้อจำกัดในเรื่องของการวัดด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ผู้สอบได้ทำข้อสอบทีละข้อ ให้เป็นการทำข้อสอบทีละชุด (หลายข้อ) ส่งผลให้การประมาณค่าความสามารถที่ได้มีความถูกต้องแม่นยำมากยิ่งขึ้น และใช้รูปแบบแยกทางแปรผัน (variable branching model) ที่ไม่ได้กำหนดเส้นทางของการตอบข้อสอบได้แน่นอนตายตัว ข้อสอบที่จะตอบข้อต่อไปนั้นขึ้นอยู่กับผลการตอบข้อสอบข้อที่ผ่านมา ซึ่งต้องมีการประมาณค่าความสามารถของผู้สอบทุกครั้งหลังจากตอบข้อสอบแต่ละข้อ จะเห็นได้ว่าการทดสอบแบบหลายขั้นตอน (MST) จะมีประสิทธิภาพมากขึ้นเมื่อนำมาใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) โดยช่วยแก้ปัญหาในเรื่องของการประมาณค่าคะแนนความสามารถที่สูงหรือต่ำกว่าความเป็นจริงของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบธรรมดาได้ อีกทั้งการทดสอบแบบหลายขั้นตอนยังช่วยในเรื่องการบริหารจัดการเพื่อควบคุมความเท่าเทียมกันของเนื้อหาที่ต้องการวัดได้ รวมทั้งยังช่วยลดความวิตกกังวลในการทดสอบด้วยการเพิ่มโอกาสในการแก้ไข ทบทวนการตอบของผู้สอบได้ด้วย

3.5 สรุป

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) มีเป้าหมายเพื่อที่จะดำเนินการสอบด้วยการคัดเลือกข้อสอบให้เหมาะสม (Tailoring) กับความสามารถของผู้สอบแต่ละคน โดยอาศัยทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) มาใช้เป็นพื้นฐานในการคำนวณค่าสารสนเทศของข้อสอบ ซึ่งจะมีการคัดเลือกข้อสอบที่สามารถให้สารสนเทศอันเป็นประโยชน์เกี่ยวกับความสามารถของผู้สอบแต่ละคน ทำให้การทดสอบเป็นไปอย่างมีประสิทธิภาพใช้จำนวนข้อสอบที่น้อยลง ประหยัดเวลาของการสอบ ผลที่ได้มีความคลาดเคลื่อนต่ำ มีความแม่นยำของการวัดความสามารถที่แท้จริงของผู้สอบได้อย่างมีประสิทธิภาพ และเหมาะที่จะนำไปใช้กับทดสอบกับผู้เรียนในยุคปัจจุบันนี้

โครงข่ายประสาทเทียม (Artificial Neuron Network)

โครงข่ายประสาทเทียม (Neural network) เป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่เน้นด้านการคำนวณ และได้รับแรงบันดาลใจจากโมเดลทางชีวภาพ โดยมีความพยายามที่จะลอกเลียนความสามารถของสมองมนุษย์ยกตัวอย่างเช่น พัฒนาการด้านสมองของเด็กตั้งแต่แรกเกิดจนถึงสามปีที่ในระยะเริ่มต้นของวัยนี้ เด็กจะเรียนรู้ว่าผู้ใดคือพ่อแม่ หรือคนแปลกหน้า ในระยะต่อมาจะเริ่มตอบสนองเมื่อถูกเรียกชื่อ รู้จักแยกแยะสี และสิ่งของต่างๆ เริ่มเปล่งเสียงพูดเป็นคำ และเริ่มแสดงความคิดเห็น เป็นต้น จากลักษณะดังกล่าวจะเห็นว่าความสามารถในการเรียนรู้ของสมองมนุษย์มีการพัฒนาขึ้นตามลำดับ และทำงานได้อย่างมีประสิทธิภาพ แต่เมื่อเปรียบเทียบกับการทำงานของเครื่องคอมพิวเตอร์จะพบว่ามีการทำงานตามคำสั่งอย่างตรงไปตรงมาเท่านั้น ยกตัวอย่าง เช่น เปิดแฟ้มข้อมูล การใช้อินเตอร์เน็ต และสืบค้นข้อมูล เป็นต้น การทำงานดังกล่าวไม่มีประสิทธิภาพพอที่จะให้เครื่องแสดงความคิดเห็น หรือวิพากย์วิจารณ์เกี่ยวกับข้อมูลต่างๆ ได้ ดังนั้น จะเห็นได้ว่าการทำงานของสมองมนุษย์มีความซับซ้อนเป็นอย่างมาก ซึ่งนักประสาทวิทยาตั้งแต่อดีตถึงปัจจุบันได้พยายามศึกษาค้นคว้ามาอย่างต่อเนื่องที่จะสร้างโปรแกรมคอมพิวเตอร์ที่มีลักษณะการประมวลผลคล้ายการทำงานของสมองมนุษย์และผลการศึกษาค้นคว้าที่ผ่านมา ส่วนหนึ่งก็สามารถนำไปประยุกต์ใช้ในสาขาวิชาต่างๆ ได้อย่างเป็นประโยชน์อย่างมากมาย

รูปแบบของโครงข่ายประสาทเทียมเป็นความก้าวหน้าอย่างหนึ่งของความพยายามที่จะพัฒนาโปรแกรมคอมพิวเตอร์ในปัจจุบันให้สามารถทำนายสิ่งต่างๆ ได้อย่างกว้างขวาง เนื่องจากโครงข่ายประสาทเทียมสามารถหาความสัมพันธ์ของข้อมูลนำเข้า (input) และข้อมูลส่งออก (output) ได้โดยไม่จำเป็นต้องรู้ความสัมพันธ์ทางคณิตศาสตร์ของข้อมูลที่เป็นข้อมูลนำเข้าและข้อมูลส่งออกมาก่อน

4.1 ประวัติความเป็นมาของโครงข่ายประสาทเทียม

ปี พ.ศ. 2486 (Garson, 1998) ถือได้ว่าเป็นปีแห่งการกำเนิดของสาขาโครงข่ายประสาทเทียมในวงการวิทยาศาสตร์ โดยแม็คคัลลอช (Mc Culloch) และพิทส์ (Pitts) ได้เสนอแบบจำลองของเซลล์ประสาท และได้แสดงให้เห็นว่าในทางทฤษฎีนั้น โครงข่ายของแบบจำลองเซลล์ประสาทสามารถทำงานร่วมกับโปรแกรมคอมพิวเตอร์ได้

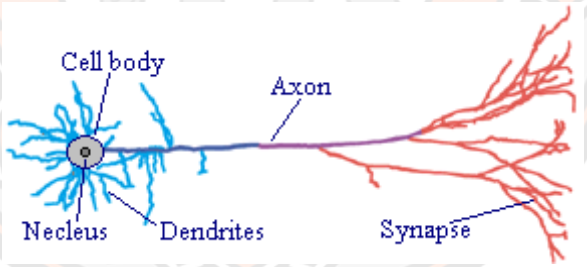
ต่อมาปี พ.ศ. 2492 โดนัลด์ เฮบบ์ (Donald Hebb) ได้เสนอผลงานวิจัยว่า การเรียนรู้ของสมองสามารถอธิบายได้ด้วยรูปแบบของการประกอบเซลล์ประสาทเข้าด้วยกันเป็นโครงข่ายและได้เสนอกฎการเรียนรู้ของเฮบบ์ (Hebb's rule) ที่ทำให้โครงข่ายของเซลล์ประสาทเทียมที่แม็คคัลลอชและพิทส์เสนอไว้ สามารถเรียนรู้ปัญหาง่ายๆ ได้สำเร็จ การเรียนรู้ในรูปแบบของเฮบบ์บนเซลล์ประสาทเทียมของแม็คคัลลอชและพิทส์นั้นเป็นการเรียนรู้แบบ “ไมมีผู้สอน” ซึ่งในทางปฏิบัติแล้วโครงข่ายประสาทเทียมที่เรียนรู้จะพยายามจัดกลุ่มข้อมูลที่โครงข่ายมองว่าคล้ายคลึงกันไปไว้ในกลุ่มเดียวกัน ซึ่งไม่เหมาะสมกับปัญหาประเภทที่ต้องมีการควบคุมกระบวนการเรียนรู้ขณะที่ในช่วงปี พ.ศ. 2490 คอมพิวเตอร์ที่ทำงานเลียนแบบสมองเครื่องแรกของโลกถูกสร้างและทดสอบโดยมินสกี้ (Minsk) แฟรงค์ โรเซ็นแบลทท์ (Frank Rosenblatt) ได้พัฒนาลักษณะโครงข่ายประสาทเทียมขึ้นโดยใช้แบบจำลองของแม็คคัลลอชและพิทส์เป็นแนวทางรวมทั้งเสนอวิธีการเรียนรู้แบบใหม่สำหรับลักษณะโครงข่ายประสาทเทียมด้วย โครงข่ายประสาท เทียมดังกล่าวเรียกว่า เพอร์เซพตรอน (Perceptron) ซึ่งมีการเรียนรู้แบบ “มีผู้สอน” (supervised learning) โดยใช้การปรับความแข็งแรงของการเชื่อมโยง ซึ่งจะพิจารณาได้จากการเปรียบเทียบความรู้ของโครงข่ายประสาทเทียมกับความรู้ของ “ผู้สอน” (teacher) เพอร์เซพตรอนมีความเหมาะสมกับงานประเภท “การระบุชนิด” ซึ่งในระหว่างการเรียนรู้นั้น เพอร์เซพตรอนจะถูกสอนว่าข้อมูลตัวอย่างที่สอนเข้าไปแต่ละแบบนั้นจัดเป็นชนิดใดบ้างหากปัญหาและข้อมูลตัวอย่างมีความเหมาะสมเพอร์เซพตรอนจะสามารถระบุชนิดของข้อมูลที่ไม่เคยเห็นมาก่อนได้ถูกต้อง

ในช่วงปี พ.ศ. 2500 เบอร์นาร์ด วิโดรว (Bernard Widrow) และมาร์เซียน ฮอฟฟ์ (Marcian Hoff) ได้พัฒนาอุปกรณ์ที่เรียกว่า อดาไลน์ (ADALINE; Adaptive Linear combiner) และกฎการเรียนรู้แบบใหม่ที่มีประสิทธิภาพสูงเรียกว่า กฎการเรียนรู้ของวินโดรว-ฮอฟฟ์ (Windrow-Hoff learning rule) ที่เป็นการเรียนรู้แบบ “มีผู้สอน” ซึ่งในเวลาต่อมาอุปกรณ์ดังกล่าวได้รับการขยายแนวคิดไปเป็นมาดาไลน์ (MADALINE; Many ADALINEs) และได้ถูกนำไปประยุกต์ใช้ในการรู้จำรูปแบบ (pattern recognition) การพยากรณ์อากาศ และระบบควบคุมที่จำเป็นต้องมีการปรับเปลี่ยนระบบไปตามสภาพแวดล้อมต่าง ๆ

4.2 ความหมายและหลักการของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม คือ แบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์ โดยที่โครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาท (Neuron) และจุดประสานประสาทหรือไซแนปส์ (Synapses) โดยโครงสร้างของการส่งสัญญาณประสาทประกอบจากการเชื่อมต่อระหว่างเซลล์ประสาทหลายพันล้านเซลล์ เซลล์ประสาทแต่ละเซลล์ประกอบด้วยแขนงรับสัญญาณประสาทซึ่งเป็นเสมือนหน่วยรับข้อมูลป้อนเข้า เรียกว่า “เดนไดรท์” (dendrites) และส่วนปลายของเซลล์ประสาทในการส่งสัญญาณประสาทซึ่งเป็นเสมือนหน่วยส่งข้อมูลออกของเซลล์ เรียกว่า “แอคซอน” (axon) โดยการส่งสัญญาณประสาทดังกล่าว อาจทำให้เกิดได้ทั้งการกระตุ้นและยับยั้ง ทั้งนี้นอกจากลักษณะดังกล่าวแล้ว วิธีการประมวลผลภายในเซลล์ประสาทแต่ละเซลล์ยังมีการขยายหรือลดขนาดของสัญญาณอีกด้วย โดยสัญญาณจากเดนไดรท์ต่างๆ จะรวมกันเข้าสู่เซลล์ประสาท และหากสัญญาณรวมมีความแรงเกินค่าระดับ (threshold) ของเซลล์ประสาทนั้นๆ เซลล์ประสาทก็จะส่งสัญญาณออกทางแอคซอนต่อไป

กระบวนการเรียนรู้ในสิ่งมีชีวิตจะมีผลให้เกิดการสร้างไซแนปส์ระหว่างเซลล์ประสาทขึ้นมาใหม่หรือทำให้เกิดการเปลี่ยนสภาพของไซแนปส์ต่างๆ ในโครงข่ายของเซลล์ประสาท โครงข่ายประสาทของสิ่งมีชีวิตจึงไม่ได้ทำงานแบบเป็นลำดับขั้นตอน (sequential) แต่เพียงอย่างเดียวลักษณะโดยทั่วไปของโครงสร้างระบบประสาทแสดงดังภาพ 2.19



าพที่ 20 ระบบเซลล์ประสาทของมนุษย์

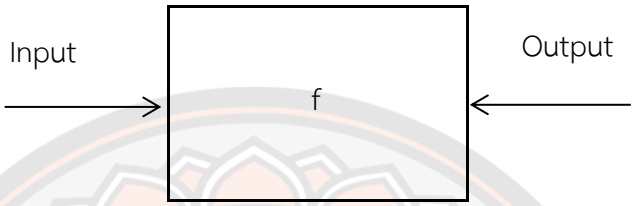
4.3 แนวคิดเบื้องต้น

ปัจจุบันคอมพิวเตอร์มีขีดความสามารถและประโยชน์ในการใช้งานสูง มีหน่วยประมวลผลที่มีความรวดเร็วและแม่นยำในการคำนวณเชิงคณิตศาสตร์ ซึ่งสมองของมนุษย์ไม่อาจเทียบได้ แต่คอมพิวเตอร์ก็ยังมีส่วนด้อยกว่าสมองของมนุษย์อยู่มากในหลายๆ ด้าน เช่น การตีความหมายของรูปภาพ การจดจำใบหน้าคน การจดจำตัวอักษร การแยกแยะลักษณะที่แตกต่างของสิ่งที่มีคุณลักษณะใกล้เคียงกัน และความสามารถในการเรียนรู้ เป็นต้น

แนวคิดการสร้างโครงข่ายประสาทเทียมที่สำคัญ คือ การสร้างโครงข่าย (Network) ที่มีการทำงานใน 2 ลักษณะ ดังนี้

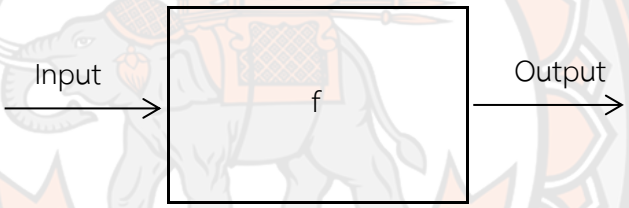
1) Training Mode

- ฝึกสอน โครงข่ายด้วยตัวเอง
- ปรับตัวเอง ให้เป็นฟังก์ชั่นตามลักษณะตัวอย่างของคู่ Input-Output



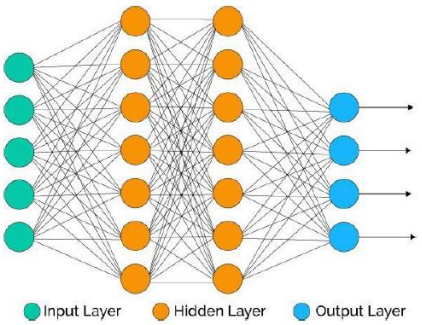
2) Application Mode

- ทำงานตามฟังก์ชั่นที่ผ่านการฝึกสอนมาแล้ว



4.4 องค์ประกอบของโครงข่ายประสาทเทียม

วิธีการสร้างแบบจำลองข้อมูลหรือโมเดลข้อมูล (Data modeling) โดยวิธีโครงข่ายประสาทเทียมอัจฉริยะ (Artificial Neuron Network: ANN) เป็นวิธีการที่มีพื้นฐานมาจากการเลียนแบบการทำงานของสมองมนุษย์ซึ่งประกอบด้วยนิวรอน (Neurons) ที่ใช้ในการเชื่อมโยงเพื่อแก้ปัญหาต่างๆ โดยมีโครงสร้างการทางานตามรูปที่ 4.1 (Zou, Han, and So 2009) ซึ่งประกอบด้วยข้อมูลเข้า (Input) 1 ชั้น และข้อมูลออก (Output) 1 ค่าจะเรียกว่าเพอร์เซพทรอน (Perceptron)



ภาพที่ 21 ส่วนประกอบของโครงข่ายประสาทเทียม

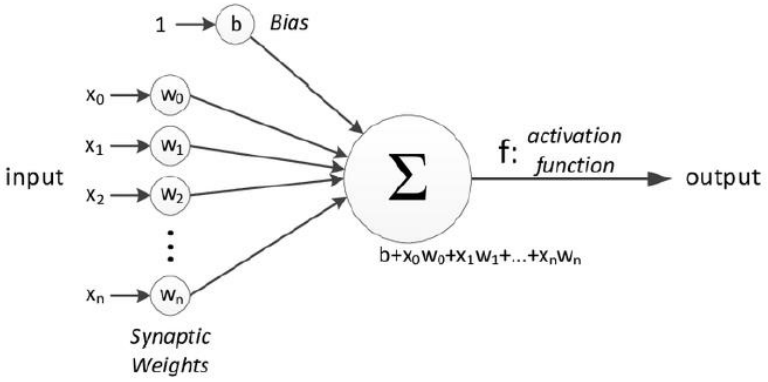
ที่มา: ไกรศักดิ์ เกษร (2564, น. 264)

ส่วนประกอบของโครงข่ายประสาทเทียม จากภาพที่ 20 จะเห็นว่าส่วนประกอบของโครงข่ายประสาทเทียมประกอบด้วย 3 ส่วน

1) ชั้นข้อมูลเข้า (Input Layer) ที่ชั้นนี้จะเป็นข้อมูลเข้า จำนวนของโหนดในชั้นนี้ขึ้นอยู่กับจำนวนของข้อมูลเข้าว่ามีข้อมูลอะไรบ้างที่จะนำเข้ามาคิดในแบบจำลอง เช่น ถ้าข้อมูลของลูกค้าเป็นข้อมูลเข้าที่ประกอบด้วย อายุ เพศ จังหวัดที่อาศัย รวมทั้งสิ้น 4 อย่าง ดังนั้นชั้นข้อมูลเข้าก็จะมี 4 โหนด ซึ่งอาจจะเรียกปัจจัยที่นำมาวิเคราะห์เหล่านี้ว่าคุณลักษณะ (Feature)

2) ชั้นซ่อน (Hidden Layer) เป็นชั้นที่อยู่ระหว่างกลาง ซึ่งจะมีผลอย่างมากต่อประสิทธิภาพในการเรียนรู้ของแบบจำลอง ซึ่งในชั้นซ่อนนั้นจะมีกี่ชั้นก็ได้และแต่ละชั้นจะมีจำนวนของนิวรอนจำนวนเท่าไหร่ก็ได้ เช่นกัน ซึ่งการเพิ่มชั้นและจำนวนนิวรอนจะส่งผลต่อการทางานของโมเดลในส่วนของชั้นซ่อนมีการทำงาน เปรียบเสมือนส่วนที่เรียนรู้ข้อมูลเชิงลึก หรือ deep learning นั่นเอง โดยสิ่งสำคัญในชั้นซ่อนอีกประการหนึ่งคือ ทุกๆ โหนดต้องประกอบด้วยฟังก์แบบไม่เป็นเชิงเส้นซึ่งจะอธิบายต่อไป

3) ชั้นข้อมูลออก (Output Layer) ชั้นที่จะนำเอาข้อมูลจากการคำนวณไปใช้ และจำนวนของโหนดในชั้นนี้ขึ้นอยู่กับรูปแบบของข้อมูลออกที่จะเอาไปใช้ ตัวอย่างเช่น ถ้างานที่ทำเป็นสมการถดถอย (Regression) ก็กำหนดให้ชั้นข้อมูลออกเป็นแบบ 1 โหนด เพราะต้องการคาตอบเพียงค่าเดียว ถ้าเป็นหลายค่าก็เพิ่มไปตามที่ภาพประกอบที่ 8 ประกอบด้วยข้อมูลเข้า x_0 x_1 x_2 x_3 ,....,x_n ซึ่งเข้าทางปลายประสาทหรือเรียกว่า “Synapse” ที่ปลายของประสาทเทียมนี้จะมีค่าน้ำหนักต่างๆ กัน แสดงถึงความแข็งแรงของแต่ละโหนดของปลายประสาท ซึ่งแทนด้วย w_0, w_1, w_2,... w_n และ ค่า b คือค่าความลาเอียง (bias) เพื่อทำให้ฟังก์ชันกระตุ้น(Activation function) ทำงานได้มีประสิทธิภาพมากยิ่งขึ้นต้องการ เช่น ในบางงานอาจจะทำนายหาตำแหน่งของภาพในแกน x และ y พร้อมๆ กัน ในกรณีนี้ก็ต้องกำหนดชั้นข้อมูลออกเป็น 2 โหนด เป็นต้น



ภาพที่ 22 โครงสร้างการทำงานของโครงข่ายประสาทเทียม

ที่มา : ไกรศักดิ์ เกษร (2564, น. 265)

โครงข่ายประสาทเทียมจะนำข้อมูลเข้าและค่าน้ำหนักแต่ละปลายประสาทมาคูณกัน และนำค่าทั้งหมดมารวมกัน หากค่าที่รวมกันนี้เกินกว่าค่าแบ่งที่กำหนดไว้ (threshold: T) จะทำการส่งต่อข้อมูลไปยังฟังก์ชันกระตุ้น ดังสมการ

$$y = f\left[\left(b + \sum_{i=1}^{n}(x_i w_i)\right) - T\right]$$

ฟังก์ชันกระตุ้นหรือเรียกอีกชื่อว่า “ฟังก์ชันการส่งต่อ (Transfer function)” เป็นฟังก์ชันในการคำนวณเพื่อทำนายค่าของข้อมูลออก รูปแบบที่นิยมมากที่สุดและมีประโยชน์คือแบบที่ไม่เป็นฟังก์ชันสมการเส้นตรง (Non-linear function) เนื่องจากปัญหาในโลกความเป็นจริงมีลักษณะเป็นแบบสมการเส้นตรงน้อยมาก (Minsky, Papert and Bottou 2017) ฟังก์ชันกระตุ้นทำหน้าที่ในการตัดสินใจว่านิวรอนควรจะถูกกระตุ้นหรือไม่ โดยดูค่าผลรวมของข้อมูลเข้าและค่าน้ำหนัก ฟังก์ชันกระตุ้นจะถูกนำไปใช้ทั้งโหนดซ่อน (Hidden node) และโหนดข้อมูลออก (Output node) ซึ่งทั้งสองโหนดอาจจะใช้ฟังก์ชันกระตุ้นที่เหมือนหรือต่างกันก็ได้แต่ส่วนมากจะใช้ฟังก์ชันแบบไม่เป็นเชิงเส้น เนื่องจากในโหนดซ่อนจะมีการคำนวณแบบการรวมเชิงเส้น (Linear combination) ถ้าฟังก์ชันกระตุ้นของโหนดซ่อนจะมีการคำนวณแบบเชิงเส้นอีก จะเป็นการทางานซ้ำซ้อนกับการคำนวณแบบการรวมเชิงเส้นในชั้นข้อมูลออก และจะทำให้ผลลัพธ์เทียบเท่ากับสมการถดถอยลอจิสติก ฟังก์ชันกระตุ้นจะมีหลากหลายรูปแบบ ดังต่อไปนี้

1) ฟังก์ชันกระตุ้นค่าแบ่ง (Threshold Activation Function) ฟังก์ชันนี้เรียกอีกชื่อว่าฟังก์ชั่นไบนารี่สเตป (Binary step function) ซึ่งจะพิจารณาค่าข้อมูลเข้าว่ามากกว่าหรือน้อยกว่าค่าแบ่งที่กำหนดไว้(threshold) หรือไม่ เพื่อส่งค่าต่อไปยังชั้นถัดไป

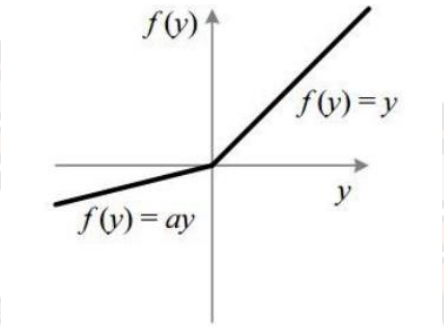
2) ฟังก์ชั่นกระตุ้นซิกมอยด์ (Sigmoid Activation Function) เป็นฟังก์ชันทางคณิตศาสตร์ที่มีลักษณะเป็นตัวเอส “S-curve” หรือเรียกว่า “Sigmoid curve” จะมีค่าระหว่าง 0 และ 1 ฟังก์ชันนี้จะใช้เมื่อต้องการทำนายความน่าจะเป็น (Probability) ของข้อมูลออก

3) ฟังก์ชันไฮเปอร์โบลิคแทงก์เจนท์ (Hyperbolic Tangent Function: tanh) มีการทำงานคล้ายฟังก์ชันกระตุ้นซิกมอยด์ แต่มีประสิทธิภาพดีกว่าจะมีค่าระหว่าง [-1, 1] ข้อดีของฟังก์ชันไฮเปอร์โบลิคแทงก์เจนท์ คือ สามารถแปลงค่าข้อมูลเข้าที่มีค่าเป็นลบมากๆ ให้เป็นข้อมูลออกที่ติดลบได้ และข้อมูลที่ค่าเป็นศูนย์จะถูกแปลงเป็นข้อมูลออกที่มีค่าใกล้ศูนย์ (near-zero output)

4) ฟังก์ชันเรคติไฟด์ลินเนียยูนิต (Rectified Linear Units, ReLu) เป็นฟังก์ชันที่นิยมใช้งานมากที่สุดในโครงข่ายประสาทเทียมแบบบิด (Convolutional Neural Networks: CNN) และโครงข่ายประสาทเทียมอัจฉริยะ (ANN) ฟังก์ชันนี้จะมีค่าอยู่ระหว่าง [∞) หมายถึงถ้าข้อมูลเข้า

มีค่ามากกว่าศูนย์ ข้อมูลออกจากเป็นค่าบวก และถ้าข้อมูลเข้ามีค่าศูนย์หรือติดลบ ข้อมูลออกจะมีค่าเป็นศูนย์

นักวิทยาศาสตร์พิสูจน์แล้วว่าฟังก์ชันเรคติไฟด์ลินเนียยูนิตมีประสิทธิภาพดีกว่าฟังก์ชันไฮเปอร์โบลิคแทงก์เจนท์สูงถึง 6 เท่า (Zou et al., 2009) ข้อเสียของฟังก์ชันเรคติไฟด์ลินเนียยูนิตคือ อาจจะเกิดปัญหานิวรอนตาย (dead neurons) กล่าวคือ ฟังก์ชันกระตุ้นไม่ทำงานไม่ว่าข้อมูลเข้ามีค่าเท่าไหร่ก็ตาม จึงมีการพัฒนาฟังก์ชันเรคติไฟด์ลินเนียยูนิตเป็นฟังก์ชัน “ลีคกี้เรคติไฟด์ลินเนียยูนิต (Leaky ReLu)” ซึ่งจะมีค่าจาก -∞ ถึง +∞ แสดงดังภาพที่ 23



ภาพที่ 23 ฟังก์ชันลีคกี้เรคติไฟด์ลินเนียยูนิต
ที่มา : ไกรศักดิ์ เกษร (2564, น. 267)

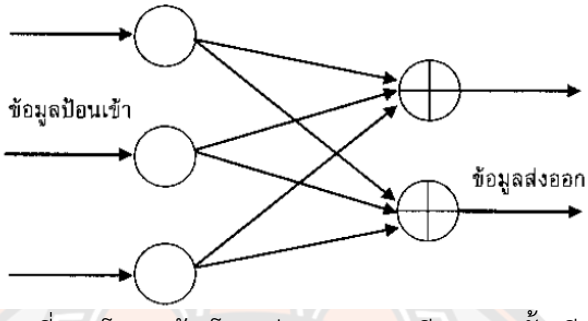
การเลือกฟังก์ชันกระตุ้นสำหรับโหนดข้อมูลออกต้องเลือกสิ่งที่เข้ากับคำตอบที่เราต้องการ เช่น เราต้องการทำนายความน่าจะเป็น ซึ่งคำตอบที่เป็นไปได้คือ เป็น (0) และไม่เป็น (1) เราจึงควรเลือกฟังก์ชันกระตุ้นซิกมอยด์ เพราะฟังก์ชันนี้ให้ผลลัพธ์ระหว่าง 0 ถึง 1

4.5 ลักษณะของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมประกอบด้วยเซลล์ประสาทเทียม หรือโหนดจำนวนมากเชื่อมต่อกัน ซึ่งการเชื่อมต่อแบ่งออกเป็นกลุ่มย่อย เรียกว่า ชั้น (layer) ชั้นแรก เป็นชั้นนำข้อมูลเข้า เรียกว่า ชั้นรับข้อมูลป้อนเข้า (input layer) ส่วนชั้นสุดท้ายเรียกว่าชั้นส่งข้อมูลออก (output layer) และชั้นที่อยู่ระหว่างชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเรียกว่า ชั้นแอบแฝง (hidden layer) ซึ่งโดยทั่วไปชั้นแอบแฝงอาจมีมากกว่า 1 ชั้นก็ได้ ด้วยเหตุนี้จึงสามารถแบ่งประเภทของโครงข่ายประสาทเทียมตามจำนวนชั้นของโครงข่ายแบบกว้างๆ ได้ 2 แบบ ได้แก่ โครงข่ายแบบชั้นเดียว (single layer) และโครงข่ายแบบหลายชั้น (multi-layer)

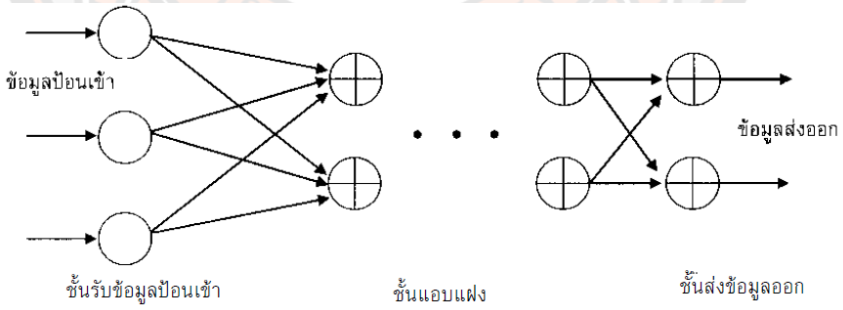
1. โครงข่ายแบบชั้นเดียวโครงข่ายแบบชั้นเดียว เป็นโครงข่ายประสาทเทียมอย่างง่ายที่มีเพียงชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเท่านั้น โหนดในชั้นรับข้อมูลป้อนเข้าทำหน้าที่รับข้อมูลเข้า (input value) แล้วส่งข้อมูลผ่านเส้นเชื่อมโยงต่างๆ ไปให้โหนดในชั้นส่งข้อมูลออกความเข้มของสัญญาณ หรือปริมาณข้อมูลที่นำเข้าสู่โหนดในชั้นส่งข้อมูลออกจะขึ้นอยู่กับค่าน้ำหนักที่อยู่บนเส้น

เชื่อมโยงโหนดในชั้นส่งข้อมูลออกจะนำข้อมูลที่ได้รับมาคำนวณโดยใช้ฟังก์ชันทางคณิตศาสตร์ที่เรียกว่า ฟังก์ชันการแปลง (transfer function) ที่เหมาะสมกับปัญหาแล้วส่งผลลัพธ์ที่ได้ออกมาเป็นข้อมูลส่งออก เช่น โครงข่ายแบบชั้นเดียวแบบเพอเซบตรอนอย่างง่าย (simple perceptron) และโครงข่ายโฮบฟิลด์ (Hopfield networks) ลักษณะโครงข่ายแบบชั้นเดียวแสดงดังภาพที่ 24



ภาพที่ 24 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว

2. โครงข่ายแบบหลายชั้นโครงข่ายแบบหลายชั้น เป็นโครงข่ายที่มีชั้นแอบแฝงตั้งแต่ 1 ชั้นขึ้นไป โครงข่ายแบบหลายชั้นจะใช้ในกรณีที่ปัญหามีความซับซ้อน ซึ่งโครงข่ายแบบชั้นเดียวไม่สามารถแก้ปัญหาได้ จึงเพิ่มจำนวนโหนดที่มีการคำนวณ หรือชั้นแอบแฝงให้กับโครงข่าย ตัวอย่างของโครงข่ายแบบหลายชั้น เช่น การแพร่ย้อนกลับ (back propagation) เซลฟ์ออร์แกนไนซิ่งแมปซ์ (self organizing maps) และเคาน์เตอร์พรอพะเกชัน (counter propagation) เป็นต้น ลักษณะโครงสร้างโครงข่ายแบบหลายชั้นแสดงดังภาพที่ 25



ภาพที่ 25 โครงข่ายประสาทเทียมแบบหลายชั้น

4.6 ประเภทของการเรียนรู้ของโครงข่ายประสาทเทียม

4.6.1 การเรียนรู้แบบมีผู้สอน (supervised learning) ข้อมูลจะประกอบด้วยตัวอย่างข้อมูลที่ต้องการสอนและผลลัพธ์ที่ต้องการให้โครงข่ายสร้าง เมื่อมีการนำข้อมูลในลักษณะเดียวกันมาเป็นข้อมูลป้อนเข้าโครงข่ายจะกำหนดค่าผลลัพธ์ที่เป็นเป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัวโครงข่ายจะนำค่าผิดพลาดระหว่างค่าเป้าหมายกับค่าผลลัพธ์ที่ได้ มาใช้ในการปรับค่าน้ำหนัก เพื่อให้ค่าผลลัพธ์ที่ใกล้เคียงกับเป้าหมายมากที่สุด ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการสอนนักเรียนโดยมี

ครูผู้สอนคอยให้คำแนะนำตัวอย่างแบบจำลองนี้ ได้แก่ การแพร่ย้อนกลับ และเพอเซบตรอน (perceptron) เป็นต้น

4.6.2. การเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) การเรียนรู้แบบนี้จะสอนโครงข่ายโดยการนำข้อมูลป้อนเข้าอย่างต่อเนื่องเพียงอย่างเดียว ไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัว การปรับน้ำหนักจะใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่า โดยค่าน้ำหนักจะปรับตามกลุ่มที่ข้อมูลป้อนเข้าที่มีรูปแบบคล้ายคลึงกันถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการที่เราสามารถแยกแยะพันธุ์พืช พันธุ์สัตว์ ตามลักษณะรูปร่างของมันได้ด้วยตนเอง ตัวอย่างแบบจำลองนี้ได้แก่ เคาน์เตอร์พรอพะเกชัน (counter propagation : CPN) แบบจำลองอะแดปทีฟรีโซแนนซ์เทียรี่ (Adaptive Resonance Theory neural networks : ART) เป็นต้น

4.7 การเรียนรู้ของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมใช้กระบวนการเรียนรู้ข้อมูลโดยการปรับค่าน้ำหนักเป็นค่าที่เหมาะสมที่สุดโครงข่ายประสาทเทียมอัจฉริยะมีการเรียนรู้ 2 แบบ คือ

1) การเรียนแบบมีการสอน (Supervised Learning) เป็นการเรียนแบบที่มีการตรวจคำตอบเพื่อให้โครงข่ายประสาทเทียมปรับตัว ชุดข้อมูลที่ใช้สอนโครงข่ายประสาทเทียมจะมีคำตอบไว้คอยตรวจดูว่าโครงข่ายประสาทเทียมให้คาตอบที่ถูกหรือไม่ ถ้าตอบไม่ถูก โครงข่ายประสาทเทียมก็จะปรับตัวเองเพื่อให้ได้คำตอบที่ดีขึ้น

2) การเรียนแบบไม่มีการสอน (Unsupervised Learning) เป็นการเรียนแบบไม่มีผู้แนะนำไม่มีการตรวจคำตอบว่าถูกหรือผิด โครงข่ายประสาทเทียมจะจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูลผลลัพธ์ที่ได้ โครงข่ายประสาทเทียมจะสามารถจัดหมวดหมู่ของข้อมูลได้ กระบวนการเรียนรู้ของโครงข่ายประสาทเทียมอัจฉริยะถูกพัฒนาขึ้นหลากหลายวิธี เพื่อรองรับจุดประสงค์ในการใช้งานต่างๆ วิธีการที่นิยมใช้มากที่สุดคือวิธีการแก้ไขข้อผิดพลาด (Error correction) และวิธีเพื่อนบ้านใกล้สุด (Nearest neighbor) วิธีการแก้ไขข้อผิดพลาดจะเป็นวิธีการแพร่กลับ (Back propagation) ซึ่งมีการทำงาน (Zou et al., 2009) คือ การเรียนรู้ของโมเดลเกิดขึ้นเมื่อเอาค่าที่ได้จากการคำนวณในของการแพร่ไปข้างหน้า (forward propagation) มาเทียบกับค่าของข้อมูลออกที่เกิดขึ้นจริง (Ground Truth) ค่าความผิดพลาดที่เกิดขึ้นเรียกว่าค่าต้นทุน (Cost) ค่าสูญเสีย (Loss) ค่าผิดพลาด (Error) หรือ ค่าเหลือ (Residual)

ดังนั้นกำหนดให้ข้อผิดพลาดของโหนด k (error: e_k) สามารถคำนวณได้จากค่าความต่างของผลลัพธ์ (y) ของโหนด k ในรอบที่ n แทนด้วยสัญลักษณ์ yk ,n และข้อมูลออกที่เกิดขึ้นจริงของโหนด k แทนด้วยสัญลักษณ์ yk ดังนั้นค่าความผิดพลาดคำนวณได้จากสมการ

$$e_k = y_{k,n} - y_k^*$$

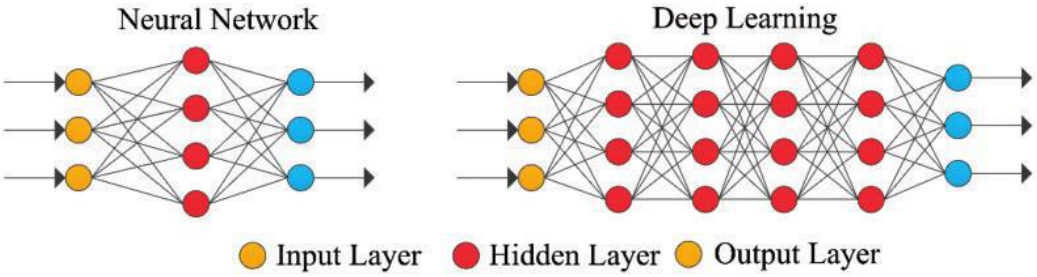
โดยค่าความผิดพลาด e_k ยิ่งมีค่าใกล้ศูนย์ยิ่งดี เป็นตัวอย่างของการคำนวณแบบแพร่ไปข้างหน้า แต่วิธีการแพร่กลับทางานเหมือนสมองคน คือการเรียนรู้จากความผิดพลาด นั่นคือเมื่อรู้ค่าผิดพลาดของโหนด *k* แล้ว จะนำค่าผิดพลาดนั้นมาคำนวณหาค่าน้ำหนักใหม่ (w_{new}) ในรอบที่ n+1 ของโหนด *k* ดังสมการ (2.3)

$$w_{new} = w_{old} - \lambda \frac{\partial E}{\partial w_{old}}$$

โดย λ คือค่าคงที่ในการปรับน้ำหนักซึ่งอาจจะเรียกว่า “Step“ หรือ “Learning rate” ซึ่งทุกๆ รอบของการทำงานจะมีการปรับค่าน้ำหนักใหม่ทุกครั้งจนกว่าค่าน้ำหนักจะไม่เปลี่ยนแปลงหรือเปลี่ยนแปลงน้อย (Convergence) ถือเป็นค่าน้ำหนักที่เหมาะสมที่สุด (Optimization) จะมีผลทำให้การทำนายผลลัพธ์มีความแม่นยำมากยิ่งขึ้น

4.8 การเรียนรู้เชิงลึก

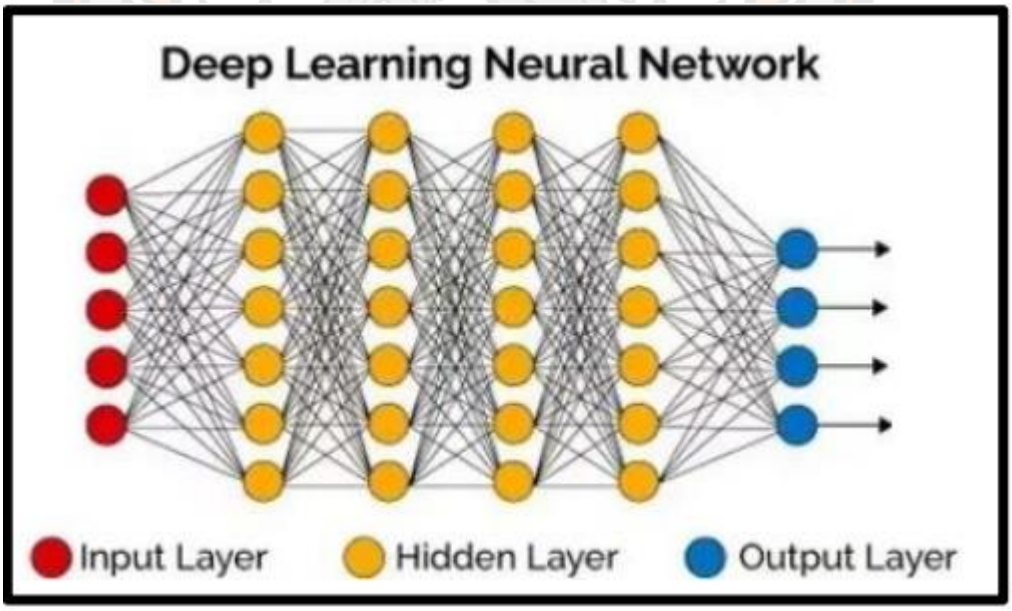
การเรียนรู้เชิงลึก (Deep Learning) คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาทเทียม (Neural Network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะถูกนำไปใช้ในการตรวจจับรูปแบบ (Pattern) หรือจำแนกข้อมูล (Classify the Data) โดยทั่วไปวิธีโครงข่ายประสาทเทียมจะเรียนรู้ได้เพียงไม่กี่ชั้นเพื่อที่จะทำให้ Neural Network ของเรานั้นสามารถคิดและประมวลผลซับซ้อนได้เหมือนสมองมนุษย์ ชั้นที่เป็น Hidden Layer จึงต้องมีหลายๆ ชั้นให้มันส่งข้อมูลประมวลผลต่อๆ กันไป ทำให้มันสามารถคำนวนอะไรที่ซับซ้อนได้เยอะขึ้น เนื่องจากยังไม่มีข้อมูลสอน (Training Data) หรือความสามารถด้านคอมพิวเตอร์ยังไม่สูงพออย่างไรก็ดีไม่นานนี้เทคโนโลยีได้มีการพัฒนามากขึ้นคอมพิวเตอร์มีประสิทธิภาพมากยิ่งขึ้น จึงทำให้มีข้อมูลชั้นของโครงข่ายได้ง่ายขึ้นและมากขึ้น ยิ่งมีซ้อนกันหลายชั้นโครงข่ายก็ยิ่งมีความซับซ้อนและลึกขึ้น จึงเป็นที่มาของคำว่า การเรียนรู้เชิงลึก หรือ “Deep Learning” ตามรูปแบบของการเรียนรู้เครื่องจักร โดยทั่วไปเมื่อมีข้อมูลดิบเข้ามาจะไม่มีการประมวลโดยอัตโนมัติ แต่จะต้องอาศัยความรู้เฉพาะทาง (Domain Knowledge) สำหรับคุณลักษณะในการจัดหมวดหมู่ข้อมูลบางประเภท (Hand-Craft Features)



ภาพที่ 26 ความแตกต่างระหว่างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก

ที่มา: ไกรศักดิ์ เกษร (2564, น. 269)

เทคนิคในการปรับค่าน้ำหนักใหม่ของการเรียนรู้เชิงลึกจะใช้วิธีการแพร่กลับ เช่นเดียวกับโครงข่ายประสาทเทียมอัจฉริยะ (ANN) อย่างไรก็ตามวิธีการแพร่กลับไม่ได้เป็นวิธีการที่ดีที่สุดเพราะวิธีการดังกล่าวจะมีปัญหาในเรื่องของการอัพเดทข้อมูลในกรณีที่มีชั้นซ่อนหลายๆ ชั้น และตอนที่ส่งค่าผิดพลาดกลับ (เอาค่าน้ำหนักใหม่กลับไปปรับ) (กานต์, 2018) ก็คือการนำค่าผิดพลาดไปคูณกับข้อมูลใหม่อีกรอบ ยิ่งคูณกันไปเรื่อยๆ ทำให้จุดทศนิยมจะยิ่งมีค่าน้อยลงไปเรื่อยๆ ตัวอย่าง เช่น จาก 13.23 คูณไปเรื่อยๆ จะกลายเป็น 12.402383 และไปอีกเรื่อยๆ กลายเป็น 12.402391 ซึ่งจะพบว่าค่าจะเปลี่ยนไปไม่มาก คือเปลี่ยนในระดับทศนิยม ซึ่งหมายความว่าการเรียนรู้ของ ANN มันยิ่งเรียนรู้ได้น้อยลงเรื่อยๆ นั่นเอง



ภาพที่ 27 โมเดล Neural Net อัลกอริทึม Deep Learning

ที่มา: G.E. Hinton et al., (2012)

G.E. Hinton, N.Srivastava, A.Krizhevsky, I. Sutskever, and R.R. Salakhutdinov. (2012) *Improving neural networks by preventing co-adaptation of feature detectors*. ArXiv e-prints,July 2012.

4.9 การเรียนรู้แบบแพร่ย้อนกลับ

Werbos (1974) ได้เสนอแนวคิดของการเรียนรู้แบบแพร่ย้อนกลับ จากนั้น Parker (1982) ได้นำเสนออีกครั้ง และเป็นที่รู้จักมากยิ่งขึ้นโดย Rumelhart และคณะ (1986) ในหนังสือ Parallel Distributed Processing ซึ่งได้กล่าวถึงศักยภาพของโครงข่ายประสาทเทียม และการเรียนรู้แบบแพร่ย้อนกลับการเรียนรู้แบบแพร่ย้อนกลับ สามารถแก้ปัญหาที่ต้องการรูปแบบโดยการป้อนรูปแบบเข้าไป โครงข่ายประสาทเทียมจะให้รูปแบบผลลัพธ์ที่เกี่ยวข้องกัน (Dayhoff, 1990) การเรียนรู้แบบแพร่ย้อนกลับเป็นวิธีการหนึ่งของโครงข่ายประสาทเทียมที่ง่ายต่อการเข้าใจ เนื่องจากกระบวนการเรียนรู้ และปรับปรุงแก้ไขนั้นเป็นไปด้วยตนเองถ้าโครงข่ายประสาทเทียมให้คำตอบที่ผิดค่าน้ำหนักจะถูกปรับจนกว่าค่าความผิดพลาดจะน้อยลง หรืออยู่ในเกณฑ์ที่ยอมรับได้ นั่นคือ ค่าที่ได้ในครั้งถัดไปจะมีความถูกต้องมากยิ่งขึ้น ด้วยโครงสร้างประสาทเทียมที่มีลักษณะเป็นชั้น แต่ละชั้นเชื่อมโยงกันอย่างทั่วถึง เมื่อโครงข่ายประสาทเทียมได้รับข้อมูลป้อนเข้า จะคำนวณค่าน้ำหนักของหน่วยรับข้อมูลป้อนเข้าไปยังชั้นแอบแฝง และจากชั้นแอบแฝงไปยังชั้นส่งข้อมูลออกเมื่อเกิดผลต่างระหว่างค่าผลลัพธ์จริงกับค่าผลลัพธ์เป้าหมาย โครงข่ายประสาทเทียมจะปรับค่าความผิดพลาดจากชั้นส่งข้อมูลออก และแพร่ย้อนกลับไปยังชั้นแอบแฝงจากนั้นจึงแพร่ย้อนกลับไปยังชั้นรับข้อมูลป้อนเข้าตามลำดับ

จากปัญหาดังกล่าวนักวิจัยจึงปรับเปลี่ยนวิธีการเทรนใหม่ขึ้นมาใหม่โดยมีแนวคิดคือ ให้ชั้นแรกๆ เรียนรู้ข้อมูลในระดับต่ำ (Low Level feature) เท่านั้น และในชั้นต่อๆ ไปก็จะเอาข้อมูลมาประกอบกัน ถ้าเปรียบเทียบเป็นข้อมูลรูปภาพสุนัข ในชั้นแรกๆ ก็อาจจะเห็นข้อมูลเช่น จมูกสุนัข หูสุนัข ปากสุนัข และชั้นถัดๆ ไปเมื่อประกอบกันไปเรื่อย ก็จะเห็นเป็นรูปหน้าสุนัข หรือสรุปคือ การแยกส่วนกันเรียนรู้ในชั้นแรกๆ และนำมาประกอบกันไปเรื่อยๆ ในชั้นที่ลึกขึ้นนั่นเอง รูปแบบอื่นๆ ของโครงข่ายประสาทเทียม (มัสกานต์ เสน่หา, 2559) แต่ในงานวิจัยนี้จะขอยกตัวอย่างเฉพาะบางวิธีที่นิยมใช้กันในกลุ่มนักวิจัย

1) โครงข่ายประสาทเทียมคอนโวลูชัน

โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN) เป็นวิธีการที่ถูกนำมาใช้อย่างกว้างขวางทั้งในคณิตศาสตร์ สถิติ การประมวลผลสัญญาณ (Signal processing) รวมถึงการประมวลผลภาพ (Computer vision) คอนโวลูชัน เป็นวิธีการทางคณิตศาสตร์เพื่อศึกษาการเปลี่ยนแปลงของฟังก์ชัน (f) เมื่อมีฟังก์ชัน (g) เข้ามา CNN ถูกนำมาใช้ในรูปแบบโมเดลที่สามารถเรียนรู้และเลือกใช้ลักษณะเด่น (Feature extraction) ของรูปภาพได้ด้วยตัวเอง ข้อดีคือ โมเดลจะเรียนรู้และเลือกลักษณะเด่นของภาพได้ดีกว่ามนุษย์

ดังนั้นจะทำให้ได้การประมวลผลที่แม่นยำมากกว่า ข้อแตกต่างจาก ANN คือ ค่าน้ำหนักที่ใช้เป็นค่าชุดเดียวกันหมด และนี่คือหนึ่งเหตุผลว่าทำไมถึงนักวิจัยจึงเลือกใช้ CNN เพราะถ้าใช้ ANN

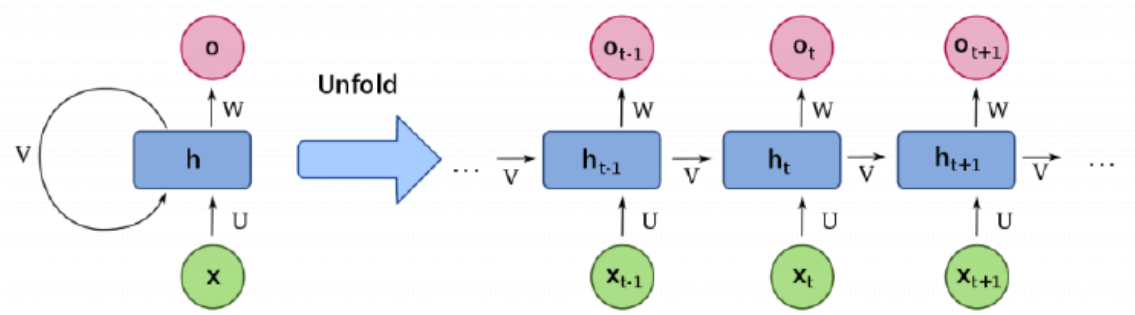
ปกตินั้น โมเดลจะต้องเรียนรู้ค่าน้ำหนักจำนวนมากโดยใช้วิธีการแพร่กลับ ข้อมูลเข้าของ CNN จะเป็นเมทริกซ์จากการแปลงมาจากรูปภาพ และโครงสร้างของ CNN ประกอบไปด้วยชั้นการทำงานจำนวน 2 ชั้น คือ 1) ชั้นคอนโวลูชั่น (Convolutional layer) และ 2) ชั้นพูลิ่ง (Pooling layer)

2) โครงข่ายประสาทเทียมแบบวนกลับ

โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN) เป็นวิธีการที่ถูกนำมาใช้ในการวิจัยเกี่ยวกับการรู้จำเสียง (Speech recognition) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing) การทำงานของ RNN ก็ตรงตัวตามชื่อ เพราะการเอาผลลัพธ์ที่ได้จากการคำนวณย้อนกลับมาใช้เป็นข้อมูลขาเข้าอีกครั้ง ซึ่งมีประโยชน์อย่างมากในข้อมูลที่มีความต่อเนื่อง เช่น ข้อมูลเสียง ข้อความ หรือแม้แต่รูปภาพเองก็ตาม

ANN เป็นฟังก์ชันที่รับข้อมูลเข้าและประมวลผลออกมาเป็นข้อมูลออก โดย ANN จะมองข้อมูลเข้าแต่ละตัวแยกกันชัดเจนไม่ขึ้นต่อกัน ข้อมูลเข้าที่รับเข้าไปจะเข้ามีการเรียงลำดับอย่างไรก็ได้ ANN จะไม่สนใจและจะประมวลผลให้ข้อมูลออกมาเหมือนเดิมตามข้อมูลเข้า 1 ตัวอย่างนั้น ซึ่ง ANN จะมีปัญหากับข้อมูลที่เป็นลำดับ เช่น ข้อความ (ลำดับตัวอักษร), เสียง (ลำดับแรงดันอากาศ), วิดีโอ (ลำดับของภาพและเสียง) และข้อมูลที่เป็นแบบ Time Series ต่างๆ เป็นต้น

RNN ถูกออกแบบมาเพื่อแก้ปัญหาสำหรับงานที่ข้อมูลที่มีลำดับ โดยใช้หลักการนำสถานะภายในของโมเดล กลับมาเป็นข้อมูลเข้าใหม่คู่กับข้อมูลเข้าแบบปกติ เรียกว่า สถานะซ่อน (Hidden State) หรือสถานะภายใน (Internal State) ช่วยให้โมเดลรู้จารูปแบบ (Pattern) ของลำดับข้อมูลเข้า (Input Sequence) ได้แสดงดังภาพที่ 28



ภาพที่ 28 โครงสร้างของโครงข่ายประสาทเทียมแบบวนกลับ

ที่มา : ไกรศักดิ์ เกษร (2564, น. 271)

ในแต่ละโหนดของ RNN จะมีข้อมูลเข้าสองอย่าง ได้แก่ 1) ข้อมูลเข้า ณ โหนดนั้นๆ และ 2) ผลลัพธ์ที่ได้จากการคำนวณในโหนดก่อนหน้า ซึ่งทั้งสองข้อมูลจะถูกนำมารวมเข้าด้วยกัน และออกผลลัพธ์มาเป็นสองทางคือ 1) ผลลัพธ์ที่ออก ณ โหนดนั้นๆ และออกเพื่อไปเข้าเป็นข้อมูลขาเข้าในโหนดถัดไป ข้อดีของ RNN คือ มันมีการใช้ข้อมูลก่อนหน้าในการทำนายสิ่งที่อาจจะเกิดขึ้น

ในอนาคต ซึ่งหมายถึงอะไรที่เคยเกิดขึ้นในอดีตย่อมส่งผลต่อเหตุการณ์ที่จะเกิดขึ้นในอนาคตด้วย แม้ RNN จะมีข้อดีในการทำงานของข้อมูลที่มีความต่อเนื่อง แต่ข้อเสียของ RNN คือ มันสามารถดูย้อนกลับได้แค่เพียงในช่วงระยะเวลาสั้นๆ เท่านั้น ซึ่งปัญหาหลักๆ ของ RNN เกิดมาจากค่าเกรเดียนที่เริ่มน้อยลงในข้อมูลที่มีความยาวมากขึ้น จนแทบจะไม่สามารถเห็นความเปลี่ยนแปลงของเกรเดียนได้เลย ซึ่งปัญหานี้ถูกเรียกว่าปัญหาการสูญหายของเกรเดียน (Vanishing Gradient Problem: VGP) ซึ่งปัญหานี้ถูกแก้ไขโดยใช้เกทแบบวนกลับ (Gated Recurrent Unit: GRU) และหน่วยความจาระยะสั้นแบบยาว (Long Short-Term Memory: LSTM)

4.10 การประยุกต์ใช้งานโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมเป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ที่ประยุกต์ใช้ความรู้จากหลากหลายสาขามารวมเข้าด้วยกัน นับว่าเป็นศาสตร์ที่กำลังมีบทบาทอย่างยิ่งในปัจจุบันกล่าวโดยสรุปโครงข่ายประสาทเทียมสามารถนำมาประยุกต์ใช้กับงานด้านต่างๆ ได้ดังนี้

1. การจำแนกรูปแบบ (pattern recognition) เช่น การมองเห็นวัตถุ หรือการวิเคราะห์เสียงพูดเพื่อแปลความหมาย

2. การทำนาย (prediction) หรือการพยากรณ์ (forecasting) เช่น การทำนายราคาหุ้นของตลาดหลักทรัพย์ การพยากรณ์อัตราการไหลของน้ำ การพยากรณ์ราคาสินค้า เป็นต้น

3. การควบคุม (control) เช่น การควบคุม ระบบของเครื่องปรับอากาศ การควบคุมระบบเครื่องยนต์ และการควบคุมหุ่นยนต์ เป็นต้น

4. การหาความเหมาะสม (optimization) เช่น การเลือกระยะทางที่ใกล้ หรือสั้นที่สุดในการเดินทาง (shortest path)

5. การจัดกลุ่ม (clustering) และการจัดหมู่ (categorization) เช่น การวิเคราะห์ข้อมูลจากภาพถ่ายดาวเทียม หรือภาพถ่ายทางอากาศ

4.11 การประยุกต์ใช้เครือข่ายใยประสาทกับการวิจัยทางสังคมศาสตร์

การประยุกต์ใช้เครือข่ายใยประสาทกับการวิจัยทางสังคมศาสตร์แบ่งการนำเสนอเป็น 4 ด้านดังต่อไปนี้ ด้านแรกคือ การคำนวณ ด้านที่สองคือ ลักษณะข้อมูล ด้านที่สามคือ การวิเคราะห์ทางสถิติ และด้านสุดท้ายคือ ปัญหาและอุปสรรค

1. ด้านการคำนวณ

Kai Hwang (1989 อ้างถึงใน Garson, 1998) กล่าวว่าเครือข่ายใยประสาทมีความแตกต่างกับการคำนวณโดยใช้สถิติแบบเดิม 7 ประการคือ 1) สามารถใช้กับข้อมูลจำนวนมาก (massive parallelism) โดยเครือข่ายใยประสาทสามารถใช้กับการคำนวณที่มีจำนวนมากๆ ได้ดีและใช้เวลาในการคำนวณน้อย 2) ข้อมูลที่มีความเกี่ยวเนื่องกันสูง (high inter connectivity) เนื่องจากเครือข่ายใยประสาทประกอบด้วย neuron จำนวนมาก ดังนั้นเมื่อจัดการกับข้อมูลจึงมีความแม่นยำ

มาก 3) กระบวนการไม่ซับซ้อน (simple processing) การประมวลผลโดยเครือข่ายใยประสาทเป็นการทำงานที่ง่าย และเป็นการใช้ผลรวมของข้อมูลเข้ามาสร้างฟังก์ชั่นเป็นค่าน้ำหนักแล้วคำนวณหาค่าข้อมูลออกมา 4) แยกการเสนอออกเป็นส่วน ๆ (distributed representation) โดยค่าน้ำหนักที่เชื่อมต่อเป็นค่าที่ได้จากค่าความคลาดเคลื่อนหรือข้อมูลที่ใส่เข้าไปใหม่ เพื่อนำมาคำนวณเป็นข้อมูลออก ผลการคำนวณขึ้นอยู่กับข้อมูลเข้า และข้อมูลที่มีอยู่เดิม 5) แก้ปัญหาข้อมูลที่ผิดพลาดได้ (fault tolerance) โดยเครือข่ายใยประสาทสามารถทำงานโดยที่มีค่า outlier หรือค่าที่ผิดปกติมากๆ ใน model ได้โดยไม่มีผลต่อข้อมูลออก 6) การคำนวณจากข้อมูลทั้งหมด (collective computation) การแก้ปัญหาหรือการคำนวณเป็นการประมวลผลจาก neuron ทั้งหมดในระบบ และ 7) จัดโครงสร้างการทำงานได้เอง (self-organization) การที่เครือข่ายใยประสาทสามารถปรับโครงสร้างการคำนวณได้ ขึ้นอยู่กับแบบแผนของข้อมูลเข้าที่แตกต่างกัน

2. ด้านลักษณะข้อมูล

Haykin (1994 อ้างถึงใน Garson, 1998) สรุปถึงลักษณะข้อมูลที่เหมาะสมกับการทำงานของเครือข่ายใยประสาทไว้ดังนี้ 1) ข้อมูลที่มีความสัมพันธ์แบบไม่ใช่เชิงเส้นตรง เครือข่ายใยประสาทสามารถใช้กับการวิเคราะห์ข้อมูลที่ไม่เป็นเชิงเส้นตรงได้ โดยเฉพาะการวิเคราะห์ด้วยสถิติขั้นสูง 2) ข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลเข้า-ข้อมูลออก การวิเคราะห์ด้วยเครือข่ายใยประสาทไม่ต้องมีการตรวจสอบการแจกแจงของข้อมูล อีกทั้งยังไม่มีข้อตกลงเกี่ยวกับลักษณะการแจกแจงใดๆ ของข้อมูลเข้า และข้อมูลออก ในกรณีนี้จะเหมือนกับสถิตินอนพาราเมทริกซ์ 3) ข้อมูลมีความยืดหยุ่นสูง เครือข่ายใยประสาทสามารถทำงานกับข้อมูลที่มีความแตกต่างมากๆ ได้ หรือมีการรบกวนจากสิ่งแวดล้อม (ใช้ในกรณีแยกเสียง หรือการพยากรณ์อากาศ เป็นต้น) 4) evidential response โดยเครือข่ายใยประสาทไม่เพียงแต่ใช้ในการแยกประเภทเท่านั้นยังให้ผลของข้อมูลที่เชื่อถือได้อีกด้วย ซึ่งข้อมูลที่ได้จะเป็นประโยชน์สำหรับการแยกประเภทในอนาคตได้ 5) contextual information กล่าวคือ เครือข่ายใยประสาทมีการทำงานที่รวมถึงความสามารถในการหาปฏิสัมพันธ์ (interaction) ด้วย 6) แก้ปัญหาข้อมูลที่ผิดพลาดได้ (fault tolerance) หมายความว่าเครือข่ายใยประสาทสามารถทำงานโดยที่มีค่า outlier หรือค่าที่ผิดปกติมากๆ ใน model ได้โดยไม่มีผลต่อข้อมูลออก 7) VLSI (very large scale implementation) การทำงานของเครือข่ายใยประสาทสามารถทำงานกับข้อมูลที่มีจำนวนมากๆ หรือมีลักษณะการวัดที่มากๆ ได้ และใช้เวลาในการคำนวณน้อยกว่าวิธีเดิมๆ 8) uniformity of analysis and design เครือข่ายใยประสาทมีรูปแบบการวิเคราะห์และการทำงานเป็นเอกลักษณ์ และ 9) neurobiological analogy เนื่องจากเครือข่ายใยประสาทจำลองการทำงานเช่นเดียวกับการทำงานของสมองจึงมีประสิทธิภาพสูงทั้งในเชิงความเร็ว และการคาดเดา หรือการคำนวณความน่าจะเป็นที่เกิดขึ้นในขณะที่โปรแกรมเครือข่ายใยประสาทมีขนาดเล็ก (สามารถคำนวณโดยใช้คอมพิวเตอร์ตั้งโต๊ะทั่วไปได้) เช่นเดียวกับสมองของมนุษย์ที่มีขนาดเล็กแต่มีประสิทธิภาพสูงมาก

3. ด้านการวิเคราะห์ทางสถิติ

การนำเครือข่ายใยประสาทมาประยุกต์ใช้ในการวิเคราะห์ทางสังคมศาสตร์และพฤติกรรมศาสตร์ เริ่มขึ้นในช่วงต้นศตวรรษที่ 20 (Ader & Bramsen, 1998) ระยะแรกเป็นการเปรียบเทียบผลการวิเคราะห์ด้วยเครือข่ายใยประสาทกับสถิติอื่นๆ เช่น การวิเคราะห์ถดถอย การวิเคราะห์ถดถอยโลจิสติก การวิเคราะห์จำแนก การวิเคราะห์องค์ประกอบ (Wilson & Hardgrave, 1995; Sinha & McKim, 2000; Gonzalez & DesJardins, 2002) ระยะต่อมา มีการศึกษาและพัฒนาการวิเคราะห์ด้วยเครือข่ายใยประสาทกว้างขวางขึ้นอาทิเช่น งานวิจัยของ Yang, Browne and Picton (2002) ที่นำเครือข่ายใยประสาทมาประยุกต์ใช้เป็นการวิเคราะห์พหุลำดับ (multistage neural network) Ader & Bramsen (1998) แสดงความเห็นเกี่ยวกับการวิเคราะห์รูปแบบสมการเชิงโครงสร้าง (SEM) กับเครือข่ายใยประสาทว่ามีความคล้ายคลึงกันโดยที่ข้อมูลเข้าและข้อมูลออกของเครือข่ายใยประสาทเปรียบเสมือนโมเดลการวัดของ SEM ขณะที่ hidden layers เปรียบเสมือนตัวแปรแฝงภายใน ทั้งนี้ Ader และ Bramsen ทดลองนำผลการวิจัยของ Bramsen (1995 อ้างถึงใน Ader and Bramsen, 1998) ที่ศึกษาความสัมพันธ์ของผลที่เกิดขึ้นหลังเหตุการณ์สงครามโลก ได้แก่ อาการที่เกิดขึ้นเรื้อรังหลังเกิดความเครียด (post-traumatic stress disorder) ความไม่สุขสบายเกี่ยวกับกระเพาะอาหาร และคุณลักษณะส่วนบุคคล กับระยะเวลาที่เผชิญสงคราม วิเคราะห์ด้วย SEM กับเครือข่ายใยประสาท ผลการวิเคราะห์พบว่าค่าน้ำหนักที่เกิดจากการเรียนรู้ของเครือข่ายใยประสาท มีค่าใกล้เคียงกับการประมาณค่าพารามิเตอร์ของ SEM และผลของการวิเคราะห์ด้วย SEM ให้ข้อมูลที่เป็นประโยชน์ในการกำหนดข้อมูลแอบแฝงสำหรับการวิเคราะห์ด้วยเครือข่ายใยประสาท Ader และ Bramsen กล่าวว่าการวิเคราะห์ด้วย SEM และเครือข่ายใยประสาทยังมีอีกหลายจุดที่น่าสนใจ ได้แก่ ความสอดคล้องของผลการประมาณค่าพารามิเตอร์กับค่าน้ำหนักความสัมพันธ์ระหว่างค่าที่ประมาณได้กับจำนวน hidden layers เป็นต้น แต่อุปสรรคสำคัญของการวิเคราะห์ด้วยเครือข่ายใยประสาทคือ การใส่ข้อมูลที่ผู้วิจัยจะต้องแปลงข้อมูลที่เป็นค่าต่อเนื่องให้เป็นข้อมูลที่มีค่าไม่ต่อเนื่องก่อน

4. ปัญหาและอุปสรรค

Garson (1998) สรุปอุปสรรคที่ทำให้เครือข่ายใยประสาทไม่ได้นำมาใช้อย่างแพร่หลายทางสังคมศาสตร์ไว้ 4 ประการ ดังนี้ ประการแรก neuron model เป็น model สำหรับพยากรณ์ จึงไม่ได้อธิบายความเป็นเหตุผล ขณะที่ชั้นแอบแฝงเปรียบเสมือนกล่องดำ (black box) ที่ไม่สามารถจินตนาการถึงสิ่งที่อยู่ภายในได้ และการวิเคราะห์ของเครือข่ายใยประสาทใช้อัลกอริทึมซึ่งให้ค่าน้ำหนักที่ไม่สามารถอธิบายที่มาได้ นักสังคมศาสตร์ส่วนใหญ่ต้องการรู้เหตุผลว่า การคิดเพื่อกำหนดค่าน้ำหนักมีหลักการหรือวิธีการอย่างไร ประการที่สอง รูปแบบเครือข่ายใยประสาทมีหลายรูปแบบ และมีความซับซ้อนในแต่ละรูปแบบ บางครั้งรูปแบบที่มีก็ไม่เหมาะสมกับข้อมูลทางสังคมศาสตร์ การสร้าง

หรือการพัฒนาโปรแกรมจึงเป็นเรื่องยากสำหรับนักสังคมศาสตร์ ประการที่สาม การใช้ศัพท์ที่แตกต่างจากทางสังคมศาสตร์ ทำให้ผู้ใช้ใหม่เกิดความสับสน ความจริงศัพท์บางคำของเครือข่ายใยประสาท มีความหมายเหมือนกับศัพท์ทางสังคมศาสตร์ เช่น case หรือ observation ศัพท์ทาง เครือข่ายใยประสาท เรียกว่า patterns ส่วนคำว่า ตัวแปรต้น ศัพท์ทางเครือข่ายใยประสาท เรียกว่า ข้อมูลเข้า (input) คำว่า ตัวแปรตาม ศัพท์ทางเครือข่ายใยประสาทเรียกว่า ข้อมูลออก (output) หรือ targets และคำว่า การประมาณค่า (estimation) ศัพท์ทางเครือข่ายใยประสาท เรียกว่า training, learning หรือ self-organization เป็นต้น การศึกษาโปรแกรมสำเร็จรูปที่ใช้ในการวิเคราะห์เครือข่ายใยประสาทพบว่า มีผู้พัฒนาโปรแกรมสำเร็จรูปแบบสาธารณะ (free or shareware) ส่วนใหญ่ใช้อัลกอริทึ่มแบบแพร่กระจายย้อนกลับ (backpropagation) ผู้วิจัยศึกษาวิธีใช้งาน 2 โปรแกรมได้แก่ โปรแกรม RapidMiner Studio Educational 10.1 และ Weka version 8 พบว่า ทั้งสองโปรแกรมให้ความสะดวกในการวิเคราะห์ข้อมูลพอควร โดยโปรแกรมเป็นสาธารณะ (free or shareware) การใช้งานง่ายเนื่องจากเป็นการโยงสัญลักษณ์ (symbol) ต่อกันไม่ต้องเรียนรู้คำสั่งที่ซับซ้อน ผลการวิเคราะห์แสดงในรูปแบบตารางอีกทั้งยังสามารถวิเคราะห์ด้วยวิธีอื่นๆ ที่สอดคล้องกับเครือข่ายใยประสาทเพื่อศึกษาความสัมพันธ์ของการทำงานของโปรแกรมได้

การวิจัยครั้งนี้ผู้วิจัยเลือกใช้การวิเคราะห์โครงข่ายประสาทเทียม (ANN) แบบเพอร์เซฟตรอนหลายชั้น (Multi-layer Perceptron Neural Network) หรือโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Feed-forward Backpropagation Neural Network) ซึ่งเป็นโครงข่ายประสาทเทียมแบบต้องมีผู้สอน (Supervised Learning) เท่านั้น และผู้วิจัยเลือกใช้โปรแกรม Rapid miner studio 10.1 เป็นซอฟต์แวร์เวอร์ชันฟรี (Free Version) ที่สามารถดาวน์โหลด (Download) มาใช้งานได้ โดยไม่เสียค่าใช้จ่าย เป็นเครื่องมือที่ช่วยในการคำนวณวิเคราะห์ประมวลผลโดยมีโอเปอร์เรเตอร์ต่างๆ สามารถนำมาใช้ในขั้นตอนต่างๆ ตั้งแต่การเตรียมข้อมูล จนถึงขั้นตอนการสร้างแบบจำลองให้เป็นไปตามที่ผู้ใช้ต้องการ ซึ่งในงานวิจัยนี้ผู้วิจัยเลือกใช้ Rapid miner studio ที่เป็นประเภทสำหรับการศึกษา (RapidMiner Studio Educational 10.1.003) ดังภาพประกอบที่ 8 เพื่อให้สามารถดำเนินการกับชุดข้อมูลที่มากกว่า 10,000 Rows ได้

ภาพที่ 29 รูปโปรแกรม Rapid miner studio 10.1 ประเภทสำหรับการศึกษา (Educational Program)

งานวิจัยที่เกี่ยวข้อง

1. งานวิจัยในประเทศ

โสพส สุขานนท์สวัสดิ์ (2556) ได้วิจัยเรื่องการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้ทฤษฎีการตัดสินใจในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ได้เสนอวิธีการคัดเลือกข้อสอบข้อถัดไปที่ใช้ทฤษฎีการตัดสินใจในสภาวการณ์ความเสี่ยง (Decision Making Under Risk) ตามเกณฑ์ทางเลือกที่มีค่าความคาดหวังสูงสุด (Maximum Expected Monetary Value: EMV) โดยจะเลือกข้อสอบข้อที่มีค่าสัมประสิทธิ์การกระจาย (Coefficient of Variation) สูงสุดมาเป็นข้อสอบข้อถัดไป ซึ่งจากการศึกษาเปรียบเทียบวิธีการนี้กับวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดกับกลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญา จำนวน 280 คน ผลการวิจัยสรุปได้ว่าจำนวนข้อสอบที่ใช้และเวลาในการทดสอบของวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจมีค่าน้อยกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

สมประสงค์ เสนารัตน์ (2555, น. 188-192) ได้พัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อวินิจฉัยกระบวนการพุทธิปัญญาในการเรียนพีชคณิต ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติประกอบด้วย 4 ขั้นตอน คือ 1) การพัฒนาคลังข้อสอบพีชคณิตโดยประมาณค่าพารามิเตอร์ของข้อสอบ ด้วยโมเดลปกติสะสมแบบพหุมิติ ผลการวิจัยในขั้นตอนนี้ ทำให้ได้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์พร้อมคู่มือการ

ใช้งานที่สามารถนำไปใช้ในการทดสอบเพื่อวินิจฉัยกระบวนการพุทธิปัญญาของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ได้อย่างถูกต้องแม่นยำและเชื่อถือได้ ผลการทดลองใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จากการทดลองใช้จำนวน 2 ครั้ง พบว่า มีประเด็นปัญหาที่ต้องปรับปรุงเล็กน้อย เกี่ยวกับความพึงพอใจในการจัดวางตำแหน่งของปุ่มต่างๆ บนหน้าจอคอมพิวเตอร์ผลการประเมินกระบวนการพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สรุปได้ดังนี้

1. คลังข้อสอบที่พัฒนาขึ้นมีคุณภาพสามารถนำมาใช้เป็นฐานข้อมูลในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้

2. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถนำไปใช้ประกอบการในการเรียนการสอนได้

3. กลุ่มผู้ทดลองใช้งานเห็นว่าการกระบวนการทดสอบทำให้ได้รับรู้ความสามารถของผู้ทดสอบได้ทันที ให้ผลการทดสอบที่ถูกต้องน่าเชื่อถือ ให้ผลการทดสอบได้หลากหลายมิติ และเป็นการทดสอบที่ให้ความสนุกเพลิดเพลินกับผู้ทดสอบมากกว่าการทดสอบแบบดั้งเดิม และนักเรียนส่วนใหญ่อยากให้ครูที่สอนนำเอาวิธีการทดสอบรูปแบบนี้ไปใช้ในเนื้อหาอื่นๆ และวิชาอื่นๆ

4. ผลการประเมินการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีประโยชน์ในการวินิจฉัยกระบวนการพุทธิปัญญาในการเรียนพีชคณิต และให้ผลการวินิจฉัยที่ถูกต้องครอบคลุมกระบวนการพุทธิปัญญาในการเรียนพีชคณิต ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยมีผลการประเมินอยู่ในระดับมากทั้งในภาพรวมและรายด้าน

ประพล เปรมทองสุข (2560) ได้วิจัยเรื่องพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผลการวิจัยปรากฏว่า วิธีการคัดเลือกข้อสอบข้อถัดไปและมีการควบคุมการใช้ข้อสอบที่พัฒนาขึ้นมีประสิทธิภาพด้านการประมาณค่าความสามารถของผู้สอบและลดความยาวของแบบทดสอบได้โดยพิจารณาจากค่าความลำเอียงเฉลี่ยปรากฏว่า วิธีการ RDM วิธีการ HC และวิธีการ HC-Ex มีประสิทธิภาพสูงสุดเทียบเท่ากัน เมื่อเปรียบเทียบด้านความยาวของแบบทดสอบปรากฏว่าวิธีการ HC มีประสิทธิภาพสูงสุดเทียบเท่ากับวิธีการ RDM และเมื่อเปรียบเทียบด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.2 ปรากฏว่าวิธีการ MIC มีประสิทธิภาพสูงสุดและโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น (วิธีการ HC-Ex) มีความเหมาะสมในการใช้งานระดับมากถึงมากที่สุด ในด้านความสะดวกในการใช้โปรแกรมความถูกต้อง ลักษณะทั่วไปของโปรแกรมและความชัดเจนของคู่มือ และค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น (วิธีการ HC-Ex) กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่ามีความสัมพันธ์กันทางบวก ($p < .01$)

รัชกฤช ธนพัฒนดล (2561) ได้วิจัยเรื่องการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้วิธีระบบอาณานิคมมด สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ การคัดเลือกข้อสอบข้อถัดไปเป็นหัวใจสำคัญของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจากส่งผลต่อประสิทธิภาพของการทดสอบ การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมด 2) เปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3 วิธี คือ วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (วิธี MIC) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และมีการควบคุมการใช้ข้อสอบ (วิธี HC-Ex) และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมด (วิธี ACS) ในด้านการประมาณค่าความสามารถของผู้สอบด้านความยาวของแบบทดสอบ และด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.20 3) พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมด และ 4) ศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมดกับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี ปีการศึกษา 2561 จำนวน 30 คน และวิเคราะห์ข้อมูลด้วยการวิเคราะห์สหสัมพันธ์แบบเพียร์สัน ผลการวิจัยปรากฏว่า 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมดมี 3 ขั้นตอน ดังนี้ (1) จัดกลุ่มข้อสอบในคลังตามค่าความยากของข้อสอบออกเป็น 5 ระดับ และเลือกข้อสอบข้อถัดไปตามกฎการเลือกข้อสอบเพื่อจัดเรียงบนโครงสร้างต้นไม้ตัดสินใจแบบสามเหลี่ยม (2) ออกแบบกระบวนการของวิธีอาณานิคมมดโดยการจำลองผลการตอบข้อสอบเพื่อวัดความเหมาะสมของการเลือกข้อสอบข้อถัดไป และ (3) เชื่อมโยงกระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ให้สัมพันธ์กับวิธีระบบอาณานิคมมด 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี ACS มีประสิทธิภาพสูงกว่าวิธี MIC และวิธี HC-Ex ในด้านการประมาณค่าความสามารถของผู้สอบ และด้านจำนวนข้อสอบที่มีอัตราการใช้ข้อสอบมากกว่า 0.20 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในขณะที่วิธี HC-Ex มีประสิทธิภาพในด้านความยาวของแบบทดสอบสูงสุด 3) โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นในภาพรวมมีความเหมาะสมในการใช้งานระดับมากที่สุด และ 4) ค่าประมาณความสามารถของผู้สอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น กับคะแนนรวมของผู้สอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์มีความสัมพันธ์กันทางบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

บูรพา วิถีปัญญา (2563) ได้การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้วิธีการจำลองข้อมูล ผลการวิจัยสรุปได้ดังนี้ 1) ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียมพบว่า โมเดลโครงข่ายประสาทเทียมที่มีชั้นข้อมูลเข้าจำนวน 20 Node ชั้นข้อมูลซ่อน 1 Node และชั้นข้อมูลออก

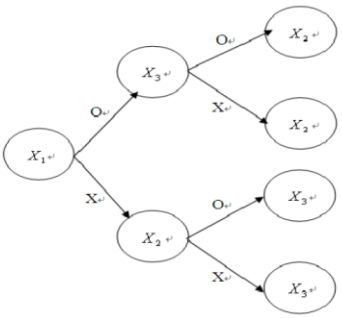
1 Node 2) ผลการพัฒนาโปรแกรมจำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ได้โปรแกรมจำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม การประมาณค่าความสามารถด้วยวิธีความน่าจะเป็นสูงสุด และการยุติการทดสอบด้วยเกณฑ์ค่ามาตรฐานความผิดพลาด (SE) น้อยกว่าหรือเท่ากับ 0.40 3) การนำโครงข่ายประสาทเทียม (ANN) มาใช้ในการคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้วิธีการจำลองข้อมูล ใช้จำนวนข้อสอบน้อยกว่าวิธีการสุ่ม (R) และวิธีการใช้ค่าสารสนเทศสูงสุด (MI)

2. งานวิจัยต่างประเทศ

Han (2012) ได้นำเสนอวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ประสิทธิภาพของข้อมูลเพื่อหาวิธีการคัดเลือกข้อสอบที่มีการรักษาความสมดุลของเนื้อหาโดยใช้การเปรียบเทียบ 2 วิธีคือ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศของฟิชเชอร์สูงสุด (Maximum Fisher Information: MFI) และ 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยวิธีการแบ่งชั้น (Stratification Methods) ผลการศึกษาปรากฏว่า วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ประสิทธิภาพของข้อมูลมีการรักษาความสมดุลของเนื้อหามากกว่าวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้สารสนเทศของฟิชเชอร์สูงสุด และวิธีการคัดเลือกข้อสอบข้อถัดไปโดยวิธีการแบ่งชั้น

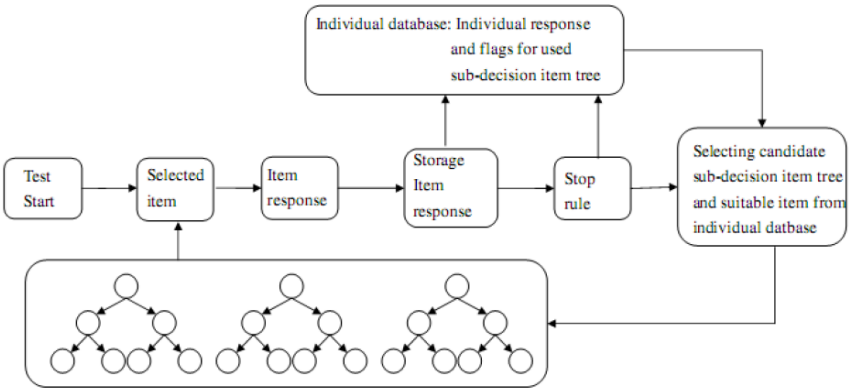
Karahora and Ince (2009) ได้นำเสนอวิธีจัดกลุ่มข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยประยุกต์ใช้วิธี Adaptive Network Based Fuzzy Inference System (ANFIS) ซึ่งในกระบวนการจะทำการจัดข้อสอบในคลังข้อสอบทั้งหมดออกเป็น 5 ระดับดังนี้ 1) ง่ายมาก (Very Easy) แทนด้วยตัวเลข –1, 2) ง่าย (Easy) แทนด้วยตัวเลข -0.5, 3) ปานกลาง (Medium) แทนด้วยตัวเลข 0, 4) ยาก (Hard) แทนด้วยตัวเลข 0.5 และ 5) ยากมาก (Very Hard) แทนด้วยตัวเลข 1 จากนั้นนำผลการจัดกลุ่มข้อสอบด้วยวิธีดังกล่าวไปเปรียบเทียบกับความเร็วและความถูกต้องกับอีก 2 วิธี คือ Artificial Neural Network (ANN) และ Support Vector Machines (SVM) พบว่ามีผลลัพธ์ของแต่ละวิธีเป็นดังนี้ ANFIS = 0.99, ANN = 0.74 และ SVM = 0.79 จึงสรุปได้ว่าวิธี ANFIS ที่นำเสนอมีประสิทธิภาพในการจัดกลุ่มข้อสอบมากกว่าอีก 2 วิธีดังกล่าว

Yu Hsieh and Chen Kuo (2010) ได้นำเสนอการเลือกข้อสอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้วิธี Blocking Strategy โดยอาศัยทฤษฎีของ Bayesian Network ในหลักการที่สำคัญคือจะจัดกลุ่มข้อสอบโดยใช้ทฤษฎีของ Bayesian Network ให้อยู่ในรูปแบบของ Decision Tree ประเภท Binary Tree ดังภาพที่ 30



ภาพที่ 30 *Decision Item Tree Network* (Yu Hsieh & Chen Kuo, 2010, p. 1161)

จากภาพที่ 30 หากผู้สอบทำข้อสอบข้อที่ X1 ถูกข้อสอบข้อถัดไปที่ต้องทำคือ ข้อที่ X3 ในทางตรงกันข้ามหากทำข้อที่ X1 ผิด ข้อสอบข้อถัดไปที่ต้องทำคือข้อที่ X2 ทำลักษณะนี้ไปเรื่อยๆ จนครบเงื่อนไขการทดสอบ ในส่วนของ Blocking Strategy ที่นำเสนอแสดงได้ดังภาพที่ 31



ภาพที่ 31 *CAT Based on Decision Item Subtree* (Yu Hsieh & Chen Kuo, 2010, p. 1164)

จากภาพที่ 31 พบว่าผู้วิจัยได้ออกแบบการแบ่งข้อสอบในคลังข้อสอบออกเป็น 100 block แต่ละ block ดังกล่าว จะให้ผู้สอบทำข้อสอบ 3 ข้อต่อ 1 ชุดข้อสอบ ซึ่งจะต้องจัดวางข้อสอบบน Decision Item Tree Network ซึ่งเป็น Subtree ย่อยๆ จำนวน 7 ข้อต่อ 1 ชุดข้อสอบจากนั้นนำผลการการทดสอบของวิธีที่นำเสนอเปรียบเทียบกับวิธี Experts's Classification พบว่าวิธี Blocking Strategy โดยอาศัยทฤษฎีของ Bayesian Network ให้ประสิทธิภาพดีกว่า โดยการใช้จำนวนข้อสอบในการทดสอบได้น้อยกว่า

Costa, Karino, Moura, and Andrade (2009) ได้เปรียบเทียบวิธีการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 3 วิธีการ ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์คูลแบค-ไลเบลอร์ (Kullback-Leibler Criterion: KL) และ 3) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าคาดหวังของค่าสารสนเทศสูงสุด (Maximum Expected Information Criterion: MEI) โดยใช้คลังข้อสอบวิชาภาษาอังกฤษ จำนวน 246 ข้อ และศึกษา

เปรียบเทียบในสถานการณ์จำลอง 5 กรณีคือ กรณีที่ 1 เปรียบเทียบในด้านความยาวของแบบทดสอบ โดยใช้เกณฑ์ยุติการทดสอบที่ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถน้อยกว่า 0.4 ปรากฏว่า ทั้ง 3 วิธีการ ใช้ข้อสอบโดยเฉลี่ยใกล้เคียงกัน คือ 23, 22 และ 21 ข้อ ตามลำดับ กรณีที่ 2 ศึกษาความสัมพันธ์ระหว่างค่าความสามารถที่แท้จริงกับค่าประมาณความสามารถของผู้สอบเมื่อกำหนดค่าประมาณความสามารถเริ่มต้นเท่ากับ 0.00 และใช้เกณฑ์ยุติการทดสอบที่ 25 ข้อ ปรากฏว่าทั้ง 3 วิธีการ มีค่าความสัมพันธ์สูงมากใกล้เคียงกัน กรณีที่ 3) เปรียบเทียบผลการทดสอบ เมื่อกำหนดค่าประมาณความสามารถเริ่มต้นแตกต่างกัน 3 ค่า คือ -1.50, 0.00 และ 1.50 และใช้เกณฑ์ยุติการทดสอบที่ 25 ข้อ ปรากฏว่าเมื่อเสร็จสิ้นการทดสอบทั้ง 3 วิธีการ ให้ค่าประมาณความสามารถใกล้เคียงกัน กรณีที่ 4 เปรียบเทียบผลการทดสอบ เมื่อกำหนดข้อสอบข้อเริ่มต้นมีค่าความยากของข้อสอบแตกต่างกัน 3 แบบ คือ มีค่าน้อยกว่า, เท่ากับ และมากกว่าค่าความสามารถจริงของผู้สอบ ปรกฏว่าทั้ง 3 วิธีการ ให้ค่าประมาณความสามารถใกล้เคียงกัน และกรณีที่ 5 ศึกษาประสิทธิภาพการประมาณค่าความสามารถของผู้สอบ ซึ่งพิจารณาจากค่าความลำเอียงเฉลี่ย และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเมื่อกำหนดค่าความสามารถที่แท้จริงของผู้สอบแตกต่างกัน 10 ค่า ปรากฏว่าทั้ง 3 วิธีการให้ค่าความลำเอียงเฉลี่ย และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน

Fan, Wang, Chang, and Douglas (2012) ได้เสนอวิธีการคัดเลือกข้อสอบข้อถัดไป วิธีการใหม่ 2 วิธีการ ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุดต่อหน่วยเวลา (Maximum Information per Time Unit) ซึ่งพัฒนาจากวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) และ 2) วิธีการ Time Weighted a-Stratification ซึ่งพัฒนามาจากวิธีการควบคุมการใช้ข้อสอบ a-Stratified Strategy (ASTR) โดยวิธีการใหม่ที่ได้พัฒนาขึ้น จะคัดเลือกข้อสอบข้อถัดไป โดยพิจารณาจากเวลาที่ใช้ในการทำข้อสอบควบคู่กับค่าสารสนเทศของข้อสอบ เพื่อแก้ไขปัญหาที่ผู้สอบบางคนได้รับข้อสอบที่ใช้เวลาในการทำข้อสอบนานกว่าคนอื่นๆ หลังจากนั้น นำวิธีการที่พัฒนาขึ้นมาเปรียบเทียบกับวิธีการดั้งเดิมด้วยการศึกษาในสถานการณ์จำลอง จำนวน 1,000 ครั้ง และใช้คลังข้อสอบขนาด 500 ข้อ ปรากฏว่าการดำเนินการทดสอบด้วยวิธีการใหม่มีสมดุลด้านเวลาเพิ่มขึ้น แต่ยังคงมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบ และการควบคุมการใช้ข้อสอบใกล้เคียงกับวิธีการดั้งเดิม

Murphy, Dodd, and Vaughn (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3 วิธี ได้แก่ 1) วิธี Maximum Fisher's Information (MFI) 2) วิธี Maximum Posterior Weighted Information (MPWI) และ 3) วิธี Minimum Expected Posterior Variance (MEPV) ในกรณที่คลังข้อสอบถูกแบ่งเป็นคลังข้อสอบย่อย (Testlet) ภายใต้โมเดลการตอบสนองข้อสอบ (IRT) และโมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย (Testlet Response Theory: TRT) ดำเนินการวิจัยโดยใช้การศึกษาในสถานการจำลอง โดยจำลองแบบ

แผนการตอบของผู้สอบ จำนวน 10 กลุ่ม แต่ละกลุ่มมีผู้สอบ 1,000 คน ผลการศึกษาปรากฏว่า เมื่อใช้โมเดล IRT กับข้อมูลที่ฝ่าฝืนข้อตกลงเบื้องต้น ความเป็นอิสระระหว่างข้อสอบ ผลการประมาณค่าความสามารถของผู้สอบมีแนวโน้มที่จะมีค่ามากกว่าค่าจริง และผลการเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ปรากฏว่าวิธีการทั้ง 3 วิธี มีประสิทธิภาพไม่แตกต่างกัน เนื่องจากทั้ง 3 วิธี ต่างทำให้ผู้สอบแต่ละคนได้รับข้อสอบจำนวนมาก (ประมาณ 50 ข้อ) เป็นเหตุให้ค่าประมาณความสามารถของผู้สอบลู่เข้าหาค่าความสามารถที่แท้จริงของผู้สอบ

Wang, Kuo, Chao, and Tsai (2012) ได้พัฒนาและประเมินผลระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความสามารถทางภาษาจีน ภายใต้กรอบมาตรฐานการประเมินความสามารถทางภาษาในกลุ่มสหภาพยุโรป (The Common European Framework of Reference: CEFR) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ ถูกพัฒนาในรูปแบบระบบการทดสอบบนเว็บ (Web – Base Test System) สำหรับวัดความสามารถทางภาษาจีนในระดับ A1 และ A2 (จาก 6 ระดับ ได้แก่ A1, A2, B1, B2, C1 และ C2 ตามลำดับ) โดยเป็นการวัดทักษะการฟังของผู้สอบ ประกอบด้วยคำถาม 2 ลักษณะ ได้แก่ 1) คำถามวัดความเข้าใจในการฟัง (Listening Comprehension Item) มีลักษณะให้ผู้สอบฟังวลี หรือบทสนทนา แล้วให้เลือกคำตอบที่ถูกที่สุดเพียงข้อเดียว จากตัวเลือกทั้งหมด 4 ตัวเลือก ซึ่งคำถามแต่ละข้อถูกจำกัดเวลาในการตอบ หากหมดเวลา ผู้สอบจะได้รับคำถามข้อถัดไปทันที 2) คำถามวัดความเข้าใจในการฟัง-มอง (Visual-Listening Comprehension Item) มีลักษณะให้ผู้สอบฟังวลี หรือบทสนทนา แล้วให้เลือกรูปภาพที่สอดคล้องกับวลี หรือบทสนทนานั้น ซึ่งมีตัวเลือก (รูปภาพ) ทั้งหมด 4 ตัวเลือก กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนระดับมัธยม เกรด 5 ถึงเกรด 10 ของโรงเรียน Grace Christian Chinese School ในประเทศฟิลิปปินส์ ซึ่งนักเรียนเกรด 5 ถึงเกรด 7 ได้ทดสอบในระดับ A1 ส่วนนักเรียนเกรด 8 ถึงเกรด 10 ได้ทดสอบในระดับ A2 ผลการวิเคราะห์ความเที่ยงของคำถาม ปรากฏว่า คำถามใน A1 และ A2 มีค่าสัมประสิทธิ์แอลฟามากกว่า 0.8 แสดงว่า คำถามใน A1 และ A2 มีความเที่ยงสูง

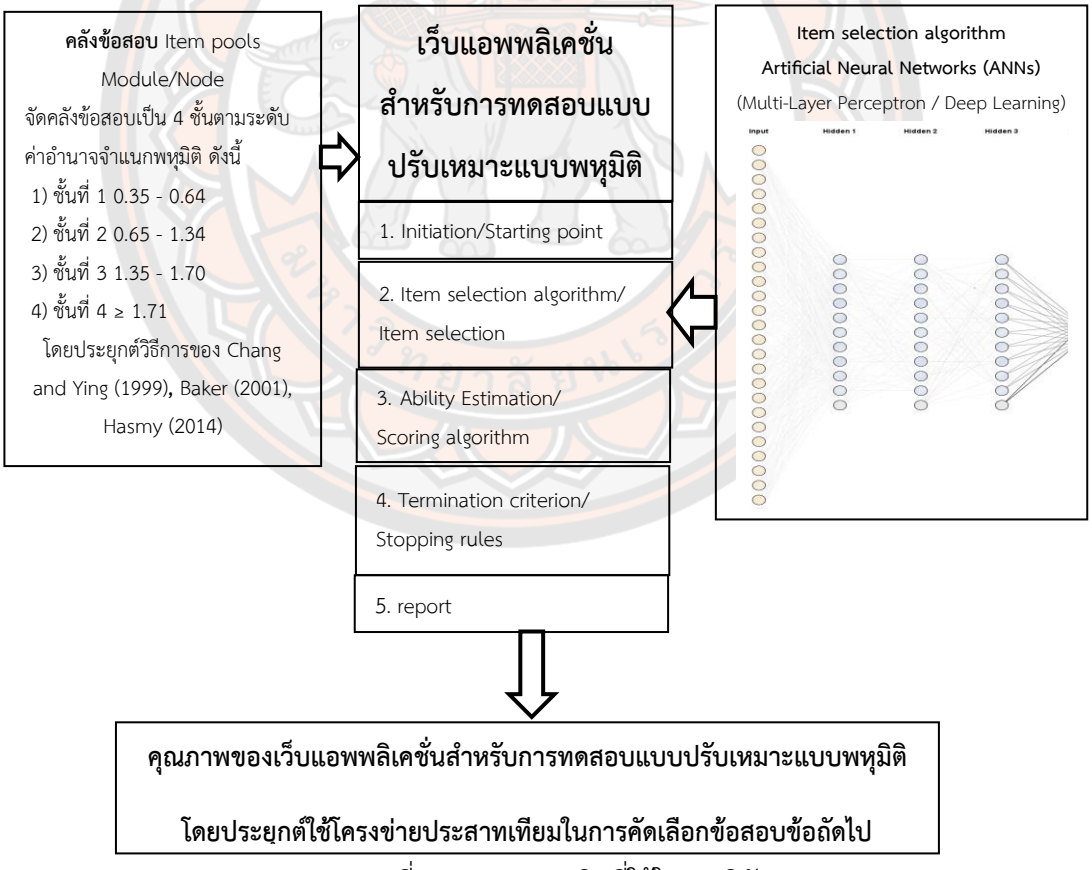
Koedsri, Lawthong and Ngudgratoke, (2014) ได้ศึกษาประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่มีความยาวแบบยืดหยุ่น (Variable-Length) โดยใช้โมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย (Testlet Response Model: TRT) และใช้วิธีการแบ่งกลุ่มค่าอำนาจจำแนกของข้อสอบแบบถ่วงน้ำหนักที่มีการบังคับ (The Constraint-Weighted a-Stratification Method: CWA) เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป เนื่องจากความเป็นอิสระระหว่างข้อสอบเป็นข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบแบบดั้งเดิม (Traditional Item Response Theory) การละเมิดข้อตกลงนี้มีผลต่อการประมาณค่าความสามารถของผู้สอบ ดังนั้น โมเดล TRT จึงถูกพัฒนาขึ้นเพื่อแก้ข้อบกพร่องนี้ด้วยเหตุนี้ การวิจัยนี้จึงมีวัตถุประสงค์ เพื่อศึกษาความถูกต้องของวิธีการคัดเลือกข้อสอบข้อถัดไปด้วยวิธี CWA

ในโมเดล TRT ที่ใช้กับแบบทดสอบที่มีความยาวยืดหยุ่น ซึ่งวิธี CWA เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป ที่มีประสิทธิภาพสูงสามารถควบคุมการใช้ข้อสอบ และควบคุมสัดส่วนเนื้อหาของข้อสอบ ได้พร้อมกันด้วยวิธีเดียว ดำเนินการวิจัยด้วยการศึกษาในสถานการณ์จำลองแบบมอนติคาร์โล ประกอบด้วยคลังข้อสอบจำลอง 1,000 ข้อ และค่าความสามารถที่แท้จริงของผู้สอบ 1,000 คน โดยศึกษาเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไป 3 วิธี ได้แก่ 1) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้การคัดเลือกอย่างสุ่ม (Randomization Criterion: RAN) 2) วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ค่าสารสนเทศสูงสุด (Maximum Information Criterion: MIC) และวิธีการแบ่งกลุ่มค่าอำนาจจำแนกของข้อสอบแบบถ่วงน้ำหนักที่มีการบังคับ (The Constraint-Weighted a-Stratification Method: CWA) ผลปรากฏว่า สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีความยาวแบบยืดหยุ่น และใช้โมเดลการตอบสนองข้อสอบที่ใช้แบบทดสอบย่อย วิธี CWA เป็นวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีประสิทธิภาพสูงสุด เมื่อเทียบกับวิธีการ RAN และวิธีการ MIC

จากการศึกษางานวิจัยที่เกี่ยวข้องกับวิธีการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่าในช่วงเริ่มแรกส่วนใหญ่เป็นงานวิจัยที่เกี่ยวข้องกับการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปขึ้นใหม่ ส่วนงานวิจัยในช่วงต่อมาส่วนใหญ่เป็นการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของวิธีการคัดเลือกข้อสอบข้อถัดไปที่มีผู้พัฒนาเอาไว้แล้ว โดยมีจุดมุ่งหมายเพื่อให้ทราบข้อดี หรือข้อจำกัดที่เกิดขึ้นในการนำวิธีการนั้นๆ ไปใช้งานจริง ซึ่งการศึกษาโดยใช้ข้อมูลจริง หรือศึกษาในสถานการณ์จำลอง ควบคู่กับการนำวิธีการคัดเลือกข้อสอบข้อถัดไปที่พัฒนาขึ้นมาแล้วไปใช้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อวัดความรู้ความสามารถของผู้สอบในด้านใดด้านหนึ่ง ซึ่งมีงานวิจัยวิธีการในการคัดเลือกข้อสอบข้อถัดไป 2 กลุ่ม กลุ่มที่ 1 เป็นงานวิจัยวิธีการคัดเลือกข้อสอบข้อถัดไปจากการสร้างสมการหรือคำนวณความน่าจะเป็นของข้อสอบเพื่อเป็นสารสนเทศในการเลือกใช้ข้อสอบข้อถัดไป กลุ่มที่ 2 เป็นงานวิจัยวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีการทางปัญญาประดิษฐ์ ซึ่งวิธีนี้มีการใช้งานอย่างกว้างขวาง เพราะเป็นวิธีการที่ไม่ต้องใช้การคำนวณที่ต้องซ้ำซ้อนและมีการยืดหยุ่นกับปัญหาได้ดี และผู้วิจัยยังไม่พบว่ามีการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ดังนั้นผู้วิจัยจึงมีความสนใจที่จะริเริ่มนำเทคนิคทางด้านปัญญาประดิษฐ์มาพัฒนาต่อยอดองค์ความรู้ในการพัฒนาการคัดเลือกข้อสอบข้อถัดไปที่ประยุกต์ใช้โครงข่ายประสาทเทียมเพื่อเป็นแนวทางในการวิจัยและพัฒนาต่อไป

กรอบแนวคิดในการวิจัย

การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปที่นำเสนอในงานวิจัยนี้ ได้นำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์เข้ากับทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory) และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) โดยใช้อัลกอรึทึมโครงข่ายประสาทเทียม ทำการคัดเลือกข้อสอบและจัดเรียงข้อสอบแบบอัตโนมัติให้เหมาะสมที่สุดเพื่อนำไปใช้ในการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะแบบพหุมิติเลือกใช้โครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron: MLP) การเรียนรู้แบบมีผู้สอน (supervised learning) โครงข่ายประสาทเทียมนี้ประกอบด้วยเซลล์ประสาท 3 ชั้น ได้แก่ 1) ชั้นข้อมูลป้อนเข้า 2) ชั้นซ่อน และ 3) ชั้นข้อมูลส่งออก สามารถเขียนเป็นกรอบแนวคิดในการวิจัย ดังภาพที่



ภาพที่ 32 กรอบแนวคิดที่ใช้ในการวิจัย

บทที่ 3
วิธีดำเนินการวิจัย

การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ผู้วิจัยได้แบ่งวิธีดำเนินการวิจัยได้ 4 ระยะ ได้แก่

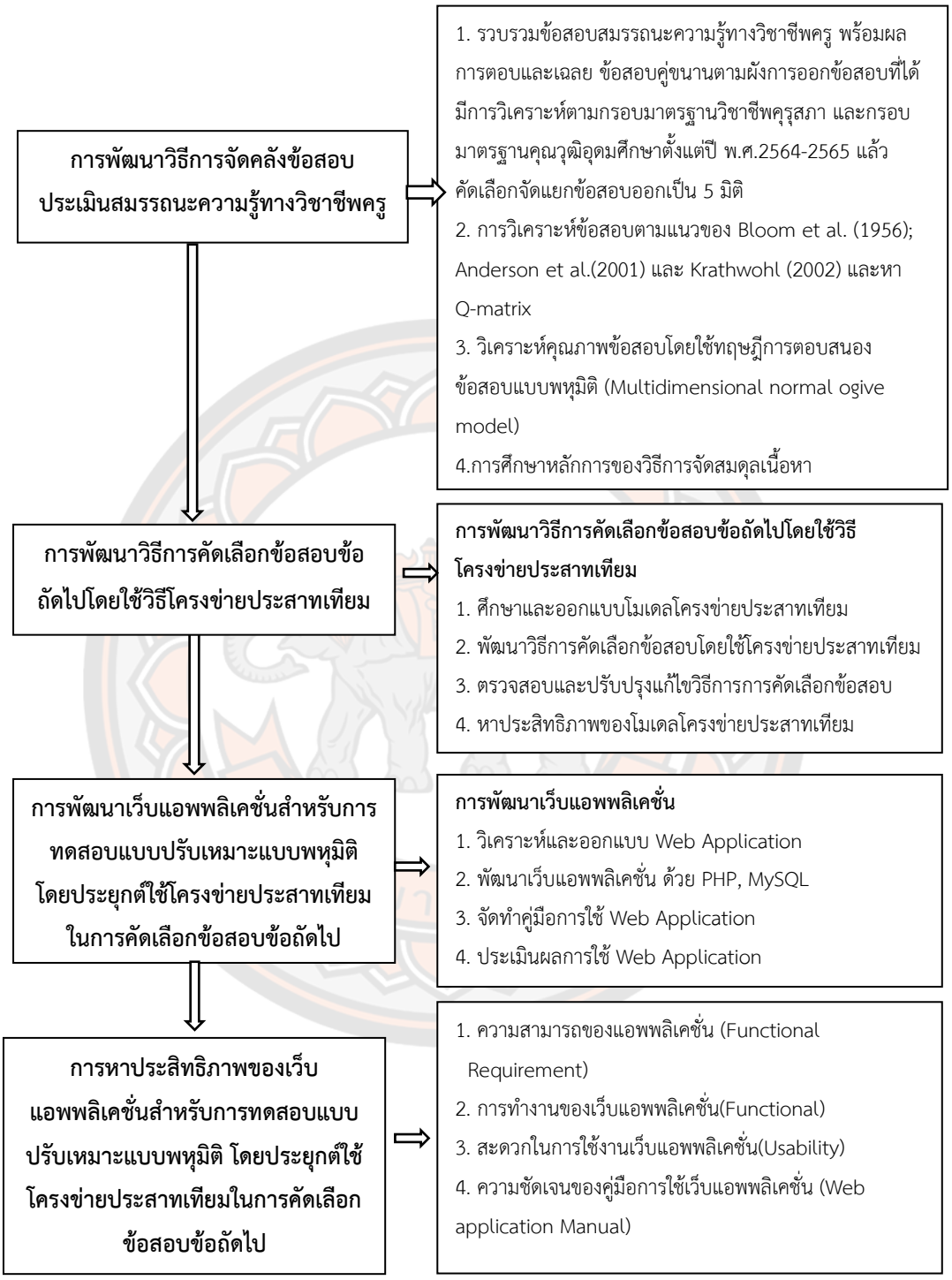
ระยะที่ 1 การพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู

ระยะที่ 2 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

ระยะที่ 3 การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ระยะที่ 4 การหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

การวิจัยนี้สามารถแสดงขั้นตอนการดำเนินการวิจัยได้ ดังภาพที่ 33



ภาพที่ 33 ขั้นตอนการดำเนินการวิจัย

ระยะที่ 1 การพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู

1. การพัฒนาวิธีการจัดคลังข้อสอบ โดยประยุกต์โมเดลการตอบสนองข้อสอบแบบพหุมิติ มีรายละเอียดในการดำเนินการดังนี้

ขั้นที่ 1 การศึกษาเอกสารที่เกี่ยวข้องกับแนวคิดของ Bloom et al., (1956); Anderson et al., (2001, p. 21); Krathwohl (2002, p. 213)

กำหนดโครงสร้างสองมิติของจุดประสงค์ทางการศึกษาของ Bloom et al., (1956) และแบบปรับใหม่ โดย Anderson (Anderson et al., 2001, pp. 67-68; Krathwohl, 2002, pp. 213-215) วิเคราะห์เนื้อหาการออกข้อสอบ การวิเคราะห์เนื้อหาการออกข้อสอบ พร้อมระบุขอบข่ายประเด็น และระดับพฤติกรรม ทั้ง 260 ข้อ ตัวอย่างการวิเคราะห์ข้อสอบ แสดงดังภาพที่ 34

☐ การเปลี่ยนแปลงบริบทฯ ☑ จิตวิทยาฯ ☐ หลักสูตรฯ ☐ วัดประเมิน ☐ ประกันคุณภาพ

ขอบข่าย/ประเด็น 3. ให้คำแนะนำช่วยเหลือผู้เรียนให้มีคุณภาพชีวิตที่ดีขึ้นได้
3.2 วิเคราะห์ปัจจัยที่เกี่ยวข้องกับปัญหาและการแก้ไขปัญหาได้อย่างสมเหตุสมผล

ระดับพฤติกรรม ☐ เข้าใจ ☑ **นำไปใช้** ☐ วิเคราะห์ ☐สังเคราะห์/สร้างสรรค์ ☐ ประเมิน

ระดับความยากง่าย ☐ ง่าย ☑ **ปานกลาง** ☐ ยาก

โจทย์คำถาม (สถานการณ์)

นางสาวใจดี มีผลการเรียนอยู่ในเกณฑ์ดี กำลังศึกษาอยู่ในชั้นมัธยมศึกษาปีที่ 5 ต้องการขอคำปรึกษาในการวางแผนการศึกษาต่อ ในฐานะที่ท่านเป็นครูประจำชั้น จะให้คำปรึกษากับนางสาวใจดีอย่างไร

ตัวเลือกที่	ตัวเลือก
1	เลือกเรียนในสาขาที่ตนเองชอบและมีความถนัด
2	เลือกเรียนในสาขาที่มีโอกาสเข้าทำงานที่มีค่าตอบแทนสูงๆ
3	เลือกเรียนในสาขาที่มีโอกาสรับราชการเพราะมีความมั่นคง
4	เลือกเรียนในสาขาที่ครอบครัวให้การสนับสนุน ครอบครัวต้องการให้เรียน
5	เลือกเรียนในสาขาที่มีคนเลือกเรียนน้อย เมื่อมีคนจบน้อยก็มีโอกาสเข้าทำงาน

เกณฑ์การให้คะแนนตัวเลือก และเหตุผล

ตัวเลือกที่	คะแนน	เหตุผล
1	1	การให้คำปรึกษากับนักเรียน แนะแนวการศึกษาต่อ นักเรียนจะมีโอกาสประสบความสำเร็จในการเรียนจากสิ่งที่ตนเองชอบและมีความถนัด
2	0	
3	0	
4	0	
5	0	

ภาพที่ 34 การวิเคราะห์เนื้อหาการออกข้อสอบ พร้อมระบุขอบข่ายประเด็น และระดับพฤติกรรม

ขั้นที่ 2 ผู้วิจัยนำข้อสอบสมรรถนะความรู้ทางวิชาชีพครู และเฉลย เป็นข้อสอบการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครู ที่ได้มาจากสภาคณบดีคณะครุศาสตร์มหาวิทยาลัยราชภัฏ (สครภ) ลักษณะข้อสอบคู่ขนานตามผังการออกข้อสอบที่ได้มีการวิเคราะห์ตามกรอบ

มาตรฐานวิชาชีพคุรุสภา และกรอบมาตรฐานคุณวุฒิอุดมศึกษาที่ผ่านการวิพากษ์และแก้ไขปรับปรุงโดยผู้ทรงคุณวุฒิแล้ว จำนวน 270 ข้อ มาวิเคราะห์ตามจุดประสงค์ทางการศึกษาใหม่ตามแนวคิดของ Bloom et al., (1956); Anderson et al., (2001, p. 21) และ Krathwohl (2002, p. 213) โดยการปรับลำดับขั้นและคำศัพท์ที่ใช้ในกระบวนการพุทธิปัญญา ซึ่งมี 6 กระบวนการเหมือนเดิม แต่ 3 กระบวนการแรกเปลี่ยนชื่อเป็น จำ (Remember) เข้าใจ (Understand) และประยุกต์ (Apply) ส่วน 3 กระบวนการหลังเปลี่ยนชื่อที่มีลักษณะเป็นคำนามไปเป็นคำกริยาและสลับที่กระบวนการที่ 5 กับ 6 และสร้างสรรค์ (Create) เปลี่ยนชื่อมาจาก การสังเคราะห์ (Synthesis) (Anderson et al., 2001, p. 21; Krathwohl, 2002, p. 213) และเปลี่ยนโครงสร้างจากมิติเดียวเป็นสองมิติ โดยได้เพิ่มโครงสร้างในมิติด้านความรู้ (Knowledge Dimension) เข้ามาในโครงสร้างของจุดประสงค์ทางการศึกษาด้านพุทธิปัญญา ทำให้โครงสร้างใหม่มีลักษณะเป็นสองมิติประกอบด้วยมิติด้านกระบวนการพุทธิปัญญา และมิติด้านความรู้ รวมทั้งอธิบายความหมายของมิติด้านความรู้โดยมิติด้านความรู้ แบ่งเป็น 4 ส่วน คือ Factual Knowledge Conceptual Knowledge Procedural Knowledge และ Metacognitive Knowledge และมิติด้านกระบวนการพุทธิปัญญา แบ่งเป็น 6 กระบวนการ คือ จำ (Remember) เข้าใจ (Understand) ประยุกต์ (Apply) วิเคราะห์ (Analyze) ประเมินค่า (Evaluate) และสร้างสรรค์ (Create)

ขั้นที่ 3 วิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติเป็นโมเดลการวัดที่ถือว่าการตอบสนองข้อสอบของผู้สอบเกิดจากคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบ ซึ่งการพิจารณาความสามารถของคนจากหลายองค์ประกอบ จะช่วยให้โมเดลมีความสอดคล้องกับข้อมูลได้ดียิ่งขึ้น ทั้งนี้โมเดลการตอบสนองข้อสอบแบบพหุมิติมีอยู่หลากหลายโมเดล เช่น โมเดลโลจิสติกแบบพหุมิติ โมเดลปกติสะสมแบบพหุมิติ โมเดลสำหรับการวัดมิติที่ไม่สามารถทดแทนได้ โมเดลสำหรับการวัดการเรียนรู้และการเปลี่ยนแปลง โมเดลในการระบุโครงสร้างของระดับคุณลักษณะ และโมเดลสำหรับการจำแนกกลุ่มบุคคลในการวิเคราะห์ข้อมูลทางสถิติ สามารถแบ่งโมเดลดังกล่าวได้เป็น 2 ลักษณะคือ Exploratory MIRT Model และ Confirmatory MIRT Model ในการวิจัยนี้โมเดลที่ใช้คือ โมเดลปกติสะสมแบบพหุมิติ เป็นโมเดลที่มีการประมาณค่าพารามิเตอร์ได้ใกล้เคียงกับโมเดลโลจิสติกแบบพหุมิติ (ค่าพารามิเตอร์มีความแตกต่างกันไม่เกิน .01) และวิธีที่เหมาะสมสำหรับโมเดลปกติสะสมแบบพหุมิติ คือ วิธีการวิเคราะห์ข้อมูลแบบ Confirmatory MIRT Model มีความเหมาะสมในการยืนยันความสัมพันธ์ระหว่างข้อสอบ(ตัวแปรสังเกตได้) กับมิติด้านความรู้ แบ่งเป็น 4 ส่วน คือ ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge) ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge) ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge) และความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge) และมิติด้านกระบวนการพุทธิปัญญา แบ่งเป็น 6

กระบวนการ คือ จำ (Remember) เข้าใจ (Understand) ประยุกต์ (Apply) วิเคราะห์ (Analyze) ประเมินค่า (Evaluate) และสร้างสรรค์ (Create) แสดงดังตารางที่ 6 – 10

ตารางที่ 6 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหาเรื่อง การเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ

มิติความรู้ (The Knowledge Dimension)	**มิติกระบวนการพุทธิปัญญา (The Cognitive Process Dimension)**						**จำนวนข้อ**
	จำ (Remember)	**เข้าใจ (Understand)**	**ประยุกต์ (Apply)**	**วิเคราะห์ (Analyze)**	**ประเมินค่า (Evaluate)**	**สร้างสรรค์ (Create)**	
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge)	2	5	5	5	3		20
2. ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge)		5	5	5	5		20
3. ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge)		5	5	5	5		20
4. ความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge)							
รวม							**60**

ตารางที่ 7 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหาเรื่อง จิตวิทยาการศึกษา จำนวน 50 ข้อ

มิติความรู้ (The Knowledge Dimension)	**มิติกระบวนการพุทธิปัญญา (The Cognitive Process Dimension)**						**จำนวนข้อ**
	จำ (Remember)	**เข้าใจ (Understand)**	**ประยุกต์ (Apply)**	**วิเคราะห์ (Analyze)**	**ประเมินค่า (Evaluate)**	**สร้างสรรค์ (Create)**	
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge)	2	2	2	2	2		10
2. ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge)	4	4	4	4	4		20
3. ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge)		4	4	4	4	4	20
4. ความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge)							
รวม							**50**

ตารางที่ 8 การกำหนดโครงสร้างพฤติกรรม(มิติ) ที่วัดเนื้อหาเรื่อง การประกันคุณภาพการศึกษา จำนวน 50 ข้อ

มิติความรู้ (The Knowledge Dimension)	**มิติกระบวนการพุทธิปัญญา (The Cognitive Process Dimension)**						**จำนวนข้อ**
	จำ (Remember)	**เข้าใจ (Understand)**	**ประยุกต์ (Apply)**	**วิเคราะห์ (Analyze)**	**ประเมินค่า (Evaluate)**	**สร้างสรรค์ (Create)**	
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge)	2	2	2	2	2		10
2. ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge)		4	4	4	4	4	20
3. ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge)		4	4	4	4	4	20
4. ความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge)							
รวม							**50**

ตารางที่ 9 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่วัดเนื้อหา เรื่อง วัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ

มิติความรู้ (The Knowledge Dimension)	**มิติกระบวนการพุทธิปัญญา (The Cognitive Process Dimension)**						**จำนวนข้อ**
	จำ (Remember)	**เข้าใจ (Understand)**	**ประยุกต์ (Apply)**	**วิเคราะห์ (Analyze)**	**ประเมินค่า (Evaluate)**	**สร้างสรรค์ (Create)**	
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge)	4	4	4	4	4		20
2. ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge)		4	4	4	4	4	20
3. ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge)		4	4	4	4	4	20
4. ความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge)							
รวม							**60**

ตารางที่ 10 การกำหนดโครงสร้างพฤติกรรม (มิติ) ที่เนื้อหาเรื่อง หลักสูตรและการสอน จำนวน 50 ข้อ

มิติความรู้ (The Knowledge Dimension)	มิติกระบวนการพุทธิปัญญา (The Cognitive Process Dimension)						จำนวนข้อ
	จำ (Remember)	เข้าใจ (Understand)	ประยุกต์ (Apply)	วิเคราะห์ (Analyze)	ประเมินค่า (Evaluate)	สร้างสรรค์ (Create)	
1. ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge)	2	2	2	2	2		10
2. ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge)		4	4	4	4	4	20
3. ความรู้เกี่ยวกับวิธีดำเนินการ (Procedural Knowledge)		4	4	4	4	4	20
4. ความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge)							
รวม							**50**

1.1.1 สร้าง Q-matrix โดยผู้วิจัยจัดแยกข้อสอบ ออกเป็น 5 มิติ คือ มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ มิติการประกันคุณภาพการศึกษา จำนวน 50 ข้อมิติจิตวิทยาการศึกษา

จำนวน 50 ข้อมิติการเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ และมิติหลักสูตรและการสอน จำนวน 50 ข้อ รวม 270 ข้อ แสดงดังตารางที่ 11

ตารางที่ 11 โครงสร้างของข้อสอบจำแนกตามจุดประสงค์ทางการศึกษาด้านพุทธิปัญญาและสมรรถนะความรู้ทางวิชาชีพครูที่ออกข้อสอบ จำแนกตามมิติของข้อสอบ

จุดประสงค์ทางการศึกษาด้านพุทธิปัญญาที่มุ่งวัด	**มิติความรู้ทางวิชาชีพครู**					**รวม**
	หลักสูตรและการสอน	**จิตวิทยา**	**การประกันคุณภาพการศึกษา**	**วัดผลและวิจัยการเรียนรู้**	**การเปลี่ยนแปลงบริบทโลก**	
1. วิเคราะห์เกี่ยวกับข้อเท็จจริง	10	10	5	10	10	45
2. เข้าใจเกี่ยวกับมโนทัศน์	10	10	10	10	10	50
3. ประยุกต์เกี่ยวกับมโนทัศน์	10	10	5	10	10	45
4. วิเคราะห์เกี่ยวกับมโนทัศน์	10	10	10	10	10	50
5. ประยุกต์เกี่ยวกับวิธีดำเนินการ	5	5	10	10	10	40
6. วิเคราะห์เกี่ยวกับวิธีดำเนินการ	5	5	10	10	10	40
รวม	**50**	**50**	**50**	**60**	**60**	**270**

1.1.2 นำข้อสอบ ข้อสอบสมรรถนะความรู้ทางวิชาชีพครู ที่อยู่ในข้อที่ 1.1.1 ไปให้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ตรวจสอบความสอดคล้องระหว่างข้อสอบกับมิติของข้อสอบ (ตรวจสอบความเหมาะสมของ Q-matrix) แล้วนำผลที่ได้มาวิเคราะห์ปรับปรุงตามข้อเสนอแนะของผู้เชี่ยวชาญ และคัดเลือกข้อสอบที่มีค่า IOC ของความตรงเชิงเนื้อหา และความเหมาะสมของการะบุกระบวนการพุทธิปัญญาตั้งแต่ 0.50 ขึ้นไป เพื่อนำไปใช้ในการจัดทำแบบทดสอบแบบเขียนตอบหาค่าความเที่ยงตรงเชิงเนื้อหา

โดยการพิจารณาจากค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (Index of Consistency)(สมบัติ ท้ายเรือคำ, 2552, น. 107)

1.1.3 นำแบบประเมินความสอดคล้องระหว่างข้อสอบกับมิติของข้อสอบ (ตรวจสอบความเหมาะสมของ Q-matrix) ที่สร้างขึ้นให้ผู้เชี่ยวชาญตรวจสอบความสอดคล้อง ระหว่างข้อสอบกับมิติของข้อสอบ โดยมีเกณฑ์ในการให้คะแนนความสอดคล้องระหว่างข้อสอบกับมิติของข้อสอบ ดังนี้

ให้ +1 คะแนน เมื่อแน่ใจว่าข้อสอบนั้นวัดได้สอดคล้องกับมิติที่ระบุ

ให้ 0 คะแนน เมื่อไม่แน่ใจว่าข้อสอบนั้นวัดได้สอดคล้องกับมิติที่ระบุ

ให้ -1 คะแนน เมื่อแน่ใจว่าข้อสอบนั้นไม่สอดคล้องกับมิติที่ระบุ

ผลการตรวจสอบความเหมาะสมของ Q-matrix โดยผู้เชี่ยวชาญ พบว่ามีข้อสอบผ่านเกณฑ์ จำนวน 270 ข้อ ประกอบด้วยข้อสอบที่วัดมิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ มิติการประกันคุณภาพการศึกษา จำนวน 50 ข้อ มิติจิตวิทยาการศึกษา จำนวน 50 ข้อ มิติการเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ และมิติหลักสูตรและการสอน จำนวน 50 ข้อ รวม 270 ข้อ ตามลำดับ ทั้งนี้ผู้วิจัยได้นำข้อสอบที่ผ่านเกณฑ์ ไปทดลองใช้กับนักศึกษาที่กำลังเรียนคณะครุศาสตร์ ชั้นปีที่ 3-4 คำนวณค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติต่อไป

1.1.4 ผู้วิจัยรวบรวมข้อมูลผลการตอบข้อสอบของนักศึกษา จำนวน 500 คน ซึ่งเป็นกลุ่มทดลองใช้เครื่องมือเป็นนักศึกษาคณะครุศาสตร์ ชั้นปีที่ 3-4 ที่กำลังศึกษาหลักสูตรทางการศึกษาของกลุ่มมหาวิทยาลัยราชภัฏ จำนวน 40 สถาบัน เพื่อนำไปวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามแนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติโดยใช้โปรแกรมสำเร็จรูป NOHARM 4.0

1.1.5 กำหนดเงื่อนไขการวิเคราะห์คุณภาพข้อสอบ ดังต่อไปนี้

เกณฑ์การคัดเลือกข้อสอบ

1) ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) มีค่าตั้งแต่ 0.00 ขึ้นไป

2) ค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) มีค่าในช่วง -4.00 ถึง 4.00

3) ค่าการเดาข้อสอบถูก (c) กำหนดให้ทุกข้อมีค่าคงที่เท่ากับ 0.20 (กรณีข้อสอบเป็นชนิด 5 ตัวเลือก)

รายละเอียดเกี่ยวกับประชากรและการกลุ่มตัวอย่าง มีดังนี้

ประชากรและตัวอย่าง

1. ประชากรที่ใช้ในการศึกษา ได้แก่ นักศึกษาวิชาชีพครู ชั้นปีที่ 3 และ 4 ที่กำลังศึกษาหลักสูตรทางการศึกษาของกลุ่มมหาวิทยาลัยราชภัฏ จำนวน 40 สถาบัน สังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม (อว.) ปีการศึกษา 2566

2. กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักศึกษาหลักสูตรครุศาสตร์บัณฑิต ชั้นปีที่ 3 และ 4 ที่กำลังศึกษาหลักสูตรทางการศึกษาของกลุ่มมหาวิทยาลัยราชภัฏ สังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม ได้มาจากการสุ่มตัวอย่างแบบหลายขั้นตอน (Multi-stage Random Sampling) โดยการแบ่งมหาวิทยาลัยราชภัฏ ออกเป็น 4 ภูมิภาค ได้แก่ ภาคใต้ ภาคกลาง ภาคเหนือ และภาคตะวันออกเฉียงเหนือ โดยแบ่งกลุ่มตัวอย่างเป็น 2 กลุ่ม ดังนี้

กลุ่มที่ 1 เป็นนักศึกษาที่ทดลองใช้ (Tryout) แบบวัดมาตรฐานความรู้วิชาชีพครู เพื่อตรวจสอบคุณภาพของแบบวัดด้านความเที่ยง ความตรง และความยาก ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยได้กำหนดขนาดตัวอย่างโดยใช้ rule of thumb (Hair, Black, Babin, Anderson and Tatham, 2010) ซึ่งกล่าวไว้ว่าการกำหนดขนาดตัวอย่างที่ใช้กันมากในการวิจัยตัวแปรพหุ ควรมีขนาด 10-20 คนต่อหนึ่งตัวแปร (Hair et al., 1998) ซึ่งในการวิจัยนี้มีตัวแปรทั้งหมดจำนวน 5 ตัวแปร ซึ่งเป็นตัวชี้วัดมาตรฐานความรู้วิชาชีพครู

ดังนั้น ผู้วิจัยจึงได้กำหนดขนาดตัวอย่าง 20 คนต่อ 1 ตัวแปร ขนาดตัวอย่างขั้นต่ำอย่างน้อยจำนวน 100 คน และเพื่อต้องการให้เกิดความแม่นยำในการประมาณ ค่าพารามิเตอร์มากขึ้นควรกำหนดขนาดตัวอย่างเป็น 500 คน และเพื่อให้มีความครอบคลุมและเป็นตัวแทนที่ดีของประชากรจึงสุ่มตัวอย่างจากทั้ง 4 ภูมิภาค ได้แก่ ภาคใต้ ภาคกลาง ภาคเหนือ และภาคตะวันออกเฉียงเหนือ

ขั้นตอนที่ 1 จากระดับภูมิภาค สุ่มมหาวิทยาลัยราชภัฏมา ภูมิภาคละ 3 มหาวิทยาลัยโดยใช้วิธีการสุ่มอย่างง่าย (simple random sampling) ด้วยวิธีการจับสลาก

ตารางที่ 12 วิธีการสุ่มแบบตัวอย่างแบบหลายขั้นตอน (Multi-stage random Sampling)

ภูมิภาค	**มหาวิทยาลัยราชภัฏ**	**n**
ภาคใต้	มหาวิทยาลัยราชภัฏสุราษฎร์ธานี	30
	มหาวิทยาลัยราชภัฏภูเก็ต	40
	มหาวิทยาลัยราชภัฏนครศรีธรรมราช	30
	รวม	**100**
ภาคกลาง	มหาวิทยาลัยราชภัฏนครปฐม	40
	มหาวิทยาลัยราชภัฏสวนสุนันทา	30
	มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง	30
	รวม	**120**
ภาคเหนือ	มหาวิทยาลัยราชภัฏอุตรดิตถ์	44
	มหาวิทยาลัยราชภัฏพิบูลสงคราม	43

ภูมิภาค	มหาวิทยาลัยราชภัฏ	n
	มหาวิทยาลัยราชภัฏลำปาง	43
	รวม	**130**
ภาคตะวันออกเฉียงเหนือ	มหาวิทยาลัยราชภัฏมหาสารคาม	44
	มหาวิทยาลัยราชภัฏนครราชสีมา	43
	มหาวิทยาลัยราชภัฏบุรีรัมย์	43
	รวม	**130**
	รวมทั้งสิ้น	**500**

ขั้นตอนที่ 2 สุ่มนักศึกษาจากสุ่มมหาวิทยาลัยราชภัฏ ในแต่ละภูมิภาคทั้ง 4 ภูมิภาค ด้วยการสุ่มแบบแบ่งชั้น (stratified random sampling) มาอย่างละเท่าๆ กัน เพื่อให้ได้ตัวแทนทุกภูมิภาค และจำนวนใกล้เคียงกันทุกภูมิภาค รวมทั้งหมด จำนวน 500 คน

แต่อย่างไรก็ตามในการเก็บข้อมูลจริงผู้วิจัยได้เก็บข้อมูลตัวอย่างถึง 500 คน โดยเก็บข้อมูลเป็นนักศึกษาชั้นปีที่ 3-4 ที่กำลังฝึกประสบการณ์วิชาชีพครู ซึ่งตัวอย่างที่ใช้ในการหาคุณภาพของแบบวัดครั้งนี้ควรเป็นนักศึกษาที่เรียนรายวิชาชีพครูและวิชาชีพพื้นฐานมาแล้ว ดังนั้น นักศึกษาชั้นปีที่ 3 - 4 ป.บัณฑิต (วิชาชีพครู) และ ป.โท (วิชาชีพครู) จึงมีคุณสมบัติดังกล่าว

โดยตัวอย่างทั้งหมดนี้จะใช้สำหรับหาคุณภาพของแบบวัดมาตรฐานความรู้วิชาชีพครู ด้านความตรงตามโครงสร้าง (construct validity) และวิธีการวิเคราะห์ข้อมูลแบบ Confirmatory MIRT Model เพื่อหาคุณภาพข้อสอบรายข้อเพื่อคัดเลือกเข้าคลังข้อสอบ

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยครั้งนี้

1. การจัดเตรียมข้อมูลเอกสารที่เกี่ยวข้องในการเก็บข้อมูล

1.1 จัดทำหนังสือขอความร่วมมือในการเก็บข้อมูลจาก คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร เพื่อขอความอนุเคราะห์ในการเก็บข้อมูล

1.2 นำหนังสือขอความร่วมมือในการเก็บข้อมูลไปติดต่อกับมหาวิทยาลัยราชภัฏที่เป็นกลุ่มตัวอย่าง เพื่อขอความร่วมมือในการเก็บข้อมูล ซึ่งจัดการทดสอบผ่านทาง Website (http://examadaptive.com/) เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปที่ผู้วิจัยพัฒนาขึ้น ซึ่งนักศึกษาสามารถทำการทดสอบที่ใดก็ได้ ทุกที่ ทุกเวลา และตามความพร้อม ระหว่างวันที่ 1 กันยายน ถึง 1 ตุลาคม พ.ศ. 2566

2. หลังจากเสร็จสิ้นการทดสอบตามเวลาที่กำหนด ผู้วิจัยก็นำผลการทดสอบจากนักศึกษามาวิเคราะห์หาประสิทธิภาพของการคัดเลือกข้อสอบข้อถัดไปโดยวิธีการใช้โครงข่ายประสาทเทียม

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยครั้งนี้ แบ่งการวิเคราะห์ข้อมูลออกเป็น 2 ตอน ได้แก่ 1) การวิเคราะห์คุณภาพของเครื่องมือ และ 2) วิเคราะห์ข้อมูลประสิทธิภาพของการคัดเลือกข้อสอบข้อถัดไปโดยวิธีการใช้โครงข่ายประสาทเทียม (ANN) โดยมีรายละเอียดดังนี้

1. การวิเคราะห์คุณภาพของเครื่องมือ

ในขั้นตอนนี้ผู้วิจัยได้วิเคราะห์คุณภาพของข้อสอบเพื่อนำมาจัดทำคลังข้อสอบแบบวัดสมรรถนะวิชาชีพครู เพื่อนำไปบรรจุไว้ในฐานข้อมูลของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยมีรายละเอียดของการวิเคราะห์ข้อมูล ดังนี้

1.1 หาค่าความสอดคล้องระหว่างข้อสอบกับมิติของข้อสอบโดยการพิจารณาจากค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (Index of consistency)

1.2 วิเคราะห์คุณภาพของข้อสอบรายข้อ ตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response theory: MIRT) ด้วยโมเดลปกติสะสมแบบพหุมิติกับโอกาสการเดาข้อสอบ (Multidimensional normal ogive model guessing) (Embretson & Reise, 2000) โดยใช้โปรแกรมสำเร็จรูป NOHARM 4.0

1.3 กำหนดเงื่อนไขการวิเคราะห์คุณภาพข้อสอบ ดังต่อไปนี้

เกณฑ์การคัดเลือกข้อสอบ

1) ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) มีค่าตั้งแต่ 0.00 ขึ้นไป

2) ค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) มีค่าในช่วง -4.00 ถึง 4.00

3) ค่าการเดาข้อสอบถูก (c) กำหนดให้มีค่าคงที่เท่ากับ 0.20

1.4 ความตรงเชิงโครงสร้างของโมเดลพหุมิติของแบบทดสอบ เพื่อบอกความพอเพียงของตัวอย่างที่นำมาใช้ โดยพิจารณาจาก Sum of squares of residuals (SSR) Root mean square of residuals (RMSR) และค่า Tanaka index of goodness of fit (GFI) (McDonald & Mok, 1995)

Sum of squares of residuals (SSR) หมายถึง ค่าดัชนีรากของกำลังสองเฉลี่ยของส่วนที่เหลือ (Root mean square of residuals) ซึ่ง RMSR มีค่าอยู่ระหว่าง 0 ถึง 1 ถ้ามีค่าต่ำกว่า 0.06 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

Tanaka index of goodness of fit (GFI) หมายถึง ค่าดัชนีวัดความเหมาะสมของโมเดลข้อสอบพหุมิติที่ปรับแก้ของ Tanaka (Tanaka index of goodness of fit) ซึ่งแสดงให้เห็นถึงความตรงเชิงโครงสร้างของข้อสอบแบบพหุมิติว่า มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

ซึ่งควรมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้ามีค่าใกล้ 1.00 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

1.5 หาคุณภาพของแบบทดสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response theory: MIRT) โดยค่าพารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ คือ ค่าพารามิเตอร์ของข้อสอบ (item parameter) ค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบแบบตรวจให้คะแนน 2 ค่า ได้จากการวิเคราะห์องค์ประกอบด้วยโมเดลปกติสะสมแบบพหุมิติ (Multidimensional normal ogive model) (Embretson & Reise, 2000) โดยใช้โปรแกรม Noharm 4.0 เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก (a_{ik}) ค่าจุดตัดของความยาก (d_i) ตามรูปแบบทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ดังสมการ

$$p(x_i=1)=\frac{\sum_{k=1}^{m} a_{ik}\theta_k+d_i}{1{,}0+e^{\sum_{k=1}^{m} a_{ik}\theta_k+d_i}}$$

เมื่อ X_i แทน คะแนน(0,1) ของข้อสอบข้อที่ i

a_{ij} แทน ส่วนประกอบของ a_i – เวกเตอร์ของมิติที่ k

di แทน พารามิเตอร์ที่มีความสัมพันธ์กับความความยากของข้อสอบในแบบทดสอบ

θ_k แทน เวกเตอร์ค่าพารามิเตอร์ความสามารถ

1.6 ค่าพารามิเตอร์อำนาจจำแนกแบบพหุมิติของข้อสอบ ($MDISC_i$) ควรมีค่าอยู่ระหว่าง -4.00 ถึง 4.00 ถ้ามีค่าอยู่ช่วงพิสัยดังกล่าว แสดงว่าข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติที่เหมาะสม ดังสมการ

$$MDISC_i=\sqrt{\sum_{k=1}^{m} a_{ik}^2}$$

เมื่อ $MDISC_i$ แทน ค่าอำนาจจำแนกรวมของข้อสอบข้อที่ i ใน m มิติ

m แทน จำนวนของมิติใน θ-space

a_{ik} แทน ส่วนประกอบของ ai –เวกเตอร์ของมิติที่ k

จากสมการจะเห็นได้ว่า $MDISC_i$ เป็นสัดส่วนของความชันพื้นผิวการตอบสนองข้อสอบตรงตำแหน่งที่มีความชันสูงสุดจึงสังเกตได้ว่ามีลักษณะที่คล้ายกับค่าพารามิเตอร์อำนาจจำแนกในโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Reckase, 2009)

1.7 ค่าพารามิเตอร์ความยากแบบพหุมิติของข้อสอบ (MDIFFi) ควรมีค่าอยู่ระหว่าง -4.00 ถึง 4.00 ถ้ามีค่าอยู่ช่วงพิสัยดังกล่าว แสดงว่าข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติที่เหมาะสม ดังสมการ

$$MDIFF_i = \frac{-d_i}{\sqrt{\sum_{k-1}^{m} a_{ik}^2}}$$

เมื่อ d_i แทน พารามิเตอร์ที่มีความสัมพันธ์กับค่าความยากของข้อสอบในแบบทดสอบ

จากสมการ พบว่าค่า d_i เป็นพารามิเตอร์ที่มีความสัมพันธ์กับค่าความยากของข้อสอบในแบบทดสอบอย่างไรก็ตาม $MDIFF_i$ ไม่สามารถแปลความหมายเหมือนค่าความยากของข้อสอบในโมเดลการตอบสนองข้อสอบแบบเอกมิติซึ่งค่าของ $MDIFF_i$ บ่งชี้ถึงระยะทางจากจุดกำเนิดของ θ-space ไปยังจุดที่มีความชันสูงสุด (Steepest Slope)

เกณฑ์คุณภาพของแบบทดสอบแบบพหุมิติ

เกณฑ์คุณภาพของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (Multidimension Discriminant : MDISC) และค่าความยากแบบพหุมิติของข้อสอบ (Multidimension Difficulties : MDIFF) จำแนกตามเกณฑ์พิจารณาของ Baker (2001) และ Hasmy (2014) ดังตารางที่ 13 และ 14 ดังนี้

ตารางที่ 13 เกณฑ์พิจารณาค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC)

เกณฑ์	**แปลความหมาย**
MDISC ≥ 1.7	สูงมาก (Very high)
1.35 ≤ MDISC < 1.7	สูง (High)
0.65 ≤ MDISC < 1.35	ปานกลาง (Moderate)
0.35 ≤ MDISC < 0.65	ต่ำ (Low)
MDISC < 0.35	ต่ำมาก (Very low)

ตารางที่ 14 เกณฑ์พิจารณาค่าความยากแบบพหุมิติของข้อสอบ (MDIFF)

เกณฑ์	แปลความหมาย
MDIFF ≥ 2	ยากมาก (Very hard)
0.5 ≤ MDIFF < 2	ยาก (Hard)
-0.5 ≤ MDIFF < 0.5	ปานกลาง (Medium)
-2 ≤ MDIFF < -0.5	ง่าย (Easy)
MDIFF < -2	ง่ายมาก (Very easy)

1.1.6 คัดเลือกข้อสอบที่ตรงตามเกณฑ์การคัดเลือกข้อสอบในข้อ 1.1.5 ทั้ง 3 เกณฑ์ เข้าคลังข้อสอบ ข้อสอบที่ได้รับการคัดเลือกเข้าคลังข้อสอบ แสดงดังตารางที่ 15

ตารางที่ 15 จำนวนข้อสอบที่ผ่านเกณฑ์จากการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติที่ได้รับคัดเลือกเข้าคลังข้อสอบ จำแนกตามมิติของข้อสอบ

จุดประสงค์ทางการศึกษาด้านพุทธิปัญญาที่มุ่งวัด	**มิติความรู้ทางวิชาชีพครู**					**รวม**
	หลักสูตรและการสอน	**จิตวิทยา**	**การประกันคุณภาพการศึกษา**	**วัดผลและวิจัยการเรียนรู้**	**การเปลี่ยนแปลงบริบทโลก**	
1. วิเคราะห์เกี่ยวกับข้อเท็จจริง	8	8	5	10	7	38
2. เข้าใจเกี่ยวกับมโนทัศน์	7	7	7	10	10	41
3. ประยุกต์เกี่ยวกับมโนทัศน์	8	7	8	10	8	41
4. วิเคราะห์เกี่ยวกับมโนทัศน์	8	5	8	10	9	40
5. ประยุกต์เกี่ยวกับวิธีดำเนินการ	8	7	9	10	10	44
6. วิเคราะห์เกี่ยวกับวิธีดำเนินการ	8	6	8	10	6	38
รวม	**47**	**40**	**45**	**60**	**50**	**242**

จากตารางที่ 15 ผลการวิเคราะห์คุณภาพข้อสอบตามหลักการของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยใช้โปรแกรมสำเร็จรูป NOHARM 4.0 มีจำนวนข้อสอบที่ผ่านเกณฑ์ 242 ข้อ จำแนกตามจุดประสงค์ทางการศึกษาด้านพุทธิปัญญาและสมรรถนะความรู้ทางวิชาชีพครูที่ออกข้อสอบ ได้ดังนี้ 1) วิเคราะห์เกี่ยวกับข้อเท็จจริง จำนวน 38 ข้อ 2) เข้าใจเกี่ยวกับมโนทัศน์ จำนวน 41 ข้อ

3) ประยุกต์เกี่ยวกับมโนทัศน์ จำนวน 41 ข้อ 4) วิเคราะห์เกี่ยวกับมโนทัศน์ จำนวน 40 ข้อ 5) ประยุกต์เกี่ยวกับวิธีดำเนินการ จำนวน 44 ข้อ และ 6) วิเคราะห์เกี่ยวกับวิธีดำเนินการ จำนวน 38 ข้อ

1.1.7 นำข้อสอบที่ได้รับการคัดเลือกจัดทำเป็นคลังข้อสอบ โดยใช้โปรแกรม MySQL ในการจัดทำฐานข้อมูลคลังข้อสอบ

ขั้นที่ 4 การศึกษาหลักการของวิธีการจัดสมดุลเนื้อหาและการแสดงข้อสอบการจัดสมดุลเนื้อหาของข้อสอบ (Content Balancing) เป็นสิ่งสำคัญอย่างยิ่ง โดยเฉพาะการวัดผลสัมฤทธิ์ทางการศึกษาในการทดสอบแบบดั้งเดิม แบบทดสอบมาตรฐานส่วนใหญ่ได้สร้างตามรายละเอียดในตารางการกำหนดคุณลักษณะของข้อสอบ (Table of Content Specification) ผู้สอบทุกคนจะได้รับแบบทดสอบชุดเดียวกัน ประกอบด้วยข้อสอบครบทุกคุณลักษณะที่ต้องการวัดตามที่ออกแบบไว้ ด้วยเหตุนี้จึงเป็นข้อจำกัดในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เนื่องจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จะคัดเลือกข้อสอบที่ให้ค่าสารสนเทศสูงสุดที่ค่าประมาณความสามารถของผู้สอบ ณ ขณะนั้น ซึ่งไม่ได้พิจารณาขอบเขตเนื้อหา (Content Area) ของข้อสอบ เป็นผลให้ผู้สอบบางคนไม่ได้รับข้อสอบทุกขอบเขตเนื้อหาของข้อสอบ ตัวอย่างเช่นในการวัดความสามารถด้านการคำนวณ การบวก การลบ การคูณ และการหาร ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้สอบอาจได้รับข้อสอบเฉพาะการบวก การลบ แต่ไม่ได้รับข้อสอบการคูณและการหาร ส่งผลให้ผู้ที่ไม่มีความรู้เรื่องการคูณและการหาร จะได้รับการประเมินความสามารถสูงเกินความจริง ในขณะเดียวกัน ผู้สอบที่มีความรู้เรื่องการคูณและการหาร จะถูกประเมินต่ำเกินจริง ดังนั้น เพื่อรับประกันว่าผู้สอบแต่ละคนจะได้รับข้อสอบที่เป็นตัวแทนตามสัดส่วนของแต่ละขอบเขตของเนื้อหาครบถ้วน

ขั้นที่ 5 จัดคลังข้อสอบตามจัดคลังข้อสอบตามระดับชั้น โดยประยุกต์วิธีการของ Chang and Ying, (1996); Baker, (2001); Hasmy, (2014)

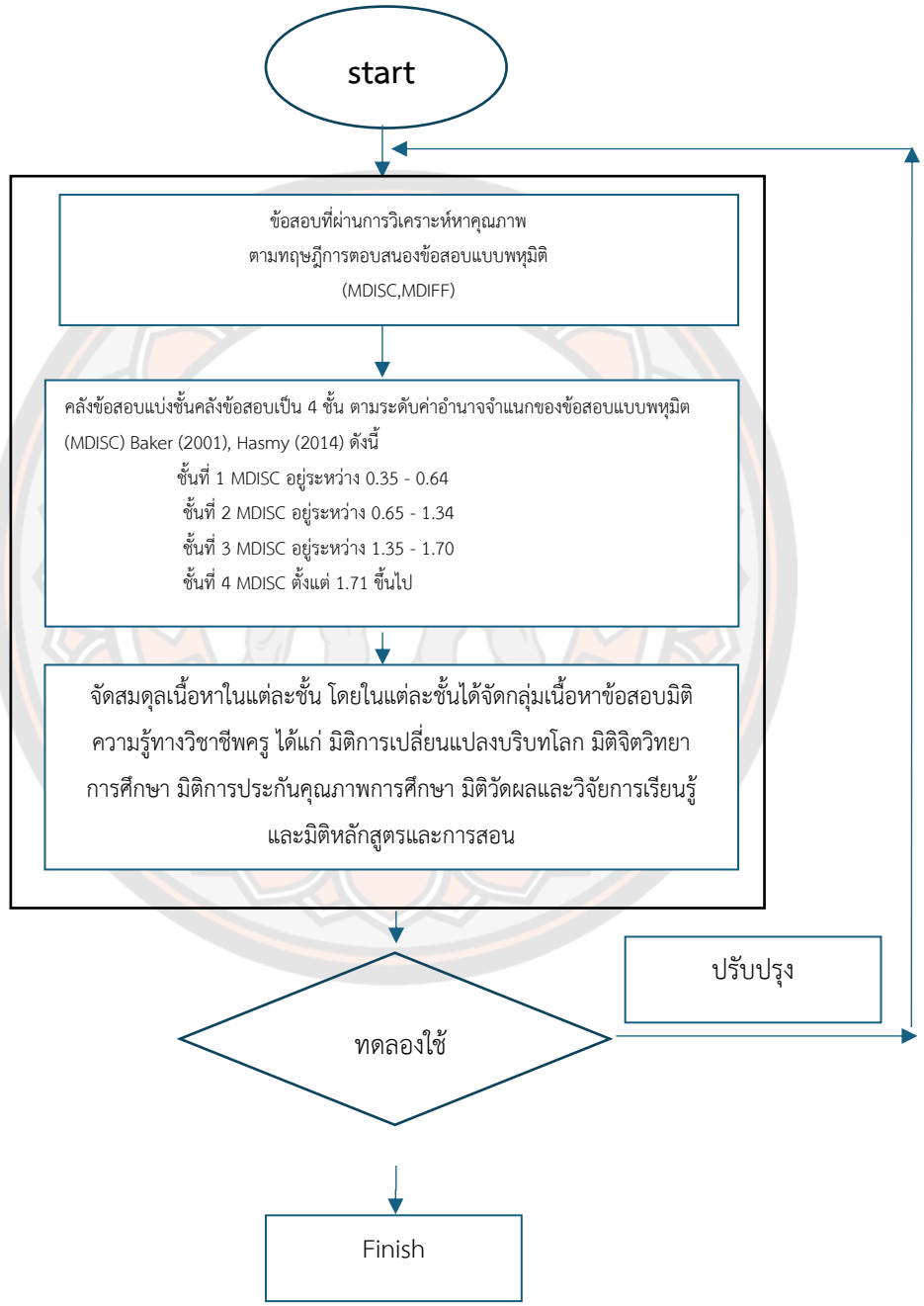
วิธีการจัดคลังข้อสอบตามอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) ของข้อสอบ แบบเป็นช่วงของแต่ละระดับชั้นร่วมกับการจัดสมดุลเนื้อหามีขั้นตอนดังนี้

5.1 การพัฒนาคลังข้อสอบเริ่มจากแบ่งชั้นคลังข้อสอบเป็น 4 ชั้น ตามระดับค่าอำนาจจำแนกของข้อสอบแบบพหุมิต (MDISC) Baker, (2001); Hasmy, (2014) ดังนี้ ชั้นที่ 1 MDISC อยู่ระหว่าง 0.35-0.64 ชั้นที่ 2 MDISC อยู่ระหว่าง 0.65-1.34 ชั้นที่ 3 MDISC อยู่ระหว่าง 1.35-1.70 และชั้นที่ 4 MDISC ตั้งแต่ 1.71 ขึ้นไป

5.2 การพัฒนาคลังข้อสอบเริ่มจากแบ่งชั้นคลังข้อสอบเป็น 4 ชั้น ตามระดับค่าอำนาจจำแนกข้อสอบร่วมกับการจัดสมดุลเนื้อหาในแต่ละชั้น โดยในแต่ละชั้นได้จัดกลุ่มเนื้อหาข้อสอบ

สมรรถนะความรู้ทางวิชาชีพครู ได้แก่ มิติการเปลี่ยนแปลงบริบทโลก มิติจิตวิทยาการศึกษา มิติการประกันคุณภาพการศึกษา มิติวัดผลและวิจัยการเรียนรู้ และมิติหลักสูตรและการสอน

5.3 จัดคลังข้อสอบตามสมรรถนะความรู้ทางวิชาชีพครู ของข้อสอบแบบเป็นช่วงของแต่ระดับชั้นร่วมกับการจัดสมดุลเนื้อหา ตามรายละเอียดดังภาพ 3.3



ภาพที่ 35 ขั้นตอนการพัฒนาวิธีการจัดคลังข้อสอบประเมินมิติความรู้ทางวิชาชีพครู

ระยะที่ 2 การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

เนื่องจากวิธีโครงข่ายประสาทเทียมไม่ได้ถูกพัฒนาขึ้นมา เพื่อใช้ในการคัดเลือกข้อสอบข้อถัดไปในการทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยเฉพาะ จึงส่งผลให้ไม่สามารถนำหลักการของวิธีดังกล่าวมาใช้กับการเลือกข้อสอบได้โดยตรง ดังนั้นจึงต้องทำการคิดวิเคราะห์และออกแบบขั้นตอนการดำเนินงานต่างๆ เพื่อให้วิธีโครงข่ายประสาทเทียมเลือกข้อสอบได้อย่างมีประสิทธิภาพซึ่งการดำเนินการดังกล่าว ผู้วิจัยได้ออกแบบขั้นตอนการพัฒนาเป็น 3 ขั้นตอนมีรายละเอียดดังต่อไปนี้

1. ศึกษาและออกแบบโมเดลโครงข่ายประสาทเทียม

1.1 ศึกษาวิธีการโครงข่ายประสาทเทียม ทฤษฎีที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ การทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ จากบทความและงานวิจัยที่เกี่ยวข้องเพื่อค้นหารูปแบบหรือวิธีการของโครงข่ายประสาทเทียมที่สามารถนำมาใช้ในการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปของการทดสอบแบบปรับเหมาะแบบพหุมิติ

1.2 วิเคราะห์และออกแบบโมเดลโครงข่ายประสาทเทียม โครงข่ายประสาทเทียมประกอบด้วย 3 ชั้นได้แก่ ชั้นข้อมูลเข้า ชั้นข้อมูลซ่อน และชั้นข้อมูลออก ในการพัฒนาโมเดลวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม เพื่อนำไปใช้ในขั้นตอนของการคัดเลือกข้อสอบข้อถัดไปของการทดสอบแบบปรับเหมาะแบบพหุมิติ พบว่าโครงข่ายประสาทเทียมที่เหมาะสมกับการนำมาใช้ในการคัดเลือกข้อสอบข้อถัดไป และเป็นเทคนิควิธีการจำแนกประเภทข้อมูลวิธีการหนึ่งที่นำมาใช้ในการแก้ปัญหาอย่างแพร่หลาย ซึ่งให้ผลลัพธ์การจำแนกประเภทได้อย่างแม่นยำ อีกทั้งยังสามารถแก้ปัญหาที่มีความสัมพันธ์ซับซ้อน และข้อมูลที่มีความสัมพันธ์ไม่เป็นเชิงเส้นได้ โดยการทดลองฝึกเรียนรู้ในรูปแบบการนำเข้าข้อมูลและการส่งออกข้อมูล ในงานวิจัยนี้เลือกใช้โครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron: MLP) การเรียนรู้แบบมีผู้สอน (supervised learning) แบบการแพร่ย้อนกลับ (back propagation Learning) เนื่องจากมีความสามารถในการเรียนรู้ความสัมพันธ์ไม่เป็นเชิงเส้นได้ดี (Mohammadi & Mirabedini, 2014, pp. 112-113) โครงข่ายประสาทเทียมนี้ประกอบด้วยเซลล์ประสาท 3 ชั้น ได้แก่ 1) ชั้นข้อมูลป้อนเข้า 2) ชั้นข้อมูลส่งออก และ 3) ชั้นซ่อน ซึ่งกำหนดสถาปัตยกรรมของโครงข่ายประสาทเทียม (ANN) พิจารณาตามขั้นตอน ดังนี้

1.3 ในงานวิจัยนี้ได้ออกแบบออกแบบโมเดลโครงข่ายประสาทเทียมจากข้อสอบ 5 ข้อ X 5 มิติ (5-DIMENSIONAL) ฉะนั้นจะได้ตัวแปร 5 x 5 = 25 ตัว เหล่านี้เป็นตัวแปรเชิงปริมาณทำหน้าที่เป็นชั้นข้อมูลนำเข้า ประกาศเป็นตัวแปรข้อมูลนำเข้า (Input Layer) 25 ตัวแปร จำนวนชั้นซ่อน (Hidden Layer) 10 โหนด จำนวนชั้นข้อมูลออก (Output Layer) 2 โหนด คือ ค่าอำนาจจำแนก

(ann) เป็นตัวแปรเชิงปริมาณ การกำหนดจำนวนโหนดในชั้นซ่อนยังไม่มีเกณฑ์ที่แน่นอน โดยทั่วไปจำนวนโหนดในชั้นซ่อนจะได้จากการทดลอง สำหรับการกำหนดจำนวนโหนดในชั้นซ่อนจะใช้กฎของบาม เฮาสเลอร์ (Baum-Haussler rule) ซึ่งได้เสนอไว้ในปี ค.ศ.1998 ดังนี้ (สายชล สินสมบูลณ์ทอง, 2560, น. 464)

$$n_{hidden} \leq n_{hmax} = \frac{n_{dataset} \times n_{input}}{n_{input} \times n_{output}}$$

โดย n_{hidden} คือ จำนวนโหนดในชั้นซ่อน

n_{hmax} คือ จำนวนโหนดที่มากที่สุด

n_{input} คือ จำนวนโหนดในชั้นข้อมูลเข้า

n_{output} คือ จำนวนโหนดในชั้นข้อมูลออก

$n_{dataset}$ คือ จำนวนโหนดในการฝึกหัด

จะได้ว่า $n_{hidden} \leq n_{hmax} = \frac{n_{dataset} \times n_{input}}{n_{input} \times n_{output}}$

$$= \frac{10 \times 25}{25+2}$$

$$= \frac{250}{27}$$

$$= 9.259$$

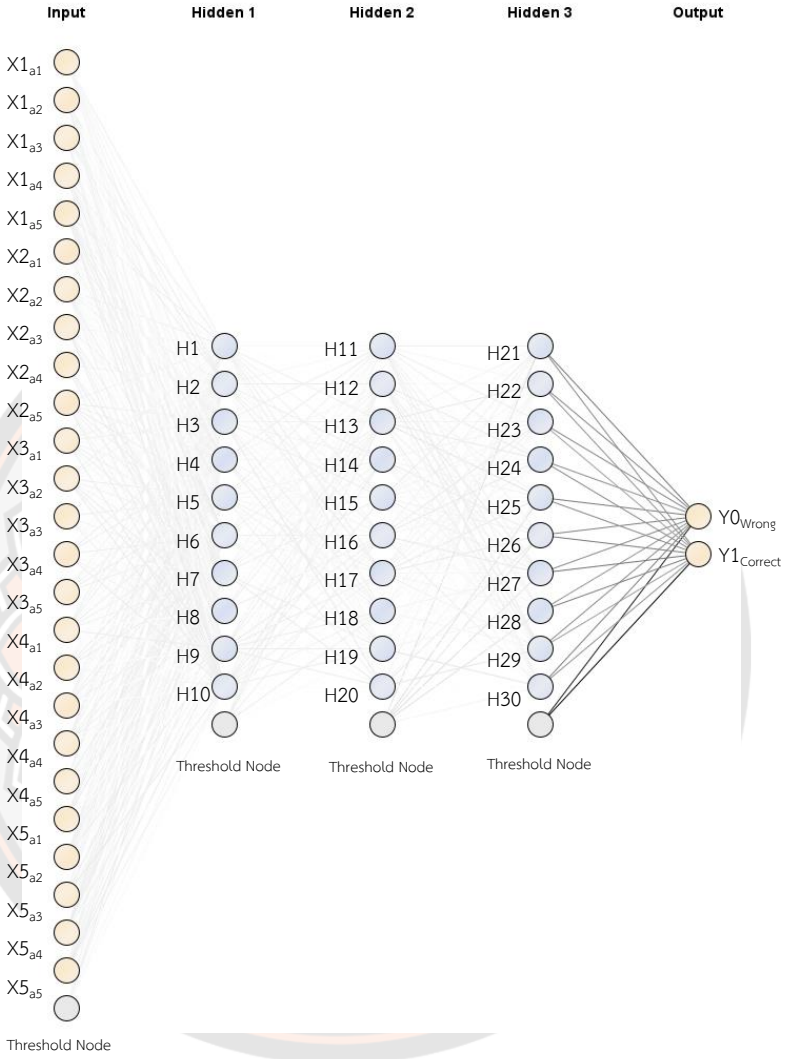
1.4 จากการคำนวณทดลองหาค่าเริ่มต้นของจำนวนโหนดชั้นข้อมูลซ่อนของโครงข่าย ประสาทเทียมเท่ากับ 9.259 เพื่อให้เกิดความแม่นยำในการประมาณค่าพารามิเตอร์มากขึ้น ผู้วิจัยจึงกำหนดจำนวนโหนดชั้นข้อมูลซ้อนเท่ากับ 10 โหนด ในการเริ่มต้นทดลองหาโมเดลที่เหมาะสม กำหนดจำนวนโหนดโดยใช้เทคนิคโครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron : MLP) ในชั้นข้อมูลป้อนเข้า (n_{input}) เท่ากับ 25 โหนด และจำนวนโหนดในชั้นซ่อน $(n_{hidden1})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden2})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden3})$ จำนวน 10 โหนด และจำนวนโหนดชั้นข้อมูลส่งออก (n_{output}) จำนวน 2 โหนด (รูปแบบข้อมูล 25-10-10-10-2) และค่าพารามิเตอร์ในกระบวนการของโครงข่ายประสาทเทียม และการหาค่าเหมาะสมที่สุด ประกอบด้วย อัตราการเรียนรู้เท่ากับ 0.95 โมเมนตัม (Momentum) เท่ากับ 0.05 (Mohammadi & Mirabedini, 2014, p. 118) ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ โดยมีส่วนประกอบที่สำคัญ 3 ส่วน ได้แก่

ส่วนที่ 1 จำนวนโหนดของชั้นข้อมูลป้อนเข้ามี จำนวน 25 โหนด (Node) ซึ่งมีค่าเท่ากับค่าอำนาจจำแนกของข้อสอบแต่ละมิติ จำนวน 25 ตัวแปร ได้แก่

$X1_{a1}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 1
$X1_{a2}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 2
$X1_{a3}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 3
$X1_{a4}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 4
$X1_{a5}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 5
$X2_{a1}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 1
$X2_{a2}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 2
$X2_{a3}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 3
$X2_{a4}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 4
$X2_{a5}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 5
$X3_{a1}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 1
$X3_{a2}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 2
$X3_{a3}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 3
$X3_{a4}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 4
$X3_{a5}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 5
$X4_{a1}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 1
$X4_{a2}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 2
$X4_{a3}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 3
$X4_{a4}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 4
$X4_{a5}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 5
$X5_{a1}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 1
$X5_{a2}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 2
$X5_{a3}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 3
$X5_{a4}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 4
$X5_{a5}$ แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 5

ส่วนที่ 2 จำนวนโหนดของชั้นซ่อนมี 30 โหนด (Node)

ส่วนที่ 3 จำนวนโหนดของชั้นข้อมูลส่งออกมี 2 โหนด (Node) ซึ่งแสดงผลการคัดเลือกข้อสอบข้อถัดไปมีผลลัพธ์จำนวนสองค่า คือ 1 แทน ตอบถูก (Correct answer) และ 0 แทน ตอบผิด (Wrong answer)



ภาพที่ 36 โมเดลโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไป

การวิเคราะห์ประสิทธิภาพของโมเดลโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไป มีขั้นตอนดังนี้

1. แบ่งข้อมูลเป็น 2 ส่วน ได้แก่ ข้อมูลการฝึกเรียนรู้และข้อมูลทดสอบโดยใช้อัตราส่วน 80% และ 20% ตามลำดับ

2. กำหนดจำนวนครั้งในการแบ่งชุดข้อมูลเป็น 100 ครั้ง

3. นำชุดข้อมูลที่ได้มาวิเคราะห์สร้างโมเดลโครงข่ายประสาทเทียมในการวิจัยครั้งนี้ได้ใช้เทคนิคการการจำแนกประเภทข้อมูล (Classification) โดยใช้อัลกอริธึมโครงข่ายประสาทเทียม

4. สร้างตัวแบบการจำแนกประเภทข้อมูล (Classification) ด้วยชุดข้อมูลฝึกการเรียนรู้แบบมีผู้สอน (Supervised Learning)

5. ประเมินผลประสิทธิภาพของโมเดลโครงข่ายประสาทเทียมโดยวัดประสิทธิภาพของโมเดล มีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่างๆ อยู่ 4 ค่า คือ

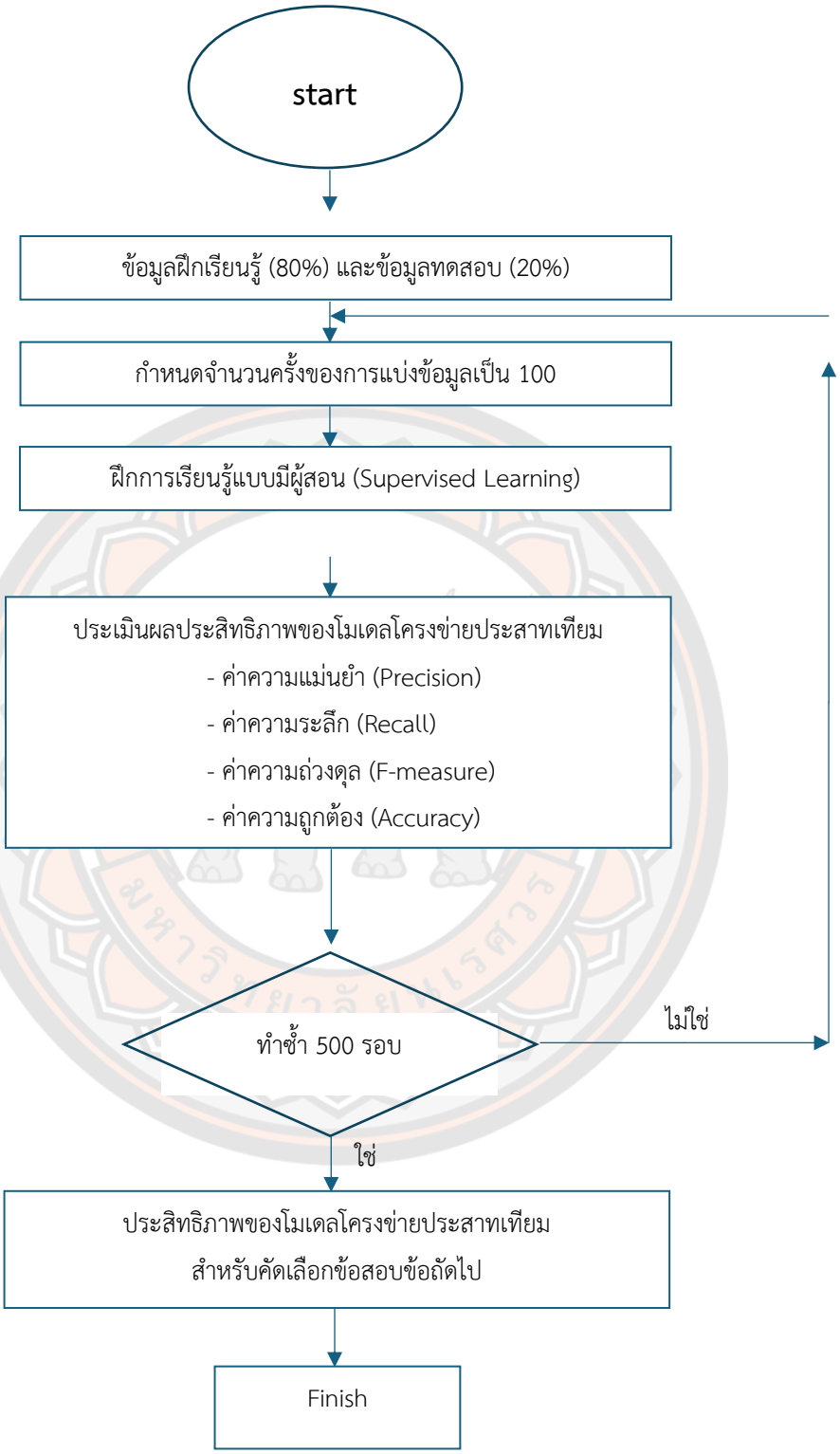
5.1 ค่าความแม่นยำ (Precision) คือค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์

5.2 ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล

5.3 ค่าความถ่วงดุล (F-measure) คือค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก

5.4 ค่าความถูกต้อง (Accuracy) คือจำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส

6. การแบ่งข้อมูลเพื่อใช้ในการวัดประสิทธิภาพของโมเดลโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไปด้วยวิธี Cross-validation Test เป็นวิธีที่นิยมใช้ในการทดสอบประสิทธิภาพของโมเดลเนื่องจากผลที่ได้มีความน่าเชื่อถือ การวัดประสิทธิภาพด้วยวิธี Cross-validation แบ่งข้อมูลออกเป็น 10-fold cross-validation คือ การแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่ากัน หลังจากนั้นข้อมูลหนึ่งส่วนจะใช้เป็นตัวทดสอบประสิทธิภาพของโมเดล ทำวนไปเช่นนี้จนครบจำนวนที่แบ่งไว้



ภาพที่ 37 การทดสอบประสิทธิภาพโครงข่ายประสาทเทียม สำหรับการคัดเลือกข้อสอบข้อถัดไป

2. พัฒนาวิธีการคัดเลือกข้อสอบโดยใช้โครงข่ายประสาทเทียม

2.1 วิธีการคัดเลือกข้อสอบข้อถัดไปที่นิยมใช้กันเป็นมาตรฐานในการเลือกข้อสอบ คือจะเลือกข้อสอบที่มีค่าความยากของข้อสอบ (d) ใกล้เคียงกับค่าประมาณความสามารถ (θ) ณ ขณะนั้น (ศักดิ์ชัย จันทะแสง 2562; บูรพา วิถีปัญญา, 2563) ดังนั้นโมเดลโครงข่ายประสาทเทียมที่ผู้วิจัยได้พัฒนาขึ้นโดยนำข้อมูลเข้าค่าพารามิเตอร์ของข้อสอบแบบพหุมิติเข้าฐานข้อมูล MySQL ในระบบคลังข้อสอบเป็นข้อสอบที่ผู้วิจัยได้คัดเลือกข้อสอบเข้าคลังข้อสอบ จำนวน 242 ข้อ โดยพิจารณาจากข้อสอบที่มีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) มีค่าตั้งแต่ 0.00 ขึ้นไปค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) มีค่าในช่วง -4.00 ถึง 4.00 และค่าโอกาสการเดาข้อสอบมีค่าคงที่เท่ากับ 0.20 มีดังนี้

2.2 ผู้วิจัยได้พัฒนาเว็บแอพพลิเคชั่นภาษา PHP ข้อมูลในการคัดเลือกข้อสอบไว้ในฐานข้อมูล ฐานข้อมูล MySQL และ PhpMyAdmin โดยให้เว็บแอพพลิเคชั่นสุ่มคัดเลือกข้อสอบจากค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) ใกล้กับค่าประมาณความสามารถ (θ) ขณะนั้นมากที่สุด แล้วใช้เว็บแอพพลิเคชั่นสุ่มเลือกข้อสอบอีกครั้งจากค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) เก็บค่าพารามิเตอร์ของข้อสอบ และสร้างเว็บแอพพลิเคชั่นและเชื่อมต่อฐานข้อมูลด้วย PHP + MySQL เพื่อการคัดเลือกข้อสอบข้อถัดไป

2.3 การเลือกข้อสอบข้อแรกกำหนดให้ Initity Ability = 0 โดยสุ่มข้อสอบที่มีความสามารถใกล้เคียงกับความสามารถของผู้สอบ การสุ่มข้อสอบเพื่อเลือกข้อสอบข้อแรกนั้น จะสุ่มค่าความยากของข้อสอบที่มีค่าน้อยที่สุดในแต่ละมิติละ 1 ข้อ ดังนั้นจะมีข้อสอบเพื่อคัดเลือกเป็นข้อสอบข้อแรก รวม 5 ข้อ แล้วเลือกข้อสอบที่มีระดับค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) มีค่าในช่วง –4.00 ถึง 4.00 ระดับปานกลาง (ค่า MDIFF มีค่าอยู่ระหว่าง -0.50 ถึง 0.50) (Baker, 2001; Hasmy, 2014) มาพิจารณาในการคัดเลือกข้อสอบข้อแรกในแต่ละชั้น

2.4 เมื่อได้ข้อสอบข้อแรก หาผลการตอบข้อสอบ ซึ่งข้อสอบจะให้ผลการตอบข้อสอบเป็น 0 และ 1 ถ้าผู้สอบตอบถูกผลการตอบจะเป็น 1 ถ้าผู้สอบตอบผิดผลการตอบจะเป็น 0 แล้วคำนวณหาค่า Pi (θ) และประมาณค่า (θ) ด้วยวิธีการประมาณค่าความสามารถผู้สอบด้วยวิธีเบส์ปรับใหม่ (Bayesian Updating)

2.5 การคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) จะทำการคัดเลือกข้อสอบที่มีค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) ที่ใกล้เคียงกับค่าประมาณความสามารถขณะนั้นที่สุด โดยกำหนดเป็นข้อมูลรับเข้าได้ในแต่ละโหนด โครงข่ายประสาทเทียมประมวลผลและทำนายค่าอำนาจจำแนก (ANN) นำค่าอำนาจจำแนกที่ได้ไปเปรียบเทียบกับค่าอำนาจ

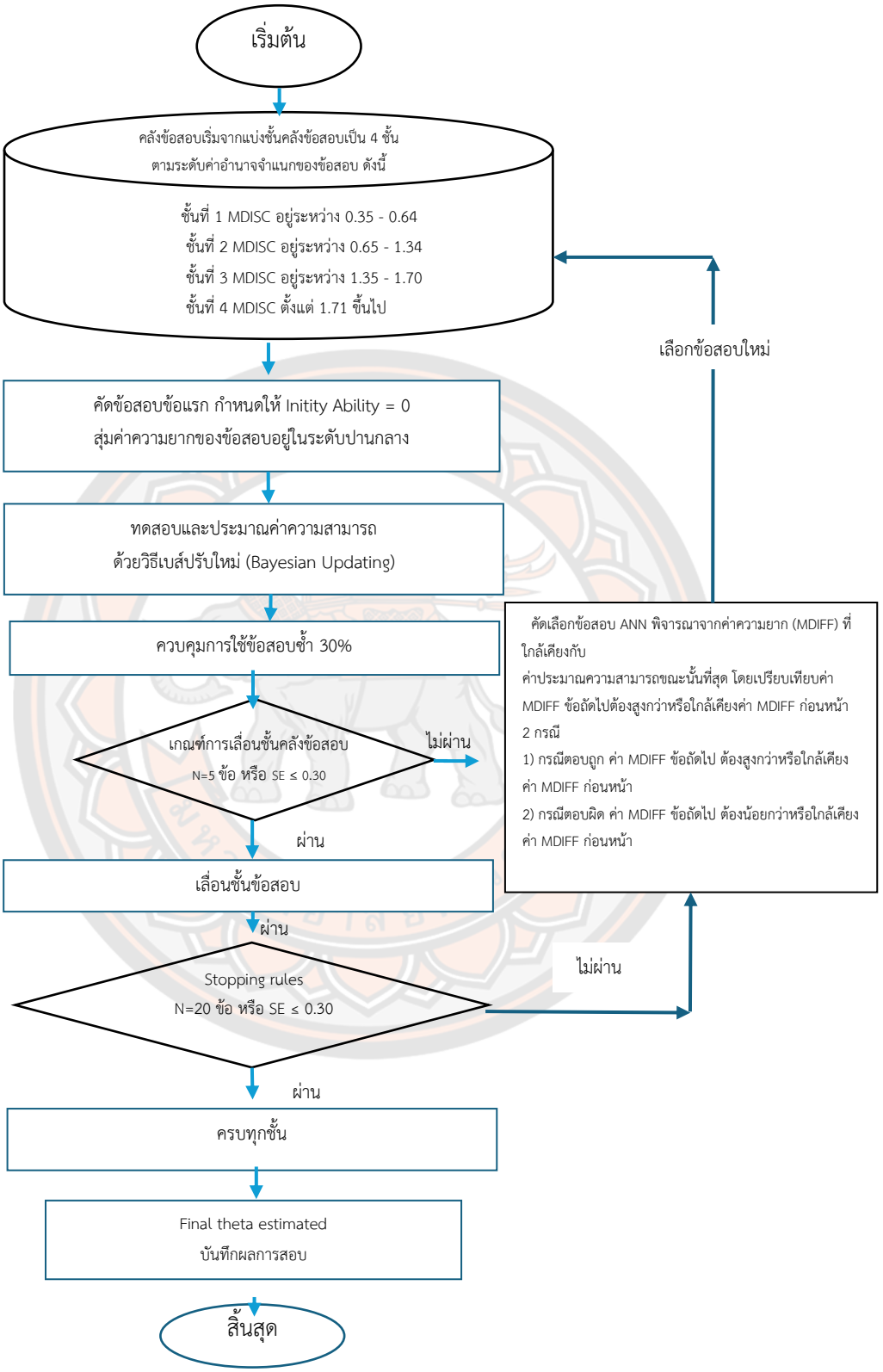
จำแนกของข้อสอบ ในข้อนั้นๆ ข้อสอบข้อไหนที่มีค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) ใกล้เคียงกับค่า Sigmoid ของโหนด Output (Output Node) โครงข่ายประสาทเทียม (ANN) มากที่สุดข้อสอบข้อนั้นจะถูกเลือกเป็นข้อสอบข้อถัดไปอยู่ 2 กรณี 1) กรณีตอบถูก ค่า MDIFF ข้อถัดไปต้องสูงกว่าหรือใกล้เคียงค่า MDIFF ข้อก่อนหน้า 2) กรณีตอบผิด ค่า MDIFF ข้อถัดไปต้องน้อยกว่าหรือใกล้เคียงค่า MDIFF ข้อก่อนหน้า

2.6 ตรวจสอบและปรับปรุงแก้ไขวิธีการการคัดเลือกข้อสอบเพื่อเป็นการตรวจสอบความเชื่อมั่นของโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียมที่ผู้วิจัยได้พัฒนาขึ้น

2.7 ทดสอบและประมาณค่าความสามารถ การตอบข้อสอบโดยการฝึกสอน (Train) แล้วก็จะนำไปใช้ทดสอบ (Test) และนำไปคำนวณด้วยฟังก์ชันการแปลงถ่ายทอดข้อมูล (Threshold Function) เพื่อให้ได้คำตอบที่เป็นจริงหรือเท็จ หรือ “0” หรือ “1” ถ้าสุ่มได้ 1 หมายถึง ตอบถูก สุ่มได้ 0 หมายถึง ตอบผิด เมื่อผู้ตอบตอบแล้วนำผลมาประมาณค่าความสามารถซึ่งเลือกใช้วิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีเบส์ที่ปรับใหม่ (Bayesian Updating)

2.8 เกณฑ์การยุติการทดสอบในแต่ละชั้นและเลื่อนชั้นการทดสอบ ผู้สอบจะต้องทดสอบในชั้นนั้นๆ ให้ครบ 5 ข้อ ตามเนื้อหาสมรรถนะความรู้ทางวิชาชีพครูที่ออกข้อสอบ หรือค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าหรือเท่ากับ 0.30 (n=5, หรือ SE ≤ 0.30) และรวบรวมผลการทดสอบในแต่ละชั้นการเลื่อนชั้นการทดสอบไปทดสอบในคลังข้อสอบในชั้นที่ 2, 3 และ 4 จะต้องคำนึงถึงความยากแบบพหุมิติของข้อสอบ (MDIFF) ของข้อสอบข้อก่อนหน้าในชั้นที่แล้วเป็นเกณฑ์

2.9 เกณฑ์การยุติการทดสอบ (Termination criterion/Stopping rules) โดยเกณฑ์ที่กำหนดไว้คือค่ามาตรฐานความผิดพลาด (Standard Error: SE) มีค่าน้อยกว่าหรือ (n=20, หรือ SE ≤ 0.30) จะทำหยุดทำการทดสอบแต่ถ้ามากกว่า 0.30 ทำการทดสอบต่อไป



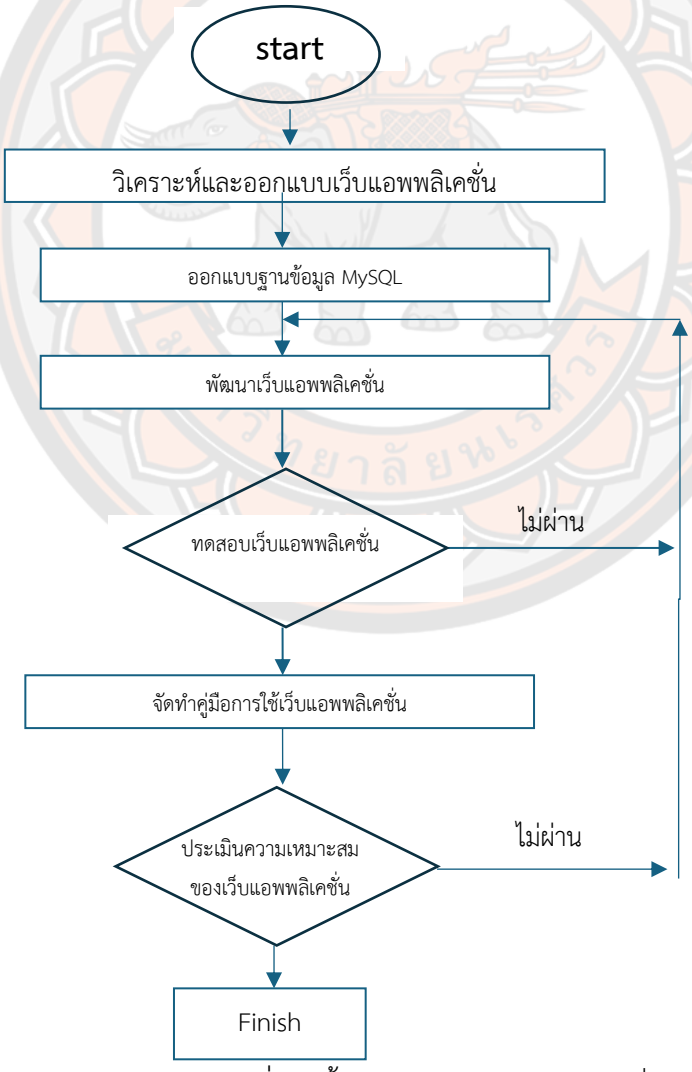
ภาพที่ 38 ขั้นตอนการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียม (ann)

ระยะที่ 3 การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ

ขั้นตอนการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป เป็นการบูรณาการแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) หลอมรวมกับโครงข่ายประสาทเทียม (Neural network) มาใช้ทดสอบเพื่อประยุกต์ใช้กับการคัดเลือกข้อสอบข้อถัดไปมีขั้นตอนการพัฒนา ดังนี้

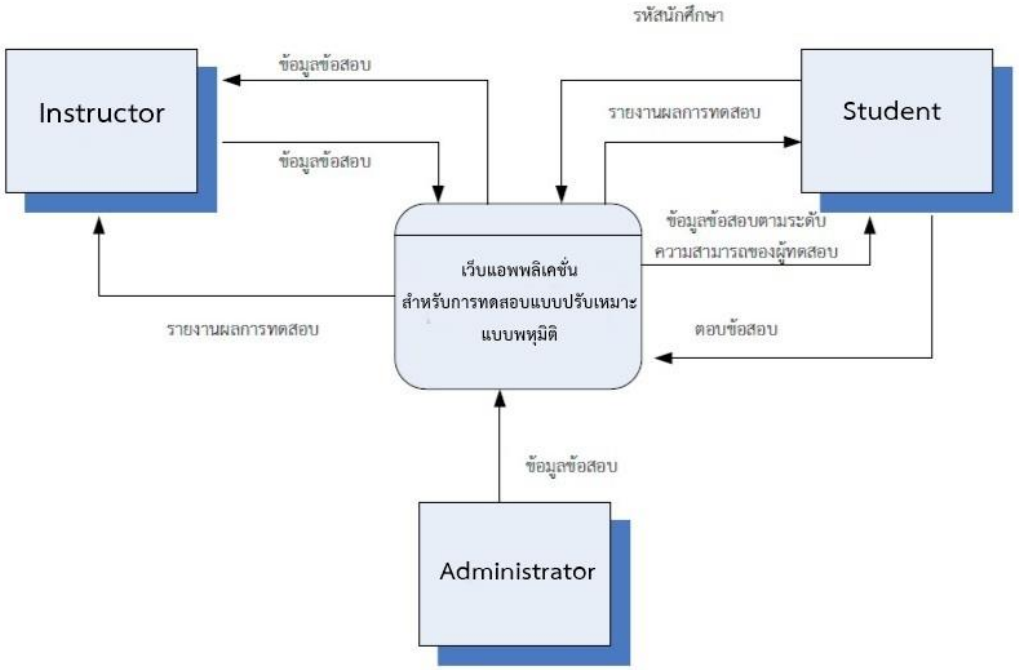
แนวทางของการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ซึ่งมีขั้นตอนการพัฒนาเว็บแอพพลิเคชั่นตามภาพที่ 39



ภาพที่ 39 ขั้นตอนพัฒนาเว็บแอพพลิเคชั่น

1. การวิเคราะห์และออกแบบระบบเว็บแอพพลิเคชั่น

1.1 แผนผังบริบท (Context Diagram) การวิเคราะห์และออกแบบเริ่มต้นจากการเขียนแบบจำลองกระบวนการ (Process Model) ในรูปแบบของแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD)



ภาพที่ 40 แผนผังบริบท (Context Diagram) ของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ

1.2 การจัดทำผังการไหลของข้อมูล (Data Flow Diagram: DFD) เป็นการวิเคราะห์ให้เห็นภาพรวมในการทำงานของเว็บแอพพลิเคชั่น โดยจะแสดงรายละเอียดการรับหรือส่งข้อมูลภายในเว็บแอพพลิเคชั่น

2. ออกแบบฐานข้อมูล

การออกแบบฐานข้อมูลของเว็บแอพพลิเคชั่นการทดสอบในการวิจัยนี้มี 7 แฟ้มข้อมูล ได้แก่ 1) แฟ้มข้อมูลข้อสอบ 2) แฟ้มข้อมูลรายการคำตอบ 3) แฟ้มข้อมูลผู้สอบ 4) แฟ้มข้อมูลการทดสอบ 5) แฟ้มข้อมูลสาขาวิชา 6) แฟ้มข้อมูลคณะวิชา และ 7) แฟ้มข้อมูลข้อสอบการทดสอบด้วยคอมพิวเตอร์

3. พัฒนาเว็บแอพพลิเคชั่น ด้วย PHP, MySQL

การพัฒนาเว็บแอพพลิเคชั่นการทดสอบในการวิจัยนี้ เป็นการพัฒนาในรูปแบบของ Web Application โดยใช้เว็บแอพพลิเคชั่น Notepad++ ที่รองรับการเขียนเว็บแอพพลิเคชั่นด้วย

ภาษา PHP และใช้ MySQL เป็นฐานข้อมูล และใช้ภาษา SQL เขียนคำสั่งในการเชื่อมโยงข้อมูลในฐานข้อมูลการออกแบบโครงสร้างหน้า จอของเว็บแอพพลิเคชั่นการทดสอบมี ดังนี้

3.1 โครงสร้างหน้าจอทั่วไปของเว็บแอพพลิเคชั่นการทดสอบ เป็นการแจ้งรายละเอียดต่าง ๆ ที่เกี่ยวข้องกับเว็บแอพพลิเคชั่นการทดสอบ แสดงได้ตามภาพที่ 41

Header
Menu
Detail

ภาพที่ 41 โครงสร้างหน้าจอทั่วไปของเว็บแอพพลิเคชั่นการทดสอบ

3.2 โครงสร้างหน้าจอการทดสอบของผู้สอบ เป็นการแสดงหน้าจอการดำเนินการทดสอบสำหรับผู้สอบ มีรายละเอียดดังนี้

3.2.1 โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ เป็นหน้าจอที่ให้ผู้สอบกรอกข้อมูลเบื้องต้นของตนเอง เช่น ชื่อ นามสกุล เลขที่ประจำตัวสอบ เป็นต้น แสดงได้ตามภาพที่ 41

Header
ข้อมูลของผู้สอบ

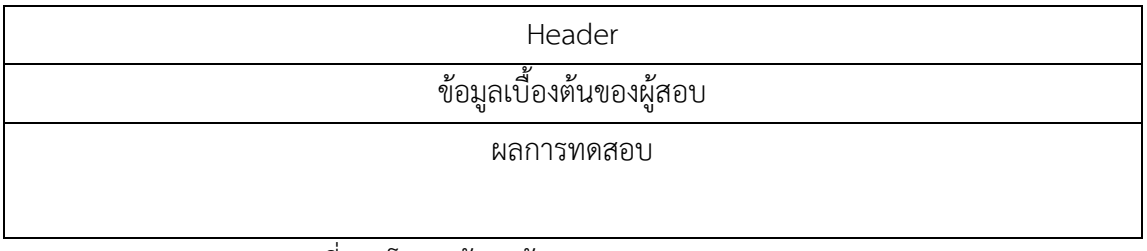
ภาพที่ 42 โครงสร้างหน้าจอการกรอกข้อมูลเบื้องต้นของผู้สอบ

3.2.2 โครงสร้างหน้าจอการทดสอบ เป็นหน้าจอที่แสดงข้อสอบให้แก่ผู้สอบซึ่งจะแสดงข้อสอบทีละข้อ แสดงได้ตามภาพที่ 42

Header
ข้อสอบ

ภาพที่ 43 โครงสร้างหน้าจอการทดสอบ

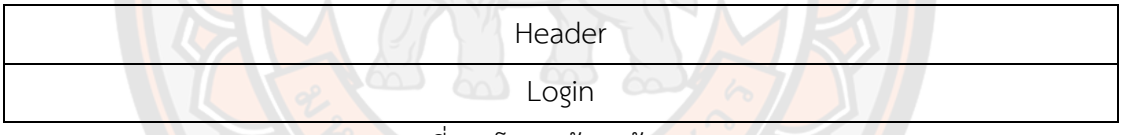
3.2.3 โครงสร้างหน้าจอการรายงานผลการทดสอบ เป็นหน้าจอที่แสดงรายงานผลการทดสอบให้แก่ผู้สอบ เช่น คะแนนการทดสอบ จำนวนข้อที่ใช้ในการทดสอบหรือเวลาที่ใช้ในการทดสอบ แสดงได้ตามภาพที่ 44



ภาพที่ 44 โครงสร้างหน้าจอการรายงานผลการทดสอบ

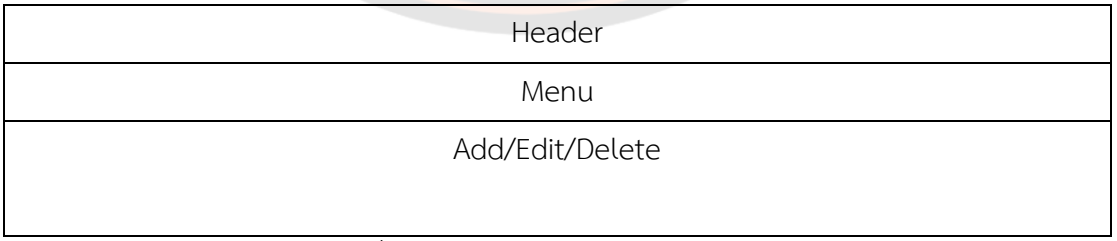
3.3 โครงสร้างหน้าจอการจัดการข้อสอบ เป็นหน้าจอแสดงการปรับปรุงข้อสอบในคลังข้อสอบ สำหรับอาจารย์ผู้สอนหรือบุคคลที่ได้รับมอบหมายให้สามารถปรับปรุงข้อสอบได้ซึ่งมีรายละเอียดดังนี้

3.3.1 โครงสร้างหน้าจอการ Login เป็นหน้าจอที่ให้อาจารย์ผู้สอน หรือบุคคลที่ได้รับมอบหมาย กรอกชื่อผู้ใช้ (Username) และรหัสผ่าน (Password) เพื่อเข้าสู่ส่วนการปรับปรุงข้อสอบ แสดงได้ตามภาพที่ 45



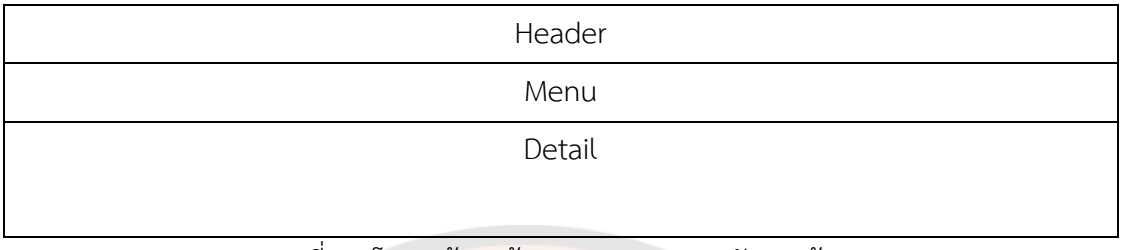
ภาพที่ 45 โครงสร้างหน้าจอ Login

3.3.2 โครงสร้างหน้าจอการจัดการข้อสอบ เป็นหน้าจอที่แสดงรายละเอียดในการเพิ่มข้อสอบ ลบข้อสอบ หรือแก้ไขข้อสอบ ในคลังข้อสอบ แสดงได้ตามภาพที่ 46



ภาพที่ 46 โครงสร้างหน้าจอการจัดการข้อสอบ

3.3.3 โครงสร้างหน้าจอการรายงานผลการจัดการข้อสอบ เป็นหน้าจอที่แสดงรายงานผลที่เกี่ยวข้องกับการจัดการข้อสอบครั้งล่าสุด เช่น จำนวนข้อสอบทั้งหมด ค่าต่ำสุด-สูงสุดของค่าพารามิเตอร์ของข้อสอบทั้งหมด แสดงได้ตามภาพที่ 47



ภาพที่ 47 โครงสร้างหน้าจอรายงานผลการจัดการข้อสอบ

4. การทดสอบและแก้ไขเว็บแอพพลิเคชั่น

เมื่อพัฒนาเว็บแอพพลิเคชั่นการทดสอบเป็นที่เรียบร้อยแล้ว ผู้วิจัยร่วมกับเว็บแอพพลิเคชั่นเมอร์ได้ตรวจสอบเว็บแอพพลิเคชั่นการทดสอบ เพื่อหาจุดบกพร่อง ได้แก่พิมพ์เครื่องหมายการดำเนินการทางคณิตศาสตร์ผิด (+ − × ÷) พิมพ์ตัวแปรผิด รวมทั้งการใช้คำสั่งในการเขียนเว็บแอพพลิเคชั่นผิดพลาด ซึ่งได้ตรวจสอบทีละคำสั่งแล้วแก้ไขเว็บแอพพลิเคชั่นให้ถูกต้อง นอกจากนี้ เพื่อให้มั่นใจว่าเว็บแอพพลิเคชั่นการทดสอบที่พัฒนาขึ้นมีความถูกต้อง สามารถใช้งานได้จริง จึงรันเว็บแอพพลิเคชั่นการทดสอบแล้วนำผลการทดสอบเทียบกับผลการทดสอบที่ได้จากการศึกษาในสถานการณ์จำลอง ปรากฏว่า เว็บแอพพลิเคชั่นการทดสอบที่พัฒนาขึ้นให้ผลการทดสอบเท่ากับผลการทดสอบจากการศึกษาในสถานการณ์จำลอง จึงมั่นใจได้ว่าเว็บแอพพลิเคชั่นการทดสอบที่พัฒนาขึ้นไม่มีความผิดพลาด จากนั้นจึงนำเว็บแอพพลิเคชั่นเสนอต่ออาจารย์ที่ปรึกษาเพื่อขอความคิดเห็น แล้วนำข้อคิดเห็นที่ได้มาปรับปรุงเว็บแอพพลิเคชั่นให้มีความสมบูรณ์ยิ่งขึ้นต่อไป

5. จัดทำคู่มือการใช้เว็บแอพพลิเคชั่น

หลังจากทดสอบและแก้ไขเว็บแอพพลิเคชั่นเป็นที่น่าพอใจแล้ว ได้จัดทำคู่มือในการใช้เว็บแอพพลิเคชั่นการทดสอบนี้ เพื่ออธิบายวิธีการใช้เว็บแอพพลิเคชั่น โดยแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 การจัดการข้อสอบ ประกอบด้วย การตั้งรหัสกลุ่มผู้สอบ การจัดการข้อสอบ การรายงานผลสอบ (ต่ออาจารย์)และการออกจากระบบ และส่วนที่ 2 การจัดการสอบ ประกอบด้วย เริ่มต้นการทดสอบ กรอกข้อมูลผู้สอบและเลือกแบบทดสอบ การดำเนินการสอบ และการรายงานผลสอบ (ต่อผู้สอบ)

ระยะที่ 4 การหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ในขั้นตอนนี้จะดำเนินการประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยการนำเอาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติไปใช้ในสถานการณ์จริง มีขั้นตอนการดำเนินการ ดังนี้

1. ศึกษาวิธีการสร้างแบบประเมินจากเอกสารและงานวิจัยที่เกี่ยวข้องกับความเป็นประโยชน์ ความเป็นไปได้ ความถูกต้องครอบคลุมและความเหมาะสม

2. สร้างเครื่องมือสำหรับการประเมินที่ประกอบด้วย 4 ด้าน ดังนี้

1) ด้านความต้องการของผู้ใช้เว็บแอพพลิเคชั่น (Functional Requirement)

2) ด้านการทำงานของโปรแกรม (Functional)

3) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability)

4) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual)

3. การประเมินการใช้เว็บแอพพลิเคชั่นและคู่มือการการใช้เว็บแอพพลิเคชั่น

การประเมินเว็บแอพพลิเคชั่นและคู่มือการใช้เว็บแอพพลิเคชั่น ดำเนินการหลังจากทดสอบ ปรับปรุงแก้ไขและจัดทำคู่มือการใช้เว็บแอพพลิเคชั่นเป็นที่เรียบร้อยแล้ว โดยผู้วิจัยนำเว็บแอพพลิเคชั่นและคู่มือการใช้เว็บแอพพลิเคชั่น ไปประเมินเว็บแอพพลิเคชั่นและคู่มือการใช้เว็บแอพพลิเคชั่นซึ่งแบ่งได้ 2 ส่วน คือ การประเมินโดยผู้เชี่ยวชาญ และการประเมินโดยผู้สอบโดยมีรายละเอียดดังนี้

3.1 การประเมินโดยผู้เชี่ยวชาญ

การประเมินโดยผู้เชี่ยวชาญ จำนวน 5 คน โดยกำหนดเกณฑ์ว่าผู้เชี่ยวชาญต้องมีความรู้ทางด้านการวัดผล หรือด้านการพัฒนาเว็บแอพพลิเคชั่นและมีประสบการณ์ทางด้านการวัดผลหรือด้านการพัฒนาเว็บแอพพลิเคชั่น ไม่น้อยกว่า 10 ปี และมีวุฒิการศึกษาไม่ต่ำกว่าระดับปริญญาโท ผู้เชี่ยวชาญได้ประเมินเว็บแอพพลิเคชั่นและคู่มือการใช้เว็บแอพพลิเคชั่นใน 4 ด้าน ได้แก่ 1) ด้านความต้องการของผู้ใช้เว็บแอพพลิเคชั่น (Functional Requirement) 2) ด้านการทำงานของโปรแกรม (Functional) 3) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น(Usability) 4) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น(Web application Manual) และข้อเสนอแนะเพื่อการพัฒนาเว็บแอพพลิเคชั่นโดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้น ซึ่งมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ ซึ่งมีเกณฑ์การให้คะแนน ดังนี้

5 หมายถึง เหมาะสมมากที่สุด

4 หมายถึง เหมาะสมมาก

3 หมายถึง เหมาะสมปานกลาง

2 หมายถึง เหมาะสมน้อย

1 หมายถึง เหมาะสมน้อยที่สุด

เกณฑ์การแปลความหมายของคะแนนเฉลี่ย ดังนี้

คะแนนเฉลี่ยตั้งแต่ 4.50 - 5.00 หมายถึง มีความเหมาะสมมากที่สุด

คะแนนเฉลี่ยตั้งแต่ 3.50 - 4.49 หมายถึง มีความเหมาะสมมาก

คะแนนเฉลี่ยตั้งแต่ 2.50 - 3.49 หมายถึง มีความเหมาะสมปานกลาง

คะแนนเฉลี่ยตั้งแต่ 1.50 - 2.49 หมายถึง มีความเหมาะสมน้อย

คะแนนเฉลี่ยตั้งแต่ 1.00 - 1.49 หมายถึง มีความเหมาะสมน้อยที่สุด

3.2 การประเมินโดยผู้ใช้งาน

การประเมินโดยผู้สอบ เป็นการประเมินความคิดเห็นในด้านความสะดวกในการนำไปใช้และด้านลักษณะทั่วไปของเว็บแอพพลิเคชั่น โดยกลุ่มตัวอย่างที่ทดลองใช้ เป็นนักศึกษาคณะครุศาสตร์ มหาวิทยาลัยราชภัฏเลย จำนวน 200 คน คัดเลือกด้วยวิธีการเลือกตัวอย่างตามสะดวก โดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้นในลักษณะมาตรประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้

5 หมายถึง เหมาะสมมากที่สุด

4 หมายถึง เหมาะสมมาก

3 หมายถึง เหมาะสมปานกลาง

2 หมายถึง เหมาะสมน้อย

1 หมายถึง เหมาะสมน้อยที่สุด

เกณฑ์การแปลความหมายของคะแนนเฉลี่ย ดังนี้

คะแนนเฉลี่ยตั้งแต่ 4.50 - 5.00 หมายถึง มีความเหมาะสมมากที่สุด

คะแนนเฉลี่ยตั้งแต่ 3.50 - 4.49 หมายถึง มีความเหมาะสมมาก

คะแนนเฉลี่ยตั้งแต่ 2.50 - 3.49 หมายถึง มีความเหมาะสมปานกลาง

คะแนนเฉลี่ยตั้งแต่ 1.50 - 2.49 หมายถึง มีความเหมาะสมน้อย

คะแนนเฉลี่ยตั้งแต่ 1.00 - 1.49 หมายถึง มีความเหมาะสมน้อยที่สุด

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การวิจัยเรื่อง การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลตามหัวข้อ ดังนี้

1. สัญลักษณ์ที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล
2. ลำดับขั้นตอนในการนำเสนอผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู

ตอนที่ 2 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

ตอนที่ 3 ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ตอนที่ 4 ผลการหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

สัญลักษณ์ที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล

ผู้วิจัยได้กำหนดสัญลักษณ์ที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล เพื่อให้เกิดความเข้าใจตรงกันในการแปลความหมาย ดังนี้

a_1 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 1 (หลักสูตรและการสอน)

a_2 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 2 (จิตวิทยา)

a_3 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 3 (การวัดผลและวิจัยการเรียนรู้)

a_4 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 4 (การประกันคุณภาพการศึกษา)

a_5 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 5 (การเปลี่ยนแปลงบริบทโลก)

MDIFF แทน ค่าความยากแบบพหุมิติของข้อสอบ

MDISC แทน ค่าอำนาจจำแนกแบบพหุมิติ

c แทน ค่าการเดาของข้อสอบ

d แทน ค่าจุดตัดความยากแบบพหุมิติของข้อสอบ (Easiness Intercept)

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปและเพื่อประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปมีขั้นตอนการดำเนินการ 4 ระยะ ผู้วิจัยได้นำเสนอรายละเอียดผลการวิเคราะห์ข้อมูลการวิจัย ตามลำดับดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู

ตอนที่ 2 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

ตอนที่ 3 ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ตอนที่ 4 ผลการหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

ตอนที่ 1 ผลการพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู

ผลการพัฒนาวิธีการจัดคลังข้อสอบข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูโดยประยุกต์โมเดลการตอบสนองข้อสอบแบบพหุมิติ รายละเอียดดังนี้

1. ผลการพัฒนาคลังข้อสอบ

1.1. ผลการคัดเลือกข้อสอบเข้าคลังข้อสอบ

คลังข้อสอบในเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ที่ผู้วิจัยได้คัดเลือกข้อสอบ

เข้าคลังข้อสอบ โดยพิจารณาจากข้อสอบที่มีค่าอำนาจจำแนกแบบพหุมิติอยู่ในช่วง -4.0 ถึง +4.0 และค่าโอกาสการเดาข้อสอบมีค่าคงที่เท่ากับ 0.20

ผลการวิจัยพบว่าจากข้อสอบจำนวน 270 ข้อ เมื่อนำมาวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม NOHARM 4.0 พบว่าได้ข้อสอบตามเกณฑ์การคัดเลือก จำนวน 242 ข้อ แสดงดังตารางที่ 16

ตารางที่ 16 จำนวนข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูที่ได้รับการคัดเลือกเข้าคลังข้อสอบ จำแนกตามมิติของข้อสอบ

มิติความรู้ทางวิชาชีพครู	จำนวนข้อสอบ (ข้อ)		คัดเลือกไว้
	วิเคราะห์หาค่าพารามิเตอร์	ค่าพารามิเตอร์เป็นไปตามเกณฑ์	
การเปลี่ยนแปลงบริบทโลก	60	50	50
จิตวิทยาการศึกษา	50	40	40
การประกันคุณภาพการศึกษา	50	45	45
วัดผลและวิจัยการเรียนรู้	60	60	60
หลักสูตรและการสอน	50	47	47
รวม	**270**	**242**	**242**

จากตารางที่ 16 เมื่อพิจารณาผลการคัดเลือกข้อสอบเข้าคลังข้อสอบความรู้ทางวิชาชีพครู พบว่าได้ข้อสอบที่เป็นไปตามเกณฑ์การคัดเลือกข้อสอบจำนวน 242 ข้อ ประกอบด้วยข้อสอบที่วัดมิติการเปลี่ยนแปลงบริบทโลก จำนวน 50 ข้อ มิติจิตวิทยาการศึกษา จำนวน 40 ข้อ มิติการประกันคุณภาพการศึกษา จำนวน 45 ข้อ มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ และ มิติหลักสูตรและการสอน จำนวน 47 ข้อ ตามลำดับ

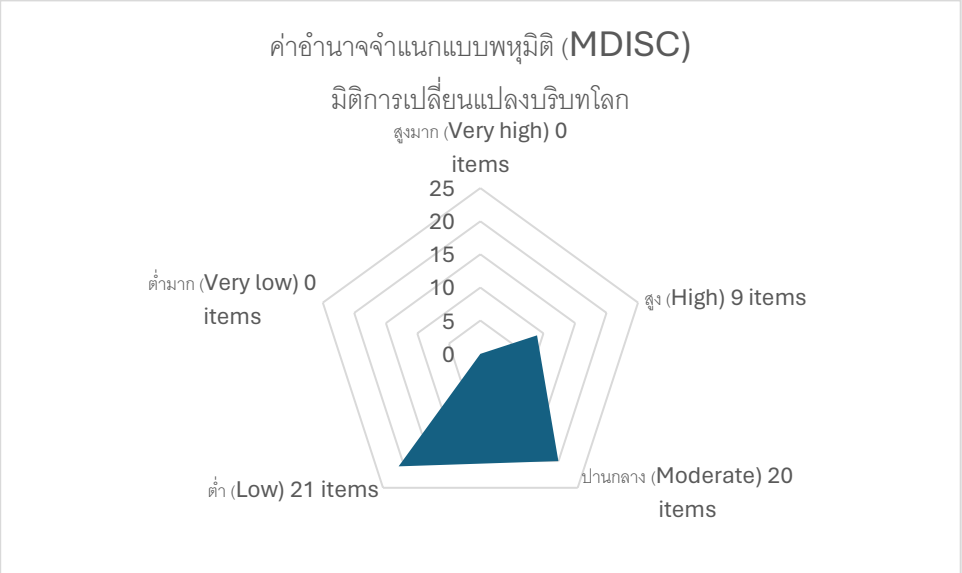
1.2 ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครูที่ผู้วิจัยรวบรวมได้จำนวน 270 ข้อ เมื่อนำมาวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม NOHARM 4.0 พบว่ามีข้อสอบที่เป็นไปตามเกณฑ์การคัดเลือกข้อสอบเข้าคลังข้อสอบ จำนวน 242 ข้อ จำแนกออกเป็น 5 มิติ ดังนี้ คือ

ตารางที่ 17 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 1 การเปลี่ยนแปลงบริบทโลก (a_1)

ค่าพารามิเตอร์	มิติข้อสอบการเปลี่ยนแปลงบริบทโลก (a_1)					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0.725	0	0	0	0	-0.399	0.725	-0.388
SD	0.167	0	0	0	0	0.697	0.167	0.911
MIN	0.448	0	0	0	0	-2.00	0.448	-2.00
MAX	1.00	0	0	0	0	1.030	1.000	1.930

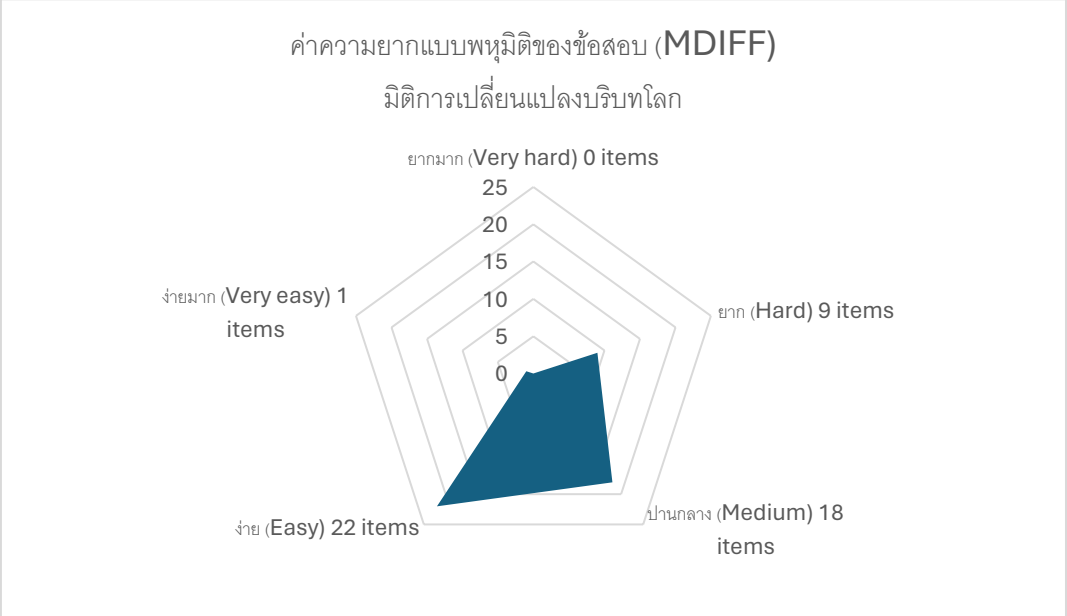
จากตารางที่ 17 พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก (a_1) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.448 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.725 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.448 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.725 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง 2.000 ถึง 1.930 ค่าเฉลี่ยมีค่าเท่ากับ -0.388 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.911 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ยมีค่าเท่ากับ -0.399 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.697

เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก ดังภาพที่ 48 และ 49



ภาพที่ 48 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก

จากภาพที่ 48 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก พบว่า มีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำ (Low) จำนวน 21 ข้อ รองลงมา เป็นข้อสอบที่มีค่าอำนาจปานกลาง (Moderate) จำนวน 20 ข้อ และมีค่าอำนาจจำแนกสูง (High) จำนวน 9 ข้อตามลำดับ



ภาพที่ 49 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก

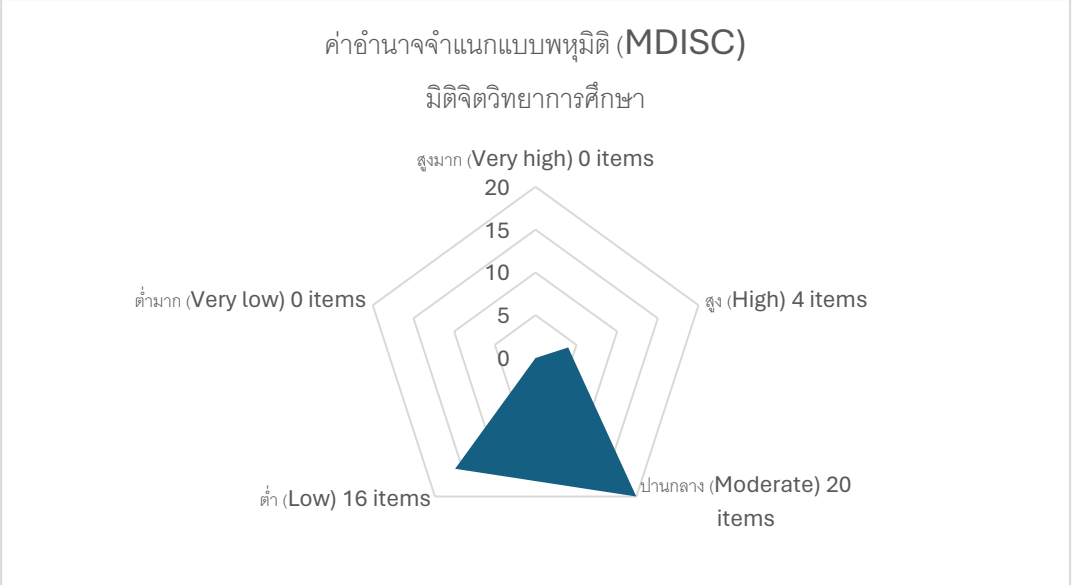
จากภาพที่ 49 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก พบว่ามีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่า (Easy) จำนวน 22 ข้อ รองลงมาเป็นข้อสอบที่มีค่าปานกลาง (Medium) จำนวน 18 ข้อ มีค่ายาก (Hard) จำนวน 9 ข้อ และมีค่าง่ายมาก (Very easy) จำนวน 1 ข้อ ตามลำดับ

ตารางที่ 18 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 2 จิตวิทยาการศึกษา (a_2)

ค่าพารามิเตอร์	มิติของข้อสอบจิตวิทยาการศึกษา (a_2)					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0	0.717	0	0	0	-0.438	0.717	-0.458
SD	0	0.172	0	0	0	0.681	0.172	0.895
MIN	0	0.430	0	0	0	-2.000	0.430	-2.000
MAX	0	1.00	0	0	0	1.030	1.000	2.000

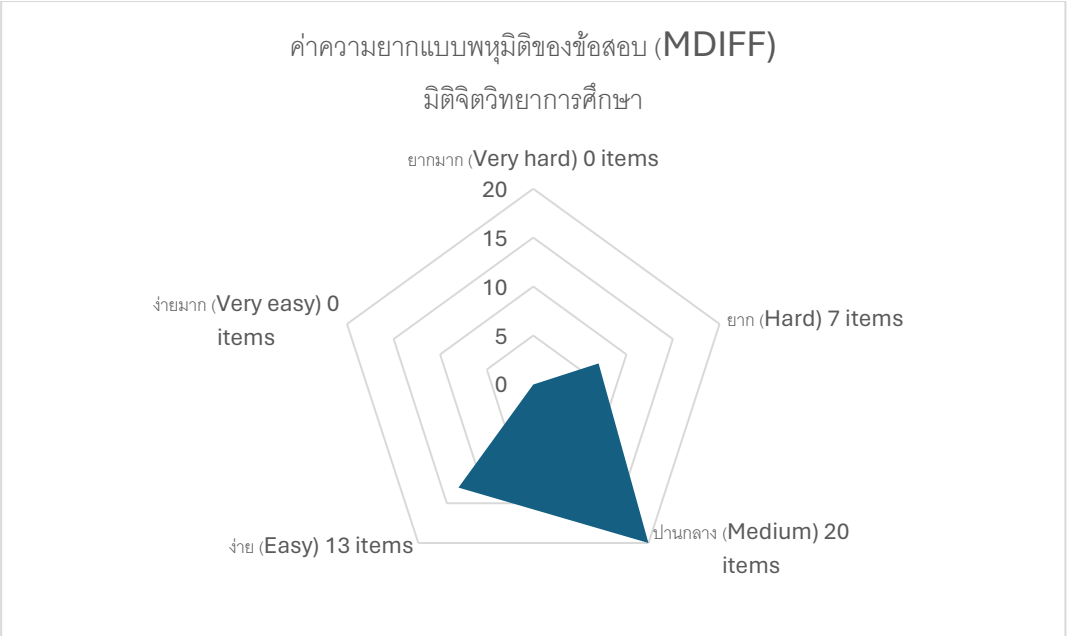
จากตารางที่ 18 พบว่า ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา (a_2) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.717 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.172 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.717 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.172 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.000 ถึง 2.000 ค่าเฉลี่ยมีค่าเท่ากับ -0.458 และส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.895 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ยมีค่าเท่ากับ -0.438 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.681

เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา ดังภาพที่ 50 และ 51



ภาพที่ 50 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา

จากภาพที่ 50 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา พบว่า มีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกปานกลาง (Moderate) จำนวน 20 ข้อ รองลงมาเป็นข้อสอบที่มีค่าอำนาจต่ำ (Low) จำนวน 16 ข้อ และมีค่าอำนาจจำแนกสูง (High) จำนวน 4 ข้อ ตามลำดับ



ภาพที่ 51 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา

จากภาพที่ 51 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา พบว่า มีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าปานกลาง (Medium) จำนวน 20 ข้อ รองลงมาเป็นข้อสอบที่มีค่าง่าย (Easy) จำนวน 13 ข้อ และมีค่ายาก (Hard) จำนวน 7 ข้อ ตามลำดับ

ตารางที่ 19 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 3 การประกันคุณภาพการศึกษา (a_3)

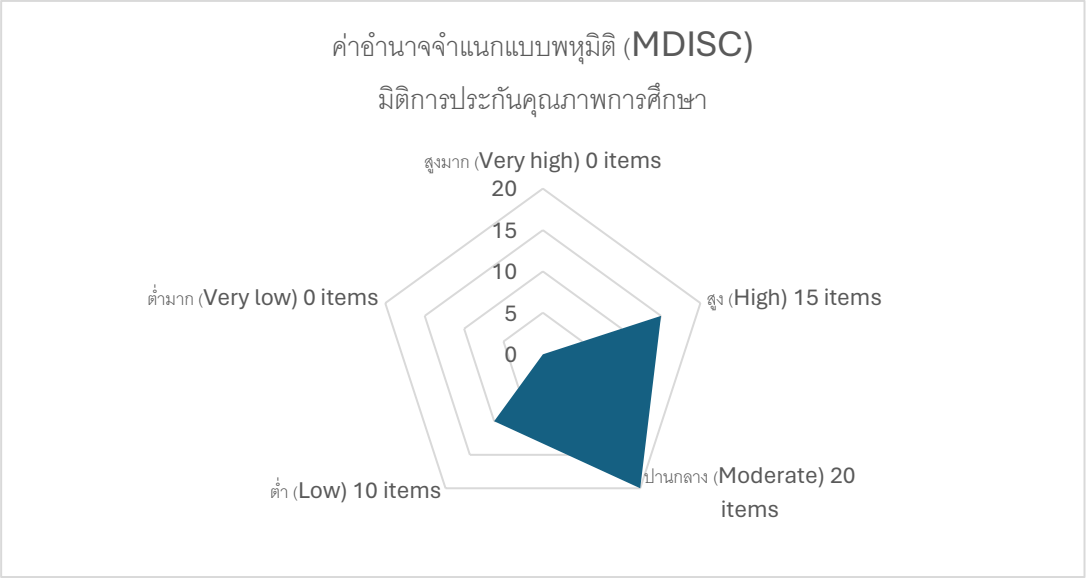
ค่าพารามิเตอร์	มิติของข้อสอบการประกันคุณภาพการศึกษา (a_3)					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0	0	0.752	0	0	-0.515	0.752	-0.527
SD	0	0	0.179	0	0	0.719	0.179	0.904
MIN	0	0	0.440	0	0	-2.430	0.440	-2.430
MAX	0	0	1.000	0	0	0.994	1.000	2.200

จากตารางที่ 19 พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา (a_3) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.440 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.752 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.179 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.440 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.752 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.179 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.430 ถึง 2.200 ค่าเฉลี่ยมีค่าเท่ากับ -0.527 และส่วนเบี่ยงเบนมาตรฐาน

มีค่าเท่ากับ 0.904 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.430 ถึง 0.994 ค่าเฉลี่ยมีค่าเท่ากับ

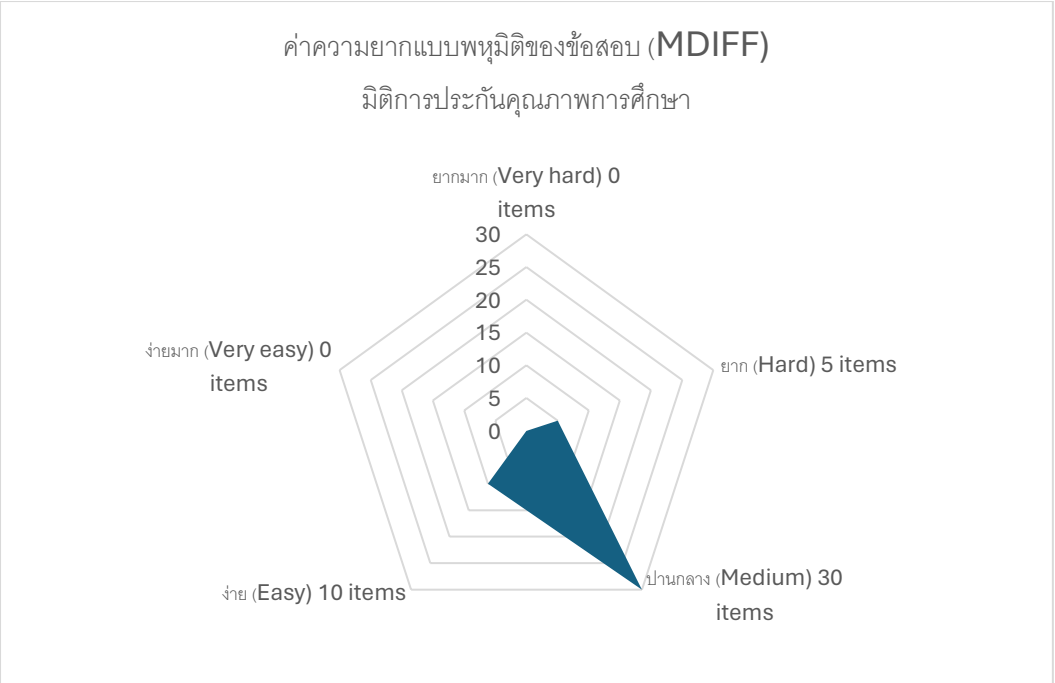
-0.515 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.719

เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา ดังภาพที่ 52 และ 53



ภาพที่ 52 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา

จากภาพที่ 52 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา พบว่ามีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกปานกลาง (Moderate) จำนวน 20 ข้อ รองลงมาเป็นข้อสอบที่มีค่าอำนาจจำแนกสูง (High) จำนวน 15 และค่าอำนาจจำแนกต่ำ (Low) จำนวน 10 ข้อ ข้อตามลำดับ



ภาพที่ 53 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา

จากภาพที่ 53 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา พบว่ามีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าปานกลาง (Medium) จำนวน 30 ข้อ รองลงมา เป็นข้อสอบที่มีค่าง่าย (Easy) จำนวน 10 ข้อ และมีค่ายาก (Hard) จำนวน 5 ข้อ ตามลำดับ

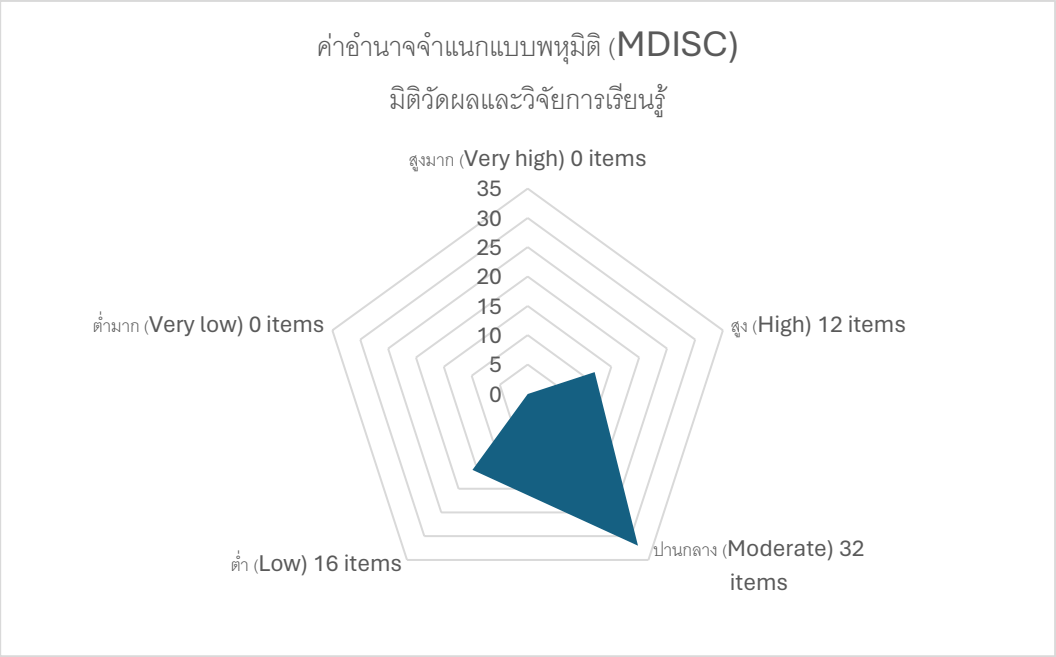
ตารางที่ 20 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 4 วัดผลและวิจัยการเรียนรู้ (a_4)

ค่าพารามิเตอร์	มิติของข้อสอบวัดผลและวิจัยการเรียนรู้ (a_4)					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0	0	0	0.764	0	-0.460	0.764	-0.476
SD	0	0	0	0.167	0	0.681	0.167	0.825
MIN	0	0	0	0.467	0	-2.000	0.467	-2.000
MAX	0	0	0	1.000	0	1.030	1.000	1.830

จากตารางที่ 20 พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ (a_4) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.467 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.764 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.467 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.764 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าความยากแบบพหุมิติ (MDIFF)

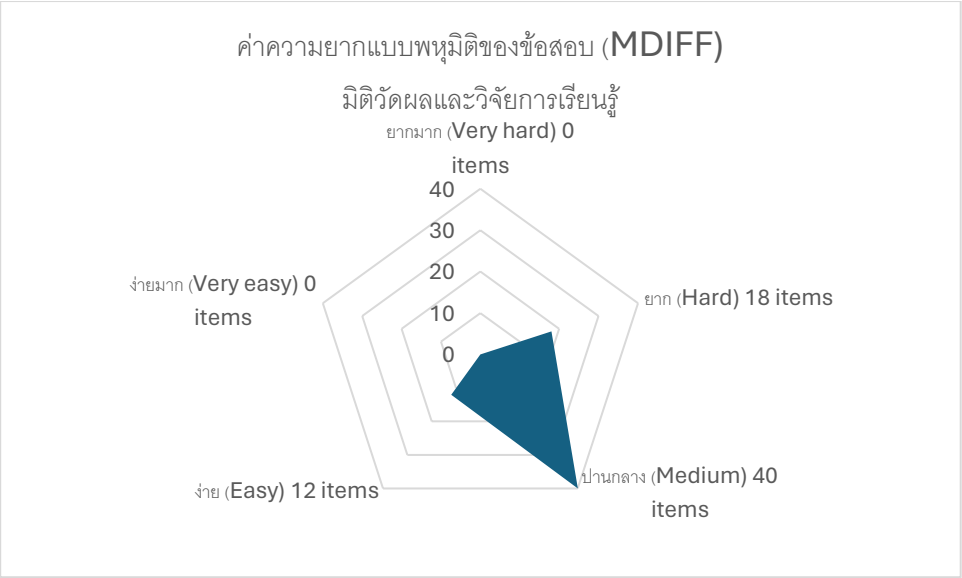
มีค่าอยู่ระหว่าง -2.000 ถึง 1.830 ค่าเฉลี่ยมีค่าเท่ากับ -0.476 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.825 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ยมีค่าเท่ากับ -0.460 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.681

เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ ดังภาพที่ 54 และ 55



ภาพที่ 54 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้

จากภาพที่ 54 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ พบว่ามีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกปานกลาง (Moderate) จำนวน 32 ข้อ รองลงมาเป็นข้อสอบมีค่าอำนาจจำแนกต่ำ (Low) จำนวน 16 ข้อ และค่าอำนาจจำแนกสูง (High) จำนวน 12 ข้อตามลำดับ



ภาพที่ 55 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้

จากภาพที่ 55 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ พบว่ามีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าปานกลาง (Medium) จำนวน 40 ข้อ รองลงมาเป็นข้อสอบที่มีค่ายาก (Hard) จำนวน 18 ข้อ และข้อสอบที่มีค่าง่าย (Easy) จำนวน 12 ข้อ ตามลำดับ

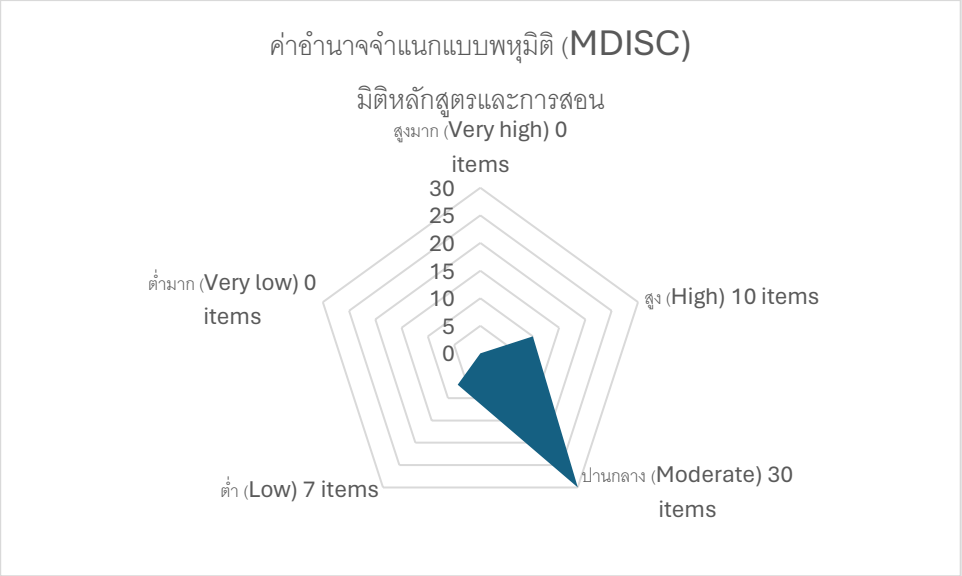
ตารางที่ 21 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 5 หลักสูตรและการสอน (a_5)

ค่าพารามิเตอร์	มิติของข้อสอบหลักสูตรและการสอน (a_5)					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0	0	0	0	0.750	-0.318	0.750	-0.301
SD	0	0	0	0	0.159	0.719	0.159	0.884
MIN	0	0	0	0	0.454	-2.00	0.454	-2.000
MAX	0	0	0	0	1.000	1.250	1.000	1.890

จากตารางที่ 21 พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติหลักสูตรและการสอน (a_5) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.454 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.750 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.159 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.454 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.750 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.159 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.000 ถึง 1.890 ค่าเฉลี่ยมีค่าเท่ากับ -0.301 และส่วนเบี่ยงเบนมาตรฐาน

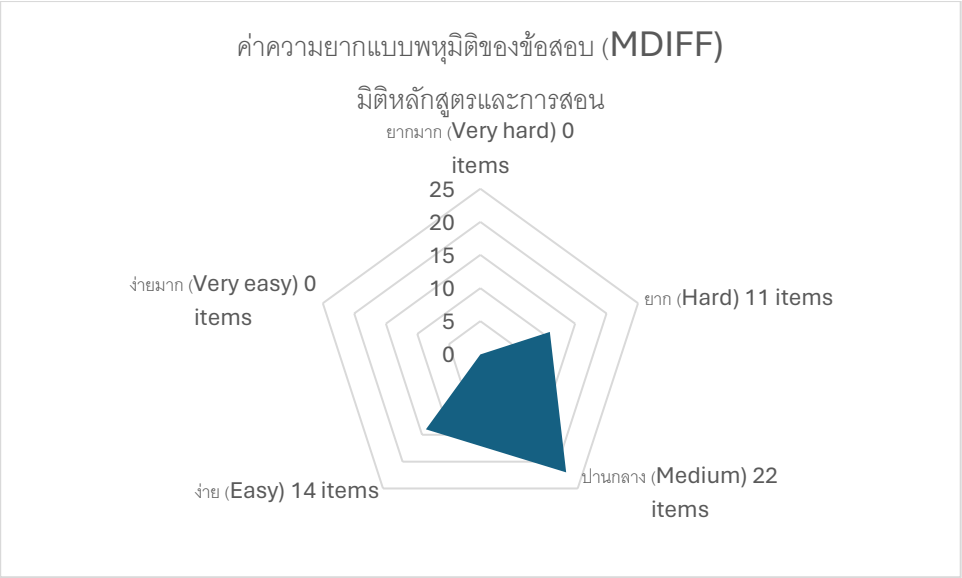
มีค่าเท่ากับ 0.884 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.250 ค่าเฉลี่ยมีค่าเท่ากับ -0.318 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.719

เมื่อนำมาทำเป็นแผนภูมิ เรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติ (MDISC) และค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติหลักสูตรและการสอน ดังภาพที่ 56 และ 57



ภาพที่ 56 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้

จากภาพที่ 56 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ พบว่า มีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกปานกลาง (Moderate) จำนวน 30 ข้อ รองลงมาเป็นข้อสอบมีค่าอำนาจจำแนกสูง (High) จำนวน 10 ข้อ และค่าอำนาจจำแนกต่ำ (Low) จำนวน 7 ข้อ ตามลำดับ



จากภาพที่ 57 ค่าความยากแบบพหุมิติ (MDIFF) ของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ พบว่ามีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าปานกลาง (Medium) จำนวน 22 ข้อ รองลงมาเป็นข้อสอบที่มีค่าง่าย (Easy) จำนวน 14 ข้อ และข้อสอบที่มีค่ายาก (Hard) จำนวน 11 ข้อ ตามลำดับ

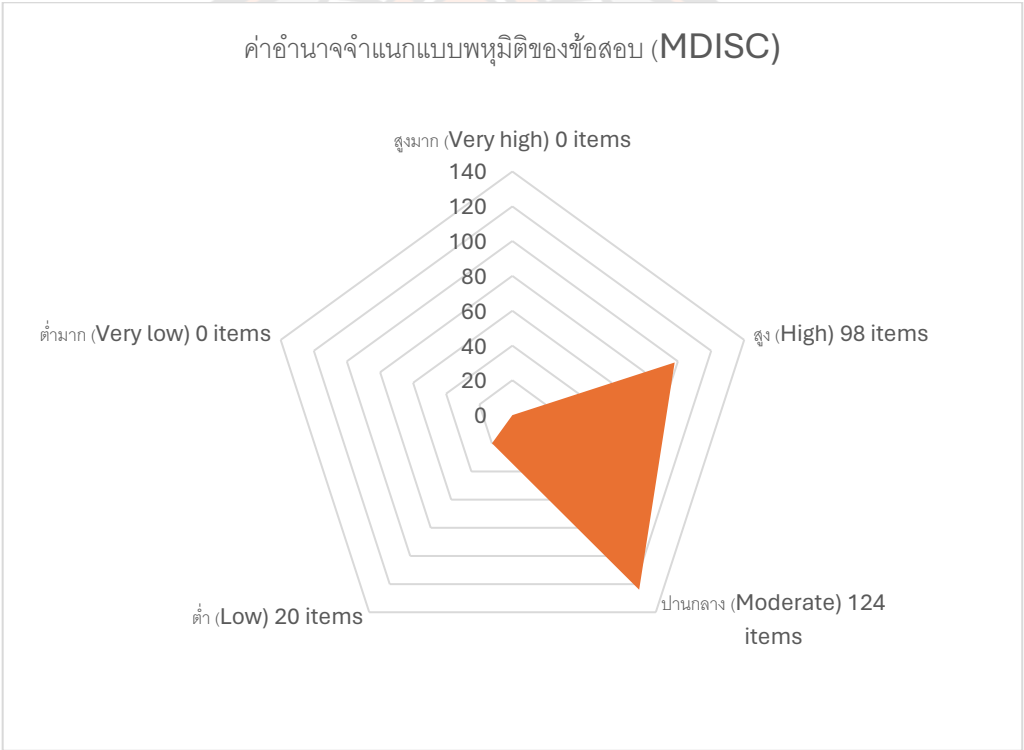
ตารางที่ 22 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู ที่ผ่านการคัดเลือกเข้าคลังข้อสอบ

ค่าพารามิเตอร์	มิติของข้อสอบ					d	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
MEAN	0.150	0.118	0.140	0.189	0.146	-0.427	0.743	-0.430
SD	0.304	0.276	0.303	0.341	0.306	0.696	0.168	0.877
MIN	0.00	0.00	0.00	0.00	0.00	-2.432	0.430	-2.432
MAX	1.00	1.00	1.00	1.00	1.00	1.254	1.000	2.199

จากตารางที่ 22พบว่าค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบมีค่าค่าอำนาจจำแนกในมิติที่ 1 มีค่าอยู่ระหว่าง 0.000 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.150 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.304 ค่าอำนาจจำแนกในมิติที่ 2 มีค่าอยู่ระหว่าง 0.000 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.118 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.276 ค่าอำนาจจำแนกในมิติที่ 3 มีค่าอยู่ระหว่าง 0.000 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.140 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.303 ค่าอำนาจจำแนกในมิติที่ 4 มีค่าอยู่ระหว่าง 0.000 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.189 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.341 ค่าอำนาจจำแนกในมิติที่ 5 มีค่าอยู่ระหว่าง 0.000 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.146 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.306 ความยากแบบพหุมิติ (MDIFF) มีค่า

อยู่ระหว่าง -2.432 ถึง 2.199 ค่าเฉลี่ยมีค่าเท่ากับ -0.430 และส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.877 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.743 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.168 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.430 ถึง 1.254 ค่าเฉลี่ยมีค่าเท่ากับ -4.427 และส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.696

ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (Multidimensional Discrimination: MDISC) ในคลังข้อสอบ มีค่าเฉลี่ยเท่ากับ 0.743 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.168 และมีค่าอยู่ระหว่าง 0.430 ถึง 1.000 เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ ดังภาพที่ 57

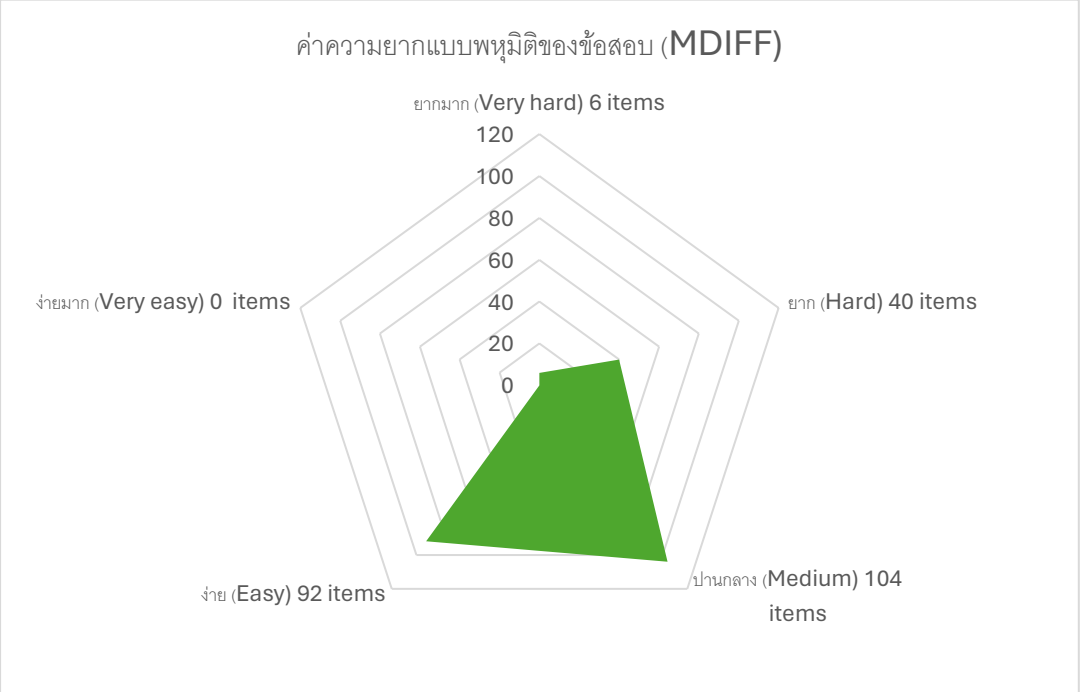


ภาพที่ 57 ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC)

จากภาพที่ 57 แนวโน้มโดยรวมของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) เมื่อพิจารณาจากแผนภูมิพบว่าคลังข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจปานกลาง (Moderate) จำนวน 124 ข้อ รองลงมา คือ มีค่าอำนาจจำแนกสูง (High) จำนวน 98 ข้อ และมีค่าอำนาจจำแนกต่ำ (Low) จำนวน 20 ข้อตามลำดับ

ค่าความยากแบบพหุมิติของข้อสอบ (Multidimensional Difficulties : MDIFF) ในคลังข้อสอบ มีค่าเฉลี่ยเท่ากับ -0.430 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.877 และมีค่าอยู่ระหว่าง -2.432

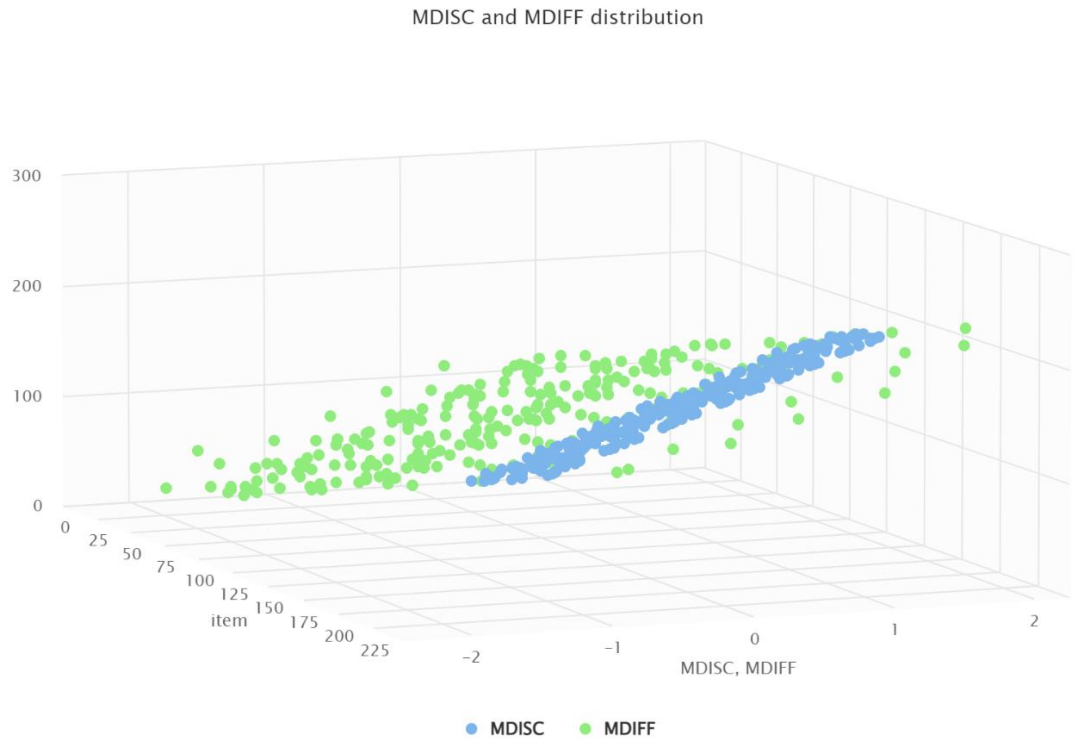
ถึง 2.199 เมื่อนำมาทำเป็นแผนภูมิเรดาห์ (Radar Charts) แสดงลักษณะแนวโน้มโดยรวมของค่าความยากแบบพหุมิติของข้อสอบ ดังภาพที่ 58



ภาพที่ 58 ค่าความยากแบบพหุมิติของข้อสอบ (MDIFF)

จากภาพที่ 58 แนวโน้มโดยรวมของค่าความยากแบบพหุมิติของข้อสอบในคลังข้อสอบ เมื่อพิจารณาจากแผนภูมิพบว่าคลังข้อสอบมีค่าความยากแบบพหุมิติของข้อสอบส่วนใหญ่เป็นข้อสอบมีค่าความยากปานกลาง (Medium) จำนวน 104 ข้อ รองลงมาข้อสอบง่าย (Easy) จำนวน 92 ข้อ ข้อสอบยาก (Hard) จำนวน 40 ข้อ และยากมาก (Very Hard) จำนวน 6 ข้อ ตามลำดับ

นำมาเขียนกราฟแบบจุด (scatter plot) เพื่อดูลักษณะกลุ่มข้อมูล (Cluster) ของค่าความยากแบบพหุมิติของข้อสอบ และค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ แสดงดังภาพที่ 59



ภาพที่ 59 แสดงการกระจายระหว่างค่าความยากแบบพหุมิติของข้อสอบ
กับค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ

ตอนที่ 2 ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

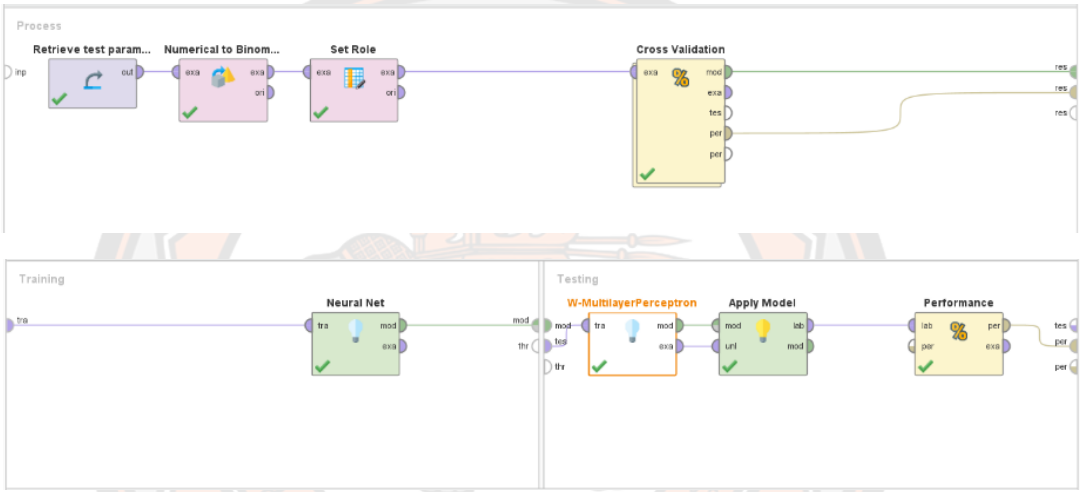
2.1 วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียมซึ่งเป็นเลือกใช้โครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron : MLP) ใช้การเรียนรู้แบบมีผู้สอน (supervised learning) ซึ่งประกอบด้วยเซลล์ประสาท 3 ชั้น ได้แก่ชั้นชั้นข้อมูลป้อนเข้า ($\mathbf{n_{input}}$) เท่ากับ 25 โหนด และจำนวนโหนดในชั้นซ่อน ($\mathbf{n_{hidden1}}$) จำนวน 10 โหนด โหนดในชั้นซ่อน ($\mathbf{n_{hidden2}}$) จำนวน 10 โหนด โหนดในชั้นซ่อน ($\mathbf{n_{hidden3}}$) จำนวน 10 โหนดและจำนวนโหนดชั้นข้อมูลส่งออก ($\mathbf{n_{output}}$) จำนวน 2 โหนด (รูปแบบข้อมูล 25-10-10-10-2) ทำการออกแบบโมเดลโดยใช้เทคนิคดาต้าไมน์นิง (Data mining) หรือการทำเหมืองข้อมูล ในการจำแนกประเภทข้อมูล (Classification) และใช้โปรแกรม RapidMiner Studio Educational 10.1 วิเคราะห์โดยใช้อัลกอริธึม Neural Net เพื่อเป็นการเพิ่มประสิทธิภาพความแม่นยำในการคัดเลือกข้อสอบข้อถัดไปให้มีความถูกต้องมากที่สุด และผู้วิจัยได้กำหนดค่าพารามิเตอร์ในกระบวนการของโครงข่ายประสาทเทียม ประกอบด้วยอัตราการเรียนรู้เท่ากับ 0.95 โมเมนตัม (Momentum) เท่ากับ 0.05 ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ

ตารางที่ 23 รายละเอียดข้อมูลสำหรับทำเหมืองข้อมูล ข้อสอบ 5 ข้อ

Variable/ Node	Detail	Key code
$X1_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 1	Parameter value
$X1_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 2	Parameter value
$X1_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 3	Parameter value
$X1_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 4	Parameter value
$X1_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 5	Parameter value
$X2_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 1	Parameter value
$X2_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 2	Parameter value
$X2_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 3	Parameter value
$X2_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 4	Parameter value
$X2_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 5	Parameter value
$X3_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 1	Parameter value
$X3_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 2	Parameter value
$X3_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 3	Parameter value
$X3_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 4	Parameter value
$X3_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 5	Parameter value
$X4_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 1	Parameter value
$X4_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 2	Parameter value
$X4_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 3	Parameter value
$X4_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 4	Parameter value
$X4_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 5	Parameter value
$X5_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 1	Parameter value
$X5_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 2	Parameter value
$X5_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 3	Parameter value
$X5_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 4	Parameter value
$X5_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 5	Parameter value

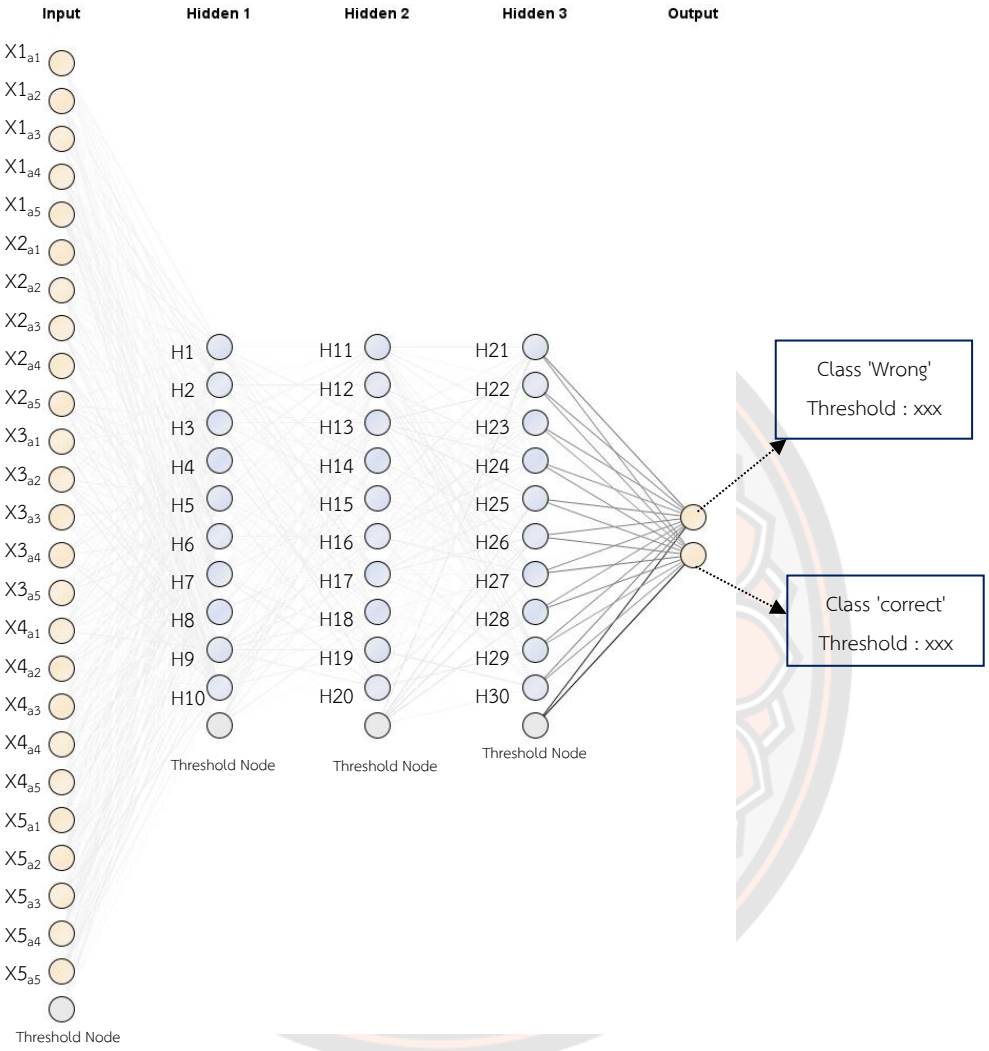
จากตารางที่ 23 นำค่าพารามิเตอร์ของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) ในคลังข้อสอบ จำนวน 5 ข้อ ที่มีค่าระดับปานกลาง (Moderate) อยู่ระหว่าง $0.65 \leq MDISC < 1.35$

(Baker, 2001; Hasmy, 2014) มาทำการทดสอบประสิทธิภาพในการจำแนกข้อมูลโดยการวัดความถูกต้องของอัลกอริธึม Neural Net ด้วยวิธีการ 10-fold cross-validation ซึ่งจะแบ่งข้อมูลออกเป็น 10 ชุดเท่า ๆ กันแล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ (Test set) 20% ส่วนที่เหลือ 9 ชุดนำมาใช้เป็นกลุ่มเรียนรู้ (Training Set) 80% แล้วทำการ วนทำเป็นจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อย ๆ จนครบ ในการวัดประสิทธิภาพของการจำแนกประเภทข้อมูลได้แก่ การหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถ่วงดุล (F-measure) และค่าความถูกต้อง (Accuracy)



ภาพที่ 60 สร้างโมเดลการคัดเลือกข้อสอบข้อถัดไป
จากโปรแกรม RapidMiner Studio Educational 10.1

จากโมเดลการคัดเลือกข้อสอบข้อถัดไปที่ใช้โปรแกรม RapidMiner Studio Educational 10.1 วิเคราะห์โดยใช้อัลกอริธึม Neural Net ผลการวิเคราะห์ปรากฏดังภาพที่ 61



ภาพที่ 61 โมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) กรณีตอบถูก และตอบผิด

การนำโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียมไปใช้จริงได้นั้น จำเป็นต้องทราบประสิทธิภาพของโมเดลก่อน โดยทั่วไปแล้วจะมีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่างๆ อยู่ 4 ค่า ดังนี้ 1) ค่าความแม่นยำ (Precision) คือค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์ 2) ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล 3) ค่าความถ่วงดุล (F-measure) คือค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก และ

4) ค่าความถูกต้อง (Accuracy) คือ จำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส โดยต้องมีค่าสูงที่สุด

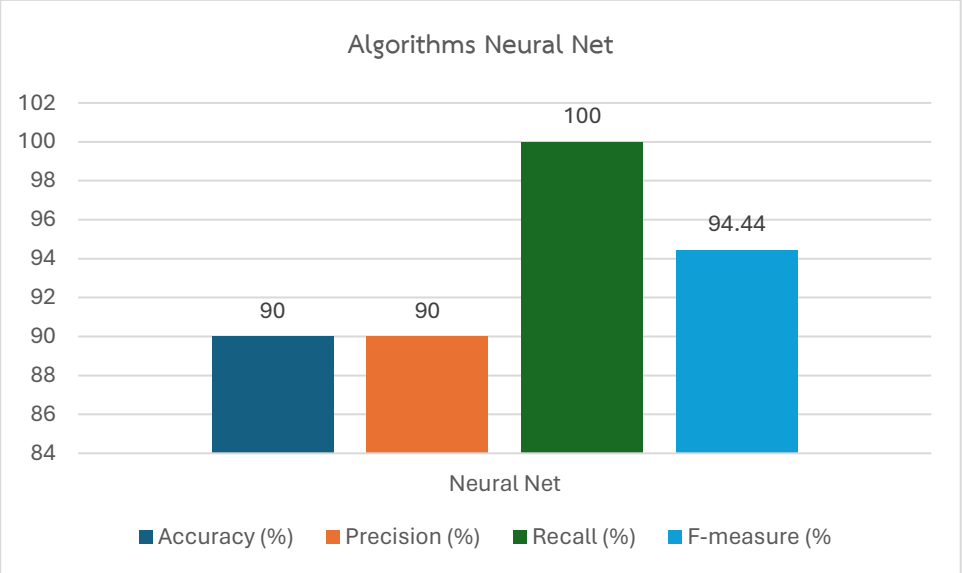
จากการทดลองโมเดล สามารถอธิบายผลการทดลองได้ ดังนี้

1. การทดสอบประสิทธิภาพ ด้วยวิธี แบบวิธี 10 Fold Cross-Validation เพื่อหาประสิทธิภาพของโมเดลโครงข่ายประสาทเทียม ดังตารางที่ 24

ตารางที่ 24 ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบ

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%
Neural Net	90	90	100	94.74

จากตารางที่ 24 พบว่าจากการประเมินประสิทธิภาพด้วยวิธีการ 10-fold cross-validation ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบข้อถัดไป ค่าความระลึก (Recall) คิดเป็นร้อยละ 100 สูงที่สุด รองลงมาคือ ค่าความถ่วงดุล (F-measure) คิดเป็นร้อยละ 94.44, ค่าความถูกต้อง (Accuracy) คิดเป็นร้อยละ 90 และการวัดค่าความแม่นยำ (Precision) คิดเป็นร้อยละ 90 ตามลำดับ



ภาพที่ 62 เปรียบเทียบค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึกและค่าความถ่วงดุลของการคัดเลือกข้อสอบข้อถัดไป

นำโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) มาเขียนสูตรคำนวณ ได้ดังนี้ :

Hidden Layer 1:

สำหรับ Nodei (โดย i=1, 2,...,10):

$a_1^{(1)}$ = Sigmoid ($\sum_{j=1}^{5}$ parameter1 _$a_j^{(1)}$ ⋅ X_j + Bias)

Hidden Layer 2:

สำหรับ Nodei (โดย i=1, 2,...,10):

$a_1^{(2)}$ = Sigmoid ($\sum_{j=1}^{10}$ parameter2 _$a_j^{(2)}$ ⋅ $a_j^{(1)}$ + Bias)

Hidden Layer 3:

สำหรับ Nodei (โดย i=1, 2,...,10):

$a_1^{(3)}$ = Sigmoid ($\sum_{j=1}^{10}$ parameter3 _$a_j^{(3)}$ ⋅ $a_j^{(2)}$ + Bias)

Output Layer:

สำหรับ Class 'wrong' : กรณีตอบผิด

o_{wrong} = Sigmoid ($\sum_{j=1}^{10}$ $parameter_{Wrong,i}$ _$a_j^{(3)}$ ⋅ $a_j^{(2)}$ + $Threshold_{wrong}$)

สำหรับ Class 'correct' : กรณีตอบถูก

$o_{Correct}$ = Sigmoid ($\sum_{j=1}^{10}$ $parameter_{correct,i}$ _$a_j^{(3)}$ ⋅ $a_j^{(2)}$ + $Threshold_{correct}$)

โดยที่:

xj คือ ข้อมูลนำเข้า (input) สำหรับ $Node_j$ ใน Hidden Layer 1.

$a_i^{(l)}$ คือ ผลลัพธ์ที่ได้จาก $Node_i$ ใน $Hidden\ Layer_l$.

$parameter_{a,j}^{(l)}$ และ $paramete_{class,i}$ คือ พารามิเตอร์ (parameters) ที่ถูกเรียนรู้ของโมเดลสำหรับ $Node_j$ ใน $Layer_l$ และ $Class_i$ ใน Output Layer ตามลำดับ.

Bias คือ พารามิเตอร์ที่เกี่ยวข้องกับการเทียบค่าเบื้องต้น (constant term) ในทุกๆ Node.

Sigmoid (x) คือ ฟังก์ชัน Sigmoid ที่นิยมใช้ใน Neural Network และมีสูตร คือ $\frac{1}{1+e^{-x}}$

การคำนวณ Output และการตัดสินใจในโมเดลการคัดเลือกข้อสอบข้อถัดไป ใช้การตัดสินใจ (decision-making) โดยนำการใช้ฟังก์ชัน sigmoid เพื่อแปลงค่าผลลัพธ์ในแต่ละโหนดของ Neural

Network ให้อยู่ในช่วงระหว่าง 0 ถึง 1 ซึ่งสามารถใช้ในการตัดสินใจว่า output Layer นั้นๆ ควรจะถูกจัดอยู่ใน Class ไหน ดังสมการ ได้ดังนี้

1. สมการของ Hidden Layer 1 :

คำนวณผลลัพธ์ใน Hidden Layer 1 ได้ด้วยฟังก์ชัน Sigmoid ตามสมการ :

$$Node_i = sigmoid(\sum_{j=1}^{5} parameter_{i,a_j} . input_{a_j} + Bias_i)$$

โดยที่

i คือ หมายเลขโหนดใน Hidden Layer 1 (เช่น Node 1, .., Node 10)

j คือ ค่าพารามิเตอร์ (parameter) ในแต่ละโหนด (จาก a_1 ถึง a_5)

ยกตัวอย่าง เช่น สำหรับ Node 1 ใน Hidden Layer 1:

$$Node_1 = sigmoid(\sum_{j=1}^{5} parameter_{1,a_j} . input_{a_j} + Bias_1)$$

และในที่นี้ a_j คือ a_1, a_2, a_3, a_4, a_5 ซึ่งแทนค่า input ที่เป็นข้อมูลนำเข้าของ Hidden Layer 1 สำหรับ Hidden Layer 2 และ Hidden Layer 3 ก็สามารถทำการคำนวณได้ตามลำดับเดียวกัน โดยนำผลลัพธ์จากชั้นก่อนหน้ามาเป็น input ของชั้นถัดไป

2. สมการของ Output Layer :

ใน Output Layer มีสอง Class คือ "ผิด (wrong)" และ "ถูก (correct)" โดยใช้ฟังก์ชัน Sigmoid เช่นกันสมการสำหรับการคำนวณใน Output Layer จะเป็นดังนี้ :

$$\text{Class 'wrong'} = \text{sigmoid}(\sum_{i=1}^{10} \text{Node}_i . \text{input}_{a_j} + \text{outputParameter}_i + \text{outputBias})$$

$$\text{Class 'correct'} = \text{sigmoid}(\sum_{i=1}^{10} \text{Node}_i . \text{input}_{a_j} + \text{outputParameter}_i + \text{outputBias})$$

โดยที่ i คือ หมายเลข Node ใน Hidden Layer3

3. สมการของตัดสินใจ (decision-making) คัดเลือกข้อสอบข้อถัดไปโดยใช้ Threshold : โดยนำ Output $Class_{'correct', 'wrong'}$ มาเปรียบเทียบกับ Threshold ได้สมการได้ดังนี้

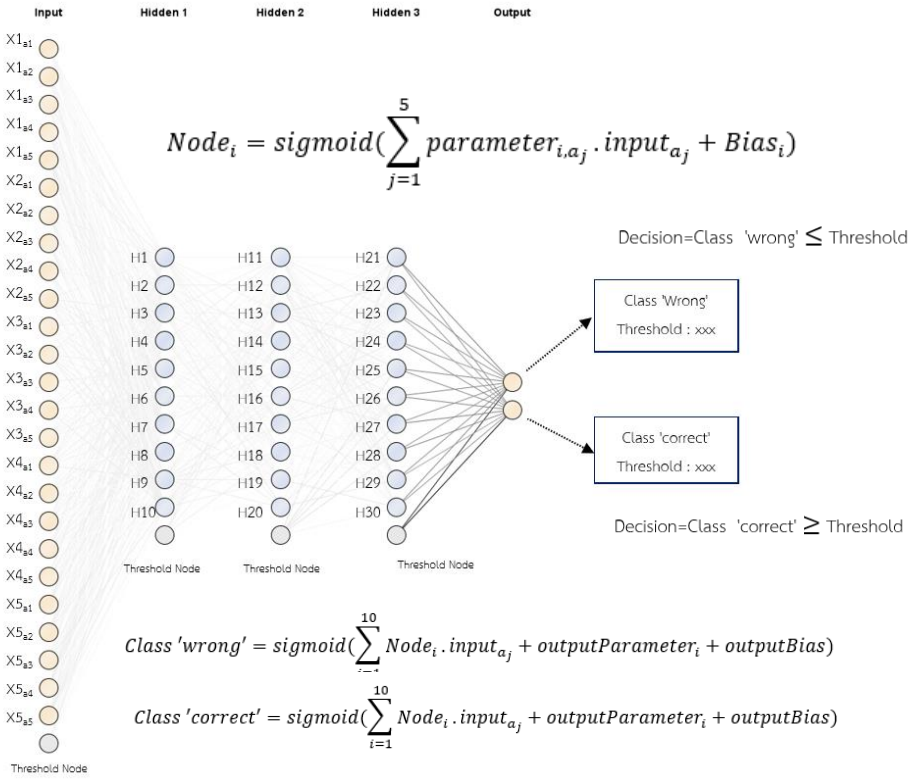
Decision = Class 'correct' $\geq$ Threshold; กรณีตอบถูก

Decision = Class 'wrong' $\leq$ Threshold; กรณีตอบผิด

การคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) จะทำการคัดเลือกข้อสอบที่มีการประมาณค่าพารามิเตอร์ความยากแบบพหุมิติ (Threshold) ของข้อสอบ ที่ใกล้เคียงกับค่าประมาณความสามารถ (θ) ขณะนั้นที่สุด ข้อสอบข้อไหนที่มีค่าพารามิเตอร์ความยาก (Threshold) ของข้อสอบ ใกล้เคียงกับการประมาณค่าของโหนด Output (Output Node) มากที่สุดข้อสอบข้อนั้นจะถูกเลือกเป็นข้อสอบข้อถัดไป อยู่ 2 กรณี 1) กรณีตอบถูก (correct) ค่า Threshold ข้อถัดไป ต้องสูงกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after answering correctly) 2) กรณีตอบผิด (wrong) ค่า Threshold ข้อถัดไป ต้องน้อยกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after answering incorrectly)

จากสมการสามารถนำมาออกแบบโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) (25-10-10-10-2) ได้ดังนี้

A model for selecting the next exam using a neural network method



ภาพที่ 63 สมการโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) (25-10-10-10-2)

นำสมการของโมเดลการคัดเลือกข้อสอบข้อถัดไปนี้มาเขียนเป็น Code PHP เนื่องจากการเขียน Code PHP มีความซับซ้อนมาก ผู้วิจัยจึงกำหนดฟังก์ชันคำนวณผลลัพธ์ โดยใช้ Forward Propagation ของแต่ละ

Node และ Layer ของ Neural Network ด้วยค่าพารามิเตอร์ที่กำหนดฟังก์ชัน Sigmoid มาทำการคำนวณผลลัพธ์ของ Neural Network ดังนี้

```php
<?php
function sigmoid($x) {
return 1 / (1 + exp(-$x));
}
function makeDecision($output, $threshold) {
return $output >= $threshold;
}
function forwardPropagation($inputValues, $parameters, $biases) {
$A = $inputValues;
foreach ($parameters as $W) {
$Z = dotProduct($W, $A) + array_pop($biases);
$A = sigmoid($Z);
}
return $A;
}
// Example usage
$inputValues = [/* parameters_a1_a5.txt */]; ##(your input values)
$outputThreshold = 0.5;
$hidden1Params = [/* parameters for hidden layer 1 */];
$hidden2Params = [/* parameters for hidden layer 2 */];
$hidden3Params = [/* parameters for hidden layer 3 */];
$outputParams = [/* parameters for output layer */];
$hiddenBiases = [
$hidden1Biases = [/* biases for hidden layer 1 */],
$hidden2Biases = [/* biases for hidden layer 2 */],
$hidden3Biases = [/* biases for hidden layer 3 */],
$outputBiases = [/* biases for output layer */]
];
$parameters = [$hidden1Params, $hidden2Params, $hidden3Params, $outputParams];
$biases = $hiddenBiases;
$output = forwardPropagation($inputValues, $parameters, $biases);
$decision = makeDecision($output, $outputThreshold);
if ($decision) {
$nextQuestion = "Next question after answering correctly";
} else {
$nextQuestion = "Next question after answering incorrectly";
}
echo $nextQuestion;
function dotProduct($array1, $array2) {
$result = 0;
$length = min(count($array1), count($array2));
for ($i = 0; $i < $length; $i++) {
$result += $array1[$i] * $array2[$i];
}
return $result;
}
?>
```

ภาพที่ 64 Code PHP สำหรับคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN)

ตอนที่ 3 ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

โดยผู้วิจัยออกแบบโปรแกรม สร้างผังงานโปรแกรม เขียนโปรแกรมโดยใช้ภาษา PHP และตรวจสอบการทำงานของโปรแกรมเบื้องต้นก่อนส่งให้ผู้เชี่ยวชาญตรวจสอบ โดยให้ผู้เชี่ยวชาญตรวจสอบความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานโปรแกรม ปรับปรุงโปรแกรมและคู่มือการใช้งานตามคำแนะนำ ของผู้เชี่ยวชาญและจัดทำ เป็นโปรแกรมและคู่มือการใช้งานโปรแกรมต้นฉบับ โดยนำ เสนอผลการวิเคราะห์ดังนี้

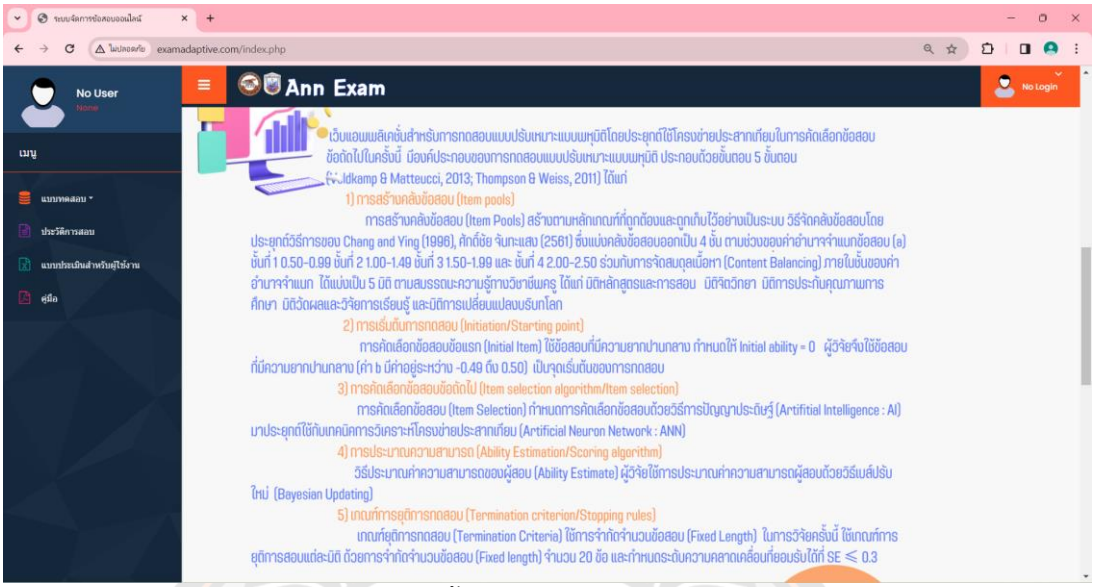
1. ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ

1.1 ผลการออกแบบโปรแกรม

เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปประกอบด้วยหน้าจอหลักและหน้าจอย่อยๆ ซึ่งมีปุ่มคำสั่งต่างๆ ดังนี้

1.1.1 หน้าจอหลัก ประกอบด้วยปุ่มคำสั่งต่างๆ และคู่มือการใช้งานโปรแกรม ซึ่งในส่วนของหน้าจอนี้ผู้ใช้งานสามารถเข้าไปศึกษาข้อมูลเบื้องต้นก่อนทำการทดสอบและศึกษาข้อมูลการใช้งานอย่างละเอียดจากคู่มือการใช้งานโปรแกรม ส่วนประกอบต่างๆ ของหน้าจอหลักแสดงดังภาพที่ 65 และภาพที่ 66

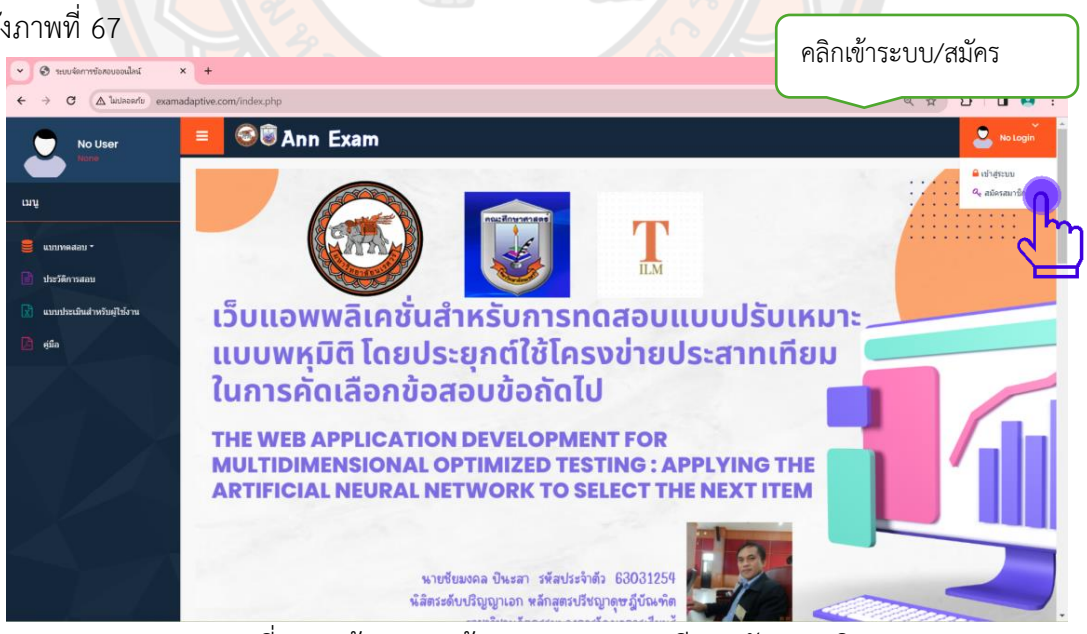
ภาพที่ 65 หน้าจอแรกของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป เมื่อเข้าใช้งานโปรแกรมที่เว็บไซต์ http://examadaptive.com/index.php



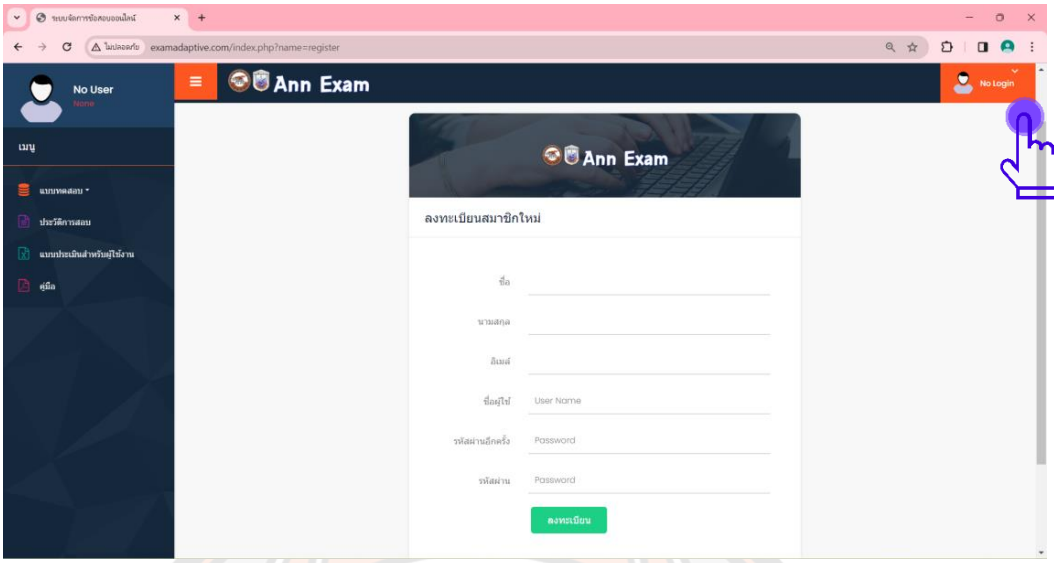
ภาพที่ 66 หน้าจอแสดงความหมายและขั้นตอนการทำงานของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยสรุป

1.1.2 หน้าจอย่อยๆ ประกอบด้วยหน้าจอต่างๆ ดังนี้

1) หน้าจอการเข้าระบบและลงทะเบียนสมัครสมาชิก มุมบนขวา มีลักษณะเป็นแบบฟอร์มให้ผู้สอบพิมพ์และเลือกข้อมูลของตนเองตามความเป็นจริงลงในช่องว่างตามหัวข้อที่กำหนด ประกอบด้วย 1) ชื่อ 2) นามสกุล 3) อีเมล์ 4) ชื่อผู้ใช้ และ 5) รหัสผ่าน/รหัสผ่านอีกครั้ง ดังภาพที่ 67

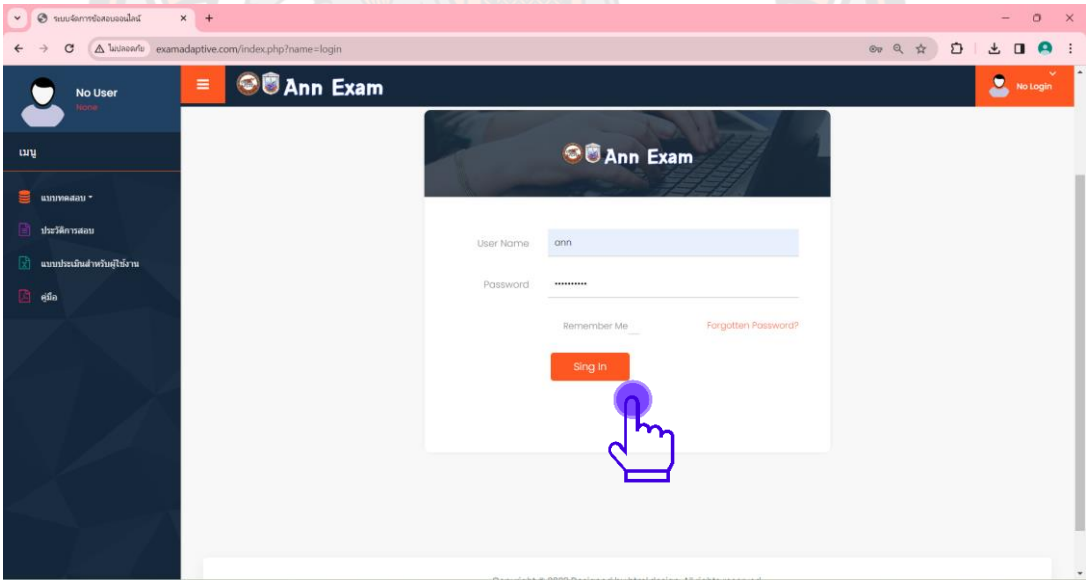


ภาพที่ 67 หน้าจอการเข้าระบบและลงทะเบียนสมัครสมาชิก



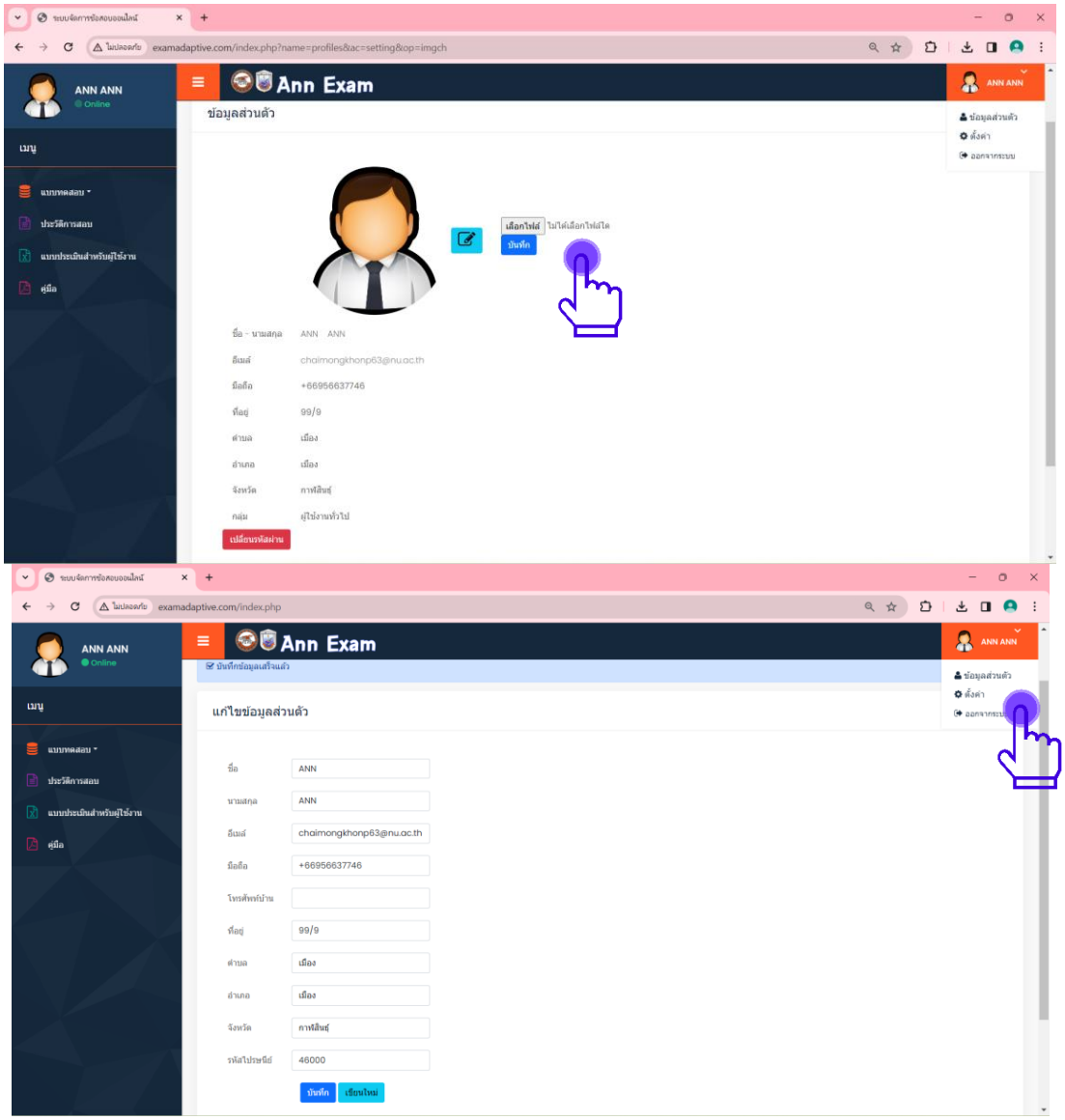
ภาพที่ 68 หน้าจอแสดงแบบฟอร์มให้ผู้สอบพิมพ์ข้อมูลและเลือกข้อมูลของตนเองตามความเป็นจริง เมื่อกดปุ่ม ลงทะเบียน

2) หน้าจอแรกของการเข้าใช้งานโปรแกรม โดยกดปุ่ม เข้าระบบ ที่แถบเมนูด้านบนเพื่อเข้าใช้งานแอพพลิเคชั่น แสดงดังภาพที่ 69



ภาพที่ 69 หน้าจอแรกของการเข้าใช้งานโปรแกรม

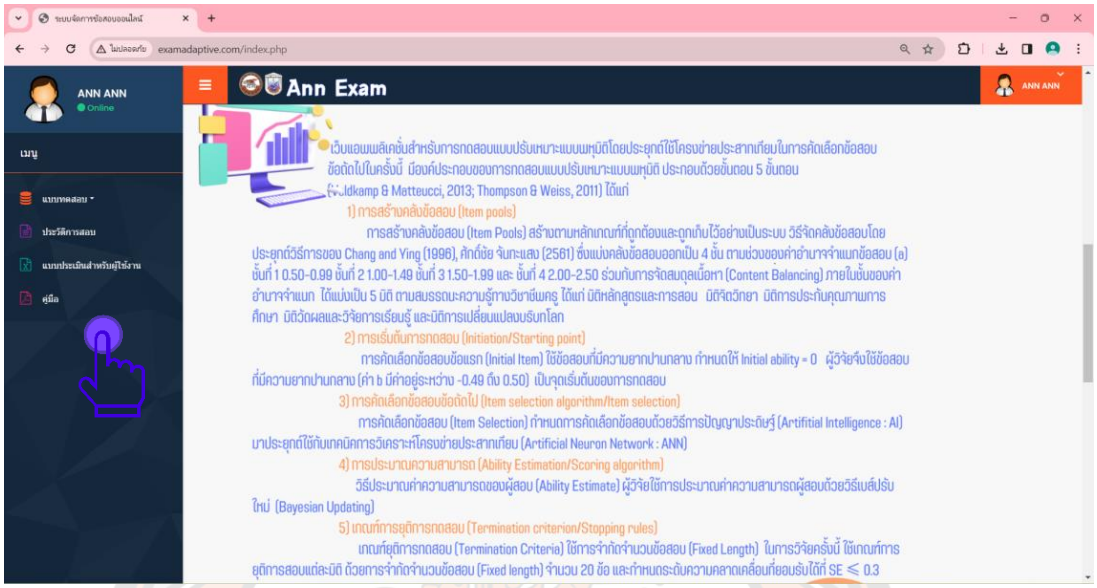
3) หน้าแสดงแก้ไขข้อมูลส่วนตัว ตั้งค่าต่างๆ และ ออกจากระบบ ที่แถบเมนูด้านบนขวา แสดงดังภาพที่ 70



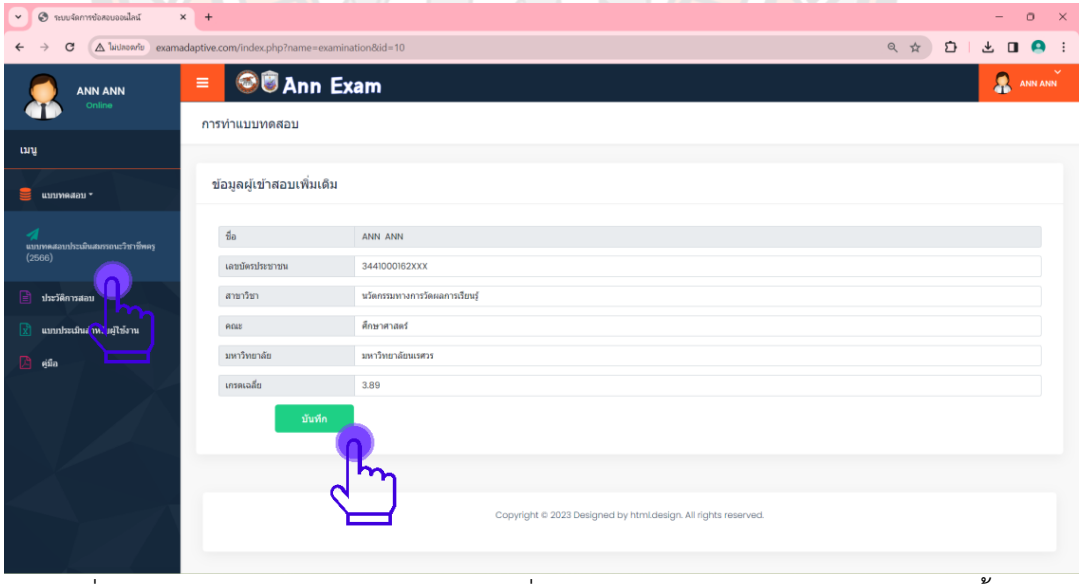
ภาพที่ 70 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้

หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้สามารถตั้งค่าการจัดการข้อมูลผู้ใช้ และแก้ไขข้อมูลส่วนตัว

4) หน้าจอเริ่มทำข้อสอบ มุมซ้ายมือ ประกอบด้วย 1) แบบทดสอบ 2) ประวัติการสอบ 3) แบบประเมินสำหรับผู้ทดสอบ และ 4) คู่มือการใช้แอพพลิเคชั่น แสดงดังภาพที่ 71



5) หน้าจอเริ่มทำข้อสอบ มุมซ้ายมือก่อนจะทำแบบทดสอบ ระบบจะให้กรอกข้อมูลเบื้องต้นของผู้สอบเพิ่มเติม แล้วกด บันทึก

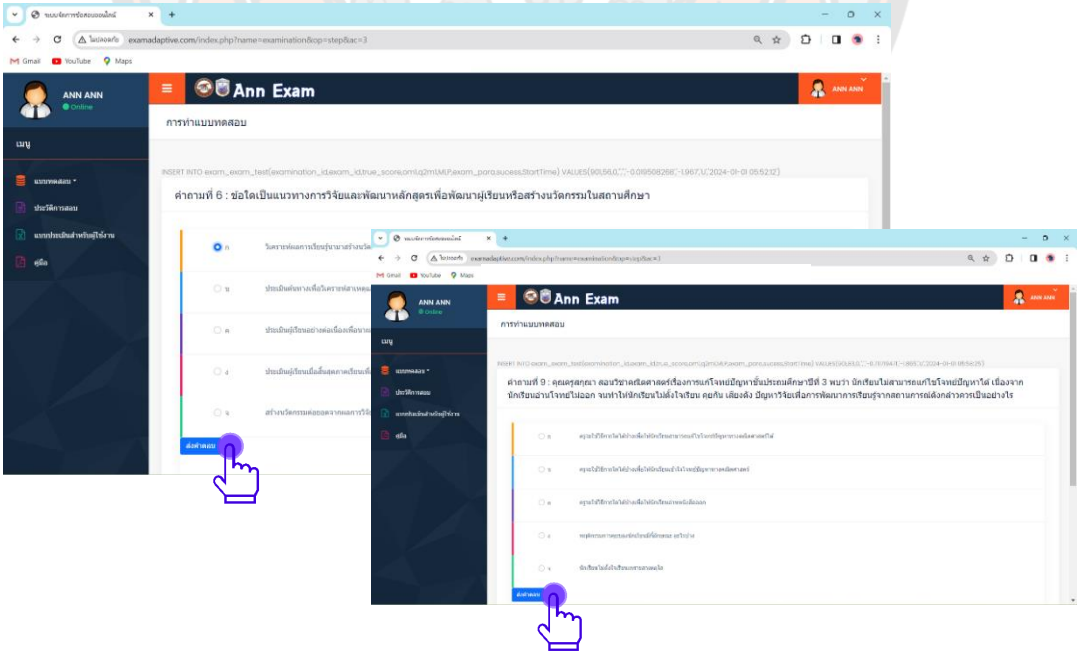


ภาพที่ 71 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้ ผู้สอบตรวจสอบข้อมูลเบื้องต้นของผู้สอบ แล้วกดปุ่ม บันทึก

6) หลังจากผู้สอบตรวจสอบข้อมูลเบื้องต้นของผู้สอบแล้ว จะปรากฏหน้าจอข้อมูลรายวิชา และรายละเอียดแบบทดสอบ ให้ผู้สอบอ่านคำชี้แจงให้ละเอียด ก่อนกด เริ่มทำแบบทดสอบ

ภาพที่ 72 หน้าจอแสดงข้อมูลรายวิชา และรายละเอียดแบบทดสอบ ผู้สอบอ่านคำชี้แจงให้ละเอียดก่อนกด เริ่มทำแบบทดสอบ

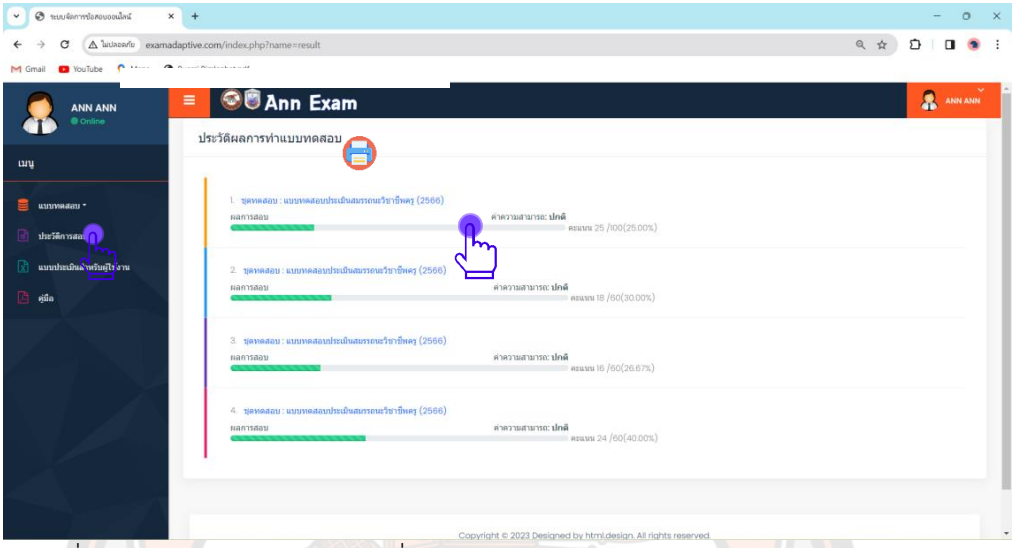
7) หน้าจอแสดงการทำแบบทดสอบ ส่วนนี้หน้าจอจะแสดงคำถาม และตัวเลือก 5 ตัวเลือก



ภาพที่ 73 หน้าจอแสดงตัวเลือกตอบ 5 ตัวเลือก

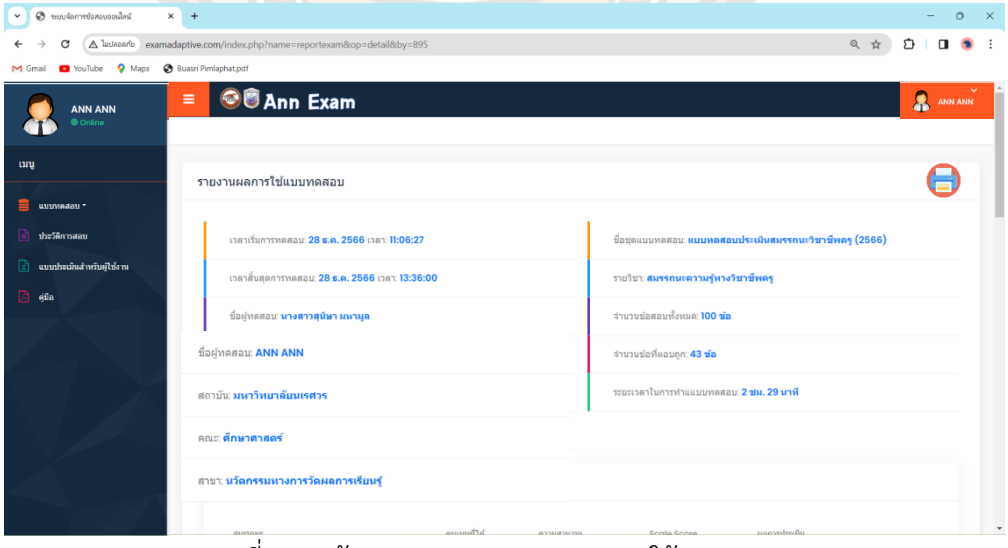
เมื่อผู้สอบอ่านคำถามจนเข้าใจแล้วผู้สอบจะต้องกดเลือกตอบเพียง 1 ตัวเลือก และกดปุ่มส่งคำตอบ

8) หน้าจอแสดงผลการทำแบบทดสอบ เมื่อผู้สอบทำข้อสอบครบทั้ง 5 มิติแล้วให้ผู้สอบกดปุ่ม ประวัติการสอบ ที่แถบเมนูด้านซ้าย เพื่อดูผลการสอบ โปรแกรมจะแสดงข้อมูลตามที่ผู้สอบลงทะเบียนไว้ และผลการทำแบบทดสอบ ดังภาพที่ 74



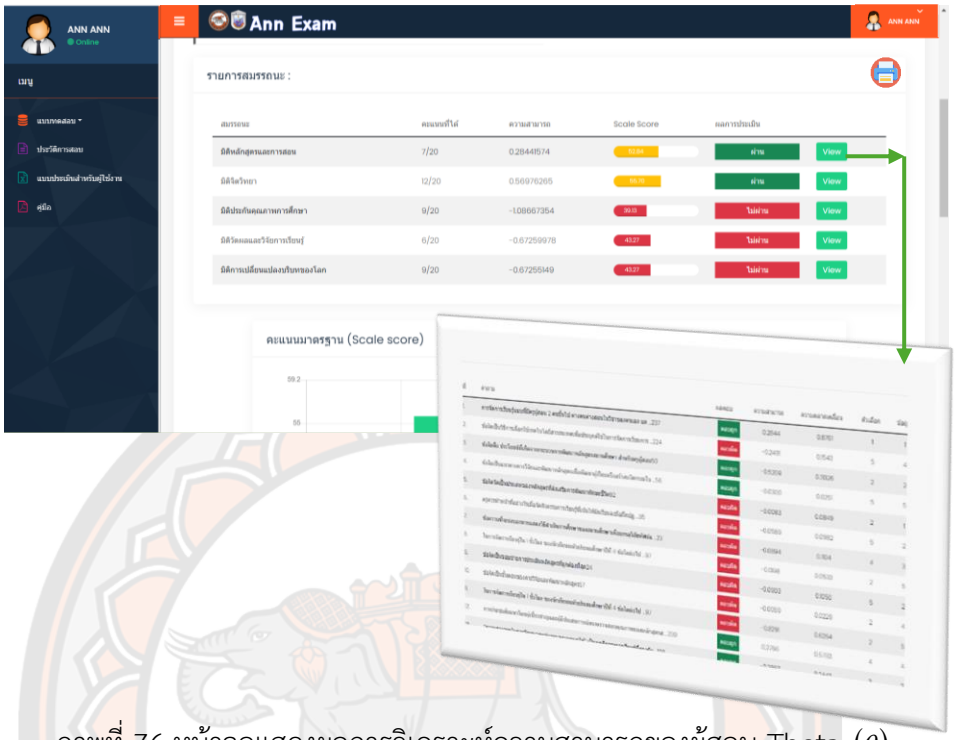
ภาพที่ 74 หน้าจอแสดงข้อมูลตามที่ผู้สอบลงทะเบียนไว้ และประวัติผลการทำแบบทดสอบ

12) หน้าจอแสดงผลประวัติผลการทำแบบทดสอบ สามารถกดดูรายงานผลการใช้แบบทดสอบ ผลการวิเคราะห์ความสามารถของผู้สอบแยกตามมิติของข้อสอบ หน้าจอผลการวิเคราะห์ความสามารถของผู้สอบแยกตามมิติของข้อสอบหน้าจอนี้จะแสดงค่าความสามารถและผลการประเมินระดับความสามารถของผู้สอบแยกตามมิติของข้อสอบ คือ มิติหลักสูตรและการสอน มิติจิตวิทยา มิติการประกันคุณภาพการศึกษา มิติวัดผลและวิจัยการเรียนรู้ และมิติการเปลี่ยนแปลงบริบทโลก และความสามารถโดยรวม ดังภาพที่ 75



ภาพที่ 75 หน้าจอแสดงผลรายงานผลการใช้แบบทดสอบ

13) หน้าจอผลการวิเคราะห์ความสามารถ Theta (θ) ของผู้สอบแยกตามมิติของข้อสอบ



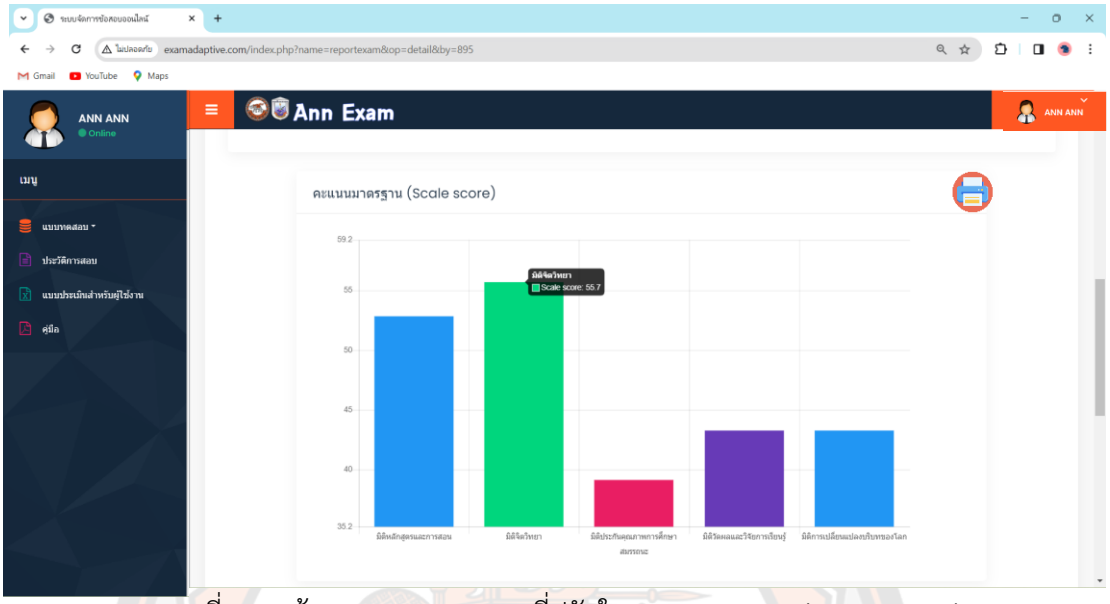
ภาพที่ 76 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบ Theta (θ), ค่า Scale Score และผลการประเมินแยกตามมิติ

จากภาพที่ 76 การรายงานผลการทดสอบ จะเป็นการรายงานผลคะแนนความสามารถในแต่ละมิติ ผลการทดสอบรายงานในรูปของคะแนนความสามารถ คือค่า Theta (θ) ซึ่งได้จากการประมาณค่าตามแนวทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยค่า Theta มีช่วงของคะแนนอยู่ระหว่าง -2.00 ถึง +2.00 และเพื่อให้การแปลความหมายคะแนนง่ายต่อการรายงานผลการทดสอบ จึงได้ปรับคะแนน Theta ให้อยู่ในรูปของคะแนนมาตรฐาน (scale score) ซึ่งเป็นการปรับค่าคะแนนให้มีค่าเฉลี่ยเท่ากับ 50 และมีค่าส่วนเบี่ยงมาตรฐานเท่ากับ 10 โดยคะแนนมาตรฐานนี้มีช่วงคะแนนอยู่ระหว่าง 20 ถึง 80 คะแนน ผู้วิจัยได้เลือกใช้เกณฑ์การแปลความหมาย ดังนี้

ตารางที่ 25 คะแนนความสามารถTheta ที่ปรับเป็นคะแนนมาตรฐาน (scale score)

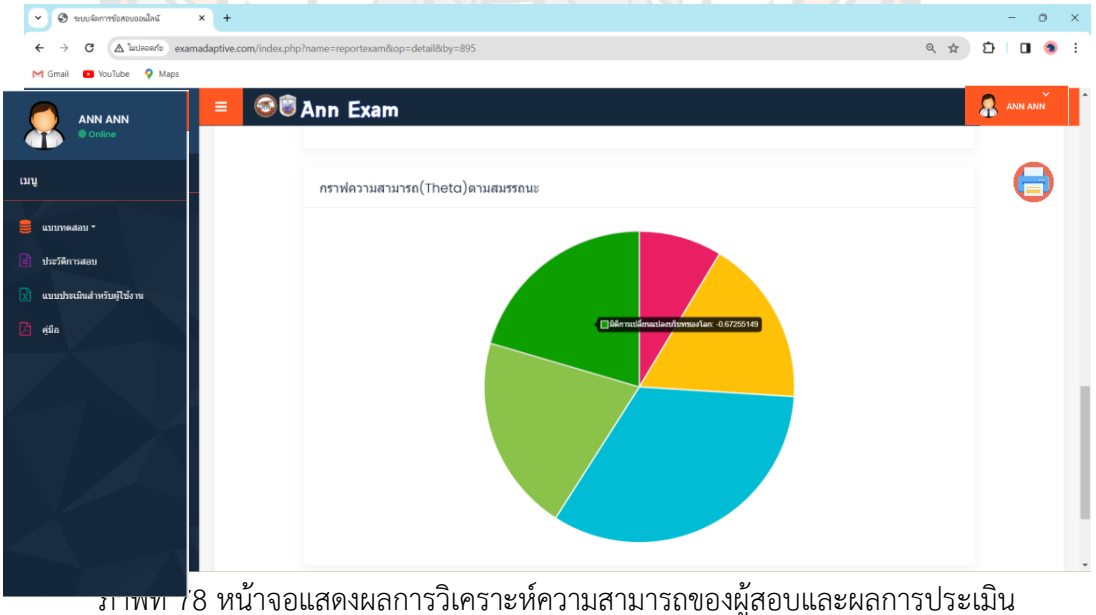
คะแนนที่ปรับสเกลมาตรฐาน (scale score) =50+(10* Theta (θ))	**ผลการประเมิน**
น้อยกว่า 50	ไม่ผ่าน
50 – 64	ผ่าน
ตั้งแต่ 65 ขึ้นไป	ดี

18) หน้าจอแสดงผลคะแนนที่ปรับในสเกลมาตรฐาน (scale score) ของผู้สอบแยกตามมิติของข้อสอบ ดังภาพที่ 77



ภาพที่ 77 หน้าจอแสดงผลคะแนนที่ปรับในสเกลมาตรฐาน (scale score)

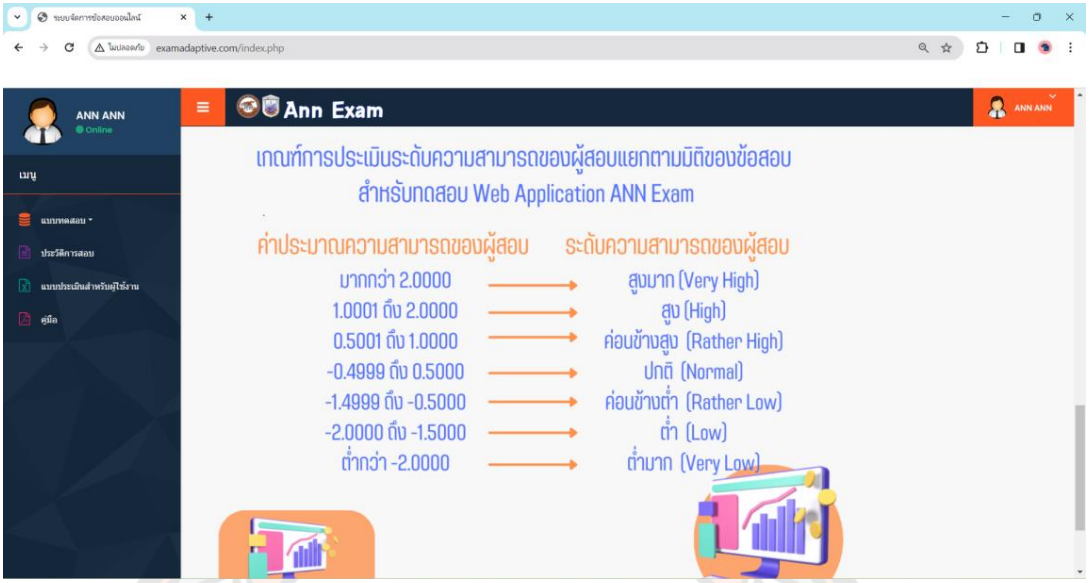
19) หน้าจอแสดงผลกราฟความสามารถ Theta (θ) ของผู้สอบแยกตามมิติของข้อสอบ ดังภาพที่ 78



ภาพที่ 78 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบและผลการประเมิน

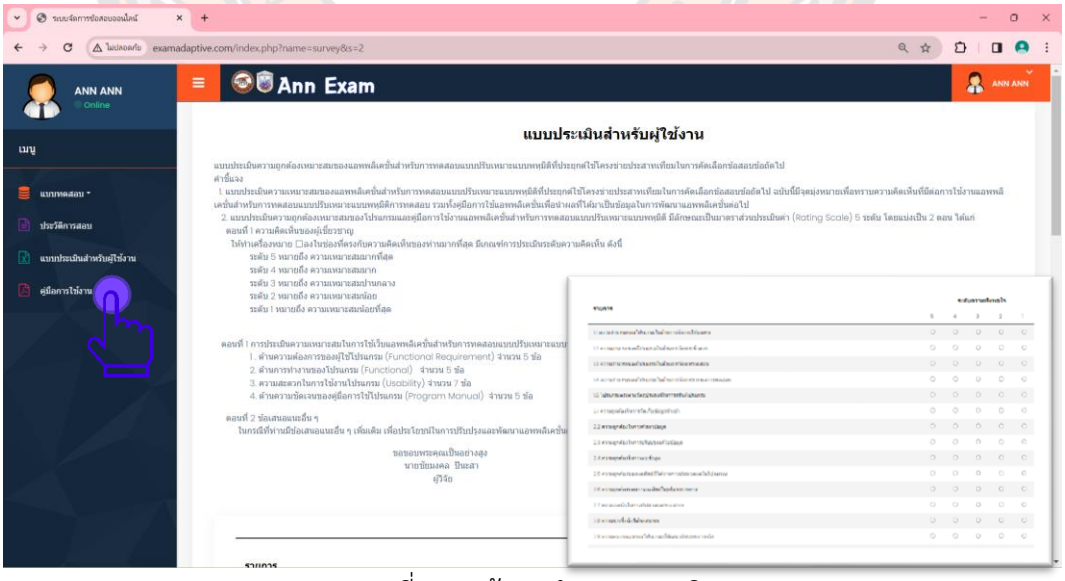
ระดับความสามารถของผู้สอบแยกตามมิติของข้อสอบ และสรุปความสามารถโดยรวมซึ่งผู้สอบสามารถเทียบระดับความสามารถของตนเองได้โดยนำค่าความสามารถที่ได้จากการทดสอบไปเทียบกับเกณฑ์ ดังภาพ 4.20

19) หน้าจอแสดงเกณฑ์การประเมินความสามารถของผู้สอบแยกตามมิติของข้อสอบ



ภาพที่ 79 หน้าจอเกณฑ์การประเมินความสามารถของผู้สอบแยกตามมิติของข้อสอบ

เมื่อผู้สอบดูรายงานผลคะแนนความสามารถในแต่ละมิติ และสั่งพิมพ์รายงานผลการทดสอบเรียบร้อยแล้ว ก่อนที่จะออกจากระบบการทดสอบนั้น ผู้สอบสามารถกดปุ่ม “ทำแบบประเมินและสิ้นสุดการทดสอบ” เพื่อทำแบบประเมินความเหมาะสมของแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปต่อการใช้ระบบการทดสอบครั้งนี้ ดังแสดงในภาพที่ 80



ภาพที่ 80 หน้าจอทำแบบประเมิน

ตอนที่ 4 ผลการหาประสิทธิภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

2.1 ผลการประเมินความเหมาะสมโดยผู้เชี่ยวชาญ

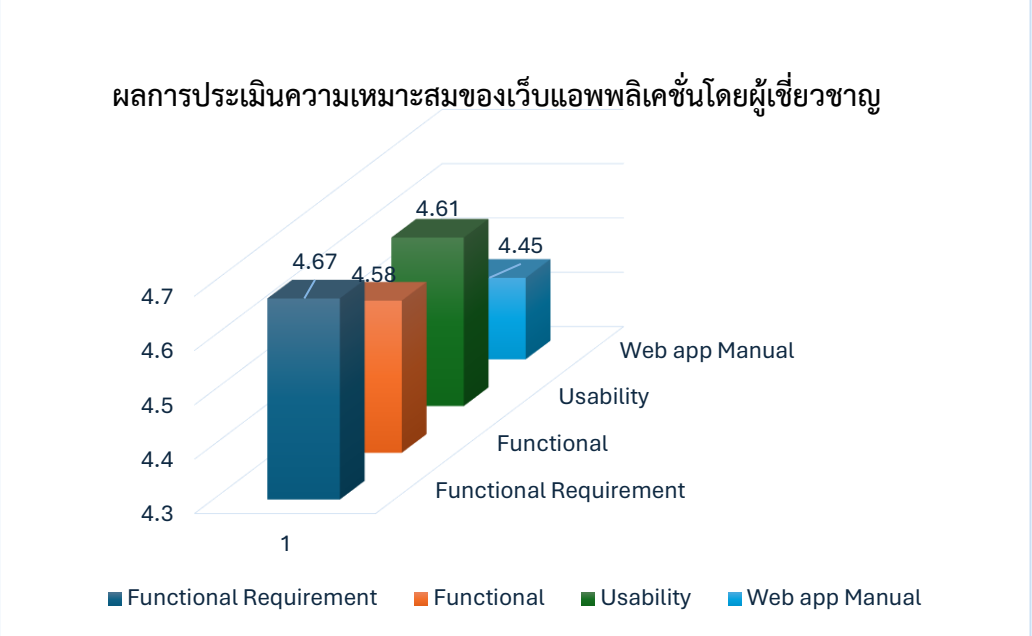
การประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้เชี่ยวชาญ ผู้วิจัยได้นำเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ พร้อมคู่มือการใช้เว็บแอพพลิเคชั่น ส่งให้ผู้เชี่ยวชาญตรวจสอบโดยผู้เชี่ยวชาญทั้ง 5 คน ได้ประเมินความเหมาะสมของเว็บแอพพลิเคชั่น ใน 4 ด้าน ได้แก่ 1) ด้านความต้องการของผู้ใช้เว็บแอพพลิเคชั่น (Functional Requirement) 2) ด้านการทำงานของโปรแกรม (Functional) 3) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น(Usability) 4) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น(Web application Manual) และข้อเสนอแนะเพื่อการพัฒนาเว็บแอพพลิเคชั่น ประเมินในระหว่างวันที่ 1–10 พฤศจิกายน พ.ศ. 2566 ผลการประเมินความเหมาะสมแสดงดังตารางที่ 26

ตารางที่ 26 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้เชี่ยวชาญ

ด้านการประเมิน	$\overline{X}$	S.D.	ระดับความเหมาะสม
1) ด้านความสามารถของแอพพลิเคชั่น (Functional Requirement)	4.67	0.50	เหมาะสมมากที่สุด
2) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional)	4.58	0.55	เหมาะสมมากที่สุด
3) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability)	4.61	0.54	เหมาะสมมากที่สุด
4) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual)	4.45	0.62	เหมาะสมมาก
รวม	**4.57**		**เหมาะสมมากที่สุด**

จากตาราง 4.6 แสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้เชี่ยวชาญ แสดงให้เห็นว่า โปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด ($\overline{X}$=4.57) เมื่อพิจารณาเป็นรายด้าน ปรากฏว่ามี 4 ด้าน ที่มีความเหมาะสมอยู่ในระดับมากที่สุด ได้แก่ ด้านความสามารถของแอพพลิเคชั่น (Functional Requirement) ($\overline{X}$= 4.67) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) ($\overline{X}$ = 4.58) และความสะดวกในการใช้งาน

เว็บแอพพลิเคชั่น (Usability) ($\overline{X}$ = 4.61) รองลงมาได้แก่ ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) มีความเหมาะสมอยู่ในระดับมาก ($\overline{X}$ = 4.45) เมื่อนำมาเขียนกราฟ ปรากฏดังภาพ 4.22

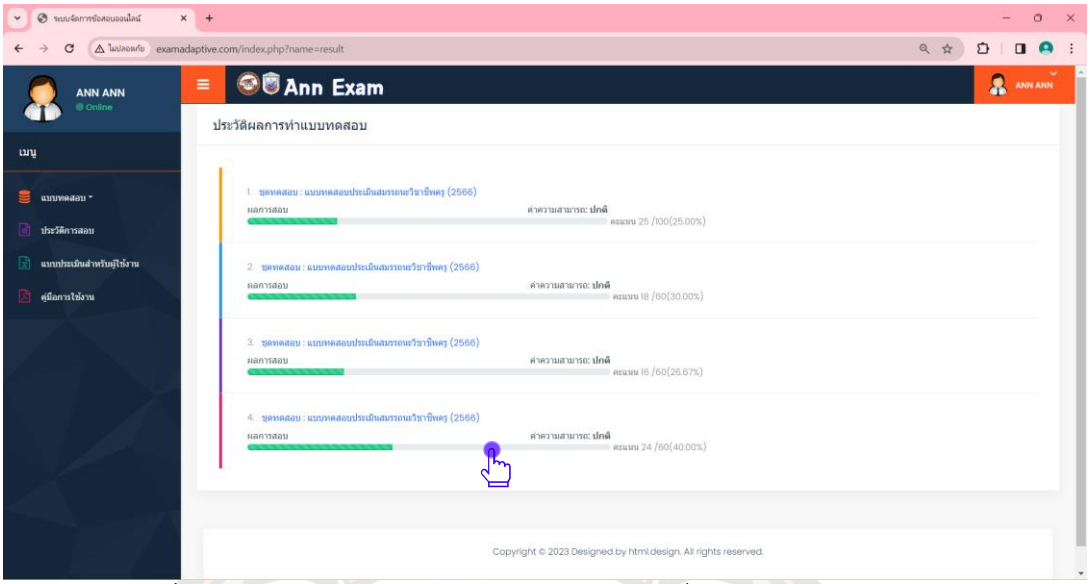


ภาพที่ 81 กราฟแสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้เชี่ยวชาญ

นอกจากนี้ ผู้เชี่ยวชาญยังมีข้อเสนอแนะเพิ่มเติม ดังนี้

1. ถ้า user คนหนึ่งทำหลายๆ ครั้ง ระบบควรเก็บข้อมูลการทำหลายครั้งไว้เพื่อเก็บสถิติและดูพัฒนาการ

ผู้วิจัยได้ปรับปรุงตามข้อเสนอแนะ โดยให้ผู้สอบสามารถเข้ามาทำการทดสอบได้หลายๆ ครั้ง และเก็บข้อมูลประวัติผลการทำแบบทดสอบด้วยเพื่อเก็บสถิติและดูพัฒนาการ



ภาพที่ 82 เก็บข้อมูลประวัติผลการทำแบบทดสอบเพื่อเก็บสถิติและดูพัฒนาการ

2.2 ผลการประเมินความเหมาะสมโดยผู้ใช้งาน

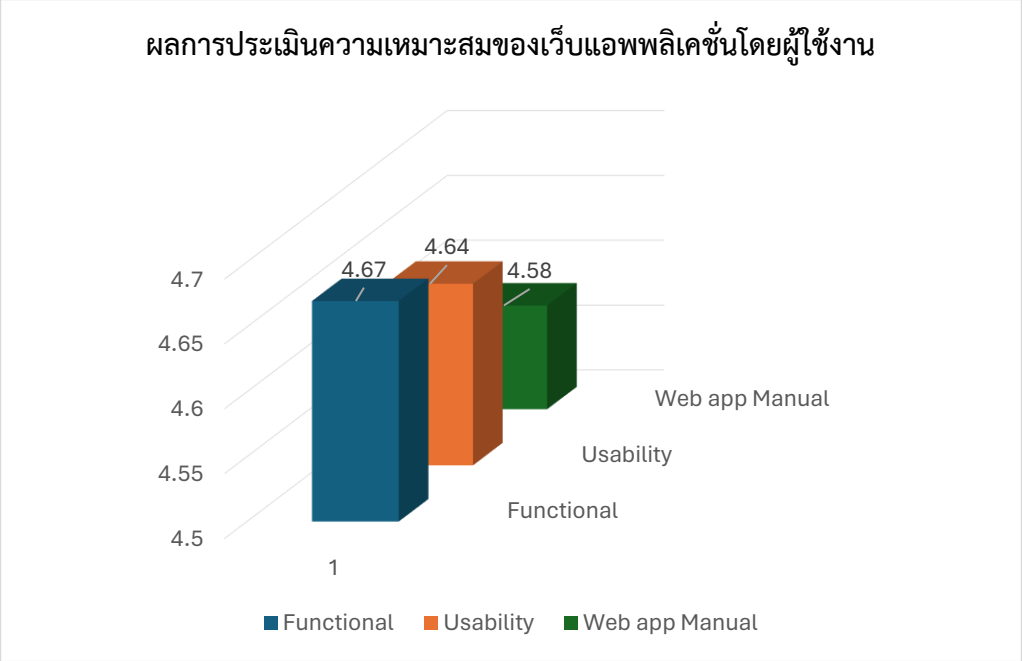
ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยโดยผู้ใช้งาน ทั้ง 3 ด้าน ได้แก่ 1) ด้านการทำงานของโปรแกรม (Functional) 2) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability) 3) ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) และข้อเสนอแนะเพื่อการพัฒนาเว็บแอพพลิเคชั่น โดยนักศึกษาที่ทดลองใช้งานเว็บแอพพลิเคชั่น จำนวน 200 คน ผลการประเมินความเหมาะสม แสดงดังตารางที่

ตารางที่ 27 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้ใช้งาน

ด้านการประเมิน	$\overline{X}$	S.D.	ระดับความเหมาะสม
1) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional)	4.67	0.55	เหมาะสมมากที่สุด
2) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability)	4.64	0.51	เหมาะสมมากที่สุด
3)ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual)	4.58	0.54	เหมาะสมมากที่สุด
รวม	**4.63**		**เหมาะสมมากที่สุด**

จากตารางที่ 27 แสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้ใช้งาน แสดงให้เห็นว่า โปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด ($\overline{X}$=4.63) เมื่อพิจารณาเป็นรายด้าน ปรากฏว่ามี 3 ด้าน ที่มีความเหมาะสมอยู่ในระดับ

มากที่สุด ได้แก่ ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) ($\overline{X}$ = 4.67) ด้านความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability) ($\overline{X}$ = 4.64) และด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) ($\overline{X}$ = 4.58) เมื่อนำมาเขียนกราฟ ปรากฏดังภาพที่ 83



ภาพที่ 83 กราฟแสดงผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นโดยผู้ใช้งาน

จากผลการประเมินความเหมาะสมของโปรแกรมทั้งผู้เชี่ยวชาญและผู้ใช้งาน ผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงแก้ไขเว็บแอพพลิเคชั่นให้มีความเหมาะสม และมีความรวดเร็วยิ่งขึ้น และเพื่อไม่ให้ระบบการทดสอบล่ม และรองรับผู้เข้าสอบพร้อมกันหลายคน ผู้วิจัยใช้ Cloud Server ข้อดีของ Cloud Server จะประกอบด้วย Server หลายๆ เครื่อง ช่วยกันทำงาน เป็นระบบที่มีประสิทธิภาพสูง สามารถทำงานได้อย่างรวดเร็ว ไม่มีสะดุด มีความปลอดภัยในการจัดเก็บข้อมูลสูงมีความเสถียรสูง สามารถใช้งานได้อย่างไร้กังวล โดยไม่ต้องกลัว Network ล่ม มีความยืดหยุ่นสูงใช้งานง่าย สามารถขยายพื้นที่ในการจัดเก็บข้อมูลได้ไม่จำกัด และง่ายต่อการ Backup และสำรองข้อมูล จากผลการประเมินความเหมาะสมโปรแกรม แสดงให้เห็นว่า โปรแกรมมีความเหมาะสมเพียงพอที่จะนำไปใช้ทดสอบจริงได้

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ผู้วิจัยได้นำเสนอผลการสรุปผลการวิจัย การอภิปรายผล และข้อเสนอแนะได้ดังนี้

1. สรุปผล
2. อภิปรายผล
3. ข้อเสนอแนะ

สรุปผล

จากผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป และเพื่อประเมินคุณภาพของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป และผลการวิเคราะห์ข้อมูลสามารถสรุปผลการวิจัย ได้ดังนี้

1. ผลการพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูโดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ สรุปผลการวิจัย ได้ดังนี้

1.1 การคัดเลือกข้อสอบเข้าคลังข้อสอบโดยวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม NOHARM 4.0 โดยกำหนดค่าพารามิเตอร์โอกาสการเดาให้มีค่าคงที่เท่ากับ 0.20 บันทึกค่าพารามิเตอร์ของข้อสอบแต่ละข้อ จากนั้นคัดเลือกข้อสอบที่มีค่า MDISC ตั้งแต่ 0.00 ขึ้นไป และมีค่า MDIFF อยู่ในช่วง -4.00 ถึง +4.00 ผลการคัดเลือกข้อสอบเข้าคลังที่ผ่านเกณฑ์การคัดเลือกจำนวน 242 ข้อ จำแนกเป็น 5 มิติ คือ 1) มิติการเปลี่ยนแปลงบริบทโลก จำนวน 50 ข้อ 2) มิติจิตวิทยาการศึกษา จำนวน 40 ข้อ 3) มิติการประกันคุณภาพการศึกษา จำนวน 45 ข้อ 4) มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ และ 5) มิติหลักสูตรและการสอน จำนวน 47 ข้อ

1.2 ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ ประกอบด้วยค่าต่างๆ ดังนี้

1.2.1 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 1 การเปลี่ยนแปลงบริบทโลก (a_1) พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติการเปลี่ยนแปลงบริบทโลก (a_1) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.448 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.725 และส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.167 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.448 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.725 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง 2.000 ถึง 1.930 ค่าเฉลี่ยมีค่าเท่ากับ -0.388 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.911 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ย มีค่าเท่ากับ -0.399 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.697

1.2.2 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 2 จิตวิทยาการศึกษา (a_2) พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติจิตวิทยาการศึกษา (a_2) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.717 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.172 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.717 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.172 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.000 ถึง 2.000 ค่าเฉลี่ยมีค่าเท่ากับ -0.458 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.895 และค่าจุดตัด

ของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ยมีค่าเท่ากับ -0.438 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.681

1.2.3 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 3 การประกันคุณภาพการศึกษา (a_3) พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติการประกันคุณภาพการศึกษา (a_3) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.440 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.752 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.179 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.440 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.752 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.179 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.430 ถึง 2.200 ค่าเฉลี่ยมีค่าเท่ากับ -0.527 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.904 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.430 ถึง 0.994 ค่าเฉลี่ยมีค่าเท่ากับ -0.515 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.719

1.2.4 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 4 วัดผลและวิจัยการเรียนรู้ (a_4) พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติวัดผลและวิจัยการเรียนรู้ (a_4) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.467 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.764 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.467 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.764 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.167 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.000 ถึง 1.830 ค่าเฉลี่ยมีค่าเท่ากับ -0.476 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.825 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.030 ค่าเฉลี่ยมีค่าเท่ากับ -0.460 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.681

1.2.5 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู มิติที่ 5 หลักสูตรและการสอน (a_5) พบว่าของข้อสอบความรู้ทางวิชาชีพครู มิติหลักสูตรและการสอน (a_5) มีค่าอำนาจจำแนกอยู่ระหว่าง 0.454 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.750 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.159 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.454 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.750 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.159 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -2.000 ถึง 1.890 ค่าเฉลี่ยมีค่าเท่ากับ -0.301 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.884 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -2.000 ถึง 1.250 ค่าเฉลี่ยมีค่าเท่ากับ -0.318 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.719

1.2.6 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู ที่ผ่านการคัดเลือกเข้าคลังข้อสอบโดยรวม พบว่าค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบมีค่าความยากแบบพหุมิติ (MDIFF) อยู่ระหว่าง -2.432 ถึง 2.199 ค่าเฉลี่ยมีค่าเท่ากับ -0.430 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.877 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) มีค่าอยู่ระหว่าง 0.430 ถึง 1.000 ค่าเฉลี่ยมีค่าเท่ากับ 0.743 ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.168 และค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่

ระหว่าง -2.430 ถึง 1.254 ค่าเฉลี่ยมีค่าเท่ากับ -4.427 และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.696

2. ผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network

2.1 การคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network ซึ่งเป็นเลือกใช้โครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron : MLP) การเรียนรู้แบบมีผู้สอน (supervised learning) แบบการแพร่ย้อนกลับ (back propagation Learning) ซึ่งประกอบด้วยเซลล์ประสาท 3 ชั้น ได้แก่ชั้นชั้นข้อมูลป้อนเข้า (n_{input}) เท่ากับ 25 โหนด และจำนวนโหนดในชั้นซ่อน $(\mathrm{n_{hidden2}})$จำนวน 10 โหนด โหนดในชั้นซ่อน $(\mathrm{n_{hidden2}})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(\mathrm{n_{hidden3}})$ จำนวน 10 โหนดและจำนวนโหนดชั้นข้อมูลส่งออก $(\mathrm{n_{output}})$ จำนวน 2 โหนด (รูปแบบ 25-10-10-10-2) ทำการออกแบบโมเดลโดยใช้เทคนิคดาต้าไมน์นิง (Data mining) หรือการทำเหมืองข้อมูล ในการจำแนกประเภทข้อมูล (Classification) และใช้โปรแกรม RapidMiner Studio Educational 10.1 วิเคราะห์โดยใช้อัลกอริธึม Neural Net เพื่อเป็นการเพิ่มประสิทธิภาพความแม่นยำในการคัดเลือกข้อสอบข้อถัดไปให้มีความถูกต้องมากที่สุด และผู้วิจัยได้กำหนดค่าพารามิเตอร์ในกระบวนการของ Artificial Neuron Network ประกอบด้วย อัตราการเรียนรู้เท่ากับ 0.95 โมเมนตัม (Momentum) เท่ากับ 0.05 ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ

2.2 นำค่าพารามิเตอร์ของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) ในคลังข้อสอบ ทั้ง 5 มิติ คัดเลือกมามิติละ 1 ข้อ (a_1-a_5) ที่มีค่าระดับปานกลาง (Moderate) อยู่ระหว่าง $0.65 \leq MDISC < 1.35$ (Baker, 2001; Hasmy, 2014) รวมจำนวน 5 ข้อ มาทำการทดสอบประสิทธิภาพในการจำแนกข้อมูลโดยการวัดความถูกต้องของอัลกอริธึม Neural Net ด้วยวิธีการ 10-fold cross-validation วิธีนี้เป็นวิธีที่นิยมในการทำงานวิจัย เพื่อใช้ในการทดสอบประสิทธิภาพของโมเดลเนื่องจากผลที่ได้มีความน่าเชื่อถือที่สุด ซึ่งจะแบ่งข้อมูลออกเป็น 10 ชุดเท่าๆ กันแล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ (Test set) 20% ส่วนที่เหลือ 9 ชุด นำมาใช้เป็นกลุ่มเรียนรู้ (Training Set) 80% แล้วทำการวนทำเป็นจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อยๆ จนครบจำนวนที่แบ่งไว้

2.3 ผลการทดสอบประสิทธิภาพของโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network ด้วยวิธี 10 Fold Cross-Validation เพื่อหาประสิทธิภาพของโมเดลโครงข่ายประสาทเทียม พบว่าจากการประเมินประสิทธิภาพด้วยวิธีการ 10-fold cross-validation ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบข้อถัดไป ค่าความระลึก (Recall) คิดเป็นร้อยละ 100 สูงที่สุด รองลงมาคือ ค่าความถ่วงดุล (F-measure) คิดเป็นร้อยละ

94.44, ค่าความถูกต้อง (Accuracy) คิดเป็นร้อยละ 90 และการวัดค่าความแม่นยำ (Precision) คิดเป็นร้อยละ 90 ตามลำดับ

นำโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) มาเขียนสูตรคำนวณ ได้ดังนี้ :

Hidden Layer 1: สำหรับ $Node_i$ (โดย i=1, 2,...,10):

$$a_1^{(1)} = \text{Sigmoid}\left(\sum_{j=1}^{5} \text{parameter1}_a_j^{(1)} \cdot X_j + \text{Bias}\right)$$

Hidden Layer 2: สำหรับ $Node_i$ (โดย i=1, 2,...,10):

$$a_1^{(2)} = \text{Sigmoid}\left(\sum_{j=1}^{10} \text{parameter2}_a_j^{(2)} \cdot a_j^{(1)} + \text{Bias}\right)$$

Hidden Layer 3: สำหรับ $Node_i$ (โดย i=1, 2,...,10):

$$a_1^{(3)} = \text{Sigmoid}\left(\sum_{j=1}^{10} \text{parameter3}_a_j^{(3)} \cdot a_j^{(2)} + \text{Bias}\right)$$

Output Layer[1]: สำหรับ Class 'wrong' : กรณีตอบผิด

$$o_{wrong} = \text{Sigmoid}\left(\sum_{j=1}^{10} \text{parameter}_{Wrong,i}_a_j^{(3)} \cdot a_j^{(2)} + \text{Threshold}_{wrong}\right)$$

Output Layer[2]: สำหรับ Class 'correct' : กรณีตอบถูก

$$o_{Correct} = \text{Sigmoid}\left(\sum_{j=1}^{10} \text{parameter}_{correct,i}_a_j^{(3)} \cdot a_j^{(2)} + \text{Threshold}_{correct}\right)$$

โดยที่ :

xj คือ ข้อมูลนำเข้า (input) สำหรับ $Node_j$ ใน Hidden Layer 1

$a_i^{(l)}$ คือ ผลลัพธ์ที่ได้จาก $Node_i$ ใน $Hidden\ Layer_l$.

$parameter_{a,j}^{(l)}$ และ $parame te_{class,i}$ คือ พารามิเตอร์ (parameters) ที่ถูกเรียนรู้ของโมเดลสำหรับ $Node_j$ ใน $Layer_l$ และ $Class_i$ ใน Output Layer ตามลำดับ

Bias คือ พารามิเตอร์ที่เกี่ยวข้องกับการเทียบค่าเบื้องต้น (constant term) ในทุกๆ Node

Sigmoid (x) คือ ฟังก์ชัน Sigmoid ที่นิยมใช้ใน Neural Network และมีสูตร คือ $\frac{1}{1+e^{-x}}$

สูตรคำนวณ Output และการตัดสินใจในโมเดลการคัดเลือกข้อสอบข้อถัดไปใช้การตัดสินใจ (decision-making) โดยนำการใช้ฟังก์ชัน sigmoid เพื่อแปลงค่าผลลัพธ์ในแต่ละโหนดของ Neural Network ให้อยู่ในช่วงระหว่าง 0 ถึง 1 ซึ่งสามารถใช้ในการตัดสินใจว่า output Layer นั้นๆ ควรจะถูกจัดอยู่ใน Class ไหน ดังสูตร ได้ดังนี้

1. สูตรคำนวณของ Hidden Layer 1:

คำนวณผลลัพธ์ใน Hidden Layer 1 ได้ด้วยฟังก์ชัน Sigmoid ตามสูตรคำนวณ :

$$Node_i = sigmoid(\sum_{j=1}^{5} parameter_{i,a_j} . input_{a_j} + Bias_i)$$

โดยที่

i คือ หมายเลขโหนดใน Hidden Layer1 (เช่น Node1, Node2, .., Node10)

j คือ ค่าพารามิเตอร์ (parameter) ในแต่ละโหนด (จาก a1 ถึง a5)

2. สูตรคำนวณของ Output Layer :

ใน Output Layer มีสอง Class คือ "ตอบผิด (wrong)" และ "ตอบถูก (correct)" โดยใช้ฟังก์ชัน Sigmoid เช่นกัน สูตรการคำนวณสำหรับ Output Layer จะเป็นดังนี้

$$Class\ 'wrong' = sigmoid(\sum_{i=1}^{10} Node_i . input_{a_j} + outputParameter_i + outputBias)$$

$$Class\ 'correct' = sigmoid(\sum_{i=1}^{10} Node_i . input_{a_j} + outputParameter_i + outputBias)$$

โดยที่ i คือ หมายเลข Node ใน Hidden Layer 3

3. สูตรคำนวณการตัดสินใจ (decision-making) คัดเลือกข้อสอบข้อถัดไปโดยใช้ Threshold : โดยนำ Output $Class_{'correct', 'wrong'}$ มาเปรียบเทียบกับ Threshold ได้สูตรคำนวณได้ดังนี้

Decision=Class 'correct' $\geq$ Threshold ; กรณีตอบถูก

Decision=Class 'wrong' $\leq$ Threshold ; กรณีตอบผิด

การคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้ Artificial Neuron Network จะทำการคัดเลือกข้อสอบที่มีการประมาณค่าพารามิเตอร์ความยากแบบพหุมิติ (Threshold) ของข้อสอบที่ใกล้เคียงกับค่าประมาณความสามารถ (θ) ขณะนั้นที่สุด ข้อสอบข้อไหนที่มีค่าพารามิเตอร์ความยาก (Threshold) ของข้อสอบ ใกล้เคียงกับการประมาณค่าของโหนด Output (Output Node) มากที่สุดข้อสอบข้อนั้นจะถูกเลือกเป็นข้อสอบข้อถัดไปอยู่ 2 กรณี 1) กรณีตอบถูก (correct) ค่า Thresholdข้อถัดไปต้องสูงกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after

answering correctly) 2) กรณีตอบผิด (wrong) ค่า Threshold ข้อถัดไป ต้องน้อยกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after answering incorrectly)

3. ผลการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป โดยพัฒนาอยู่ในรูปของ Web Application เขียนโปรแกรมด้วยภาษา PHP (Hypertext Preprocessor) เป็น Opensource พัฒนาเว็บแอบพลิเคชันนี้ร่วมกับ HTML, CSS, JavaScript, MySQL Database และ Cloud Server จะประกอบด้วย Server หลายๆ เครื่อง ช่วยกันทำงาน โดยพัฒนาให้สามารถใช้งานบนอุปกรณ์ที่เชื่อมต่อกับอินเทอร์เน็ตได้ เช่น คอมพิวเตอร์ โน๊ตบุ๊ค หรือสมาร์ทโฟน เว็บแอพพลิเคชั่นการทดสอบนี้แบ่งออกเป็น 3 ส่วน คือ 1) ข้อมูลทั่วไปเกี่ยวกับโปรแกรม เป็นที่อธิบายส่วนประกอบต่างๆ ของโปรแกรม 2) การจัดการของผู้ดูแลระบบเป็นส่วนที่อาจารย์ หรือผู้ที่เกี่ยวข้องในการจัดการข้อสอบนำข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) มาใส่ในเว็บแอพพลิเคชั่นเพื่อใช้เป็นคลังข้อสอบ รวมทั้งการลบ และการแก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ นอกจากนี้ยังมีรายงานผลการทดสอบของผู้สอบทั้งหมด เพื่อให้อาจารย์ และผู้เกี่ยวข้องนำผลการทดสอบไปใช้ในการประเมินผลการเรียนรู้ต่อไป 3) การจัดการสอบ เป็นส่วนที่ให้ผู้สอบคือ นักศึกษา ใช้เว็บแอพพลิเคชั่นนี้เพื่อทดสอบความสามารถทางความรู้วิชาชีพครู ซึ่งเมื่อผู้สอบทำการทดสอบเสร็จสิ้นจะรายงานผลการทดสอบแบบระบบเรียลไทม์ (Real-Time System) คือระบบที่สามารถรายงานผลสอบได้ทันที และตลอดเวลา

4. ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปมีดังนี้

4.1 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้เชี่ยวชาญ ปรากฏว่า เว็บแอพพลิเคชั่นในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด เมื่อพิจารณาเป็นรายด้านปรากฏว่ามี 4 ด้าน ที่มีความเหมาะสมอยู่ในระดับมากที่สุด ได้แก่ ด้านความสามารถของแอพพลิเคชั่น (Functional Requirement) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) และความสะดวกในการใช้งานเว็บแอพพลิเคชั่น(Usability) รองลงมาได้แก่ ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) มีความเหมาะสมอยู่ในระดับมาก

4.2 ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้ใช้งาน แสดงให้เห็นว่าโปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด เมื่อพิจารณาเป็นรายด้านปรากฏว่ามี 3 ด้าน ที่มีความเหมาะสมอยู่ในระดับมากที่สุด ได้แก่ ด้าน

การทำงานของเว็บแอพพลิเคชั่น (Functional) ด้านความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability) และด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual)

จากผลการประเมินความเหมาะสมของโปรแกรมทั้งผู้เชี่ยวชาญและผู้ใช้งาน ผู้วิจัยได้นำข้อเสนอแนะมาปรับปรุงแก้ไขเว็บแอพพลิเคชั่นให้มีความเหมาะสม และมีความรวดเร็วยิ่งขึ้นและเพื่อไม่ให้ระบบการทดสอบล่ม และรองรับผู้เข้าสอบพร้อมกันหลายคน ผู้วิจัยใช้ Cloud Server ข้อดีของ Cloud Server จะประกอบด้วย Server หลายๆ เครื่อง ช่วยกันทำงานเป็นระบบที่มีประสิทธิภาพสูง สามารถทำงานได้อย่างรวดเร็วไม่มีสะดุด มีความปลอดภัยในการจัดเก็บข้อมูลสูง มีความเสถียรสูง สามารถใช้งานได้อย่างไร้กังวล โดยไม่ต้องกลัว Network ล่ม มีความยืดหยุ่นสูงใช้งานง่าย สามารถขยายพื้นที่ในการจัดเก็บข้อมูลได้ไม่จำกัด และง่ายต่อการ Backup และสำรองข้อมูลจากผลการประเมินความเหมาะสมโปรแกรม แสดงให้เห็นว่า โปรแกรมมีความเหมาะสม สามารถตอบสนองความต้องการของผู้ใช้งานได้เป็นอย่างดี และเพียงพอที่จะนำไปใช้ทดสอบจริงได้

อภิปรายผล

การอภิปรายผลการวิจัย แบ่งได้ 4 ประเด็น ได้แก่ 1) การพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูโดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ 2) การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network 3) การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป และ 4) การประเมินคุณภาพของของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ดังนี้

1) การพัฒนาวิธีการจัดคลังข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครูโดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ

คลังข้อสอบในเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป เป็นข้อสอบแบบเลือกตอบ 5 ตัวเลือก ข้อสอบการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครู ที่ได้มาจากสภาคณบดีคณะครุศาสตร์มหาวิทยาลัยราชภัฏ (สครภ) ลักษณะข้อสอบคู่ขนานตามผังการออกข้อสอบที่ได้มีการวิเคราะห์ตามกรอบมาตรฐานวิชาชีพคุรุสภา และกรอบมาตรฐานคุณวุฒิอุดมศึกษาที่ผ่านการวิพากษ์และแก้ไขปรับปรุงโดยผู้ทรงคุณวุฒิแล้ว ได้ข้อสอบทั้งหมด 270 ข้อ แล้วสร้าง Q-matrix โดยผู้วิจัยจัดแยกข้อสอบออกเป็น 5 มิติ คือ 1) มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ 2) มิติการประกันคุณภาพการศึกษา จำนวน 50 ข้อ 3) มิติจิตวิทยาการศึกษา จำนวน 50 ข้อ 4) มิติการเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ และ 5) มิติหลักสูตรและการสอน จำนวน 50 ข้อ แล้วนำข้อสอบทั้ง 5 มิติ 270 ข้อ ให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อสอบกับมิติของข้อสอบ ได้ข้อสอบที่ผ่าน

เกณฑ์การพิจารณาโดยผู้เชี่ยวชาญซึ่งเป็นข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป (ศิริชัย พงษ์วิชัย, 2550) จำนวน 242 ข้อ จากนั้นนำข้อสอบมาวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโมเดลปกติสะสมแบบพหุมิติกับโอกาสการเดา โดยใช้โปรแกรม NOHARM 4.0 และกำหนดค่าพารามิเตอร์โอกาสการเดา ให้คงที่เท่ากับ 0.20 (ข้อสอบแบบเลือกตอบ 5 ตัวเลือก) แล้วคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ คือ 1) มีค่าอำนาจจำแนกแบบพหุมิติ (MDISC) ตั้งแต่ 0.00 ขึ้นไป และ 2) มีค่าความยากแบบพหุมิติ (MDIFF) อยู่ในช่วง -4.00 ถึง +4.00 เพื่อจัดทำเป็นคลังข้อสอบ (Item bank) จากการคัดเลือกพบว่ามีข้อสอบผ่านเกณฑ์การคัดเลือก จำนวน 242 ข้อ ประกอบด้วย 5 มิติ ดังนี้ 1) มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ 2) มิติการประกันคุณภาพการศึกษา จำนวน 45 ข้อ 3) มิติจิตวิทยาการศึกษา จำนวน 40 ข้อ 4) มิติการเปลี่ยนแปลงบริบทโลก จำนวน 50 ข้อ และ 5) มิติหลักสูตรและการสอน จำนวน 47 ข้อ

สำหรับในการวิจัยครั้งนี้ผู้วิจัยพัฒนาคลังข้อสอบจำนวน 242 ข้อ ซึ่งสอดคล้องกับที่ Weiss (1988) ได้เสนอกฎของจำนวนข้อสอบในคลังว่าควรมีอย่างน้อย 100 – 200 ข้อ ซึ่งเพียงพอที่จะใช้ในสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ และการกำหนดจำนวนข้อสอบที่เหมาะสมในคลังข้อสอบ แม้ที่ผ่านมาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติจะไม่มีงานวิจัยหรือทฤษฎีใดที่กำหนดจำนวนข้อสอบที่เหมาะสม แต่จากที่ผู้วิจัยได้ศึกษางานวิจัยของ สุกัญญา บุญศรี (2554); ทัศน์ศิรินทร์ สว่างบุญ (2554); สมประสงค์ เสนารัตน์ (2555); พัชราภรณ์ ไวกุณฐ์วิวรรธน์ (2561); ศักดิ์ชัย จันทะแสง (2561); รัชกฤช ธนพัฒนดล (2561); Linden (1999); Cheng (2009); Diao and Reckase, (2009) และ Veldkamp and Linden จำนวนข้อสอบในคลังข้อสอบทั้งแบบจำลองข้อมูลและเก็บข้อมูลจริงมีจำนวน 50-300 ข้อ ซึ่งคลังข้อสอบที่ผู้วิจัยพัฒนาขึ้นในงานวิจัยครั้งนี้ มีจำนวนสอดคล้องกับงานวิจัยที่ผ่านมาและเพียงพอที่จะใช้ในสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ

2) การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไป เป็นการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปของการทดสอบแบบปรับเหมาะแบบพหุมิติ โดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์เข้ากับทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) และหลักการการคัดเลือกข้อสอบมาพิจารณาร่วมกัน สามารถสรุปผลการพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ Artificial Neuron Network ได้ดังนี้

2.1 การคัดเลือกข้อสอบข้อถัดไป เพื่อหารูปแบบ Artificial Neuron Network ที่สามารถนำมาใช้ในการคัดเลือกข้อสอบข้อถัดไปได้ พบว่า Artificial Neuron Network แบบ

เพอเซปตรอนหลายชั้น (Multilayer Perceptron) ที่ใช้การเรียนรู้แบบมีผู้สอน (supervised learning) มีความเหมาะสมที่สุด ประกอบด้วยชั้น 3 ชั้นได้แก่ ชั้นข้อมูลป้อนเข้า (n_{input}) เท่ากับ 25 โหนด และจำนวนโหนดในชั้นซ่อน $(n_{hidden1})$จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden2})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden3})$จำนวน 10 โหนด และจำนวนโหนดชั้นข้อมูลส่งออก (n_{output}) จำนวน 2 โหนด (รูปแบบข้อมูล 25-10-10-10-2) ทำการออกแบบโมเดลโดยใช้เทคนิคดาต้าไมน์นิ่ง (Data mining) หรือการทำเหมืองข้อมูล ในการจำแนกประเภทข้อมูล (Classification) การคัดเลือกข้อสอบ โดยใช้โปรแกรม RapidMiner Studio Educational 10.1 วิเคราะห์โดยใช้อัลกอริธึม Neural Net เพื่อเป็นการวิเคราะห์หาประสิทธิภาพความแม่นยำในการคัดเลือกข้อสอบข้อถัดไปให้มีความถูกต้องมากที่สุด โดยผู้วิจัยได้กำหนดค่าพารามิเตอร์ในกระบวนการของ Artificial Neuron Network ประกอบด้วย อัตราการเรียนรู้เท่ากับ 0.95 โมเมนตัม (Momentum) เท่ากับ 0.05 ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ

2.2 ทำการทดสอบประสิทธิภาพโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธี Artificial Neuron Network ก่อนนำไปใช้จริงก่อนโดยการวัดความถูกต้องของอัลกอริธึม Neural Net ด้วยวิธีการ 10-fold cross-validation ซึ่งจะแบ่งข้อมูลออกเป็น 10 ชุดเท่า ๆ กันแล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ (Test set) 20% ส่วนที่เหลือ 9 ชุด นำมาใช้เป็นกลุ่มเรียนรู้ (Training Set) 80% แล้วทำการ วนทำเป็นจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อยๆ จนครบในการวัดประสิทธิภาพของการจำแนกประเภทข้อมูลพบว่า จากการประเมินประสิทธิภาพด้วยวิธีการ 10-fold cross-validation ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบข้อถัดไป ได้ค่าความระลึก (Recall) คิดเป็นร้อยละ 100 สูงที่สุด รองลงมาคือ ค่าความถ่วงดุล (F-measure) คิดเป็นร้อยละ 94.70, ค่าความถูกต้อง (Accuracy) คิดเป็นร้อยละ 90 และการวัดค่าความแม่นยำ (Precision) คิดเป็นร้อยละ 90

จากการทดลองโมเดล Artificial Neuron Network ค่าอัตราการเรียนรู้ค่าโมเมนตัม สรุปได้ว่า วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียมโมเดล 25-10-10-10-2 มีค่าความระลึก (Recall) คิดเป็นร้อยละ 100 สูงที่สุด แสดงให้เห็นว่าโมเดล Artificial Neuron Network มีความถูกต้องของโมเดลน่าเชื่อถือมากที่สุดและสามารถนำโมเดลไปใช้งานจริงกับการคัดเลือกข้อสอบข้อถัดไปของสำหรับใช้ทดสอบกับเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติได้

3) การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป

3.1 การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติในงานวิจัยนี้ เป็นการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป มีลักษณะเป็น Web

Application สามารถใช้ผ่านอุปกรณ์ทั่วไปที่เชื่อมต่อกับอินเทอร์เน็ตได้ เช่น คอมพิวเตอร์ โน๊ตบุ๊คหรือสมาร์ทโฟน การพัฒนาเว็บแอพพลิเคชั่นนี่ใช้ภาษา PHP ในการพัฒนาโปรแกรมใช้ MySQL เป็นฐานข้อมูล และใช้ภาษา SQL เป็นคำสั่ง ในการเชื่อมโยงข้อมูลในฐานข้อมูล ซึ่งผู้วิจัยพัฒนาเว็บแอพพลิเคชั่นตามหลักการของวงจรการพัฒนาซอฟต์แวร์ (Software Development Life Cycle: SDLC) (Elliott, 2004) เนื่องจาก PHP สามารถเรียนรู้ได้ง่าย สามารถเชื่อมต่อกับฐานข้อมูลชนิดต่างๆ ได้ เช่น Access, FoxPro ไปจนถึง SQL Server หรือ Oracle ได้ และสามารถเชื่อมต่อกับ Object อื่นๆ ที่มีอยู่แล้วได้ เช่น ActiveX, COM และ DCOM แต่สิ่งที่จำเป็นสำหรับการเริ่มต้นใช้ PHP คือ ภาษาสคริปต์ของเว็บ ซึ่งอาจจะเป็น HTML VBScript หรือ JavaScript สำหรับใช้ในการจัดการกับ Object และ Method ต่าง ๆ ของ PHP แต่ข้อจำกัดของการใช้ PHP คือ เป็นการเพิ่มภาระให้กับเซอร์ฟเวอร์และเครือข่ายอินเทอร์เน็ต เนื่องจากโค้ดของ PHP จะถูกประมวลผลที่ฝั่งเซอร์ฟเวอร์ จึงทำให้เวลาผู้สอบพร้อมกันหลายคนระบบอาจจะล่มได้ง่ายและเพื่อไม่ให้ระบบการทดสอบล่ม ผู้วิจัยจึงใช้ Cloud Server ข้อดีของ Cloud Server จะประกอบด้วย Server หลายๆ เครื่อง ช่วยกันทำงานพร้อมกัน เป็นระบบที่มีประสิทธิภาพสูง สามารถทำงานได้อย่างรวดเร็ว ไม่มีสะดุด มีความปลอดภัยในการจัดเก็บข้อมูลสูง มีความเสถียรสูง สามารถใช้งานได้อย่างไร้กังวล โดยไม่ต้องกลัว Network ล่ม มีความยืดหยุ่นสูง ใช้งานง่าย สามารถขยายพื้นที่ในการจัดเก็บข้อมูลได้ไม่จำกัด และง่ายต่อการ Backup และสำรองข้อมูล โดบเว็บแอพพลิเคชั่นการทดสอบนี้แบ่งออกเป็น 3 ส่วน คือ 1) ข้อมูลทั่วไปเกี่ยวกับโปรแกรม 2) การจัดการของผู้ดูแลระบบและ 3) การจัดการสอบ เมื่อผู้สอบทดสอบเสร็จสิ้น จะรายงานผลการทดสอบให้ผู้สอบทราบทันที

3.2 เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่พัฒนาขึ้นนี้แต่ละขั้นตอนมีจุดเด่นดังนี้ 1.การสร้างคลังข้อสอบ เว็บแอพพลิเคชั่นนี้รองรับข้อสอบที่มีลักษณะการตรวจให้แบบตรวจให้คะแนน 2 ค่า (Dichotomous Item Response Theory) คือ ถูก-ผิด หรือ 0-1 และแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Item Response Theory) คือ คำตอบ 4 หรือ 5 รายการได้และรองรับโมเดลเอกมิติ (unidimensional model) และโมเดลพหุมิติ (multidimensional model) ได้ทุกค่าไม่เกิน 4 พารามิเตอร์ และออกแบบคลังข้อสอบไว้เพื่อรองรับมิติได้ไม่เกิน 10-DIMENSIONAL 2. Algorithms สำหรับคัดเลือกข้อสอบข้อถัดไป สามารถใช้ Algorithm ได้ทั้ง Artificial Neural Networks และ Decision Tree

3.3 การตรวจสอบและประเมินความถูกต้องเหมาะสมของเว็บแอพพลิเคชั่นการทดลองใช้งานเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติในการวิจัยนี้ มีการตรวจสอบ 3 ขั้นตอน คือ 1) ผู้วิจัยตรวจสอบโปรแกรมด้วยตนเอง เพื่อหาจุดบกพร่องเบื้องต้น 2) การตรวจสอบความถูกต้องเหมาะสมของโปรแกรมจากผู้เชี่ยวชาญจำนวน 5 ท่าน โดยผู้เชี่ยวชาญมีความรู้ด้านการวัดผลการศึกษาหรือการพัฒนาโปรแกรมคอมพิวเตอร์ไม่น้อยกว่า 10 ปี และ 3) การตรวจสอบเว็บ

แอพพลิเคชั่นโดยผู้ทดลองใช้งานเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ เพื่อหาจุดบกพร่องในการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบ ซึ่งผู้วิจัยได้ปรับปรุง แก้ไขโปรแกรมตามข้อแนะนำของผู้เชี่ยวชาญและผู้ทดลองใช้งานในทุกประเด็น

4) การประเมินคุณภาพของของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ดังนี้

การประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป ผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้เชี่ยวชาญ ปรากฏว่า เว็บแอพพลิเคชั่นในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด คือด้านความสามารถของแอพพลิเคชั่น (Functional Requirement) ด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) และด้านความสะดวกในการใช้งานเว็บแอพพลิเคชั่น(Usability) ส่วนด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) มีความเหมาะสมอยู่ในระดับมาก และผลการประเมินความเหมาะสมของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยผู้ใช้งาน โปรแกรมในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด เมื่อพิจารณาเป็นรายด้านปรากฏว่าทุกด้าน คือด้านการทำงานของเว็บแอพพลิเคชั่น (Functional) ด้านความสะดวกในการใช้งานเว็บแอพพลิเคชั่น (Usability) และด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) มีความเหมาะสมอยู่ในระดับมากที่สุด แสดงให้เห็นว่าเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติสามารถนำไปใช้ในการวิเคราะห์ความสามารถของผู้ทดสอบได้เป็นอย่างดี ประกอบกับเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิตินี้ถูกสร้างขึ้นมาอย่างมีหลักการ แนวคิด และทฤษฎีมารองรับ รวมทั้งใช้เทคนิควิธีการที่หลากหลายโดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์เข้ากับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) เป็นเทคนิคการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แนวใหม่ที่ทันสมัยทำให้มีประสิทธิภาพยิ่งขึ้นในการคัดเลือกข้อสอบ และมีความแม่นยำในการประมาณค่าความสามารถของผู้สอบ และผ่านการพิจารณาตรวจสอบความถูกต้องเหมาะสมจากผู้เชี่ยวชาญหลายๆ ท่าน รวมทั้งผ่านการทดลองใช้งานหลายๆ รอบ ก่อนที่จะนำมาประเมินคุณภาพ จึงทำให้เว็บแอพพลิเคชั่นสามารถตอบสนองความต้องการของผู้ใช้งานได้เป็นอย่างดีที่สำคัญผู้วิจัยใช้ Cloud Server เป็น Server หลายๆ เครื่องช่วยกันทำงาน เพื่อให้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติมีประสิทธิภาพสูง สามารถทำงานได้อย่างรวดเร็ว ไม่มีสะดุดมีความปลอดภัยในการจัดเก็บข้อมูลสูง มีความเสถียรสูง สามารถใช้งานได้อย่างไร้กังวล โดยไม่ต้องกลัว Network ล่ม มีความยืดหยุ่นสูง ใช้งานง่าย สามารถขยายพื้นที่ในการจัดเก็บข้อมูลได้ไม่จำกัด และง่ายต่อการ Backup และสำรองข้อมูล จากผลการ

ประเมินความเหมาะสมโปรแกรม แสดงให้เห็นว่าโปรแกรมมีความเหมาะสม สามารถตอบสนองความต้องการของผู้ใช้งานได้เป็นอย่างดี และเพียงพอที่จะนำไปใช้ทดสอบจริงได้

ข้อเสนอแนะ

1. ข้อเสนอแนะการนำผลการวิจัยไปใช้

1.1 นักศึกษาหลักสูตรครุศาสตรบัณฑิต สามารถเข้าใช้งานเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ใช้วัดความรู้ทางวิชาชีพครู ขณะที่ศึกษาในชั้นปีที่ 3 หรือ 4 เพื่อประเมินสมรรถนะความรู้วิชาชีพของตนเองว่าระดับความสามารถอยู่ในระดับใด ควรพัฒนาระดับความสามารถ Theta ของนักศึกษาในแต่ละมิติว่าอยู่ในระดับผ่าน/ ไม่ผ่าน นอกจากนี้ในรายงานผลของความสามารถในแต่ละมิตินั้นจะรายงานผลการทดสอบทันทีว่าทำข้อสอบในแต่ละมิติได้กี่ข้อจำนวนข้อสอบทั้งหมดที่ทำการทดสอบในครั้งนี้ โดยข้อมูลส่วนนี้สามารถนำผลการทดสอบที่ได้ไปพัฒนาหรือปรับปรุงแก้ไข เพื่อให้มีความรู้ทางวิชาชีพครูในด้านนั้นเพิ่มมากขึ้นเพียงใดในแต่ละปีการศึกษา

1.2 สถาบันอุดมศึกษา ซึ่งเป็นผู้ดูแลหลักสูตรครุศาสตร์/ ศึกษาศาสตร์ สามารถนำเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป นี้ไปใช้สำหรับประเมินสมรรถนะความรู้วิชาชีพครูกับนักศึกษาชั้นปีที่ 3 ก่อนออกไปสังเกตุการสอนในสถานศึกษา และนักศึกษาชั้นปีที่ 4-5 ก่อนจบการศึกษาในหลักสูตร อาจารย์หรือผู้สอนควรส่งเสริมนักศึกษาในหลักสูตรในเรื่องใดบ้าง โดยสามารถพัฒนาส่งเสริมนักศึกษาได้ทั้งแบบรายบุคคลหรือเป็นกลุ่มตามความสามารถของนักศึกษา เพื่อให้กระบวนการจัดการเรียนการสอนของหลักสูตรประสบความสำเร็จตามวัตถุประสงค์ของหลักสูตรที่เขียนไว้ นอกจากนี้หากผลการทดสอบพบว่านักศึกษาในหลักสูตรส่วนใหญ่มีผลการประเมินที่ไม่ผ่านในมิติใด อาจารย์ก็สามารถนำผลการประเมินดังกล่าวไปเป็นสารสนเทศเพื่อใช้ในการพัฒนาปรับปรุงผู้เรียนในหลักสูตรได้

1.3 ผู้ที่สนใจ web application สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติสามารถนำวิธีการประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป ที่ออกแบบไว้ไปประยุกต์ในการคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบรายวิชาต่างๆ ได้

1.4 ครู อาจารย์ หรือผู้ที่เกี่ยวข้องในการจัดการสอบสามารถนำ web application สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ที่ได้จากการวิจัยนี้ ไปใช้วัดความสามารถทางความรู้วิชาชีพครูในระดับปริญญาตรี หรือนำไปประยุกต์กับการวัดผลการศึกษาในรายวิชาอื่นได้ เนื่องจากโปรแกรมได้ออกแบบรองรับการเพิ่มข้อสอบ และเพิ่มจำนวนมิติข้อสอบเพื่อทดสอบรายวิชาอื่นๆ ไว้แล้ว

1.5 web application สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไป ที่พัฒนาขึ้นนี้อยู่ในรูปแบบของ web

application สามารถเข้าถึงได้จาก website: http://examadaptive.com/index.php ซึ่งผู้สอบหรือผู้ที่สนใจสามารถเข้ามาทดสอบได้ทุกที่และทุกเวลา แต่จะต้องมีการเชื่อมต่ออินเตอร์เน็ตทั้งนี้เพื่อประโยชน์สำหรับผู้เรียนในการทำความคุ้นเคยกับข้อสอบและทบทวนการทำข้อสอบ

2. ข้อเสนอแนะเพื่อการวิจัยครั้งต่อไป

2.1 ควรนำเทคนิคต่างๆ ทางด้านปัญญาประดิษฐ์มาประยุกต์กับการทดสอบแบบปรับเหมาะแบบพหุมิติในขั้นตอนอื่นๆ เช่น การประยุกต์ Artificial Neuron Network ในการจัดข้อสอบในคลังข้อสอบใช้เทคนิคเหมืองข้อมูล (Data Mining) มาจัดการขนาดของคลังข้อสอบขนาดใหญ่มีผลต่อประสิทธิภาพในการประมาณค่าความสามารถ เป็นต้น

2.2 ควรทำการวิจัยเกี่ยวกับวิธี Artificial Neuron Network ในการคัดเลือกข้อสอบข้อเริ่มต้น (First item selection) การประมาณความสามารถ (Ability Estimation/Scoring algorithm) และการหยุดการทดสอบ (Stopping rule) ที่เหมาะสมในการทดสอบแบบปรับเหมาะแบบพหุมิติ

2.3 การพัฒนา web application สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้ Artificial Neuron Network ในการคัดเลือกข้อสอบข้อถัดไปสามารถนำเทคนิคทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) ประเภทการเรียนรู้เชิงลึก (Deep Learning) มาเรียนรู้แบบ Supervised Learning, Unsupervised Learning และ Reinforcement Learning ในออกแบบประยุกต์ในการวิจัยได้อีก

บรรณานุกรม

บรรณานุกรม

กานต์, (2018, Oct 5). “สรุป Deep Learning by Dr.Kan — Part 1.”[บทความออนไลน์]. สืบค้นจาก https://medium.com/odds-team/

ทัศน์ศิรินทร์ สว่างบุญ. (2554). *การเปรียบเทียบวิธีการคัดเลือกข้อสอบขั้นแรกและลำดับข้อสอบที่มีต่อคุณภาพของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์.* (วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย

ธวัชชัย ศรีสเทพ. (2544). *คัมภีร์ Web design :คู่มือออกแบบเว็บไซต์ฉบับมืออาชีพ*. กรุงเทพฯ: โปรวิชั่น.

ดวงพร เกี๋ยงคำ. (2551). *คู่มือสร้างเว็บไซต์ด้วยตนเอง*. กรุงเทพฯ: โปรวิชั่น.

บุญธรรม กิจปรีดาบริสุทธิ์. (2553). *สถิติวิเคราะห์เพื่อการวิจัย.* กรุงเทพฯ: เรือนแก้วการพิมพ์.

ผจงจิต อินทสุวรรณ. (2545). *ทฤษฎีการตอบข้อคำถาม*. กรุงเทพฯ: สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

พงษ์พิชิต พรหมสิทธิ์. (2550). *การศึกษาประสิทธิภาพการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ*. (วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต). มหาวิทยาลัยขอนแก่น

พัชรี จันทร์เพ็ง. (2550). *การเปรียบเทียบคุณภาพของวิธีการเชื่อมโยงคะแนนตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ภายใต้การหมุนแกนโครงสร้างเชิงมิติและระดับความสัมพันธ์ที่แตกต่างกัน*. (วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย

บูรพา วิถีปัญญา. (2564). การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้โครงข่ายประสาทเทียมและมีการควบคุมการใช้ข้อสอบสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้วิธีการจำลองข้อมูล. *วารสารครุพิบูล (มนุษยศาสตร์และสังคมศาสตร์). 8*(1). 37-48.

ประพล เปรมทองสุข, เสรี ชัดแช้ม และปิยะทิพย์ ประดุจพรม. (2560). *การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้เกณฑ์ของเฮอร์วิคซ์และการควบคุมการใช้ข้อสอบ สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์*. *Veridian E-Journal*, *4*(6), 32-50.

มัสกานต์ เสน่หา. 2559. “*การทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพ.*” กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย

รัชกฤช ธนพัฒนดล. (2561). *การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีระบบอาณานิคมมดสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.* (วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต). วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา. มหาวิทยาลัยบูรพา

ศักดิ์ชัย จันทะแสง. (2561). *การพัฒนาวิธีการจัดคลังข้อสอบตามค่าอำนาจจำแนกของข้อสอบแบบเป็นช่วงของแต่ละระดับชั้นร่วมกับการจัดสมดุลเนื้อหา.* (วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต). วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา. มหาวิทยาลัยบูรพา

ศจีมาจ ขวัญเมือง. (2548). *ปัจจัยที่ส่งผลต่อผลิตภาพการวิจัยของอาจารย์มหาวิทยาลัยของรัฐ: การวิเคราะห์ลิสเรลและเครือข่ายใยประสาท.* กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการตอบสนองแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

โสพส สุขานนท์สวัสดิ์, เสรี ชัดแช้ม และกฤษณะ ชินสาร. (2556). การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้ทฤษฎีการตัดสินใจ ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์. *(วารสารวิทยาการวิจัยและวิทยาการปัญญา)*, 10(2), 71-85.

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Martin, K., (2012). *An Introduction to Management Science: Quantitative Approaches to Decision Making* (13th ed.). Mason, OH: South-Western Cengage Learning.

Antal, M., Eros, L., & Imre, A. (2010). Computerized adaptive testing: implementation issues. (*Informatica)*, *2*(2), 168-183.

Ader, H. J. and Bramsen, I. (1998) *"Representation of a structural equation model as a neural Network"* (Computer Modeling of social processes). London

Bobee, B. (2000) "Daily reservoir inflow forecasting using artificial neural networks with stopped training approach". (*Journal of Hydrology).* (230). 244-257.

Baker, F. B. (2001). *The Basics of Item Response Theory* (2nd ed.). USA: ERIC Clearinghouse on Assessment and Evaluation.

Barrada, J. R., Mazuela, P., & Olea, J. (2006). *Maximum information stratification method for controlling item exposure in computerized adaptive testing*. Psicothema, 18(1), 156-159.

Belov, D. I., Armstrong, R. D., & Weissman, A. (2008). A Monte Carlo Approach for Adaptive Testing with Content Constraints. (*Applied Psychological Measurement)*, *32*(6), 431-446.

Chaimongkol, N., Pasiphol, S., & Kanjanawasee, S. (2016). Computerized Adaptive Testing with Reflective Feedback: A Conceptual Framework. (*Procedia - Social and Behavioral Sciences)*, *217*, 806-812.

Chalhoub-Deville, M., Alcaya, C., & Lozier, V. M. (1996). *An Operational Framework for Constructing a Computer-Adaptive Test of L2 Reading Ability: Theoretical and Practical Issues*. Minneapolis, MN: University of Minnesota.

Chang, H. H., Qian, J., & Ying, Z. (2001). *a-Stratified Multistage Computerized Adaptive Testing with b Blocking.* (Applied Psychological Measurement), *25*(4), 333-341.

Chang, H. H., & van der Linden, W. J. (2003). *Optimal Stratification of Item Pools in a-Stratified Computerized Adaptive Testion*. (Applied Psychological Measurement)*, 27*(4), 262 – 274.

Chang, H. H., & Ying, Z. (1996). *A global information approach to computerized adaptive testing.* (Applied Psychological Measurement), *20*, 213–229.

Chang, H. H., & Ying, Z. (1999). *A-stratified multistage computerized adaptive testing.* (Applied Psychological Measurement), *23*(3), 211-222.

Chang, S. W., & Harris, D. J. (2002). *Redeveloping the exposure control parameters of CAT items when a pool is modified. Paper presented at the annual meeting of the American Educational Research Association*. New Orleans, LA: IACAT.

Chang, S. W., & Ansley, T. N. (2003). A Comparative Study of Item Exposure Control Methods in Computerized Adaptive Testing. *(Journal of Educational Measurement)*, *40*(1), 71-103.

Chen, S. Y., & Ankenmann, R. D. (2004). Effects of practical constraints on item selection rules at the early stages of computerized adaptive testing. *(Journal of Educational Measurement), 41*(2), 149-174.

Chen, S. Y., & Doong, S. H. (2003). Predicting Item Exposure Parameters in Computerized Adaptive Testing. *Paper presented at the 2003 annual meeting of the American Educational Research Association*. Chicago, IL: IACAT.

Cheng, Y., & Chang, H. H. (2009). *The maximum priority index method for severely constrained item selection in computerized adaptive testing*. Br J Math Stat Psychol, *62*(2), 369-383.

Cheng. Y., Chang, H. H., Douglas, J., & Guo, F. (2009). *Constraint-Weighted a-Stratification for Computerized Adaptive Testing With Nonstatistical Constraints*. (Educational and Psychological Measurement), *69*(1). 35-49.

Choi, S. W. (2009). Firestar: *Computerized adaptive testing (CAT) simulation program for polytomous IRT Models (Computer software).* (Applied Psychological Measurement), *33*, 644–645.

Costa, D. R., Karino, C. A., Moura, F. A. S., & Andrade, D. F. (2009). *A comparison of three methods of item selection for computerized adaptive testing.*

In, D. J., Weiss (Ed.) *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*. From www.psych.umn.edu/psylabs/CATCentral/

Davey, T., & Parshall, C. G. (1995). *New algorithms for item selection and exposure control with computerized adaptive testing. In the annual meeting of the American*. (Educational Research Association). San Francisco, CA: IACAT.

De Ayala, R. J. (2009). *The theory and practice of item response theory.* London: The Guilford Press.

Deng, H., Ansley, T., & Chang, H. H. (2010). Stratified and Maximum Information Item Selection Procedures in Computer Adaptive Testing. *(Journal of Education Measurement), 47*(2), 202-226.

Barrada, J. R., Abad, F. J., & Olea, J. (2011). Varying the valuating function and the presentable bank in computerized adaptive testing. *(The Spanish Journal of Psychology), 14*(1), 500-508.

Chang, H. H., & Ying, Z. (2008). *A-stratified multistage computerized adaptive testing*. (Applied Psychological Measurement), *23*(3), 211-222.

Chen, S. Y., & Ankenman, R. D. (2004). Effects of practical constraints on item selection rules at the early stages of computerized adaptive testing. *(Journal of Educational Measurement)*, *41*(2), 149-174.

Chen, S. Y., Ankenmann, R. D., & Spray, J. A. (2009). The relationship between item exposure and test overlap in computerized adaptive testing. *(Journal of Educational Measurement), 45*(2), 129-145.

Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using item response theory. (Computers & Education), *44*(3), 237-255.

Delurgio, S. A. (1998). *Forecasting Principles and Applications*. Singapore : McGraw-Hill.

De Matos, G. (1994). *Neural Networks for forcasting foreign exchange rates*. [Master's Science thesis, University of Manitoba]. Canada

Fausett, L. (1994). *Foundaments of Neural Networks : Architectures Algorithms and Applications*. London: Prentice-Hall International.

Garson, G. D. (1998). *Neural Networks an Introductory Guide for Social Scientists*. London: SAGE.

Gonzalez, Julie. M. B. & DesJardins, S. L. (2002). "Artificial neural networks : a new approach to predicting application behavior" *Research in Higher Education. 43*(2), p. 235-258.

Han, K. T. (2007). *WinGen: Windows software that generates IRT parameters and item responses.* (Applied Psychological Measurement), *31*(5), 457-459.

Hand, D. J. (1993). *Artificial intelligence frontiers in statistics.* London: Chapman & Hall.

Hasmy, A. (2014). Compare unidimensional & multidimensional Rasch model for test with multidimensional construct and items local dependence. *(Journal of Education and Learning)*, *8*(3), 187-194.

Hill, T. et. al. (1996). *"Neural network models for time series forecasts"* (Management Science). *42*(7) p. 1082-1092.

Jain, A. K. and Mao, J. (1996). *"Artificial neural networks : a tutorial"* Computer. *29*(3) p. 31-43.

J. E. Dayhoff, (1990). *Regularity properties in pulse transmission networks*. (National Biomedical Research Foundation), 3, p. 621-626.

Koedsri, A., Lawthong, N., & Ngudgratoke, S. (2014). *Efficiency of Item Selection Method in Variable-length Computerized Adaptive Testing for the Testlet Response Model: Constraint-weighted A-stratification Method*. (Procedia-Social and Behavioral Sciences), *16*(3), 1890-1895.

Liebrand, B. G., Nowak, A. and Hegselmann, R. (1998). *Computer Modeling of Social Processes*. London: SAGE.

Limin, F. (1994). *Neural Networks in Computer Intelligence*. New York: Mcgraw-Hill.

Parker, D. (1982). *Learning-logic*. Center for Computational Research in Economics and Management Science. MIT.

Patterson, D. W. (1996). *Artificial Neural Networks:Theory and Applications*. Singapore : Prentice Hall.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). *Parallel Distributed Processing*. MIT.

Thompson, N. A., & Weiss, D. J. (2011). *A framework for the development of computerized adaptive tests.* (Practical Assessment, Research & Evaluation), *16*(1), 1-9.

Weiss, D. J. (2011). Better Data from Better Measurements Using Computerized Adaptive Testing. *(Journal of Methods and Measurement in the Social Sciences)*, *2*(1), 1-27.

Webos, P. (1974). *Beyond regression: new tools for prediction and analysis in the Behavioral science*. (Doctor of Philosophy). dissertation United States of America: Harvard University.

Wilson, R. L. and Hardgrave, B. C. (1995). *"Predicting graduate student success in an MBA program : regression versus classification"* (Educational and Psychological Measurement). *55*(2) p. 186-195.

Minsky, Marvin, Seymour A. Papert, and Léon Bottou. (2017). *Perceptrons.* (Reissue edition). Cambridge, MA: MIT Press.

Yung-Chin, Y., Ho, R. G., Wen-Wei, L., & Li-Ju, C. (2012). Reducing the impact of inappropriate items on reviewable computerized adaptive testing. *(Journal of Educational Technology & Society), 15*(2), 231-242.

Zou, Jinming, Yi Han, and Sung-Sau So. (2009). "Overview of Artificial Neural Networks." pp. 14–22. *in Artificial Neural Networks: Methods and Applications, Methods in Molecular BiologyTM*, edited by D. J. Livingstone. Totowa, NJ: Humana

ภาคผนวก

ภาคผนวก ก
ผลการวิเคราะห์คุณภาพเครื่องมือ

ผลการประมาณค่าพารามิเตอร์ข้อสอบข้อสอบประเมินสมรรถนะความรู้ทางวิชาชีพครู
จำนวน 242 ข้อ ตามแนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ
ด้วยโปรแกรม NOHARM 4.0

กำหนดให้

a_1 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 1 (หลักสูตรและการสอน)

a_2 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 2 (จิตวิทยา)

a_3 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 3 (การวัดผลและวิจัยการเรียนรู้)

a_4 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 4 (การประกันคุณภาพการศึกษา)

a_5 แทน ค่าอำนาจจำแนกของข้อสอบในมิติที่ 5 (การเปลี่ยนแปลงบริบทโลก)

c แทน ค่าการเดาของข้อสอบ

d (threshold) แทน ค่าจุดตัดความยากแบบพหุมิติของข้อสอบ (Easiness Intercept)
หรือการประมาณค่าพารามิเตอร์ความยากในแต่ละ threshold

MDISC แทน ค่าอำนาจจำแนกแบบพหุมิติ

MDIFF แทน ค่าความยากแบบพหุมิติของข้อสอบ

ตารางที่ 28 ค่าพารามิเตอร์ของข้อสอบความรู้ทางวิชาชีพครู ที่ผ่านการคัดเลือกเข้าคลังข้อสอบ

ITEMS	มิติของข้อสอบ					d threshold	MDISC	MDIFF
	a_1	a_2	a_3	a_4	a_5			
1	0.599	0.00	0.00	0.00	0.00	0.038	0.599	0.063
2	0.785	0.00	0.00	0.00	0.00	-0.935	0.785	-1.191
3	0.476	0.00	0.00	0.00	0.00	0.739	0.476	1.553
4	0.619	0.00	0.00	0.00	0.00	0.338	0.619	0.546
5	0.827	0.00	0.00	0.00	0.00	-0.598	0.827	-0.723
6	0.619	0.00	0.00	0.00	0.00	-0.406	0.619	-0.656
7	1.000	0.00	0.00	0.00	0.00	-1.138	1.000	-1.138
8	0.587	0.00	0.00	0.00	0.00	0.935	0.587	1.593
9	0.553	0.00	0.00	0.00	0.00	-0.113	0.553	-0.204
10	1.000	0.00	0.00	0.00	0.00	-1.388	1.000	-1.388
11	0.768	0.00	0.00	0.00	0.00	-0.605	0.768	-0.788
12	1.000	0.00	0.00	0.00	0.00	-1.227	1.000	-1.227
13	1.000	0.00	0.00	0.00	0.00	-1.036	1.000	-1.036
14	0.693	0.00	0.00	0.00	0.00	-0.319	0.693	-0.46
15	0.870	0.00	0.00	0.00	0.00	-1.232	0.870	-1.416
16	0.448	0.00	0.00	0.00	0.00	-0.345	0.448	-0.77
17	0.708	0.00	0.00	0.00	0.00	-0.468	0.708	-0.661
18	0.535	0.00	0.00	0.00	0.00	-0.126	0.535	-0.236
19	0.658	0.00	0.00	0.00	0.00	-0.305	0.658	-0.464
20	1.000	0.00	0.00	0.00	0.00	-1.356	1.000	-1.356
21	0.751	0.00	0.00	0.00	0.00	-1.103	0.751	-1.469
22	0.511	0.00	0.00	0.00	0.00	0.461	0.511	0.902
23	0.764	0.00	0.00	0.00	0.00	-1.248	0.764	-1.634
24	0.626	0.00	0.00	0.00	0.00	-0.260	0.626	-0.415
25	0.547	0.00	0.00	0.00	0.00	0.221	0.547	0.404

	มิติของข้อสอบ							
26	0.682	0.00	0.00	0.00	0.00	-0.365	0.682	-0.535
27	1.000	0.00	0.00	0.00	0.00	-2.005	1.000	-2.005
28	0.555	0.00	0.00	0.00	0.00	0.365	0.555	0.658
29	0.870	0.00	0.00	0.00	0.00	-0.358	0.870	-0.411
30	1.000	0.00	0.00	0.00	0.00	-1.372	1.000	-1.372
31	0.727	0.00	0.00	0.00	0.00	0.365	0.727	0.502
32	0.935	0.00	0.00	0.00	0.00	-1.005	0.935	-1.075
33	0.964	0.00	0.00	0.00	0.00	-1.242	0.964	-1.288
34	0.664	0.00	0.00	0.00	0.00	0.151	0.664	0.227
35	0.696	0.00	0.00	0.00	0.00	-0.138	0.696	-0.198
36	0.762	0.00	0.00	0.00	0.00	-0.088	0.762	-0.115
37	0.690	0.00	0.00	0.00	0.00	-0.240	0.690	-0.348
38	0.592	0.00	0.00	0.00	0.00	0.372	0.592	0.628
39	0.783	0.00	0.00	0.00	0.00	-0.202	0.783	-0.258
40	0.746	0.00	0.00	0.00	0.00	-0.221	0.746	-0.296
41	0.599	0.00	0.00	0.00	0.00	0.038	0.599	0.063
42	0.785	0.00	0.00	0.00	0.00	-0.935	0.785	-1.191
43	0.476	0.00	0.00	0.00	0.00	0.739	0.476	1.553
44	0.619	0.00	0.00	0.00	0.00	0.338	0.619	0.546
45	0.827	0.00	0.00	0.00	0.00	-0.598	0.827	-0.723
46	0.619	0.00	0.00	0.00	0.00	-0.406	0.619	-0.656
47	1.000	0.00	0.00	0.00	0.00	-1.138	1.000	-1.138
48	0.532	0.00	0.00	0.00	0.00	1.026	0.532	1.929
49	0.564	0.00	0.00	0.00	0.00	-0.189	0.564	-0.335
50	0.599	0.00	0.00	0.00	0.00	-1.388	0.599	-1.388
51	0.00	1.000	0.00	0.00	0.00	-0.698	1.000	-0.897
52	0.00	0.778	0.00	0.00	0.00	-1.227	0.778	-1.227
53	0.00	1.000	0.00	0.00	0.00	-1.036	1.000	-1.036
54	0.00	1.000	0.00	0.00	0.00	-0.319	1.000	-0.48

	มิติของข้อสอบ							
55	0.00	0.664	0.00	0.00	0.00	-1.231	0.664	-1.624
56	0.00	0.758	0.00	0.00	0.00	-0.345	0.758	-0.802
57	0.00	0.430	0.00	0.00	0.00	-0.468	0.430	-0.687
58	0.00	0.681	0.00	0.00	0.00	-0.126	0.681	-0.245
59	0.00	0.514	0.00	0.00	0.00	-0.305	0.514	-0.483
60	0.00	0.631	0.00	0.00	0.00	-1.356	0.631	-1.356
61	0.00	1.000	0.00	0.00	0.00	-1.103	1.000	-1.528
62	0.00	0.722	0.00	0.00	0.00	0.461	0.722	0.939
63	0.00	0.491	0.00	0.00	0.00	-1.236	0.491	-1.628
64	0.00	0.759	0.00	0.00	0.00	-0.260	0.759	-0.433
65	0.00	0.600	0.00	0.00	0.00	0.221	0.600	0.422
66	0.00	0.524	0.00	0.00	0.00	-0.365	0.524	-0.558
67	0.00	0.654	0.00	0.00	0.00	-2.005	0.654	-2.005
68	0.00	1.000	0.00	0.00	0.00	0.365	1.000	0.686
69	0.00	0.532	0.00	0.00	0.00	-0.358	0.532	-0.429
70	0.00	0.834	0.00	0.00	0.00	-1.405	0.834	-1.405
71	0.00	1.000	0.00	0.00	0.00	0.352	1.000	0.522
72	0.00	0.674	0.00	0.00	0.00	-1.005	0.674	-1.119
73	0.00	0.898	0.00	0.00	0.00	-1.235	0.898	-1.351
74	0.00	0.914	0.00	0.00	0.00	0.151	0.914	0.237
75	0.00	0.637	0.00	0.00	0.00	-0.138	0.637	-0.207
76	0.00	0.667	0.00	0.00	0.00	-0.088	0.667	-0.121
77	0.00	0.730	0.00	0.00	0.00	-0.240	0.730	-0.364
78	0.00	0.660	0.00	0.00	0.00	0.372	0.660	0.656
79	0.00	0.567	0.00	0.00	0.00	-0.202	0.567	-0.269
80	0.00	0.750	0.00	0.00	0.00	-0.221	0.750	-0.309
81	0.00	0.715	0.00	0.00	0.00	0.038	0.715	0.066
82	0.00	0.578	0.00	0.00	0.00	-0.935	0.578	-1.232
83	0.00	0.759	0.00	0.00	0.00	0.739	0.759	1.61

	มิติของข้อสอบ							
84	0.00	0.459	0.00	0.00	0.00	0.338	0.459	0.566
85	0.00	0.597	0.00	0.00	0.00	-0.598	0.597	-0.748
86	0.00	0.799	0.00	0.00	0.00	-0.406	0.799	-0.679
87	0.00	0.598	0.00	0.00	0.00	-1.138	0.598	-1.138
88	0.00	1.000	0.00	0.00	0.00	1.026	1.000	2.000
89	0.00	0.513	0.00	0.00	0.00	-0.164	0.513	-0.281
90	0.00	0.584	0.00	0.00	0.00	-1.388	0.584	-1.388
91	0.00	0.00	1.000	0.00	0.00	-0.667	1.000	-0.807
92	0.00	0.00	0.827	0.00	0.00	-1.227	0.827	-1.227
93	0.00	0.00	1.000	0.00	0.00	-1.036	1.000	-1.036
94	0.00	0.00	1.000	0.00	0.00	-0.319	1.000	-0.472
95	0.00	0.00	0.676	0.00	0.00	-2.432	0.676	-2.432
96	0.00	0.00	1.000	0.00	0.00	-0.345	1.000	-0.784
97	0.00	0.00	0.440	0.00	0.00	-0.468	0.440	-0.67
98	0.00	0.00	0.699	0.00	0.00	-0.126	0.699	-0.24
99	0.00	0.00	0.526	0.00	0.00	-0.305	0.526	-0.471
100	0.00	0.00	0.648	0.00	0.00	-1.356	0.648	-1.356
101	0.00	0.00	1.000	0.00	0.00	-1.103	1.000	-1.491
102	0.00	0.00	0.740	0.00	0.00	0.461	0.740	0.917
103	0.00	0.00	0.503	0.00	0.00	-1.233	0.503	-1.668
104	0.00	0.00	0.739	0.00	0.00	-0.260	0.739	-0.422
105	0.00	0.00	0.616	0.00	0.00	0.221	0.616	0.412
106	0.00	0.00	0.537	0.00	0.00	-0.365	0.537	-0.543
107	0.00	0.00	0.672	0.00	0.00	-2.005	0.672	-2.005
108	0.00	0.00	1.000	0.00	0.00	0.365	1.000	0.668
109	0.00	0.00	0.546	0.00	0.00	-0.358	0.546	-0.418
110	0.00	0.00	0.856	0.00	0.00	-1.372	0.856	-1.372
111	0.00	0.00	1.000	0.00	0.00	0.365	1.000	0.51
112	0.00	0.00	0.715	0.00	0.00	-1.005	0.715	-1.091

	มิติของข้อสอบ							
113	0.00	0.00	0.921	0.00	0.00	-1.244	0.921	-1.381
114	0.00	0.00	0.901	0.00	0.00	0.151	0.901	0.231
115	0.00	0.00	0.653	0.00	0.00	-0.138	0.653	-0.202
116	0.00	0.00	0.683	0.00	0.00	-0.088	0.683	-0.117
117	0.00	0.00	0.750	0.00	0.00	-0.240	0.750	-0.355
118	0.00	0.00	0.677	0.00	0.00	0.372	0.677	0.639
119	0.00	0.00	0.582	0.00	0.00	-0.202	0.582	-0.262
120	0.00	0.00	0.771	0.00	0.00	-0.221	0.771	-0.302
121	0.00	0.00	0.733	0.00	0.00	0.038	0.733	0.064
122	0.00	0.00	0.593	0.00	0.00	-0.935	0.593	-1.2
123	0.00	0.00	0.779	0.00	0.00	0.739	0.779	1.569
124	0.00	0.00	0.471	0.00	0.00	0.338	0.471	0.551
125	0.00	0.00	0.613	0.00	0.00	-0.598	0.613	-0.729
126	0.00	0.00	0.820	0.00	0.00	-0.406	0.820	-0.661
127	0.00	0.00	0.614	0.00	0.00	-1.150	0.614	-1.15
128	0.00	0.00	1.000	0.00	0.00	0.994	1.000	2.199
129	0.00	0.00	0.452	0.00	0.00	-0.164	0.452	-0.274
130	0.00	0.00	0.598	0.00	0.00	-1.388	0.598	-1.388
131	0.00	0.00	1.000	0.00	0.00	-0.667	1.000	-0.807
132	0.00	0.00	0.827	0.00	0.00	-1.227	0.827	-1.227
133	0.00	0.00	1.000	0.00	0.00	-1.036	1.000	-1.036
134	0.00	0.00	1.000	0.00	0.00	-0.319	1.000	-0.472
135	0.00	0.00	0.676	0.00	0.00	-1.227	0.676	-1.394
136	0.00	0.00	0.00	0.880	0.00	-0.345	0.880	-0.739
137	0.00	0.00	0.00	0.467	0.00	-0.468	0.467	-0.633
138	0.00	0.00	0.00	0.739	0.00	-0.126	0.739	-0.226
139	0.00	0.00	0.00	0.558	0.00	-0.305	0.558	-0.445
140	0.00	0.00	0.00	0.685	0.00	-1.356	0.685	-1.356
141	0.00	0.00	0.00	1.000	0.00	-1.103	1.000	-1.414

	มิติของข้อสอบ							
142	0.00	0.00	0.00	0.780	0.00	0.461	0.780	0.867
143	0.00	0.00	0.00	0.532	0.00	-1.241	0.532	-1.532
144	0.00	0.00	0.00	0.810	0.00	-0.260	0.810	-0.399
145	0.00	0.00	0.00	0.651	0.00	0.221	0.651	0.388
146	0.00	0.00	0.00	0.569	0.00	-0.365	0.569	-0.515
147	0.00	0.00	0.00	0.709	0.00	-2.005	0.709	-2.005
148	0.00	0.00	0.00	1.000	0.00	0.365	1.000	0.633
149	0.00	0.00	0.00	0.577	0.00	-0.358	0.577	-0.396
150	0.00	0.00	0.00	0.905	0.00	-1.372	0.905	-1.372
151	0.00	0.00	0.00	1.000	0.00	0.365	1.000	0.482
152	0.00	0.00	0.00	0.757	0.00	-1.005	0.757	-1.034
153	0.00	0.00	0.00	0.972	0.00	-1.230	0.972	-1.23
154	0.00	0.00	0.00	1.000	0.00	0.151	1.000	0.219
155	0.00	0.00	0.00	0.691	0.00	-0.138	0.691	-0.191
156	0.00	0.00	0.00	0.724	0.00	-0.088	0.724	-0.11
157	0.00	0.00	0.00	0.797	0.00	-0.240	0.797	-0.333
158	0.00	0.00	0.00	0.721	0.00	0.372	0.721	0.601
159	0.00	0.00	0.00	0.619	0.00	-0.202	0.619	-0.247
160	0.00	0.00	0.00	0.819	0.00	-0.221	0.819	-0.283
161	0.00	0.00	0.00	0.780	0.00	0.038	0.780	0.06
162	0.00	0.00	0.00	0.630	0.00	-0.935	0.630	-1.132
163	0.00	0.00	0.00	0.826	0.00	0.739	0.826	1.475
164	0.00	0.00	0.00	0.501	0.00	0.338	0.501	0.518
165	0.00	0.00	0.00	0.652	0.00	-0.598	0.652	-0.687
166	0.00	0.00	0.00	0.870	0.00	-0.406	0.870	-0.624
167	0.00	0.00	0.00	0.651	0.00	-1.138	0.651	-1.138
168	0.00	0.00	0.00	1.000	0.00	1.026	1.000	1.835
169	0.00	0.00	0.00	0.559	0.00	-0.164	0.559	-0.257
170	0.00	0.00	0.00	0.637	0.00	-1.388	0.637	-1.388

	มิติของข้อสอบ							
171	0.00	0.00	0.00	1.000	0.00	-0.698	1.000	-0.822
172	0.00	0.00	0.00	0.849	0.00	-1.227	0.849	-1.227
173	0.00	0.00	0.00	1.000	0.00	-1.036	1.000	-1.036
174	0.00	0.00	0.00	1.000	0.00	-0.319	1.000	-0.441
175	0.00	0.00	0.00	0.724	0.00	-1.230	0.724	-1.374
176	0.00	0.00	0.00	0.895	0.00	-0.345	0.895	-0.739
177	0.00	0.00	0.00	0.467	0.00	-0.468	0.467	-0.633
178	0.00	0.00	0.00	0.739	0.00	-0.126	0.739	-0.226
179	0.00	0.00	0.00	0.558	0.00	-0.305	0.558	-0.445
180	0.00	0.00	0.00	0.685	0.00	-1.356	0.685	-1.356
181	0.00	0.00	0.00	1.000	0.00	-1.103	1.000	-1.414
182	0.00	0.00	0.00	0.780	0.00	0.461	0.780	0.867
183	0.00	0.00	0.00	0.532	0.00	-1.233	0.532	-1.646
184	0.00	0.00	0.00	0.749	0.00	-0.260	0.749	-0.399
185	0.00	0.00	0.00	0.651	0.00	0.196	0.651	0.375
186	0.00	0.00	0.00	0.523	0.00	-0.365	0.523	-0.515
187	0.00	0.00	0.00	0.709	0.00	-2.005	0.709	-2.005
188	0.00	0.00	0.00	1.000	0.00	0.365	1.000	0.633
189	0.00	0.00	0.00	0.577	0.00	-0.358	0.577	-0.396
190	0.00	0.00	0.00	0.905	0.00	-1.372	0.905	-1.372
191	0.00	0.00	0.00	1.000	0.00	0.365	1.000	0.482
192	0.00	0.00	0.00	0.757	0.00	-1.005	0.757	-1.034
193	0.00	0.00	0.00	0.972	0.00	-1.235	0.972	-1.235
194	0.00	0.00	0.00	1.000	0.00	0.151	1.000	0.219
195	0.00	0.00	0.00	0.691	0.00	-0.138	0.691	-0.191
196	0.00	0.00	0.00	0.00	0.724	-0.088	0.724	-0.115
197	0.00	0.00	0.00	0.00	0.766	-0.240	0.766	-0.347
198	0.00	0.00	0.00	0.00	0.691	0.372	0.691	0.627
199	0.00	0.00	0.00	0.00	0.593	-0.202	0.593	-0.257

	มิติของข้อสอบ							
200	0.00	0.00	0.00	0.00	0.786	-0.221	0.786	-0.293
201	0.00	0.00	0.00	0.00	0.753	0.038	0.753	0.062
202	0.00	0.00	0.00	0.00	0.611	-0.935	0.611	-1.166
203	0.00	0.00	0.00	0.00	0.802	0.739	0.802	1.521
204	0.00	0.00	0.00	0.00	0.486	0.338	0.486	0.536
205	0.00	0.00	0.00	0.00	0.631	-0.598	0.631	-0.709
206	0.00	0.00	0.00	0.00	0.843	-0.406	0.843	-0.643
207	0.00	0.00	0.00	0.00	0.631	-1.138	0.631	-1.138
208	0.00	0.00	0.00	0.00	1.000	1.026	1.000	1.893
209	0.00	0.00	0.00	0.00	0.542	-0.164	0.542	-0.266
210	0.00	0.00	0.00	0.00	0.617	-1.388	0.617	-1.388
211	0.00	0.00	0.00	0.00	1.000	-0.667	1.000	-0.777
212	0.00	0.00	0.00	0.00	0.858	-1.227	0.858	-1.227
213	0.00	0.00	0.00	0.00	1.000	-1.036	1.000	-1.036
214	0.00	0.00	0.00	0.00	1.000	-0.319	1.000	-0.454
215	0.00	0.00	0.00	0.00	0.702	-1.228	0.702	-1.295
216	0.00	0.00	0.00	0.00	0.948	-0.345	0.948	-0.76
217	0.00	0.00	0.00	0.00	0.454	-0.468	0.454	-0.651
218	0.00	0.00	0.00	0.00	0.719	-0.126	0.719	-0.232
219	0.00	0.00	0.00	0.00	0.544	-0.305	0.544	-0.457
220	0.00	0.00	0.00	0.00	0.667	-1.356	0.667	-1.356
221	0.00	0.00	0.00	0.00	1.000	-1.103	1.000	-1.449
222	0.00	0.00	0.00	0.00	0.761	0.461	0.761	0.89
223	0.00	0.00	0.00	0.00	0.518	-1.237	0.518	-1.433
224	0.00	0.00	0.00	0.00	0.863	-0.260	0.863	-0.411
225	0.00	0.00	0.00	0.00	0.633	0.221	0.633	0.4
226	0.00	0.00	0.00	0.00	0.552	-0.365	0.552	-0.53
227	0.00	0.00	0.00	0.00	0.689	-2.005	0.689	-2.005
228	0.00	0.00	0.00	0.00	1.000	0.365	1.000	0.649

	มิติของข้อสอบ							
229	0.00	0.00	0.00	0.00	0.562	-0.358	0.562	-0.407
230	0.00	0.00	0.00	0.00	0.880	-1.372	0.880	-1.372
231	0.00	0.00	0.00	0.00	1.000	0.365	1.000	0.496
232	0.00	0.00	0.00	0.00	0.736	-1.005	0.736	-1.062
233	0.00	0.00	0.00	0.00	0.946	-1.244	0.946	-1.244
234	0.00	0.00	0.00	0.00	1.000	0.151	1.000	0.224
235	0.00	0.00	0.00	0.00	0.673	-0.138	0.673	-0.196
236	0.00	0.00	0.00	0.00	0.705	-0.088	0.705	-0.115
237	0.00	0.00	0.00	0.00	0.766	-0.240	0.766	-0.347
238	0.00	0.00	0.00	0.00	0.691	0.372	0.691	0.627
239	0.00	0.00	0.00	0.00	0.593	-0.202	0.593	-0.257
240	0.00	0.00	0.00	0.00	0.786	1.254	0.786	1.552
241	0.00	0.00	0.00	0.00	0.808	0.755	0.808	1.016
242	0.00	0.00	0.00	0.00	0.743	0.690	0.743	0.737
MEAN	**0.150**	**0.118**	**0.140**	**0.189**	**0.146**	**-0.427**	**0.743**	**-0.430**
SD	**0.304**	**0.276**	**0.303**	**0.341**	**0.306**	**0.696**	**0.168**	**0.877**
MIN	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**-2.432**	**0.430**	**-2.432**
MAX	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.254**	**1.000**	**2.199**

วิเคราะห์คุณภาพของข้อสอบรายข้อ ตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response theory: MIRT) ด้วยโมเดลปกติสะสมแบบพหุมิติกับโอกาสการเดาข้อสอบ (Multidimensional normal ogive model guessing) โดยใช้โปรแกรมสำเร็จรูป NOHARM 4.0

1. กำหนดเงื่อนไขการวิเคราะห์คุณภาพข้อสอบ ดังต่อไปนี้

เกณฑ์การคัดเลือกข้อสอบ

1) ค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) มีค่าตั้งแต่ 0.00 ขึ้นไป

2) ค่าความยากแบบพหุมิติของข้อสอบ (MDIFF) มีค่าในช่วง -4.00 ถึง 4.00

3) ค่าการเดาข้อสอบถูก (c) กำหนดให้มีค่าคงที่เท่ากับ 0.20

2. ความตรงเชิงโครงสร้างของโมเดลพหุมิติของแบบทดสอบ เพื่อบอกความพอเพียงของตัวอย่างที่นำมาใช้ โดยพิจารณาจาก Sum of squares of residuals (SSR) Root mean square

of residuals (RMSR) และค่า Tanaka index of goodness of fit (GFI) (McDonald & Mok, 1995)

Sum of squares of residuals (SSR) หมายถึง ค่าดัชนีรากของกำลังสองเฉลี่ยของส่วนที่เหลือ (Root mean square of residuals) ซึ่ง RMSR มีค่าอยู่ระหว่าง 0 ถึง 1 ถ้ามีค่าต่ำกว่า 0.06 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

Tanaka index of goodness of fit (GFI) หมายถึง ค่าดัชนีวัดความเหมาะสมของโมเดลข้อสอบพหุมิติที่ปรับแก้ของ Tanaka (Tanaka index of goodness of fit) ซึ่งแสดงให้เห็นถึงความตรงเชิงโครงสร้างของข้อสอบแบบพหุมิติว่า มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ซึ่งควรมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้ามีค่าใกล้ 1.00 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

3. หาคุณภาพของแบบทดสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response theory: MIRT) โดยค่าพารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ คือ ค่าพารามิเตอร์ของข้อสอบ (item parameter) ค่าพารามิเตอร์ความยาก อำนาจจำแนกของข้อสอบแบบตรวจให้คะแนน 2 ค่า ได้จากการวิเคราะห์องค์ประกอบด้วยโมเดลปกติสะสมแบบพหุมิติ (Multidimensional normal ogive model) โดยใช้โปรแกรม Noharm 4.0 เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก (a_{ik}) ค่าจุดตัดของความยาก (d_i) ตามรูปแบบทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ดังสมการ

$$p(x_i=1)=\frac{\sum_{k=1}^{m} a_{ik}\theta_k+d_i}{1{,}0+e^{\sum_{k=1}^{m} a_{ik}\theta_k+d_i}}$$

เมื่อ X_i แทน คะแนน(0,1) ของข้อสอบข้อที่ i

a_{ij} แทน ส่วนประกอบของ a_i – เวกเตอร์ของมิติที่ k

di แทน พารามิเตอร์ที่มีความสัมพันธ์กับความความยากของข้อสอบในแบบทดสอบ

θ_k แทน เวกเตอร์ค่าพารามิเตอร์ความสามารถ

4. ค่าพารามิเตอร์อำนาจจำแนกแบบพหุมิติของข้อสอบ ($MDISC_i$) ควรมีค่าอยู่ระหว่าง -4.00 ถึง 4.00 ถ้ามีค่าอยู่ช่วงพิสัยดังกล่าว แสดงว่าข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติ ที่เหมาะสม ดังสมการ

$$MDISC_i=\sqrt{\sum_{k=1}^{m} a_{ik}^2}$$

เมื่อ $MDISC_i$ แทน ค่าอำนาจจำแนกรวมของข้อสอบข้อที่ i ใน m มิติ

m แทน จำนวนของมิติใน θ-space

a_{ik} แทน ส่วนประกอบของ a_i –เวกเตอร์ของมิติที่ k

จากสมการจะเห็นได้ว่า $MDISC_i$ เป็นสัดส่วนของความชันพื้นผิวการตอบสนองข้อสอบตรงตำแหน่งที่มีความชันสูงสุดจึงสังเกตได้ว่ามีลักษณะที่คล้ายกับค่าพารามิเตอร์อำนาจจำแนกในโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Reckase, 2009)

5. ค่าพารามิเตอร์ความยากแบบพหุมิติของข้อสอบ ($MDIFF_i$) ควรมีค่าอยู่ระหว่าง 4.00 ถึง 4.00 ถ้ามีค่าอยู่ช่วงพิสัยดังกล่าว แสดงว่าข้อสอบมีค่าอำนาจจำแนกแบบพหุมิติที่เหมาะสม ดังสมการ

$$MDIFF_i = \frac{-d_i}{\sqrt{\sum_{k-1}^{m} a_{ik}^2}}$$

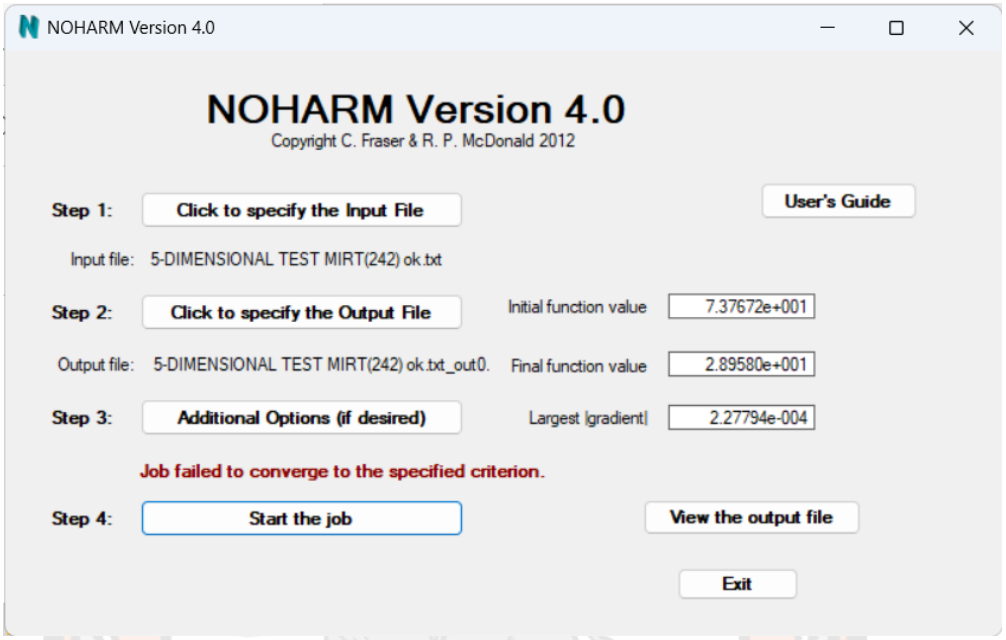
เมื่อ d_i แทน พารามิเตอร์ที่มีความสัมพันธ์กับค่าความยากของข้อสอบในแบบทดสอบ

จากสมการ พบว่าค่า d_i เป็นพารามิเตอร์ที่มีความสัมพันธ์กับค่าความยากของข้อสอบในแบบทดสอบอย่างไรก็ตาม $MDIFF_i$ ไม่สามารถแปลความหมายเหมือนค่าความยากของข้อสอบในโมเดลการตอบสนองข้อสอบแบบเอกมิติซึ่งค่าของ $MDIFF_i$ บ่งชี้ถึงระยะทางจากจุดกำเนิดของ θ-space ไปยังจุดที่มีความชันสูงสุด (Steepest Slope)

ภาคผนวก ข

คำสั่งและผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันเชิงซ้อนตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ โดยโปรแกรม NOHARM

คำสั่งวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันเชิงซ้อนตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ โดยโปรแกรม NOHARM สำหรับข้อสอบจำนวน 242 ข้อ ที่มีการวัด 5 มิติ จากกลุ่มตัวอย่างจำนวน 500 คน



TEST MIRT(242) TO ILLUSTRATE A 5-DIMENSIONAL CASE.

242 5 500 0 0 0 0 0

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20 .20

.20 .20 .20 .20

.20 .20 .20 .20 .20 .20 .20 .20

1 0 0 0 0

1 0 0 0 0

1 0 0 0 0

1 0 0 0 0

1 0 0 0 0

1 0 0 0 0

...

...

...

0 0 0 0 1

0 0 0 0 1

0 0 0 0 1

0 0 0 0 1

0 0 0 0 1

0 0 0 0 1

0 0 0 0 1

0

1 0

1 1 0

1 1 1 0

1 1 1 1 0

0 0 0 0 0 0 0 0 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0

0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0

0 1 0 0 0 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0

1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1

0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 0 0

1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0

(ข้อมูลลำดับที่ 1)

1 1 1 1 0 1 0 1 1 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0
1 0 1 1 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1
0 0 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 0 0 0 0 0
1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 0 0 0 0 0 1 1 0 0 0
0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 1
0 0 1 0 0 1 0 0 0 0 1 1 1 1 1 1 1

(ข้อมูลลำดับที่ 2)

0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0
1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0
0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1
0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 0
0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0
1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 1

(ข้อมูลลำดับที่ 3)

1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 1 1 1 0 1 1 1 0 1 0 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 0 1 1 0
1 0 1 0 0 0 1 0 0 1 1 1 1 0 1 1 1 0 1 0 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 0 1 1 0 1 0 1 0 0
0 1 0 0 1 1 1 1 0 1 1 1 0 1 0 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1
1 1 1 0 1 1 1 0 1 0 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 1 1 1 0 1
1 1 0 1 0 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 1 1 1 0 1 1 1 0 1 0
1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 1

(ข้อมูลลำดับที่ 4)

.....

.....

.....

.....

.....

.....

1 0 1 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1
0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0
1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0
1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0

0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1
0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0
1 0 1 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1
0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0
1 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0
1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0
0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 1 0 0 1
0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0

(ข้อมูลลำดับที่ 500)

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันเชิงซ้อน ตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ โดยโปรแกรม NOHARM สำหรับข้อสอบ จำนวน 242 ข้อที่มีการวัด 5 มิติ จากกลุ่มตัวอย่างจำนวน 500 คน

N O H A R M

Fitting a (multidimensional) Normal Ogive
by Harmonic Analysis - Robust Method

Input File : 5-DIMENSIONAL TEST MIRT(242) Normal Ogive.dat

Title : TEST MIRT(242) TO ILLUSTRATE A 5-DIMENSIONAL CASE.

Number of items __ = 242
Number of dimensions = 5
Number of subjects = 500

Sample Product-Moment Matrix

	1	2	3	4	5	6	7	8	9
1	0.612								
2	0.236	0.340							
3	0.542	0.304	0.816						
4	0.474	0.282	0.614	0.706					
5	0.322	0.200	0.384	0.368	0.420				
6	0.344	0.206	0.410	0.370	0.270	0.474			
7	0.228	0.164	0.264	0.250	0.168	0.186	0.302		
8	0.556	0.312	0.732	0.652	0.382	0.432	0.286	0.860	
9	0.390	0.220	0.470	0.428	0.300	0.304	0.218	0.498	0.564
10	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160

11	0.298	0.192	0.348	0.322	0.258	0.254	0.196	0.374	0.310
12	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
13	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
14	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350
15	0.154	0.090	0.154	0.152	0.144	0.154	0.086	0.164	0.140
16	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
17	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
18	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
19	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
20	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
21	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
22	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
23	0.124	0.076	0.128	0.120	0.118	0.122	0.086	0.136	0.134
24	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
25	0.434	0.268	0.586	0.516	0.310	0.332	0.246	0.604	0.408
26	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
27	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154
28	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
29	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
30	0.220	0.136	0.242	0.236	0.188	0.202	0.158	0.258	0.214
31	0.496	0.296	0.616	0.568	0.344	0.368	0.260	0.668	0.444
32	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228
33	0.142	0.120	0.152	0.140	0.128	0.142	0.116	0.158	0.130
34	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
35	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
36	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366
37	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
38	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
39	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
40	0.372	0.218	0.460	0.436	0.298	0.302	0.230	0.496	0.326
41	0.612	0.236	0.542	0.474	0.322	0.344	0.228	0.556	0.390

42	0.236	0.340	0.304	0.282	0.200	0.206	0.164	0.312	0.220
43	0.542	0.304	0.816	0.614	0.384	0.410	0.264	0.732	0.470
44	0.474	0.282	0.614	0.706	0.368	0.370	0.250	0.652	0.428
45	0.322	0.200	0.384	0.368	0.420	0.270	0.168	0.382	0.300
46	0.344	0.206	0.410	0.370	0.270	0.474	0.186	0.432	0.304
47	0.228	0.164	0.264	0.250	0.168	0.186	0.302	0.286	0.218
48	0.556	0.314	0.734	0.654	0.384	0.434	0.286	0.860	0.516
49	0.382	0.212	0.462	0.420	0.292	0.296	0.210	0.490	0.540
50	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160
51	0.290	0.184	0.340	0.314	0.250	0.246	0.188	0.366	0.286
52	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
53	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
54	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350
55	0.156	0.084	0.158	0.158	0.134	0.148	0.082	0.174	0.138
56	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
57	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
58	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
59	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
60	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
61	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
62	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
63	0.136	0.082	0.144	0.130	0.122	0.126	0.088	0.154	0.152
64	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
65	0.434	0.268	0.586	0.516	0.310	0.332	0.246	0.604	0.408
66	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
67	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154
68	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
69	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
70	0.216	0.134	0.238	0.234	0.186	0.198	0.154	0.254	0.210
71	0.492	0.294	0.612	0.566	0.342	0.364	0.256	0.664	0.440
72	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228

73	0.152	0.112	0.164	0.150	0.130	0.148	0.114	0.166	0.144
74	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
75	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
76	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366
77	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
78	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
79	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
80	0.372	0.218	0.460	0.436	0.298	0.302	0.230	0.496	0.326
81	0.612	0.236	0.542	0.474	0.322	0.344	0.228	0.556	0.390
82	0.236	0.340	0.304	0.282	0.200	0.206	0.164	0.312	0.220
83	0.542	0.304	0.816	0.614	0.384	0.410	0.264	0.732	0.470
84	0.474	0.282	0.614	0.706	0.368	0.370	0.250	0.652	0.428
85	0.322	0.200	0.384	0.368	0.420	0.270	0.168	0.382	0.300
86	0.344	0.206	0.410	0.370	0.270	0.474	0.186	0.432	0.304
87	0.228	0.164	0.264	0.250	0.168	0.186	0.302	0.286	0.218
88	0.556	0.314	0.734	0.654	0.384	0.434	0.286	0.860	0.516
89	0.390	0.220	0.470	0.428	0.300	0.304	0.218	0.498	0.548
90	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160
91	0.298	0.192	0.348	0.322	0.258	0.254	0.196	0.374	0.294
92	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
93	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
94	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350
95	0.170	0.096	0.176	0.172	0.154	0.160	0.092	0.190	0.152
96	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
97	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
98	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
99	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
100	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
101	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
102	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
103	0.140	0.078	0.142	0.134	0.110	0.124	0.104	0.156	0.142

104	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
105	0.434	0.268	0.586	0.516	0.310	0.332	0.246	0.604	0.408
106	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
107	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154
108	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
109	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
110	0.220	0.136	0.242	0.236	0.188	0.202	0.158	0.258	0.214
111	0.496	0.296	0.616	0.568	0.344	0.368	0.260	0.668	0.444
112	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228
113	0.140	0.100	0.154	0.138	0.116	0.134	0.116	0.156	0.128
114	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
115	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
116	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366
117	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
118	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
119	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
120	0.372	0.218	0.460	0.436	0.298	0.302	0.230	0.496	0.326
121	0.612	0.236	0.542	0.474	0.322	0.344	0.228	0.556	0.390
122	0.236	0.340	0.304	0.282	0.200	0.206	0.164	0.312	0.220
123	0.542	0.304	0.816	0.614	0.384	0.410	0.264	0.732	0.470
124	0.474	0.282	0.614	0.706	0.368	0.370	0.250	0.652	0.428
125	0.322	0.200	0.384	0.368	0.420	0.270	0.168	0.382	0.300
126	0.344	0.206	0.410	0.370	0.270	0.474	0.186	0.432	0.304
127	0.226	0.162	0.262	0.248	0.166	0.184	0.300	0.284	0.216
128	0.550	0.310	0.728	0.650	0.380	0.428	0.280	0.854	0.510
129	0.390	0.220	0.470	0.428	0.300	0.304	0.218	0.498	0.548
130	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160
131	0.298	0.192	0.348	0.322	0.258	0.254	0.196	0.374	0.294
132	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
133	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
134	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350

135	0.164	0.094	0.164	0.160	0.148	0.154	0.094	0.178	0.136
136	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
137	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
138	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
139	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
140	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
141	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
142	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
143	0.134	0.078	0.134	0.126	0.118	0.122	0.092	0.146	0.142
144	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
145	0.434	0.268	0.586	0.516	0.310	0.332	0.246	0.604	0.408
146	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
147	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154
148	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
149	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
150	0.220	0.136	0.242	0.236	0.188	0.202	0.158	0.258	0.214
151	0.496	0.296	0.616	0.568	0.344	0.368	0.260	0.668	0.444
152	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228
153	0.154	0.108	0.172	0.154	0.142	0.162	0.116	0.180	0.136
154	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
155	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
156	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366
157	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
158	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
159	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
160	0.372	0.218	0.460	0.436	0.298	0.302	0.230	0.496	0.326
161	0.612	0.236	0.542	0.474	0.322	0.344	0.228	0.556	0.390
162	0.236	0.340	0.304	0.282	0.200	0.206	0.164	0.312	0.220
163	0.542	0.304	0.816	0.614	0.384	0.410	0.264	0.732	0.470
164	0.474	0.282	0.614	0.706	0.368	0.370	0.250	0.652	0.428
165	0.322	0.200	0.384	0.368	0.420	0.270	0.168	0.382	0.300

166	0.344	0.206	0.410	0.370	0.270	0.474	0.186	0.432	0.304
167	0.228	0.164	0.264	0.250	0.168	0.186	0.302	0.286	0.218
168	0.556	0.314	0.734	0.654	0.384	0.434	0.286	0.860	0.516
169	0.390	0.220	0.470	0.428	0.300	0.304	0.218	0.498	0.548
170	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160
171	0.290	0.184	0.340	0.314	0.250	0.246	0.188	0.366	0.286
172	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
173	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
174	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350
175	0.162	0.090	0.160	0.160	0.146	0.154	0.092	0.176	0.136
176	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
177	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
178	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
179	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
180	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
181	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
182	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
183	0.136	0.076	0.144	0.126	0.116	0.126	0.094	0.156	0.142
184	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
185	0.426	0.266	0.578	0.510	0.304	0.324	0.240	0.596	0.400
186	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
187	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154
188	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
189	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
190	0.220	0.136	0.242	0.236	0.188	0.202	0.158	0.258	0.214
191	0.496	0.296	0.616	0.568	0.344	0.368	0.260	0.668	0.444
192	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228
193	0.146	0.114	0.162	0.144	0.134	0.144	0.124	0.170	0.134
194	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
195	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
196	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366

197	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
198	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
199	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
200	0.372	0.218	0.460	0.436	0.298	0.302	0.230	0.496	0.326
201	0.612	0.236	0.542	0.474	0.322	0.344	0.228	0.556	0.390
202	0.236	0.340	0.304	0.282	0.200	0.206	0.164	0.312	0.220
203	0.542	0.304	0.816	0.614	0.384	0.410	0.264	0.732	0.470
204	0.474	0.282	0.614	0.706	0.368	0.370	0.250	0.652	0.428
205	0.322	0.200	0.384	0.368	0.420	0.270	0.168	0.382	0.300
206	0.344	0.206	0.410	0.370	0.270	0.474	0.186	0.432	0.304
207	0.228	0.164	0.264	0.250	0.168	0.186	0.302	0.286	0.218
208	0.556	0.314	0.734	0.654	0.384	0.434	0.286	0.860	0.516
209	0.390	0.220	0.470	0.428	0.300	0.304	0.218	0.498	0.548
210	0.196	0.136	0.238	0.202	0.146	0.170	0.146	0.246	0.160
211	0.298	0.192	0.348	0.322	0.258	0.254	0.196	0.374	0.294
212	0.220	0.162	0.246	0.222	0.174	0.170	0.156	0.254	0.214
213	0.240	0.160	0.276	0.272	0.194	0.184	0.194	0.312	0.228
214	0.360	0.216	0.452	0.402	0.264	0.272	0.220	0.456	0.350
215	0.164	0.092	0.168	0.166	0.152	0.156	0.100	0.178	0.142
216	0.330	0.188	0.420	0.360	0.234	0.266	0.218	0.424	0.304
217	0.336	0.222	0.392	0.374	0.280	0.284	0.184	0.420	0.294
218	0.394	0.276	0.478	0.446	0.264	0.312	0.222	0.502	0.352
219	0.374	0.204	0.442	0.394	0.278	0.298	0.202	0.432	0.352
220	0.194	0.124	0.220	0.208	0.178	0.164	0.118	0.236	0.198
221	0.214	0.148	0.268	0.240	0.168	0.166	0.156	0.282	0.216
222	0.488	0.280	0.630	0.574	0.344	0.380	0.262	0.676	0.460
223	0.142	0.082	0.142	0.132	0.124	0.130	0.104	0.150	0.146
224	0.356	0.206	0.460	0.416	0.252	0.300	0.230	0.490	0.320
225	0.434	0.268	0.586	0.516	0.310	0.332	0.246	0.604	0.408
226	0.334	0.204	0.422	0.374	0.280	0.266	0.192	0.446	0.310
227	0.174	0.148	0.214	0.186	0.152	0.148	0.136	0.212	0.154

228	0.480	0.264	0.612	0.542	0.344	0.384	0.246	0.650	0.444
229	0.370	0.216	0.432	0.412	0.292	0.294	0.208	0.458	0.326
230	0.220	0.136	0.242	0.236	0.188	0.202	0.158	0.258	0.214
231	0.496	0.296	0.616	0.568	0.344	0.368	0.260	0.668	0.444
232	0.240	0.152	0.296	0.254	0.180	0.188	0.178	0.296	0.228
233	0.146	0.108	0.152	0.140	0.130	0.148	0.122	0.158	0.134
234	0.442	0.268	0.562	0.516	0.334	0.358	0.250	0.604	0.406
235	0.398	0.230	0.478	0.448	0.292	0.300	0.236	0.518	0.352
236	0.404	0.220	0.496	0.470	0.296	0.328	0.240	0.528	0.366
237	0.356	0.206	0.448	0.404	0.266	0.280	0.238	0.466	0.340
238	0.470	0.278	0.602	0.546	0.346	0.370	0.250	0.636	0.446
239	0.384	0.218	0.480	0.442	0.294	0.308	0.230	0.500	0.350
240	0.552	0.318	0.748	0.646	0.380	0.444	0.270	0.792	0.508
241	0.492	0.296	0.662	0.572	0.336	0.394	0.240	0.698	0.456
242	0.488	0.298	0.658	0.572	0.328	0.394	0.240	0.700	0.428

Item Covariance Matrix

	1	2	3	4	5	6	7	8	9
1	0.237								
2	0.028	0.224							
3	0.043	0.027	0.150						
4	0.042	0.042	0.038	0.208					
5	0.065	0.057	0.041	0.071	0.244				
6	0.054	0.045	0.023	0.035	0.071	0.249			
7	0.043	0.061	0.018	0.037	0.041	0.043	0.211		
8	0.030	0.020	0.030	0.045	0.021	0.024	0.026	0.120	
9	0.045	0.028	0.010	0.030	0.063	0.037	0.048	0.013	0.246
10	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010

11	0.042	0.050	0.007	0.027	0.082	0.056	0.070	0.015	0.074
12	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
13	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
14	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068
15	0.043	0.028	0.005	0.024	0.068	0.068	0.031	0.007	0.037
16	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
17	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
18	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
19	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
20	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
21	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
22	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
23	0.029	0.023	0.001	0.010	0.052	0.048	0.039	0.002	0.046
24	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
25	0.024	0.040	0.039	0.043	0.029	0.014	0.044	0.028	0.030
26	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
27	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031
28	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
29	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
30	0.056	0.045	0.023	0.047	0.075	0.075	0.077	0.028	0.063
31	0.059	0.053	0.033	0.064	0.044	0.030	0.044	0.054	0.041
32	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044
33	0.040	0.064	0.017	0.023	0.058	0.063	0.066	0.015	0.036
34	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
35	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
36	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043
37	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
38	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
39	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
40	0.048	0.038	0.028	0.062	0.075	0.051	0.070	0.040	0.027
41	0.237	0.028	0.043	0.042	0.065	0.054	0.043	0.030	0.045

42	0.028	0.224	0.027	0.042	0.057	0.045	0.061	0.020	0.028
43	0.043	0.027	0.150	0.038	0.041	0.023	0.018	0.030	0.010
44	0.042	0.042	0.038	0.208	0.071	0.035	0.037	0.045	0.030
45	0.065	0.057	0.041	0.071	0.244	0.071	0.041	0.021	0.063
46	0.054	0.045	0.023	0.035	0.071	0.249	0.043	0.024	0.037
47	0.043	0.061	0.018	0.037	0.041	0.043	0.211	0.026	0.048
48	0.019	0.015	0.018	0.034	0.015	0.018	0.021	0.105	0.021
49	0.052	0.028	0.021	0.039	0.065	0.040	0.047	0.026	0.235
50	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010
51	0.049	0.050	0.018	0.036	0.085	0.059	0.069	0.027	0.064
52	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
53	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
54	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068
55	0.043	0.021	0.008	0.028	0.057	0.061	0.026	0.016	0.034
56	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
57	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
58	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
59	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
60	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
61	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
62	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
63	0.028	0.022	3.8e-4	0.006	0.048	0.043	0.035	0.003	0.053
64	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
65	0.024	0.040	0.039	0.043	0.029	0.014	0.044	0.028	0.030
66	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
67	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031
68	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
69	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
70	0.054	0.044	0.023	0.048	0.075	0.073	0.074	0.027	0.061
71	0.057	0.053	0.033	0.065	0.044	0.027	0.042	0.053	0.040
72	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044

73	0.043	0.051	0.019	0.024	0.055	0.064	0.060	0.013	0.044
74	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
75	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
76	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043
77	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
78	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
79	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
80	0.048	0.038	0.028	0.062	0.075	0.051	0.070	0.040	0.027
81	0.237	0.028	0.043	0.042	0.065	0.054	0.043	0.030	0.045
82	0.028	0.224	0.027	0.042	0.057	0.045	0.061	0.020	0.028
83	0.043	0.027	0.150	0.038	0.041	0.023	0.018	0.030	0.010
84	0.042	0.042	0.038	0.208	0.071	0.035	0.037	0.045	0.030
85	0.065	0.057	0.041	0.071	0.244	0.071	0.041	0.021	0.063
86	0.054	0.045	0.023	0.035	0.071	0.249	0.043	0.024	0.037
87	0.043	0.061	0.018	0.037	0.041	0.043	0.211	0.026	0.048
88	0.019	0.015	0.018	0.034	0.015	0.018	0.021	0.105	0.021
89	0.055	0.034	0.023	0.041	0.070	0.044	0.053	0.027	0.239
90	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010
91	0.052	0.055	0.020	0.038	0.089	0.063	0.075	0.028	0.067
92	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
93	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
94	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068
95	0.044	0.026	0.008	0.027	0.067	0.062	0.030	0.013	0.036
96	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
97	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
98	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
99	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
100	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
101	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
102	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
103	0.030	0.017	-0.005	0.007	0.034	0.039	0.050	0.001	0.040

104	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
105	0.024	0.040	0.039	0.043	0.029	0.014	0.044	0.028	0.030
106	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
107	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031
108	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
109	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
110	0.056	0.045	0.023	0.047	0.075	0.075	0.077	0.028	0.063
111	0.059	0.053	0.033	0.064	0.044	0.030	0.044	0.054	0.041
112	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044
113	0.040	0.044	0.020	0.022	0.047	0.056	0.066	0.015	0.036
114	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
115	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
116	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043
117	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
118	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
119	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
120	0.048	0.038	0.028	0.062	0.075	0.051	0.070	0.040	0.027
121	0.237	0.028	0.043	0.042	0.065	0.054	0.043	0.030	0.045
122	0.028	0.224	0.027	0.042	0.057	0.045	0.061	0.020	0.028
123	0.043	0.027	0.150	0.038	0.041	0.023	0.018	0.030	0.010
124	0.042	0.042	0.038	0.208	0.071	0.035	0.037	0.045	0.030
125	0.065	0.057	0.041	0.071	0.244	0.071	0.041	0.021	0.063
126	0.054	0.045	0.023	0.035	0.071	0.249	0.043	0.024	0.037
127	0.042	0.060	0.017	0.036	0.040	0.042	0.209	0.026	0.047
128	0.016	0.014	0.016	0.034	0.014	0.015	0.017	0.104	0.018
129	0.055	0.034	0.023	0.041	0.070	0.044	0.053	0.027	0.239
130	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010
131	0.052	0.055	0.020	0.038	0.089	0.063	0.075	0.028	0.067
132	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
133	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
134	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068

135	0.048	0.029	0.009	0.026	0.068	0.064	0.037	0.015	0.029
136	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
137	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
138	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
139	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
140	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
141	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
142	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
143	0.031	0.021	-0.003	0.007	0.047	0.042	0.041	0.002	0.047
144	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
145	0.024	0.040	0.039	0.043	0.029	0.014	0.044	0.028	0.030
146	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
147	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031
148	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
149	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
150	0.056	0.045	0.023	0.047	0.075	0.075	0.077	0.028	0.063
151	0.059	0.053	0.033	0.064	0.044	0.030	0.044	0.054	0.041
152	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044
153	0.040	0.045	0.020	0.023	0.064	0.074	0.060	0.020	0.031
154	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
155	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
156	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043
157	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
158	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
159	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
160	0.048	0.038	0.028	0.062	0.075	0.051	0.070	0.040	0.027
161	0.237	0.028	0.043	0.042	0.065	0.054	0.043	0.030	0.045
162	0.028	0.224	0.027	0.042	0.057	0.045	0.061	0.020	0.028
163	0.043	0.027	0.150	0.038	0.041	0.023	0.018	0.030	0.010
164	0.042	0.042	0.038	0.208	0.071	0.035	0.037	0.045	0.030
165	0.065	0.057	0.041	0.071	0.244	0.071	0.041	0.021	0.063

166	0.054	0.045	0.023	0.035	0.071	0.249	0.043	0.024	0.037
167	0.043	0.061	0.018	0.037	0.041	0.043	0.211	0.026	0.048
168	0.019	0.015	0.018	0.034	0.015	0.018	0.021	0.105	0.021
169	0.055	0.034	0.023	0.041	0.070	0.044	0.053	0.027	0.239
170	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010
171	0.049	0.050	0.018	0.036	0.085	0.059	0.069	0.027	0.064
172	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
173	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
174	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068
175	0.048	0.027	0.008	0.029	0.068	0.066	0.036	0.016	0.031
176	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
177	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
178	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
179	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
180	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
181	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
182	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
183	0.026	0.015	-0.003	-0.001	0.040	0.041	0.040	0.001	0.040
184	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
185	0.021	0.041	0.038	0.043	0.026	0.010	0.040	0.027	0.027
186	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
187	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031
188	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
189	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
190	0.056	0.045	0.023	0.047	0.075	0.075	0.077	0.028	0.063
191	0.059	0.053	0.033	0.064	0.044	0.030	0.044	0.054	0.041
192	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044
193	0.037	0.053	0.017	0.018	0.059	0.060	0.070	0.017	0.034
194	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
195	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
196	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043

197	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
198	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
199	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
200	0.048	0.038	0.028	0.062	0.075	0.051	0.070	0.040	0.027
201	0.237	0.028	0.043	0.042	0.065	0.054	0.043	0.030	0.045
202	0.028	0.224	0.027	0.042	0.057	0.045	0.061	0.020	0.028
203	0.043	0.027	0.150	0.038	0.041	0.023	0.018	0.030	0.010
204	0.042	0.042	0.038	0.208	0.071	0.035	0.037	0.045	0.030
205	0.065	0.057	0.041	0.071	0.244	0.071	0.041	0.021	0.063
206	0.054	0.045	0.023	0.035	0.071	0.249	0.043	0.024	0.037
207	0.043	0.061	0.018	0.037	0.041	0.043	0.211	0.026	0.048
208	0.019	0.015	0.018	0.034	0.015	0.018	0.021	0.105	0.021
209	0.055	0.034	0.023	0.041	0.070	0.044	0.053	0.027	0.239
210	0.033	0.046	0.021	0.014	0.034	0.044	0.066	0.017	0.010
211	0.052	0.055	0.020	0.038	0.089	0.063	0.075	0.028	0.067
212	0.044	0.064	0.011	0.019	0.053	0.033	0.069	0.006	0.052
213	0.044	0.051	0.015	0.046	0.060	0.032	0.097	0.037	0.048
214	0.054	0.046	0.044	0.049	0.054	0.035	0.069	0.026	0.068
215	0.049	0.028	0.015	0.033	0.073	0.067	0.043	0.016	0.036
216	0.029	0.021	0.019	0.013	0.027	0.033	0.069	0.001	0.027
217	0.057	0.067	0.020	0.052	0.088	0.068	0.046	0.028	0.037
218	0.051	0.086	0.021	0.051	0.029	0.047	0.053	0.020	0.036
219	0.066	0.033	0.031	0.038	0.066	0.059	0.050	-0.001	0.068
220	0.029	0.032	-3.2e-4	0.017	0.065	0.036	0.036	0.004	0.046
221	0.026	0.043	0.017	0.023	0.039	0.020	0.063	0.017	0.042
222	0.034	0.028	0.025	0.050	0.032	0.028	0.038	0.038	0.042
223	0.036	0.023	1.6e-5	0.009	0.051	0.048	0.051	3.6e-4	0.048
224	0.039	0.030	0.037	0.050	0.034	0.054	0.074	0.045	0.028
225	0.024	0.040	0.039	0.043	0.029	0.014	0.044	0.028	0.030
226	0.037	0.039	0.025	0.031	0.076	0.036	0.045	0.028	0.036
227	0.041	0.074	0.036	0.032	0.060	0.045	0.070	0.025	0.031

228	0.043	0.021	0.029	0.038	0.044	0.046	0.030	0.036	0.041
229	0.071	0.050	0.034	0.067	0.087	0.063	0.061	0.038	0.051
230	0.056	0.045	0.023	0.047	0.075	0.075	0.077	0.028	0.063
231	0.059	0.053	0.033	0.064	0.044	0.030	0.044	0.054	0.041
232	0.040	0.041	0.030	0.024	0.043	0.033	0.080	0.016	0.044
233	0.046	0.052	0.018	0.024	0.061	0.070	0.072	0.017	0.042
234	0.045	0.048	0.033	0.059	0.062	0.051	0.054	0.047	0.041
235	0.058	0.041	0.024	0.055	0.058	0.036	0.068	0.040	0.038
236	0.054	0.026	0.029	0.066	0.056	0.057	0.067	0.036	0.043
237	0.035	0.028	0.020	0.034	0.046	0.032	0.080	0.015	0.044
238	0.032	0.035	0.018	0.041	0.045	0.031	0.034	0.020	0.042
239	0.056	0.036	0.043	0.064	0.069	0.054	0.068	0.039	0.048
240	-0.009	0.007	0.001	-0.001	-0.005	0.010	-0.007	0.004	-0.009
241	-0.010	0.017	-0.007	-0.007	-0.008	0.005	-0.008	-0.007	-0.006
242	-0.004	0.025	0.002	0.004	-0.010	0.013	-0.003	0.009	-0.025

Revised Fixed Guesses

1	2	3	4	5	6	7	8	9	10
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
11	12	13	14	15	16	17	18	19	20
0.200	0.200	0.200	0.200	0.082	0.200	0.200	0.200	0.200	0.200
21	22	23	24	25	26	27	28	29	30
0.200	0.200	0.056	0.200	0.200	0.200	0.200	0.200	0.200	0.200

31	32	33	34	35	36	37	38	39	40
0.200	0.200	0.066	0.200	0.200	0.200	0.200	0.200	0.200	0.200
41	42	43	44	45	46	47	48	49	50
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
51	52	53	54	55	56	57	58	59	60
0.200	0.200	0.200	0.200	0.084	0.200	0.200	0.200	0.200	0.200
61	62	63	64	65	66	67	68	69	70
0.200	0.200	0.076	0.200	0.200	0.200	0.200	0.200	0.200	0.200
71	72	73	74	75	76	77	78	79	80
0.200	0.200	0.078	0.200	0.200	0.200	0.200	0.200	0.200	0.200
81	82	83	84	85	86	87	88	89	90
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
91	92	93	94	95	96	97	98	99	100
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
101	102	103	104	105	106	107	108	109	110

0.200	0.200	0.080	0.200	0.200	0.200	0.200	0.200	0.200	0.200
111	112	113	114	115	116	117	118	119	120
0.200	0.200	0.064	0.200	0.200	0.200	0.200	0.200	0.200	0.200
121	122	123	124	125	126	127	128	129	130
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
131	132	133	134	135	136	137	138	139	140
0.200	0.200	0.200	0.200	0.090	0.200	0.200	0.200	0.200	0.200
141	142	143	144	145	146	147	148	149	150
0.200	0.200	0.068	0.200	0.200	0.200	0.200	0.200	0.200	0.200
151	152	153	154	155	156	157	158	159	160
0.200	0.200	0.086	0.200	0.200	0.200	0.200	0.200	0.200	0.200
161	162	163	164	165	166	167	168	169	170
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
171	172	173	174	175	176	177	178	179	180
0.200	0.200	0.200	0.200	0.086	0.200	0.200	0.200	0.200	0.200
181	182	183	184	185	186	187	188	189	190

0.200	0.200	0.080	0.200	0.200	0.200	0.200	0.200	0.200	0.200
191	192	193	194	195	196	197	198	199	200
0.200	0.200	0.078	0.200	0.200	0.200	0.200	0.200	0.200	0.200
201	202	203	204	205	206	207	208	209	210
0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
211	212	213	214	215	216	217	218	219	220
0.200	0.200	0.200	0.200	0.088	0.200	0.200	0.200	0.200	0.200
221	222	223	224	225	226	227	228	229	230
0.200	0.200	0.074	0.200	0.200	0.200	0.200	0.200	0.200	0.200
231	232	233	234	235	236	237	238	239	240
0.200	0.200	0.064	0.200	0.200	0.200	0.200	0.200	0.200	0.200
241	242								
0.200	0.200								

=======

Results

=======

Success. The job converged to the specified criterion.

Final Constants

1	2	3	4	5	6	7	8	9	10
0.047	-1.509	0.840	0.431	-1.064	-0.517	-5.8e+4	1.155	-0.136	-5.1e+4
11	12	13	14	15	16	17	18	19	20
-0.945	-5.4e+4	-4.0e+4	-0.442	-2.495	-0.386	-0.663	-0.149	-0.405	-4.9e+4
21	22	23	24	25	26	27	28	29	30
-1.671	0.536	-1.936	-0.333	0.264	-0.499	-3.6e+4	0.439	-0.727	-4.9e+4
31	32	33	34	35	36	37	38	39	40
0.532	-2.835	-4.662	0.202	-0.192	-0.136	-0.332	0.461	-0.325	-0.332
41	42	43	44	45	46	47	48	49	50
0.047	-1.509	0.840	0.431	-1.064	-0.517	-5.8e+4	1.212	-0.229	-5.1e+4
51	52	53	54	55	56	57	58	59	60
-1.112	-4.7e+4	-2.5e+4	-0.426	-1.886	-0.382	-0.639	-0.146	-0.394	-4.0e+4
61	62	63	64	65	66	67	68	69	70

-1.594	0.529	-1.897	-0.325	0.260	-0.483	-3.5e+4	0.431	-0.650	-4.2e+4
71	72	73	74	75	76	77	78	79	80
0.476	-2.281	-3.044	0.196	-0.186	-0.129	-0.320	0.451	-0.305	-0.316
81	82	83	84	85	86	87	88	89	90
0.046	-1.436	0.832	0.422	-0.994	-0.506	-5.0e+4	1.195	-0.202	-4.1e+4
91	92	93	94	95	96	97	98	99	100
-1.186	-5.4e+4	-3.3e+4	-0.432	-5.6e+4	-0.384	-0.654	-0.148	-0.401	-4.7e+4
101	102	103	104	105	106	107	108	109	110
-1.639	0.533	-1.831	-0.330	0.262	-0.493	-3.5e+4	0.436	-0.694	-4.6e+4
111	112	113	114	115	116	117	118	119	120
0.522	-2.573	-2.869	0.199	-0.189	-0.133	-0.327	0.457	-0.317	-0.325
121	122	123	124	125	126	127	128	129	130
0.047	-1.491	0.837	0.428	-1.044	-0.514	-5.5e+4	1.115	-0.204	-4.6e+4
131	132	133	134	135	136	137	138	139	140
-1.186	-5.4e+4	-3.3e+4	-0.432	-2.579	-0.390	-0.694	-0.151	-0.419	-5.2e+4
141	142	143	144	145	146	147	148	149	150

-1.762	0.544	-2.117	-0.342	0.269	-0.518	-3.8e+4	0.447	-0.843	-5.2e+4
151	152	153	154	155	156	157	158	159	160
0.558	-4.264	-6.7e+3	0.209	-0.201	-0.145	-0.347	0.473	-0.352	-0.353
161	162	163	164	165	166	167	168	169	170
0.048	-1.659	0.854	0.446	-1.214	-0.535	-5.6e+4	1.237	-0.212	-5.0e+4
171	172	173	174	175	176	177	178	179	180
-1.321	-5.4e+4	-4.7e+4	-0.462	-2.752	-0.390	-0.694	-0.151	-0.419	-5.2e+4
181	182	183	184	185	186	187	188	189	190
-1.762	0.544	-1.863	-0.342	0.229	-0.518	-3.8e+4	0.447	-0.843	-5.2e+4
191	192	193	194	195	196	197	198	199	200
0.558	-4.264	-8.8e+3	0.209	-0.201	-0.137	-0.333	0.462	-0.327	-0.336
201	202	203	204	205	206	207	208	209	210
0.048	-1.565	0.845	0.436	-1.112	-0.523	-5.1e+4	1.220	-0.208	-4.3e+4
211	212	213	214	215	216	217	218	219	220
-1.297	-4.9e+4	-4.6e+4	-0.447	-3.849	-0.387	-0.673	-0.150	-0.410	-4.5e+4

221	222	223	224	225	226	227	228	229	230
-1.700	0.539	-2.453	-0.336	0.265	-0.504	-3.5e+4	0.442	-0.755	-4.3e+4

231	232	233	234	235	236	237	238	239	240
0.539	-3.104	-1.9e+4	0.204	-0.195	-0.137	-0.333	0.462	-0.327	1.261

241	242
0.763	0.697

Final Coefficients of Theta

	1	2	3	4	5
1	0.748	0.0	0.0	0.0	0.0
2	1.268	0.0	0.0	0.0	0.0
3	0.541	0.0	0.0	0.0	0.0
4	0.789	0.0	0.0	0.0	0.0
5	1.472	0.0	0.0	0.0	0.0
6	0.789	0.0	0.0	0.0	0.0
7	5.1e+4	0.0	0.0	0.0	0.0
8	0.726	0.0	0.0	0.0	0.0
9	0.665	0.0	0.0	0.0	0.0
10	3.7e+4	0.0	0.0	0.0	0.0
11	1.199	0.0	0.0	0.0	0.0
12	4.4e+4	0.0	0.0	0.0	0.0
13	3.9e+4	0.0	0.0	0.0	0.0
14	0.960	0.0	0.0	0.0	0.0
15	1.761	0.0	0.0	0.0	0.0

16	0.501	0.0	0.0	0.0	0.0
17	1.004	0.0	0.0	0.0	0.0
18	0.633	0.0	0.0	0.0	0.0
19	0.873	0.0	0.0	0.0	0.0
20	3.6e+4	0.0	0.0	0.0	0.0
21	1.138	0.0	0.0	0.0	0.0
22	0.595	0.0	0.0	0.0	0.0
23	1.185	0.0	0.0	0.0	0.0
24	0.803	0.0	0.0	0.0	0.0
25	0.654	0.0	0.0	0.0	0.0
26	0.933	0.0	0.0	0.0	0.0
27	1.8e+4	0.0	0.0	0.0	0.0
28	0.667	0.0	0.0	0.0	0.0
29	1.763	0.0	0.0	0.0	0.0
30	3.5e+4	0.0	0.0	0.0	0.0
31	1.060	0.0	0.0	0.0	0.0
32	2.639	0.0	0.0	0.0	0.0
33	3.617	0.0	0.0	0.0	0.0
34	0.888	0.0	0.0	0.0	0.0
35	0.968	0.0	0.0	0.0	0.0
36	1.177	0.0	0.0	0.0	0.0
37	0.952	0.0	0.0	0.0	0.0
38	0.734	0.0	0.0	0.0	0.0
39	1.258	0.0	0.0	0.0	0.0
40	1.119	0.0	0.0	0.0	0.0
41	0.748	0.0	0.0	0.0	0.0
42	1.268	0.0	0.0	0.0	0.0
43	0.541	0.0	0.0	0.0	0.0
44	0.789	0.0	0.0	0.0	0.0
45	1.472	0.0	0.0	0.0	0.0
46	0.789	0.0	0.0	0.0	0.0

47	5.1e+4	0.0	0.0	0.0	0.0
48	0.629	0.0	0.0	0.0	0.0
49	0.684	0.0	0.0	0.0	0.0
50	3.7e+4	0.0	0.0	0.0	0.0
51	0.0	1.240	0.0	0.0	0.0
52	0.0	3.9e+4	0.0	0.0	0.0
53	0.0	2.4e+4	0.0	0.0	0.0
54	0.0	0.889	0.0	0.0	0.0
55	0.0	1.160	0.0	0.0	0.0
56	0.0	0.476	0.0	0.0	0.0
57	0.0	0.930	0.0	0.0	0.0
58	0.0	0.599	0.0	0.0	0.0
59	0.0	0.814	0.0	0.0	0.0
60	0.0	3.0e+4	0.0	0.0	0.0
61	0.0	1.043	0.0	0.0	0.0
62	0.0	0.563	0.0	0.0	0.0
63	0.0	1.164	0.0	0.0	0.0
64	0.0	0.751	0.0	0.0	0.0
65	0.0	0.615	0.0	0.0	0.0
66	0.0	0.864	0.0	0.0	0.0
67	0.0	1.7e+4	0.0	0.0	0.0
68	0.0	0.628	0.0	0.0	0.0
69	0.0	1.512	0.0	0.0	0.0
70	0.0	3.0e+4	0.0	0.0	0.0
71	0.0	0.913	0.0	0.0	0.0
72	0.0	2.038	0.0	0.0	0.0
73	0.0	2.253	0.0	0.0	0.0
74	0.0	0.826	0.0	0.0	0.0
75	0.0	0.895	0.0	0.0	0.0
76	0.0	1.068	0.0	0.0	0.0
77	0.0	0.879	0.0	0.0	0.0

78	0.0	0.688	0.0	0.0	0.0
79	0.0	1.135	0.0	0.0	0.0
80	0.0	1.022	0.0	0.0	0.0
81	0.0	0.709	0.0	0.0	0.0
82	0.0	1.167	0.0	0.0	0.0
83	0.0	0.517	0.0	0.0	0.0
84	0.0	0.745	0.0	0.0	0.0
85	0.0	1.328	0.0	0.0	0.0
86	0.0	0.746	0.0	0.0	0.0
87	0.0	4.4e+4	0.0	0.0	0.0
88	0.0	0.598	0.0	0.0	0.0
89	0.0	0.719	0.0	0.0	0.0
90	0.0	3.0e+4	0.0	0.0	0.0
91	0.0	0.0	1.471	0.0	0.0
92	0.0	0.0	4.4e+4	0.0	0.0
93	0.0	0.0	3.2e+4	0.0	0.0
94	0.0	0.0	0.918	0.0	0.0
95	0.0	0.0	2.3e+4	0.0	0.0
96	0.0	0.0	0.490	0.0	0.0
97	0.0	0.0	0.977	0.0	0.0
98	0.0	0.0	0.619	0.0	0.0
99	0.0	0.0	0.851	0.0	0.0
100	0.0	0.0	3.5e+4	0.0	0.0
101	0.0	0.0	1.099	0.0	0.0
102	0.0	0.0	0.582	0.0	0.0
103	0.0	0.0	1.098	0.0	0.0
104	0.0	0.0	0.782	0.0	0.0
105	0.0	0.0	0.636	0.0	0.0
106	0.0	0.0	0.907	0.0	0.0
107	0.0	0.0	1.8e+4	0.0	0.0
108	0.0	0.0	0.651	0.0	0.0

109	0.0	0.0	1.659	0.0	0.0
110	0.0	0.0	3.3e+4	0.0	0.0
111	0.0	0.0	1.023	0.0	0.0
112	0.0	0.0	2.358	0.0	0.0
113	0.0	0.0	2.079	0.0	0.0
114	0.0	0.0	0.863	0.0	0.0
115	0.0	0.0	0.935	0.0	0.0
116	0.0	0.0	1.132	0.0	0.0
117	0.0	0.0	0.921	0.0	0.0
118	0.0	0.0	0.715	0.0	0.0
119	0.0	0.0	1.210	0.0	0.0
120	0.0	0.0	1.079	0.0	0.0
121	0.0	0.0	0.736	0.0	0.0
122	0.0	0.0	1.244	0.0	0.0
123	0.0	0.0	0.534	0.0	0.0
124	0.0	0.0	0.775	0.0	0.0
125	0.0	0.0	1.432	0.0	0.0
126	0.0	0.0	0.779	0.0	0.0
127	0.0	0.0	4.8e+4	0.0	0.0
128	0.0	0.0	0.507	0.0	0.0
129	0.0	0.0	0.746	0.0	0.0
130	0.0	0.0	3.3e+4	0.0	0.0
131	0.0	0.0	1.471	0.0	0.0
132	0.0	0.0	4.4e+4	0.0	0.0
133	0.0	0.0	3.2e+4	0.0	0.0
134	0.0	0.0	0.918	0.0	0.0
135	0.0	0.0	1.849	0.0	0.0
136	0.0	0.0	0.0	0.528	0.0
137	0.0	0.0	0.0	1.097	0.0
138	0.0	0.0	0.0	0.672	0.0
139	0.0	0.0	0.0	0.940	0.0

140	0.0	0.0	0.0	3.8e+4	0.0
141	0.0	0.0	0.0	1.246	0.0
142	0.0	0.0	0.0	0.629	0.0
143	0.0	0.0	0.0	1.382	0.0
144	0.0	0.0	0.0	0.858	0.0
145	0.0	0.0	0.0	0.692	0.0
146	0.0	0.0	0.0	1.006	0.0
147	0.0	0.0	0.0	1.9e+4	0.0
148	0.0	0.0	0.0	0.707	0.0
149	0.0	0.0	0.0	2.129	0.0
150	0.0	0.0	0.0	3.8e+4	0.0
151	0.0	0.0	0.0	1.157	0.0
152	0.0	0.0	0.0	4.124	0.0
153	0.0	0.0	0.0	5.4e+3	0.0
154	0.0	0.0	0.0	0.957	0.0
155	0.0	0.0	0.0	1.050	0.0
156	0.0	0.0	0.0	1.319	0.0
157	0.0	0.0	0.0	1.039	0.0
158	0.0	0.0	0.0	0.788	0.0
159	0.0	0.0	0.0	1.429	0.0
160	0.0	0.0	0.0	1.247	0.0
161	0.0	0.0	0.0	0.812	0.0
162	0.0	0.0	0.0	1.467	0.0
163	0.0	0.0	0.0	0.579	0.0
164	0.0	0.0	0.0	0.859	0.0
165	0.0	0.0	0.0	1.767	0.0
166	0.0	0.0	0.0	0.859	0.0
167	0.0	0.0	0.0	4.9e+4	0.0
168	0.0	0.0	0.0	0.674	0.0
169	0.0	0.0	0.0	0.825	0.0
170	0.0	0.0	0.0	3.6e+4	0.0

171	0.0	0.0	0.0	1.606	0.0
172	0.0	0.0	0.0	4.4e+4	0.0
173	0.0	0.0	0.0	4.5e+4	0.0
174	0.0	0.0	0.0	1.051	0.0
175	0.0	0.0	0.0	2.003	0.0
176	0.0	0.0	0.0	0.528	0.0
177	0.0	0.0	0.0	1.097	0.0
178	0.0	0.0	0.0	0.672	0.0
179	0.0	0.0	0.0	0.940	0.0
180	0.0	0.0	0.0	3.8e+4	0.0
181	0.0	0.0	0.0	1.246	0.0
182	0.0	0.0	0.0	0.629	0.0
183	0.0	0.0	0.0	1.132	0.0
184	0.0	0.0	0.0	0.858	0.0
185	0.0	0.0	0.0	0.614	0.0
186	0.0	0.0	0.0	1.006	0.0
187	0.0	0.0	0.0	1.9e+4	0.0
188	0.0	0.0	0.0	0.707	0.0
189	0.0	0.0	0.0	2.129	0.0
190	0.0	0.0	0.0	3.8e+4	0.0
191	0.0	0.0	0.0	1.157	0.0
192	0.0	0.0	0.0	4.124	0.0
193	0.0	0.0	0.0	7.1e+3	0.0
194	0.0	0.0	0.0	0.957	0.0
195	0.0	0.0	0.0	1.050	0.0
196	0.0	0.0	0.0	0.0	1.190
197	0.0	0.0	0.0	0.0	0.957
198	0.0	0.0	0.0	0.0	0.737
199	0.0	0.0	0.0	0.0	1.272
200	0.0	0.0	0.0	0.0	1.145
201	0.0	0.0	0.0	0.0	0.772

202	0.0	0.0	0.0	0.0	1.343
203	0.0	0.0	0.0	0.0	0.555
204	0.0	0.0	0.0	0.0	0.813
205	0.0	0.0	0.0	0.0	1.569
206	0.0	0.0	0.0	0.0	0.814
207	0.0	0.0	0.0	0.0	4.5e+4
208	0.0	0.0	0.0	0.0	0.645
209	0.0	0.0	0.0	0.0	0.783
210	0.0	0.0	0.0	0.0	3.1e+4
211	0.0	0.0	0.0	0.0	1.669
212	0.0	0.0	0.0	0.0	4.0e+4
213	0.0	0.0	0.0	0.0	4.5e+4
214	0.0	0.0	0.0	0.0	0.985
215	0.0	0.0	0.0	0.0	2.969
216	0.0	0.0	0.0	0.0	0.509
217	0.0	0.0	0.0	0.0	1.034
218	0.0	0.0	0.0	0.0	0.648
219	0.0	0.0	0.0	0.0	0.894
220	0.0	0.0	0.0	0.0	3.3e+4
221	0.0	0.0	0.0	0.0	1.173
222	0.0	0.0	0.0	0.0	0.606
223	0.0	0.0	0.0	0.0	1.712
224	0.0	0.0	0.0	0.0	0.817
225	0.0	0.0	0.0	0.0	0.662
226	0.0	0.0	0.0	0.0	0.952
227	0.0	0.0	0.0	0.0	1.7e+4
228	0.0	0.0	0.0	0.0	0.680
229	0.0	0.0	0.0	0.0	1.852
230	0.0	0.0	0.0	0.0	3.2e+4
231	0.0	0.0	0.0	0.0	1.088
232	0.0	0.0	0.0	0.0	2.923

233	0.0	0.0	0.0	0.0	1.5e+4
234	0.0	0.0	0.0	0.0	0.911
235	0.0	0.0	0.0	0.0	0.993
236	0.0	0.0	0.0	0.0	1.190
237	0.0	0.0	0.0	0.0	0.957
238	0.0	0.0	0.0	0.0	0.737
239	0.0	0.0	0.0	0.0	1.272
240	0.0	0.0	0.0	0.0	-0.108
241	0.0	0.0	0.0	0.0	-0.144
242	0.0	0.0	0.0	0.0	-0.137

Final Correlations of Theta

	1	2	3	4	5
1	1.000				
2	1.497	1.000			
3	1.467	1.518	1.000		
4	1.386	1.438	1.406	1.000	
5	1.420	1.472	1.441	1.363	1.000

Sum of squares of residuals (lower off-diagonals) = 25.491

Root mean square of residuals (lower off-diagonals) = 0.0695662

Tanaka index of goodness of fit = 0.7452359

Threshold Values

1	2	3	4	5	6	7	8	9	10

0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.138	0.935	-0.113	-1.388
11	12	13	14	15	16	17	18	19	20
-0.605	-1.227	-1.036	-0.319	-1.232	-0.345	-0.468	-0.126	-0.305	-1.356
21	22	23	24	25	26	27	28	29	30
-1.103	0.461	-1.248	-0.260	0.221	-0.365	-2.005	0.365	-0.358	-1.372
31	32	33	34	35	36	37	38	39	40
0.365	-1.005	-1.242	0.151	-0.138	-0.088	-0.240	0.372	-0.202	-0.221
41	42	43	44	45	46	47	48	49	50
0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.138	1.026	-0.189	-1.388
51	52	53	54	55	56	57	58	59	60
-0.698	-1.227	-1.036	-0.319	-1.231	-0.345	-0.468	-0.126	-0.305	-1.356
61	62	63	64	65	66	67	68	69	70
-1.103	0.461	-1.236	-0.260	0.221	-0.365	-2.005	0.365	-0.358	-1.405
71	72	73	74	75	76	77	78	79	80
0.352	-1.005	-1.235	0.151	-0.138	-0.088	-0.240	0.372	-0.202	-0.221
81	82	83	84	85	86	87	88	89	90
0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.138	1.026	-0.164	-1.388

91	92	93	94	95	96	97	98	99	100
-0.667	-1.227	-1.036	-0.319	-2.432	-0.345	-0.468	-0.126	-0.305	-1.356
101	102	103	104	105	106	107	108	109	110
-1.103	0.461	-1.233	-0.260	0.221	-0.365	-2.005	0.365	-0.358	-1.372
111	112	113	114	115	116	117	118	119	120
0.365	-1.005	-1.244	0.151	-0.138	-0.088	-0.240	0.372	-0.202	-0.221
121	122	123	124	125	126	127	128	129	130
0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.150	0.994	-0.164	-1.388
131	132	133	134	135	136	137	138	139	140
-0.667	-1.227	-1.036	-0.319	-1.227	-0.345	-0.468	-0.126	-0.305	-1.356
141	142	143	144	145	146	147	148	149	150
-1.103	0.461	-1.241	-0.260	0.221	-0.365	-2.005	0.365	-0.358	-1.372
151	152	153	154	155	156	157	158	159	160
0.365	-1.005	-1.230	0.151	-0.138	-0.088	-0.240	0.372	-0.202	-0.221
161	162	163	164	165	166	167	168	169	170
0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.138	1.026	-0.164	-1.388

171	172	173	174	175	176	177	178	179	180
-0.698	-1.227	-1.036	-0.319	-1.230	-0.345	-0.468	-0.126	-0.305	-1.356
181	182	183	184	185	186	187	188	189	190
-1.103	0.461	-1.233	-0.260	0.196	-0.365	-2.005	0.365	-0.358	-1.372
191	192	193	194	195	196	197	198	199	200
0.365	-1.005	-1.235	0.151	-0.138	-0.088	-0.240	0.372	-0.202	-0.221
201	202	203	204	205	206	207	208	209	210
0.038	-0.935	0.739	0.338	-0.598	-0.406	-1.138	1.026	-0.164	-1.388
211	212	213	214	215	216	217	218	219	220
-0.667	-1.227	-1.036	-0.319	-1.228	-0.345	-0.468	-0.126	-0.305	-1.356
221	222	223	224	225	226	227	228	229	230
-1.103	0.461	-1.237	-0.260	0.221	-0.365	-2.005	0.365	-0.358	-1.372
231	232	233	234	235	236	237	238	239	240
0.365	-1.005	-1.244	0.151	-0.138	-0.088	-0.240	0.372	-0.202	1.254
241	242								
0.755	0.690								

Factor Loadings

	1	2	3	4	5
1	0.599	0.00	0.00	0.00	0.00
2	0.785	0.00	0.00	0.00	0.00
3	0.476	0.00	0.00	0.00	0.00
4	0.619	0.00	0.00	0.00	0.00
5	0.827	0.00	0.00	0.00	0.00
6	0.619	0.00	0.00	0.00	0.00
7	1.000	0.00	0.00	0.00	0.00
8	0.587	0.00	0.00	0.00	0.00
9	0.553	0.00	0.00	0.00	0.00
10	1.000	0.00	0.00	0.00	0.00
11	0.768	0.00	0.00	0.00	0.00
12	1.000	0.00	0.00	0.00	0.00
13	1.000	0.00	0.00	0.00	0.00
14	0.693	0.00	0.00	0.00	0.00
15	0.870	0.00	0.00	0.00	0.00
16	0.448	0.00	0.00	0.00	0.00
17	0.708	0.00	0.00	0.00	0.00
18	0.535	0.00	0.00	0.00	0.00
19	0.658	0.00	0.00	0.00	0.00
20	1.000	0.00	0.00	0.00	0.00
21	0.751	0.00	0.00	0.00	0.00
22	0.511	0.00	0.00	0.00	0.00
23	0.764	0.00	0.00	0.00	0.00
24	0.626	0.00	0.00	0.00	0.00
25	0.547	0.00	0.00	0.00	0.00
26	0.682	0.00	0.00	0.00	0.00
27	1.000	0.00	0.00	0.00	0.00

28 0.555 0.00 0.00 0.00 0.00
29 0.870 0.00 0.00 0.00 0.00
30 1.000 0.00 0.00 0.00 0.00
31 0.727 0.00 0.00 0.00 0.00
32 0.935 0.00 0.00 0.00 0.00
33 0.964 0.00 0.00 0.00 0.00
34 0.664 0.00 0.00 0.00 0.00
35 0.696 0.00 0.00 0.00 0.00
36 0.762 0.00 0.00 0.00 0.00
37 0.690 0.00 0.00 0.00 0.00
38 0.592 0.00 0.00 0.00 0.00
39 0.783 0.00 0.00 0.00 0.00
40 0.746 0.00 0.00 0.00 0.00
41 0.599 0.00 0.00 0.00 0.00
42 0.785 0.00 0.00 0.00 0.00
43 0.476 0.00 0.00 0.00 0.00
44 0.619 0.00 0.00 0.00 0.00
45 0.827 0.00 0.00 0.00 0.00
46 0.619 0.00 0.00 0.00 0.00
47 1.000 0.00 0.00 0.00 0.00
48 0.532 0.00 0.00 0.00 0.00
49 0.564 0.00 0.00 0.00 0.00
50 0.599 0.00 0.00 0.00 0.00
51 0.00 1.000 0.00 0.00 0.00
52 0.00 0.778 0.00 0.00 0.00
53 0.00 1.000 0.00 0.00 0.00
54 0.00 1.000 0.00 0.00 0.00
55 0.00 0.664 0.00 0.00 0.00
56 0.00 0.758 0.00 0.00 0.00
57 0.00 0.430 0.00 0.00 0.00
58 0.00 0.681 0.00 0.00 0.00

59	0.00	0.514	0.00	0.00	0.00
60	0.00	0.631	0.00	0.00	0.00
61	0.00	1.000	0.00	0.00	0.00
62	0.00	0.722	0.00	0.00	0.00
63	0.00	0.491	0.00	0.00	0.00
64	0.00	0.759	0.00	0.00	0.00
65	0.00	0.600	0.00	0.00	0.00
66	0.00	0.524	0.00	0.00	0.00
67	0.00	0.654	0.00	0.00	0.00
68	0.00	1.000	0.00	0.00	0.00
69	0.00	0.532	0.00	0.00	0.00
70	0.00	0.834	0.00	0.00	0.00
71	0.00	1.000	0.00	0.00	0.00
72	0.00	0.674	0.00	0.00	0.00
73	0.00	0.898	0.00	0.00	0.00
74	0.00	0.914	0.00	0.00	0.00
75	0.00	0.637	0.00	0.00	0.00
76	0.00	0.667	0.00	0.00	0.00
77	0.00	0.730	0.00	0.00	0.00
78	0.00	0.660	0.00	0.00	0.00
79	0.00	0.567	0.00	0.00	0.00
80	0.00	0.750	0.00	0.00	0.00
81	0.00	0.715	0.00	0.00	0.00
82	0.00	0.578	0.00	0.00	0.00
83	0.00	0.759	0.00	0.00	0.00
84	0.00	0.459	0.00	0.00	0.00
85	0.00	0.597	0.00	0.00	0.00
86	0.00	0.799	0.00	0.00	0.00
87	0.00	0.598	0.00	0.00	0.00
88	0.00	1.000	0.00	0.00	0.00
89	0.00	0.513	0.00	0.00	0.00

90	0.00	0.584	0.00	0.00	0.00
91	0.00	0.00	1.000	0.00	0.00
92	0.00	0.00	0.827	0.00	0.00
93	0.00	0.00	1.000	0.00	0.00
94	0.00	0.00	1.000	0.00	0.00
95	0.00	0.00	0.676	0.00	0.00
96	0.00	0.00	1.000	0.00	0.00
97	0.00	0.00	0.440	0.00	0.00
98	0.00	0.00	0.699	0.00	0.00
99	0.00	0.00	0.526	0.00	0.00
100	0.00	0.00	0.648	0.00	0.00
101	0.00	0.00	1.000	0.00	0.00
102	0.00	0.00	0.740	0.00	0.00
103	0.00	0.00	0.503	0.00	0.00
104	0.00	0.00	0.739	0.00	0.00
105	0.00	0.00	0.616	0.00	0.00
106	0.00	0.00	0.537	0.00	0.00
107	0.00	0.00	0.672	0.00	0.00
108	0.00	0.00	1.000	0.00	0.00
109	0.00	0.00	0.546	0.00	0.00
110	0.00	0.00	0.856	0.00	0.00
111	0.00	0.00	1.000	0.00	0.00
112	0.00	0.00	0.715	0.00	0.00
113	0.00	0.00	0.921	0.00	0.00
114	0.00	0.00	0.901	0.00	0.00
115	0.00	0.00	0.653	0.00	0.00
116	0.00	0.00	0.683	0.00	0.00
117	0.00	0.00	0.750	0.00	0.00
118	0.00	0.00	0.677	0.00	0.00
119	0.00	0.00	0.582	0.00	0.00
120	0.00	0.00	0.771	0.00	0.00

121	0.00	0.00	0.733	0.00	0.00
122	0.00	0.00	0.593	0.00	0.00
123	0.00	0.00	0.779	0.00	0.00
124	0.00	0.00	0.471	0.00	0.00
125	0.00	0.00	0.613	0.00	0.00
126	0.00	0.00	0.820	0.00	0.00
127	0.00	0.00	0.614	0.00	0.00
128	0.00	0.00	1.000	0.00	0.00
129	0.00	0.00	0.452	0.00	0.00
130	0.00	0.00	0.598	0.00	0.00
131	0.00	0.00	1.000	0.00	0.00
132	0.00	0.00	0.827	0.00	0.00
133	0.00	0.00	1.000	0.00	0.00
134	0.00	0.00	1.000	0.00	0.00
135	0.00	0.00	0.676	0.00	0.00
136	0.00	0.00	0.00	0.880	0.00
137	0.00	0.00	0.00	0.467	0.00
138	0.00	0.00	0.00	0.739	0.00
139	0.00	0.00	0.00	0.558	0.00
140	0.00	0.00	0.00	0.685	0.00
141	0.00	0.00	0.00	1.000	0.00
142	0.00	0.00	0.00	0.780	0.00
143	0.00	0.00	0.00	0.532	0.00
144	0.00	0.00	0.00	0.810	0.00
145	0.00	0.00	0.00	0.651	0.00
146	0.00	0.00	0.00	0.569	0.00
147	0.00	0.00	0.00	0.709	0.00
148	0.00	0.00	0.00	1.000	0.00
149	0.00	0.00	0.00	0.577	0.00
150	0.00	0.00	0.00	0.905	0.00
151	0.00	0.00	0.00	1.000	0.00

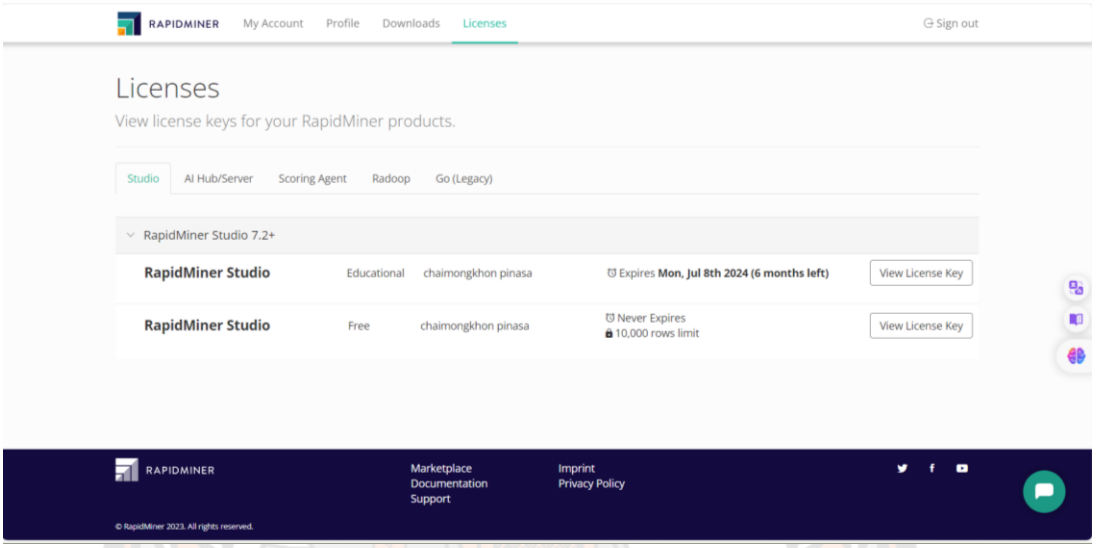
152	0.00	0.00	0.00	0.757	0.00
153	0.00	0.00	0.00	0.972	0.00
154	0.00	0.00	0.00	1.000	0.00
155	0.00	0.00	0.00	0.691	0.00
156	0.00	0.00	0.00	0.724	0.00
157	0.00	0.00	0.00	0.797	0.00
158	0.00	0.00	0.00	0.721	0.00
159	0.00	0.00	0.00	0.619	0.00
160	0.00	0.00	0.00	0.819	0.00
161	0.00	0.00	0.00	0.780	0.00
162	0.00	0.00	0.00	0.630	0.00
163	0.00	0.00	0.00	0.826	0.00
164	0.00	0.00	0.00	0.501	0.00
165	0.00	0.00	0.00	0.652	0.00
166	0.00	0.00	0.00	0.870	0.00
167	0.00	0.00	0.00	0.651	0.00
168	0.00	0.00	0.00	1.000	0.00
169	0.00	0.00	0.00	0.559	0.00
170	0.00	0.00	0.00	0.637	0.00
171	0.00	0.00	0.00	1.000	0.00
172	0.00	0.00	0.00	0.849	0.00
173	0.00	0.00	0.00	1.000	0.00
174	0.00	0.00	0.00	1.000	0.00
175	0.00	0.00	0.00	0.724	0.00
176	0.00	0.00	0.00	0.895	0.00
177	0.00	0.00	0.00	0.467	0.00
178	0.00	0.00	0.00	0.739	0.00
179	0.00	0.00	0.00	0.558	0.00
180	0.00	0.00	0.00	0.685	0.00
181	0.00	0.00	0.00	1.000	0.00
182	0.00	0.00	0.00	0.780	0.00

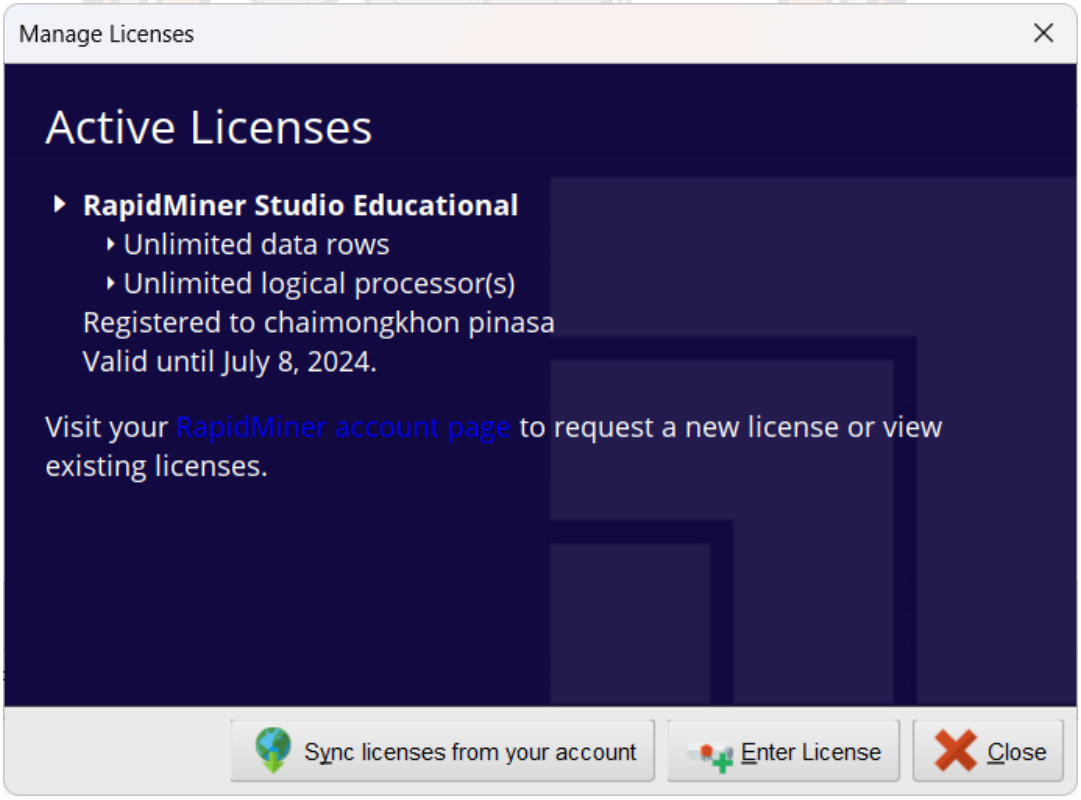
183	0.00	0.00	0.00	0.532	0.00
184	0.00	0.00	0.00	0.749	0.00
185	0.00	0.00	0.00	0.651	0.00
186	0.00	0.00	0.00	0.523	0.00
187	0.00	0.00	0.00	0.709	0.00
188	0.00	0.00	0.00	1.000	0.00
189	0.00	0.00	0.00	0.577	0.00
190	0.00	0.00	0.00	0.905	0.00
191	0.00	0.00	0.00	1.000	0.00
192	0.00	0.00	0.00	0.757	0.00
193	0.00	0.00	0.00	0.972	0.00
194	0.00	0.00	0.00	1.000	0.00
195	0.00	0.00	0.00	0.691	0.00
196	0.00	0.00	0.00	0.00	0.724
197	0.00	0.00	0.00	0.00	0.766
198	0.00	0.00	0.00	0.00	0.691
199	0.00	0.00	0.00	0.00	0.593
200	0.00	0.00	0.00	0.00	0.786
201	0.00	0.00	0.00	0.00	0.753
202	0.00	0.00	0.00	0.00	0.611
203	0.00	0.00	0.00	0.00	0.802
204	0.00	0.00	0.00	0.00	0.486
205	0.00	0.00	0.00	0.00	0.631
206	0.00	0.00	0.00	0.00	0.843
207	0.00	0.00	0.00	0.00	0.631
208	0.00	0.00	0.00	0.00	1.000
209	0.00	0.00	0.00	0.00	0.542
210	0.00	0.00	0.00	0.00	0.617
211	0.00	0.00	0.00	0.00	1.000
212	0.00	0.00	0.00	0.00	0.858
213	0.00	0.00	0.00	0.00	1.000

214	0.00	0.00	0.00	0.00	1.000
215	0.00	0.00	0.00	0.00	0.702
216	0.00	0.00	0.00	0.00	0.948
217	0.00	0.00	0.00	0.00	0.454
218	0.00	0.00	0.00	0.00	0.719
219	0.00	0.00	0.00	0.00	0.544
220	0.00	0.00	0.00	0.00	0.667
221	0.00	0.00	0.00	0.00	1.000
222	0.00	0.00	0.00	0.00	0.761
223	0.00	0.00	0.00	0.00	0.518
224	0.00	0.00	0.00	0.00	0.863
225	0.00	0.00	0.00	0.00	0.633
226	0.00	0.00	0.00	0.00	0.552
227	0.00	0.00	0.00	0.00	0.689
228	0.00	0.00	0.00	0.00	1.000
229	0.00	0.00	0.00	0.00	0.562
230	0.00	0.00	0.00	0.00	0.880
231	0.00	0.00	0.00	0.00	1.000
232	0.00	0.00	0.00	0.00	0.736
233	0.00	0.00	0.00	0.00	0.946
234	0.00	0.00	0.00	0.00	1.000
235	0.00	0.00	0.00	0.00	0.673
236	0.00	0.00	0.00	0.00	0.705
237	0.00	0.00	0.00	0.00	0.766
238	0.00	0.00	0.00	0.00	0.691
239	0.00	0.00	0.00	0.00	0.593
240	0.00	0.00	0.00	0.00	0.786
241	0.00	0.00	0.00	0.00	0.808
242	0.00	0.00	0.00	0.00	0.743

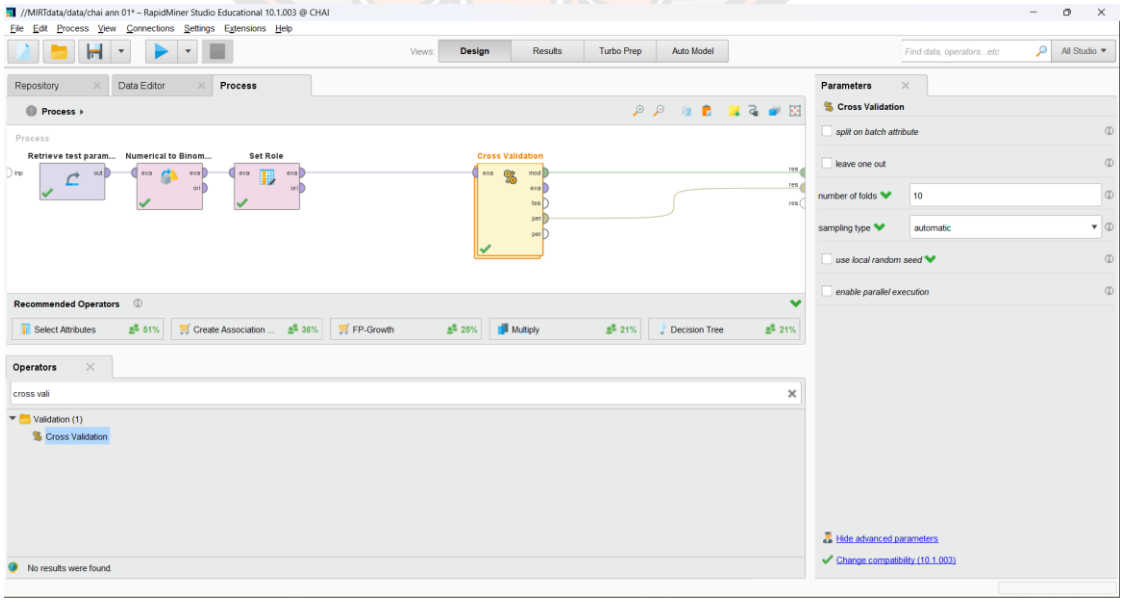
ภาคผนวก ค
การวิเคราะห์โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยใช้โปรแกรม RapidMiner Studio Educational 10.1)

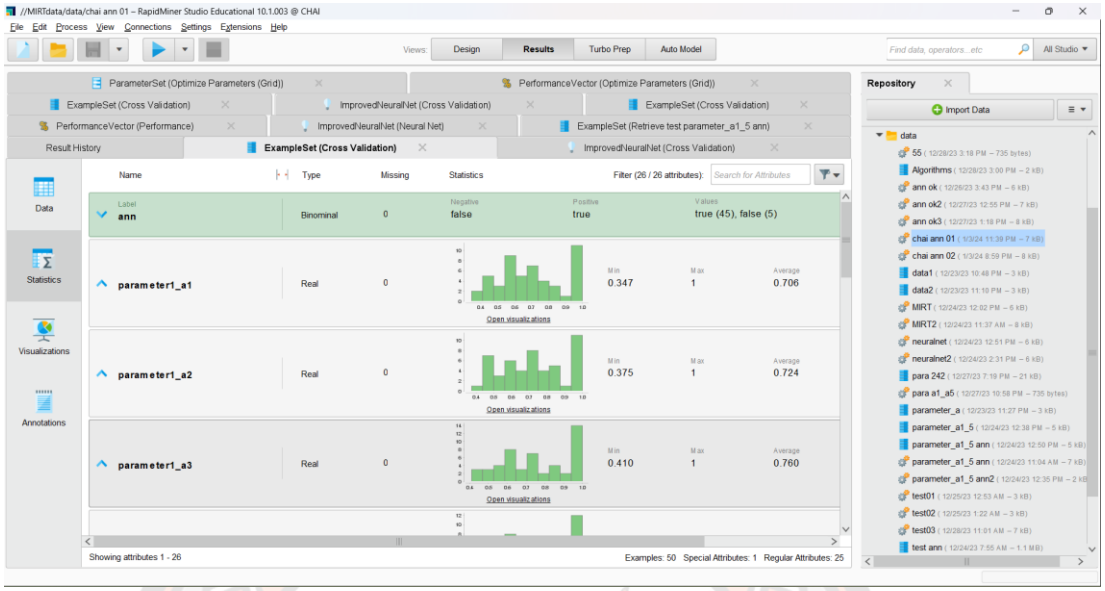
เอกสารแสดงใบอนุญาตการใช้โปรแกรม RapidMiner Studio Educational เพื่อการศึกษา Educational version จะใช้งานแบบไม่จำกัดจำนวนแถวข้อมูล ได้ 1 ปี (เมื่อครบแล้วสามารถขอต่อได้) ลงชื่อเข้าใช้ด้วยบัญชี : chaimongkhonp63@nu.ac.th

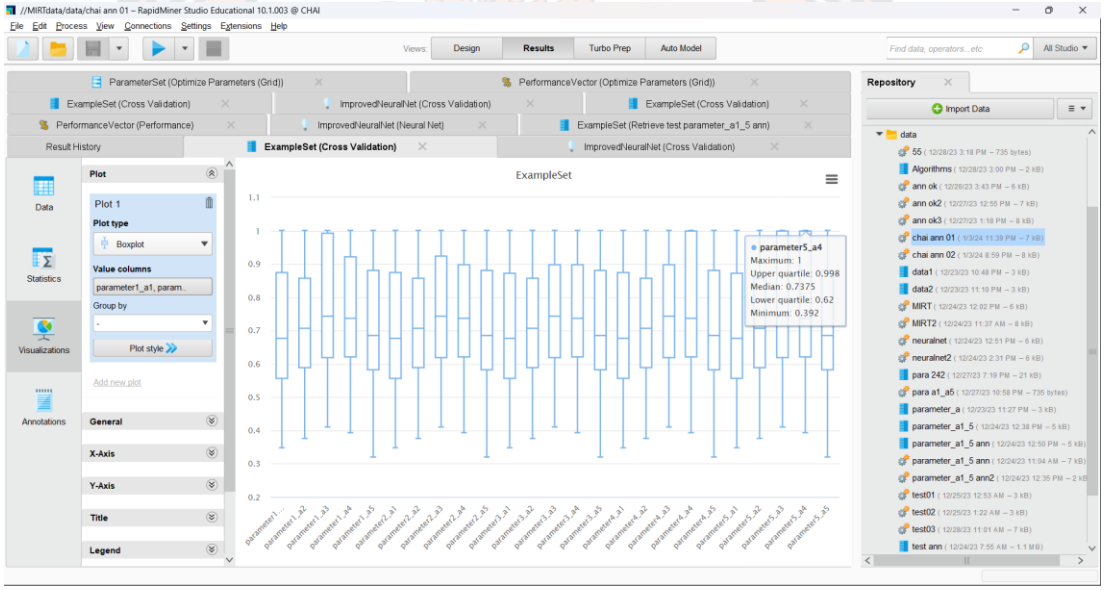




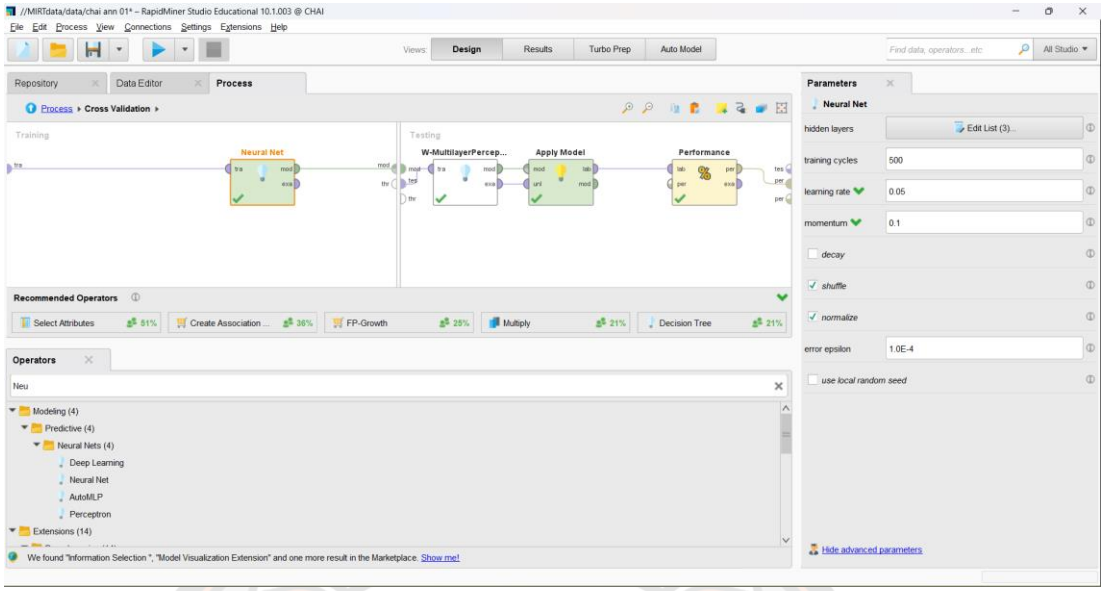
ผู้วิจัยทดสอบประสิทธิภาพในการจำแนกข้อมูลโดยการวัดความถูกต้องของอัลกอริธึม Neural Net ด้วยวิธีการ 10-fold cross-validation ซึ่งจะแบ่งข้อมูลออกเป็น 10 ชุดเท่าๆ กันแล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ (Test set) 20% ส่วนที่เหลือ 9 ชุด นำมาใช้เป็นกลุ่มเรียนรู้ (Training Set) 80% แล้วทำการ วนทำเป็นจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อยๆจนครบจำนวนที่แบ่งไว้ แสดงดังภาพ

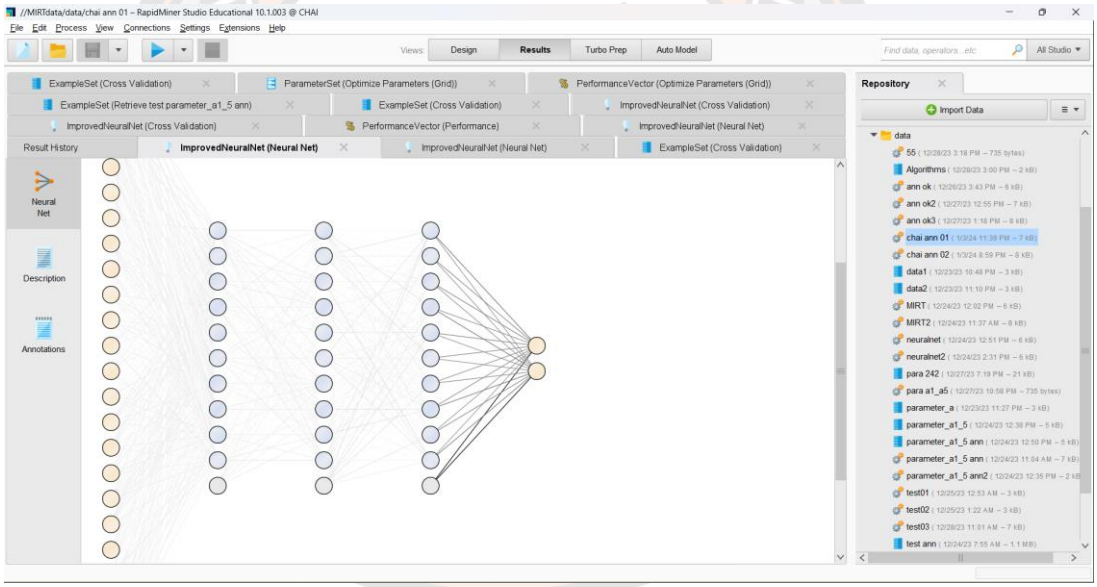


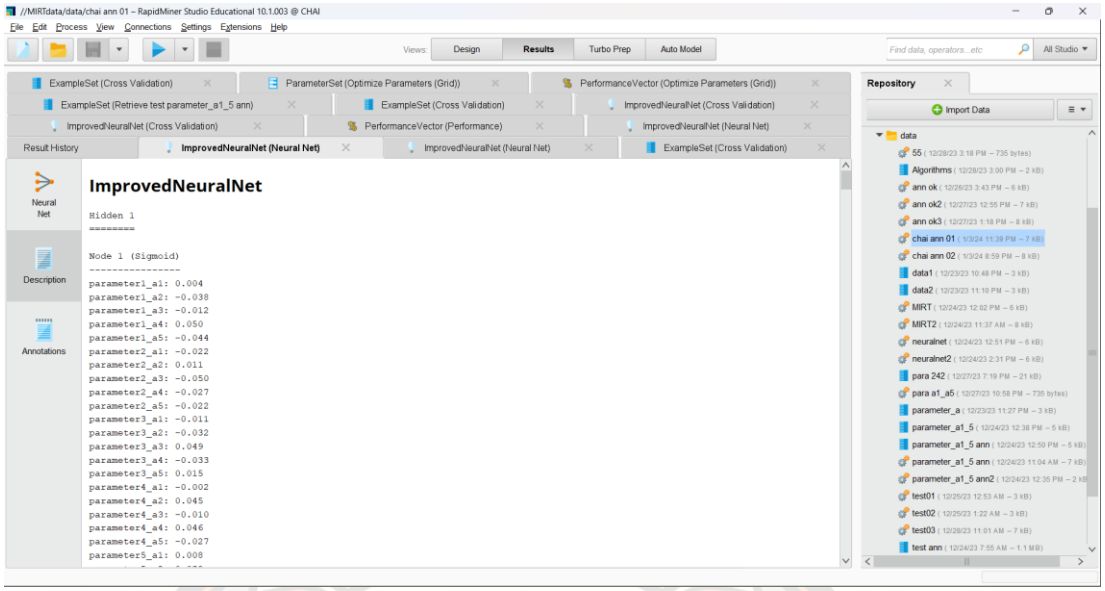




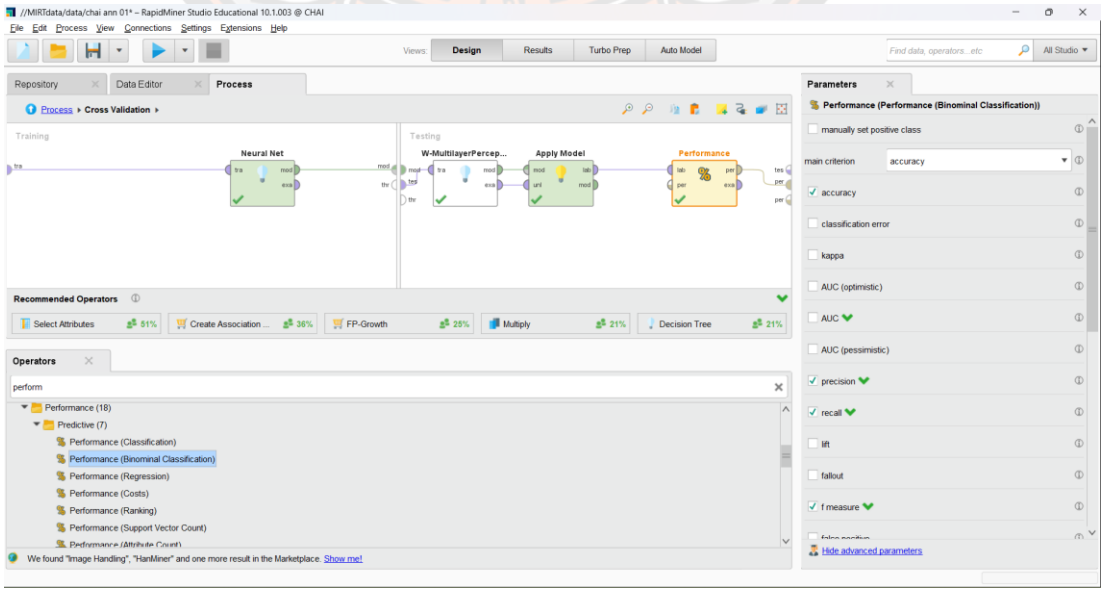
ผู้วิจัยได้กำหนดค่าพารามิเตอร์ในกระบวนการของโครงข่ายประสาทเทียม ประกอบด้วย อัตราการเรียนรู้ (learning rate) เท่ากับ 0.05 โมเมนตัม (Momentum) เท่ากับ 0.01 ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ แสดงดังภาพ


//MIRTdata/data/chai ann 01* – RapidMiner Studio Educational 10.1.003 @ CHAI
File Edit Process View Connections Settings Extensions Help
Views: Design Results Turbo Prep Auto Model
Repository
Data Editor
Process
Process ▸ Cross Validation ▸
Training
Testing
Neural Net
W-MultilayerPercep...
Apply Model
Performance
Recommended Operators
Select Attributes 51%
Create Association ... 36%
FP-Growth 25%
Multiply 21%
Decision Tree 21%
Operators
Neu
Modeling (4)
Predictive (4)
Neural Nets (4)
Deep Learning
Neural Net
AutoMLP
Perceptron
Extensions (14)
We found "Information Selection ", "Model Visualization Extension" and one more result in the Marketplace. Show me!
Parameters
Neural Net
hidden layers
Edit List (3)...
training cycles
500
learning rate
0.05
momentum
0.1
decay
shuffle
normalize
error epsilon
1.0E-4
use local random seed
Hide advanced parameters


//MIRTdata/data/chai ann 01 – RapidMiner Studio Educational 10.1.003 @ CHAI
File Edit Process View Connections Settings Extensions Help
Views: Design Results Turbo Prep Auto Model
ExampleSet (Cross Validation)
ParameterSet (Optimize Parameters (Grid))
PerformanceVector (Optimize Parameters (Grid))
ExampleSet (Retrieve test parameter_a1_5 ann)
ExampleSet (Cross Validation)
ImprovedNeuralNet (Cross Validation)
ImprovedNeuralNet (Cross Validation)
PerformanceVector (Performance)
ImprovedNeuralNet (Neural Net)
Result History
ImprovedNeuralNet (Neural Net)
ImprovedNeuralNet (Neural Net)
ExampleSet (Cross Validation)
Neural Net
Description
Annotations
Repository
Import Data
data
55
Algorithms
ann ok
ann ok2
ann ok3
chai ann 01
chai ann 02
data1
data2
MIRT
MIRT2
neuralnet
neuralnet2
para 242
para a1_a5
parameter_a
parameter_a1_5
parameter_a1_5 ann
parameter_a1_5 ann
parameter_a1_5 ann2
test01
test02
test03
test ann



การนำโมเดลไปใช้งานจริงได้นั้น จำเป็นจะต้องทราบประสิทธิภาพของโมเดลก่อน โดยทั่วไปแล้วจะมีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่างๆ อยู่ 4 ค่า ดังนี้ 1) ค่าความแม่นยำ (Precision) คือค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์ 2) ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล 3) ค่าความถ่วงดุล (F-measure) คือค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก และ 4) ค่าความถูกต้อง (Accuracy) คือ จำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส โดยต้องมีค่าสูงที่สุด แสดงดังภาพ



ผลการวิเคราะห์ประสิทธิภาพของโมเดลโครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

PerformanceVector

PerformanceVector:
accuracy: 90.00% +/- 10.54% (micro average: 90.00%)
ConfusionMatrix:
True: false true
false: 0 0
true: 5 45
precision: 90.00% +/- 10.54% (micro average: 90.00%) (positive class: true)
ConfusionMatrix:
True: false true
false: 0 0
true: 5 45
recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: true)
ConfusionMatrix:
True: false true
false: 0 0
true: 5 45
f_measure: 94.44% +/- 5.86% (micro average: 94.74%) (positive class: true)
ConfusionMatrix:
True: false true
false: 0 0
true: 5 45

ImprovedNeuralNet

ImprovedNeuralNet

Hidden 1

========

Node 1 (Sigmoid)

---

parameter1_a1: 0.004

parameter1_a2: -0.038

parameter1_a3: -0.012

parameter1_a4: 0.050

parameter1_a5: -0.044

parameter2_a1: -0.022

parameter2_a2: 0.011

parameter2_a3: -0.050

parameter2_a4: -0.027

parameter2_a5: -0.022

parameter3_a1: -0.011

parameter3_a2: -0.032

parameter3_a3: 0.049

parameter3_a4: -0.033

parameter3_a5: 0.015

parameter4_a1: -0.002

parameter4_a2: 0.045

parameter4_a3: -0.010

parameter4_a4: 0.046

parameter4_a5: -0.027

parameter5_a1: 0.008

parameter5_a2: 0.039

parameter5_a3: 0.003

parameter5_a4: -0.036

parameter5_a5: 0.046

Bias: 0.048

Node 2 (Sigmoid)

---

parameter1_a1: 0.039

parameter1_a2: 0.023

parameter1_a3: 0.000

parameter1_a4: 0.047

parameter1_a5: -0.027

parameter2_a1: -0.045

parameter2_a2: 0.010

parameter2_a3: -0.026

parameter2_a4: 0.012

parameter2_a5: 0.030

parameter3_a1: 0.026

parameter3_a2: 0.011

parameter3_a3: 0.029

parameter3_a4: -0.003

parameter3_a5: -0.001

parameter4_a1: 0.041

parameter4_a2: -0.041

parameter4_a3: -0.035

parameter4_a4: -0.017

parameter4_a5: -0.026

parameter5_a1: -0.023

parameter5_a2: 0.032

parameter5_a3: 0.022

parameter5_a4: 0.001

parameter5_a5: 0.017

Bias: 0.011

Node 3 (Sigmoid)

---

parameter1_a1: 0.040

parameter1_a2: 0.050

parameter1_a3: 0.003

parameter1_a4: 0.016

parameter1_a5: -0.048

parameter2_a1: 0.024

parameter2_a2: 0.040

parameter2_a3: 0.021

parameter2_a4: -0.008

parameter2_a5: -0.023

parameter3_a1: 0.014

parameter3_a2: -0.004

parameter3_a3: -0.019

parameter3_a4: -0.031

parameter3_a5: 0.021

parameter4_a1: 0.049

parameter4_a2: 0.023

parameter4_a3: 0.028

parameter4_a4: -0.013

parameter4_a5: 0.049

parameter5_a1: -0.024

parameter5_a2: -0.008

parameter5_a3: 0.045

parameter5_a4: -0.017

parameter5_a5: 0.009

Bias: 0.038

Node 4 (Sigmoid)

---

parameter1_a1: 0.044

parameter1_a2: -0.022

parameter1_a3: 0.004

parameter1_a4: 0.019

parameter1_a5: 0.002

parameter2_a1: 0.047

parameter2_a2: -0.028

parameter2_a3: -0.006

parameter2_a4: -0.042

parameter2_a5: -0.022

parameter3_a1: -0.042

parameter3_a2: -0.001

parameter3_a3: 0.032

parameter3_a4: 0.019

parameter3_a5: -0.025

parameter4_a1: 0.041

parameter4_a2: -0.042

parameter4_a3: -0.009

parameter4_a4: 0.017

parameter4_a5: 0.027

parameter5_a1: 0.014

parameter5_a2: -0.023

parameter5_a3: -0.030

parameter5_a4: 0.039

parameter5_a5: 0.011

Bias: -0.009

Node 5 (Sigmoid)

---

parameter1_a1: 0.035

parameter1_a2: -0.034

parameter1_a3: 0.032

parameter1_a4: 0.002

parameter1_a5: -0.033

parameter2_a1: -0.026

parameter2_a2: 0.012

parameter2_a3: -0.038

parameter2_a4: -0.012

parameter2_a5: 0.029

parameter3_a1: -0.003

parameter3_a2: -0.032

parameter3_a3: -0.003

parameter3_a4: 0.003

parameter3_a5: -0.046

parameter4_a1: -0.040

parameter4_a2: -0.011

parameter4_a3: -0.041

parameter4_a4: -0.042

parameter4_a5: 0.047

parameter5_a1: -0.038

parameter5_a2: -0.013

parameter5_a3: -0.001

parameter5_a4: -0.001

parameter5_a5: -0.015

Bias: -0.009

Node 6 (Sigmoid)

---

parameter1_a1: -0.039

parameter1_a2: -0.023

parameter1_a3: -0.011

parameter1_a4: 0.033

parameter1_a5: -0.018

parameter2_a1: 0.007

parameter2_a2: 0.006

parameter2_a3: 0.015

parameter2_a4: 0.048

parameter2_a5: -0.018

parameter3_a1: -0.043

parameter3_a2: -0.037

parameter3_a3: -0.006

parameter3_a4: 0.003

parameter3_a5: -0.026

parameter4_a1: 0.004

parameter4_a2: 0.001

parameter4_a3: -0.022

parameter4_a4: -0.017

parameter4_a5: -0.034

parameter5_a1: 0.018

parameter5_a2: 0.006

parameter5_a3: 0.011

parameter5_a4: 0.021

parameter5_a5: -0.035

Bias: 0.034

Node 7 (Sigmoid)

---

parameter1_a1: 0.038

parameter1_a2: -0.039

parameter1_a3: 0.021

parameter1_a4: -0.026

parameter1_a5: -0.018

parameter2_a1: 0.038

parameter2_a2: 0.037

parameter2_a3: -0.028

parameter2_a4: -0.045

parameter2_a5: -0.002

parameter3_a1: -0.036

parameter3_a2: 0.046

parameter3_a3: -0.036

parameter3_a4: -0.008

parameter3_a5: -0.014

parameter4_a1: -0.042

parameter4_a2: -0.039

parameter4_a3: -0.001

parameter4_a4: 0.042

parameter4_a5: 0.042

parameter5_a1: -0.015

parameter5_a2: -0.003

parameter5_a3: 0.006

parameter5_a4: -0.007

parameter5_a5: -0.010

Bias: -0.035

Node 8 (Sigmoid)

---

parameter1_a1: -0.015

parameter1_a2: 0.008

parameter1_a3: 0.006

parameter1_a4: -0.043

parameter1_a5: 0.039

parameter2_a1: -0.046

parameter2_a2: -0.041

parameter2_a3: -0.032

parameter2_a4: 0.045

parameter2_a5: -0.016

parameter3_a1: -0.024

parameter3_a2: 0.020

parameter3_a3: 0.024

parameter3_a4: -0.007

parameter3_a5: 0.034

parameter4_a1: -0.025

parameter4_a2: 0.011

parameter4_a3: -0.038

parameter4_a4: 0.037

parameter4_a5: 0.000

parameter5_a1: 0.030

parameter5_a2: 0.009

parameter5_a3: 0.028

parameter5_a4: 0.032

parameter5_a5: -0.010

Bias: -0.032

Node 9 (Sigmoid)

---

parameter1_a1: 0.011

parameter1_a2: 0.043

parameter1_a3: 0.000

parameter1_a4: -0.037

parameter1_a5: -0.007

parameter2_a1: 0.025

parameter2_a2: 0.025

parameter2_a3: -0.004

parameter2_a4: 0.035

parameter2_a5: -0.046

parameter3_a1: 0.035

parameter3_a2: 0.026

parameter3_a3: -0.013

parameter3_a4: -0.014

parameter3_a5: 0.044

parameter4_a1: 0.029

parameter4_a2: 0.036

parameter4_a3: 0.047

parameter4_a4: 0.034

parameter4_a5: -0.007

parameter5_a1: 0.018

parameter5_a2: -0.026

parameter5_a3: -0.037

parameter5_a4: -0.035

parameter5_a5: 0.039

Bias: -0.009

Node 10 (Sigmoid)

---

parameter1_a1: -0.002

parameter1_a2: 0.026

parameter1_a3: 0.005

parameter1_a4: -0.038

parameter1_a5: 0.042

parameter2_a1: -0.006

parameter2_a2: -0.037

parameter2_a3: 0.028

parameter2_a4: 0.028

parameter2_a5: 0.019

parameter3_a1: 0.002

parameter3_a2: 0.026

parameter3_a3: 0.047

parameter3_a4: 0.032

parameter3_a5: 0.033

parameter4_a1: -0.011

parameter4_a2: 0.029

parameter4_a3: -0.032

parameter4_a4: 0.041

parameter4_a5: -0.040

parameter5_a1: 0.033

parameter5_a2: -0.048

parameter5_a3: 0.014

parameter5_a4: -0.011

parameter5_a5: 0.029

Bias: -0.005

Hidden 2

========

Node 1 (Sigmoid)

Node 1: 0.048

Node 2: 0.030

Node 3: 0.002

Node 4: -0.046

Node 5: 0.024

Node 6: 0.003

Node 7: 0.008

Node 8: -0.004

Node 9: -0.013

Node 10: 0.038

Bias: 0.002

Node 2 (Sigmoid)

Node 1: 0.014

Node 2: 0.013

Node 3: 0.009

Node 4: 0.016

Node 5: -0.005

Node 6: -0.013

Node 7: -0.003

Node 8: 0.036

Node 9: -0.009

Node 10: -0.008

Bias: 0.037

Node 3 (Sigmoid)

Node 1: -0.049

Node 2: 0.039

Node 3: -0.029

Node 4: -0.005

Node 5: -0.023

Node 6: 0.036

Node 7: -0.005

Node 8: 0.018

Node 9: -0.020

Node 10: -0.001

Bias: -0.017

Node 4 (Sigmoid)

Node 1: 0.044

Node 2: 0.037

Node 3: -0.050

Node 4: 0.035

Node 5: -0.002

Node 6: -0.042

Node 7: 0.010

Node 8: 0.003

Node 9: 0.003

Node 10: 0.012

Bias: -0.025

Node 5 (Sigmoid)

Node 1: -0.033

Node 2: 0.024

Node 3: 0.011

Node 4: 0.037

Node 5: 0.027

Node 6: 0.014

Node 7: -0.020

Node 8: -0.011

Node 9: 0.020

Node 10: -0.024

Bias: 0.032

Node 6 (Sigmoid)

Node 1: 0.048

Node 2: -0.006

Node 3: 0.002

Node 4: 0.024

Node 5: 0.008

Node 6: 0.031

Node 7: -0.013

Node 8: 0.008

Node 9: -0.004

Node 10: 0.020

Bias: -0.011

Node 7 (Sigmoid)

Node 1: 0.034

Node 2: 0.003

Node 3: 0.048

Node 4: -0.005

Node 5: -0.015

Node 6: 0.047

Node 7: 0.036

Node 8: -0.047

Node 9: 0.034

Node 10: -0.015

Bias: -0.044

Node 8 (Sigmoid)

Node 1: -0.018

Node 2: 0.015

Node 3: 0.036

Node 4: -0.005

Node 5: -0.039

Node 6: -0.036

Node 7: 0.016

Node 8: 0.044

Node 9: 0.008

Node 10: 0.028

Bias: -0.033

Node 9 (Sigmoid)

Node 1: 0.026

Node 2: 0.029

Node 3: 0.016

Node 4: -0.033

Node 5: -0.018

Node 6: 0.006

Node 7: -0.033

Node 8: 0.006

Node 9: 0.034

Node 10: -0.045

Bias: -0.041

Node 10 (Sigmoid)

Node 1: -0.049

Node 2: -0.011

Node 3: 0.019

Node 4: 0.049

Node 5: -0.017

Node 6: -0.040

Node 7: -0.004

Node 8: 0.005

Node 9: 0.037

Node 10: -0.049

Bias: -0.001

Hidden 3

========

Node 1 (Sigmoid)

Node 1: 0.016

Node 2: 0.063

Node 3: 0.005

Node 4: -0.016

Node 5: -0.000

Node 6: 0.065

Node 7: -0.007

Node 8: -0.018

Node 9: -0.019

Node 10: 0.059

Bias: 0.083

Node 2 (Sigmoid)

Node 1: 0.013

Node 2: 0.040

Node 3: 0.016

Node 4: 0.003

Node 5: -0.028

Node 6: 0.004

Node 7: 0.013

Node 8: 0.011

Node 9: 0.041

Node 10: -0.008

Bias: -0.001

Node 3 (Sigmoid)

Node 1: 0.016

Node 2: 0.034

Node 3: -0.031

Node 4: 0.043

Node 5: -0.021

Node 6: -0.007

Node 7: -0.006

Node 8: -0.021

Node 9: -0.004

Node 10: 0.055

Bias: -0.006

Node 4 (Sigmoid)

Node 1: 0.067

Node 2: -0.022

Node 3: -0.020

Node 4: 0.005

Node 5: 0.025

Node 6: -0.006

Node 7: 0.038

Node 8: -0.002

Node 9: 0.038

Node 10: 0.018

Bias: 0.057

Node 5 (Sigmoid)

Node 1: 0.009

Node 2: -0.027

Node 3: 0.005

Node 4: 0.031

Node 5: 0.058

Node 6: 0.037

Node 7: 0.003

Node 8: 0.037

Node 9: 0.022

Node 10: 0.035

Bias: 0.077

Node 6 (Sigmoid)

Node 1: 0.058

Node 2: 0.063

Node 3: 0.030

Node 4: 0.025

Node 5: 0.007

Node 6: 0.039

Node 7: 0.028

Node 8: 0.030

Node 9: 0.014

Node 10: -0.026

Bias: 0.037

Node 7 (Sigmoid)

Node 1: 0.044

Node 2: -0.003

Node 3: 0.005

Node 4: 0.029

Node 5: -0.024

Node 6: -0.028

Node 7: 0.012

Node 8: 0.009

Node 9: 0.038

Node 10: 0.005

Bias: 0.068

Node 8 (Sigmoid)

Node 1: -0.019

Node 2: -0.009

Node 3: -0.018

Node 4: 0.036

Node 5: 0.028

Node 6: 0.048

Node 7: 0.005

Node 8: 0.072

Node 9: 0.068

Node 10: -0.002

Bias: 0.009

Node 9 (Sigmoid)

Node 1: 0.037

Node 2: -0.004

Node 3: 0.021

Node 4: -0.002

Node 5: -0.015

Node 6: -0.028

Node 7: -0.015

Node 8: 0.000

Node 9: 0.006

Node 10: 0.011

Bias: 0.055

Node 10 (Sigmoid)

Node 1: -0.013

Node 2: -0.008

Node 3: 0.067

Node 4: 0.046

Node 5: 0.029

Node 6: 0.058

Node 7: 0.005

Node 8: -0.022

Node 9: -0.013

Node 10: -0.018

Bias: 0.038

Output

======

Class 'false' (Sigmoid)

Node 1: -0.351

Node 2: -0.313

Node 3: -0.324

Node 4: -0.285

Node 5: -0.309

Node 6: -0.295

Node 7: -0.284

Node 8: -0.298

Node 9: -0.268

Node 10: -0.257

Threshold: -0.622

Class 'true' (Sigmoid)

Node 1: 0.268

Node 2: 0.275

Node 3: 0.276

Node 4: 0.349

Node 5: 0.270

Node 6: 0.262

Node 7: 0.283

Node 8: 0.344

Node 9: 0.307

Node 10: 0.341

Threshold: 0.628

ภาคผนวก ง

Code PHP สำหรับโครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

นำสมการของโมเดลการคัดเลือกข้อสอบข้อถัดไปนี้มาเขียนเป็น Code PHP ด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) เนื่องจากการเขียน Code PHP มีความซับซ้อนมาก ผู้วิจัยจึงกำหนดฟังก์ชันที่คำนวณผลลัพธ์ของแต่ละ Node และ Layer ของ Neural Network ด้วยค่าพารามิเตอร์ที่กำหนดฟังก์ชัน Sigmoid มาทำการคำนวณผลลัพธ์ของ Neural Network ดังนี้

```
<?php
// Sigmoid function
function sigmoid($x) {
    return 1 / (1 + exp(-$x));
}
// Function to compute the output of a node in a hidden layer
function computeHiddenNode($inputs, $parameters, $bias) {
    $sum = $bias;
    foreach ($inputs as $index => $input) {
        $sum += $parameters[$index] * $input;
    }
    return sigmoid($sum);
}
// Function to compute the output of a node in the output layer
function computeOutputNode($hiddenOutputs, $parameters, $threshold) {
    $sum = $threshold;
    foreach ($hiddenOutputs as $index => $output) {
        $sum += $parameters[$index] * $output;
    }
    return sigmoid($sum);
}
// Input values
$inputValues = array(/* Your input values here */);
// Hidden Layer 1 parameters
$hidden1Parameters = array(
    array(/* parameters for node 1 */),
```

```
    array(/* parameters for node 2 */),
    // ... parameters for other nodes
);
// Hidden Layer 2 parameters
$hidden2Parameters = array(
    array(/* parameters for node 1 */),
    array(/* parameters for node 2 */),
    // ... parameters for nodes 3-10*
);
// Hidden Layer 3 parameters
$hidden3Parameters = array(
    array(/* parameters for node 1 */),
    array(/* parameters for node 2 */),
    // ... parameters for nodes 3-10*
);
// Output Layer parameters
$outputParameters = array(
    array(/* parameters for class 'wrong' */),  กรณีตอบผิด
    array(/* parameters for class 'correct' */)   กรณีตอบถูก
);
// Compute Hidden Layer 1 outputs
$hidden1Outputs = array();
foreach ($hidden1Parameters as $nodeParameters) {
    $hidden1Outputs[] = computeHiddenNode($inputValues, $nodeParameters,
$nodeParameters['Bias']);
}
// Compute Hidden Layer 2 outputs
$hidden2Outputs = array();
foreach ($hidden2Parameters as $nodeParameters) {
    $hidden2Outputs[] = computeHiddenNode($hidden1Outputs, $nodeParameters,
$nodeParameters['Bias']);
```

```
}
// Compute Hidden Layer 3 outputs
$hidden3Outputs = array();
foreach ($hidden3Parameters as $nodeParameters) {
    $hidden3Outputs[] = computeHiddenNode($hidden2Outputs, $nodeParameters,
$nodeParameters['Bias']);
}
// Compute Output Layer outputs
$outputWrong = computeOutputNode($hidden3Outputs, $outputParameters[0],
$outputParameters[0]['Threshold']);
$outputCorrect = computeOutputNode($hidden3Outputs, $outputParameters[1],
$outputParameters[1]['Threshold']);
// Display the results
echo "Class 'wrong' output: " . $outputWrong . "\n";
echo "Class 'correct' output: " . $outputCorrect . "\n";
?>
```

ปรับโมเดล

```
<?php

// Sigmoid function
function sigmoid($x) {
    return 1 / (1 + exp(-$x));
}

// Function to compute the output of a node in a hidden layer
function computeHiddenNode($inputs, $parameters, $bias) {
    $sum = $bias;
    foreach ($inputs as $index => $input) {
        $sum += $parameters[$index] * $input;
```

```
    }
    return sigmoid($sum);
}

// Function to compute the output of a node in the output layer
function computeOutputNode($hiddenOutputs, $parameters, $threshold) {
    $sum = $threshold;
    foreach ($hiddenOutputs as $index => $output) {
        $sum += $parameters[$index] * $output;
    }
    return sigmoid($sum);
}

// Function to make a prediction using the trained neural network
function predict($inputValues) {
    // Hidden Layer 1 parameters
    $hidden1Parameters = [
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        // ... parameters for other nodes
    ];

    // Hidden Layer 2 parameters
    $hidden2Parameters = [
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        // ... parameters for other nodes
```

```
    ];

    // Hidden Layer 3 parameters
    $hidden3Parameters = [
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        ['Bias' => /* bias value */, 'Parameter1' => /* value */, /* ... */, 'ParameterN' =>
/* value */],
        // ... parameters for other nodes
    ];

    // Output Layer parameters
    $outputParameters = [
        ['Threshold' => /* threshold value */, 'Parameter1' => /* value */, /* ... */,
'ParameterN' => /* value */], // Parameters for class 'wrong'  ตอบผิด

        ['Threshold' => /* threshold value */, 'Parameter1' => /* value */, /* ... */,
'ParameterN' => /* value */], // Parameters for class 'correct'  ตอบถูก
    ];

    // Compute Hidden Layer 1 outputs
    $hidden1Outputs = [];
    foreach ($hidden1Parameters as $nodeParameters) {
        $hidden1Outputs[] = computeHiddenNode($inputValues, $nodeParameters,
$nodeParameters['Bias']);
    }

    // Compute Hidden Layer 2 outputs
    $hidden2Outputs = [];
    foreach ($hidden2Parameters as $nodeParameters) {
```

```
        $hidden2Outputs[] = computeHiddenNode($hidden1Outputs, $nodeParameters, $nodeParameters['Bias']);
    }

    // Compute Hidden Layer 3 outputs
    $hidden3Outputs = [];
    foreach ($hidden3Parameters as $nodeParameters) {
        $hidden3Outputs[] = computeHiddenNode($hidden2Outputs, $nodeParameters, $nodeParameters['Bias']);
    }

    // Compute Output Layer outputs
    $outputWrong = computeOutputNode($hidden3Outputs, $outputParameters[0], $outputParameters[0]['Threshold']);
    $outputCorrect = computeOutputNode($hidden3Outputs, $outputParameters[1], $outputParameters[1]['Threshold']);

    // Return the predicted class based on the output
    return ($outputTrue > $outputFalse) ? 'correct' : 'wrong';

}

// Example input values
$inputValues = [
    'Parameter1' => /* value */,
    'Parameter2' => /* value */,
    // ... add more input parameters as needed
];

// Make a prediction
$prediction = predict($inputValues);
```

```
// Display the prediction
echo "Prediction: Class '$prediction'\n";

?>
```

ค่าของพารามิเตอร์และข้อมูลนำเข้าในฟังก์ชัน predict ด้วยค่าที่ถูกต้องตามโครงสร้างของ Neural Network หลังจากนั้นสามารถใช้ฟังก์ชัน predict เพื่อทำนายผลลัพธ์การคัดเลือกข้อสอบข้อไปที่แม่นยำที่สุด

```
<?php

function sigmoid($x) {
    return 1 / (1 + exp(-$x));
}

function calculateLayerOutput($params, $bias, $inputs) {
    $outputs = [];
    foreach ($params as $nodeParams) {
        $nodeOutput = sigmoid(dotProduct($nodeParams, $inputs) + $nodeParams[count($nodeParams) - 1]);
        $outputs[] = $nodeOutput;
    }
    return $outputs;
}

function dotProduct($array1, $array2) {
    $result = 0;
    $length = min(count($array1), count($array2));
    for ($i = 0; $i < $length; $i++) {
        $result += $array1[$i] * $array2[$i];
```

```
    }
    return $result;
}

function makeDecision($output, $threshold) {
    return $output >= $threshold;
}

// Input
$inputValues = [/* your input values */];

// Hidden Layer 1
$hidden1 = calculateLayerOutput($hidden1Params, $hidden1Bias, $inputValues);

// Hidden Layer 2
$hidden2 = calculateLayerOutput($hidden2Params, $hidden2Bias, $hidden1);

// Hidden Layer 3
$hidden3 = calculateLayerOutput($hidden3Params, $hidden3Bias, $hidden2);

// Output Layer
$output = calculateLayerOutput($outputParams, $outputBias, $hidden3);

// Threshold for decision
$outputThreshold = 0.5;

// Make decision based on output
$classTrue = makeDecision($output[0], $outputThreshold);
$classFalse = !$classTrue;

// Select next question based on the classification
```

```
if ($classTrue) {
    // Select the next question after answering correctly
    $nextQuestion = "Next question after answering correctly";
} else {
    // Select the next question after answering incorrectly
    $nextQuestion = "Next question after answering incorrectly";
}

echo $nextQuestion;

?>
```

ใช้ฟังก์ชัน sigmoid และคำนวณผลลัพธ์ของแต่ละโหนดในทุกๆ Hidden Layer และ Output Layer ตามลำดับ จากนั้นทำการตัดสินใจว่าจะเลือกข้อสอบถัดไปหรือไม่ โดยเปรียบเทียบผลลัพธ์กับ Threshold.

ภาคผนวก จ

คู่มือการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

คู่มือการใช้
เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

โดย

นายชัยมงคล ปินะสา

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร.สำราญ มีแจ้ง

รองศาสตราจารย์ ดร.น้ำทิพย์ องอาจวาณิชย์

ผู้ช่วยศาสตราจารย์ ดร.อ้อมธจิต แป้นศรี

เว็บแอพพลิเคชั่นนี้

เป็นส่วนหนึ่งของการทำวิทยานิพนธ์หลักสูตรปรัชญาดุษฎีบัณฑิต

สาขานวัตกรรมทางการวัดผลการเรียนรู้

คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

คำนำ

คู่มือการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยนำศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์เข้ากับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) จัดทำขึ้นเพื่อช่วยให้ผู้ใช้เข้าใจขั้นตอนในการใช้งานโปรแกรมได้ดียิ่งขึ้น ซึ่งเว็บแอพพลิเคชั่นนี้เป็นส่วนหนึ่งของการทำวิทยานิพนธ์ระดับปริญญาเอก สาขานวัตกรรมทางการวัดผลการเรียนรู้ คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร เรื่อง การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป: เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยนำเสนอเนื้อหาเกี่ยวกับขั้นตอนการเรียกใช้งานเว็บแอพพลิเคชั่น ได้แก่ การเข้าสู่เว็บแอพพลิเคชั่นการลงทะเบียนใช้งานเว็บแอพพลิเคชั่น การบันทึกข้อมูลผู้สอบ การเริ่มทำการสอบการแสดงผลการสอบ การบันทึกผลการสอบการพิมพ์รายงานผลการสอบและการแปลผล ผู้วิจัยหวังเป็นอย่างยิ่งว่าคู่มือที่จัดทำขึ้นนี้จะช่วยให้ผู้ใช้งาน มีความรู้ความเข้าใจเกี่ยวกับขั้นตอนต่างๆ ในการเรียกใช้งานโปรแกรม และสามารถนำเว็บแอพพลิเคชั่นไปใช้ได้อย่างถูกต้องเต็มตามศักยภาพและก่อให้เกิดประโยชน์สูงสุดต่อการจัดการเรียนการสอน

ชัยมงคล ปินะสา

เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิตินี้ เป็น Web Application ที่พัฒนาขึ้นบนโปรแกรม PHP ร่วมกับ Cloud Server จะประกอบด้วย Server หลายๆ เครื่องช่วยกันทำงาน โดยพัฒนาให้สามารถใช้งานบนอุปกรณ์ที่เชื่อมต่อกับอินเทอร์เน็ตได้ เช่น คอมพิวเตอร์โน๊ตบุ๊คหรือสมาร์ทโฟน เว็บแอพพลิเคชั่นการทดสอบนี้ แบ่งออกเป็น 3 ส่วน คือ 1) ข้อมูลทั่วไปเกี่ยวกับโปรแกรม เป็นที่อธิบายส่วนประกอบต่างๆของโปรแกรม 2) การจัดการของผู้ดูแลระบบ เป็นส่วนที่อาจารย์ หรือผู้ที่เกี่ยวข้องในการจัดการข้อสอบ นำข้อสอบที่ผ่านการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) มาใส่ในเว็บแอพพลิเคชั่นเพื่อใช้เป็นคลังข้อสอบ รวมทั้งการลบ และการแก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ นอกจากนี้ยังมีรายงานผลการทดสอบของผู้สอบทั้งหมด เพื่อให้อาจารย์ และผู้เกี่ยวข้องนำผลการทดสอบไปใช้ในการประเมินผลการเรียนรู้ต่อไป 3) การจัดการสอบ เป็นส่วนที่ให้ผู้สอบ คือ นักศึกษา ใช้เว็บแอพพลิเคชั่นนี้เพื่อทดสอบความสามารถทางความรู้วิชาชีพครู ซึ่งเมื่อผู้สอบทำการทดสอบเสร็จสิ้น จะรายงานผลการทดสอบให้ผู้สอบทราบทันที

1) ข้อมูลทั่วไปเกี่ยวกับเว็บแอพพลิเคชั่น

เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไปในครั้งนี้ เป็นเว็บแอพพลิเคชั่นที่พัฒนาขึ้น โดยมีจุดมุ่งหมายเพื่อใช้เป็นเครื่องมือในการทดสอบและใช้วิเคราะห์ความสามารถของสมรรถนะทางวิชาชีพที่จะทดสอบและประเมินของกลุ่มวิชาชีพครู ได้แก่ หลักสูตรและการสอน จิตวิทยา การประกันคุณภาพการศึกษา วัดผลและวิจัยการเรียนรู้ และการเปลี่ยนแปลงบริบทโลก เพื่อจะได้นำข้อมูลที่ได้ไปใช้ในการวางแผนปรับปรุงเพื่อพัฒนาจุดอ่อนและส่งเสริมจุดแข็งให้กับนิสิต/นักศึกษา เน้นการนำไปใช้เพื่อการประเมินผลสรุปรวม (summative evaluation) เพื่อเป็นข้อมูลย้อนกลับให้กับหลักสูตร/ สาขาวิชา/คณะ ใช้ประโยชน์ในการพัฒนาปรับปรุงแก้ไขการจัดการเรียนการสอน ตลอดจนเป็นการฝึกให้นิสิต/นักศึกษา ได้คุ้นชินกับการทำข้อสอบก่อนที่จะดำเนินการทดสอบเพื่อขอรับใบประกอบวิชาชีพครูจริงต่อไป

การพัฒนาเว็บแอพพลิเคชั่นในครั้งนี้ อาศัยแนวคิดศาสตร์ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) ใช้โครงข่ายประสาทเทียม (Artificial Neural Network Selection : ANN)

มาประยุกต์เข้ากับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing) เพื่อใช้คัดเลือกข้อสอบข้อถัดไป ซึ่งเป็นแนวคิดที่สามารถวัดและประเมินคุณลักษณะของบุคคลได้ครั้งละหลายคุณลักษณะ ซึ่งจะช่วยให้ผู้สอนหรือผู้เรียนสามารถประเมินคุณลักษณะของตัวเองได้อย่างรวดเร็ว ใช้เวลาในการทดสอบน้อยกว่าการทดสอบแบบกระดาษเขียนตอบ รวมทั้งใช้เวลาในการทดสอบน้อยกว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบมิติเดียว หรือถ้าใช้เวลาในการทดสอบเท่ากันการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติจะให้ผลการทดสอบที่แม่นยำและรวดเร็วมากกว่า ทั้งนี้การพัฒนาเว็บแอพพลิเคชั่น มีขั้นตอนการดำเนินการดังนี้

1. การพัฒนาคลังข้อสอบ ผู้วิจัยทำการรวบรวมข้อสอบสมรรถนะความรู้ทางวิชาชีพครูและเฉลย เป็นข้อสอบการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครู ที่ได้มาจากสภาคณบดีคณะครุศาสตร์มหาวิทยาลัยราชภัฏ (สครภ) ลักษณะข้อสอบคู่ขนานตามผังการออกข้อสอบที่ได้มีการวิเคราะห์ตามกรอบมาตรฐานวิชาชีพครุสภา และกรอบมาตรฐานคุณวุฒิอุดมศึกษา ที่ผ่านการวิพากษ์และแก้ไขปรับปรุงโดยผู้ทรงคุณวุฒิแล้ว โดยผู้วิจัยคัดเลือกข้อสอบออกเป็น 5 มิติ คือ มิติวัดผลและวิจัยการเรียนรู้ มิติการประกันคุณภาพการศึกษา ข้อมิติจิตวิทยาการศึกษา ข้อมิติการเปลี่ยนแปลงบริบทโลก และมิติหลักสูตรและการสอน จากนั้นนำข้อสอบที่รวบรวมได้ทั้งหมดไปให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องของข้อสอบกับมิติของข้อสอบ (ตรวจสอบความเหมาะสมของ Q-matrix) และนำข้อสอบที่ผ่านเกณฑ์ไปวิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ ตามแนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (MIRT) ด้วยโมเดลปกติสะสมแบบพหุมิติ (Multidimensional normal ogive model) โดยใช้โปรแกรมสำเร็จรูป Noharm 4.0 ได้ข้อสอบที่มีคุณภาพเป็นไปตามเกณฑ์ (ค่าอำนาจจำแนกแบบพหุมิติตั้งแต่ 0.00 ขึ้นไปค่าความยากแบบพหุมิติมีค่าอยู่ในช่วง -4.00 ถึง +4.00 และกำหนดให้ค่าการเดาของข้อสอบทุกข้อมีค่าเท่ากับ 0.20) จำนวนทั้งสิ้น 270 ข้อ จำแนกออกเป็น 5 มิติ ดังนี้ คือ มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ มิติการประกันคุณภาพการศึกษา จำนวน 50 ข้อมิติจิตวิทยาการศึกษา จำนวน 50 ข้อมิติการเปลี่ยนแปลงบริบทโลก จำนวน 60 ข้อ และมิติหลักสูตรและการสอน จำนวน 50 ข้อ

2. การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป โดยใช้คลังข้อสอบที่ได้จากการดำเนินการในข้อที่ 1 นำมาใช้เป็นฐานข้อมูลสำหรับการพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ตามแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติที่มี

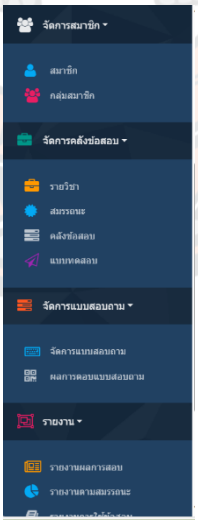
หลักการทำงานของโปรแกรม คือ เมื่อมีผู้เข้ารับการทดสอบ โปรแกรมจะคัดเลือกข้อสอบให้กับผู้สอบตามความสามารถของผู้สอบ ถ้าผู้สอบมีความสามารถสูงก็จะได้รับข้อสอบที่ยาก ส่วนผู้สอบที่มีความสามารถต่ำก็จะได้ข้อสอบที่ง่าย ซึ่งผู้สอบแต่ละคนจะได้รับข้อสอบแตกต่างกัน และในการสอบแต่ละข้อโปรแกรมจะตรวจข้อสอบและรายงานผลการสอบให้ทราบได้ทันทีโดยจะทำการวิเคราะห์ความสามารถของผู้สอบแยกตามมิติของข้อสอบ การทดสอบจะดำเนินการไปเรื่อยๆ จนได้จำนวนข้อตามที่กำหนดก็จะยุติการสอบ และสรุปผลการทดสอบให้กับผู้สอบได้ทราบทันที จากลักษณะการทำงานของโปรแกรมดังกล่าวจะช่วยให้ผู้สอนประเมินความสามารถของผู้เรียน จำแนกตามมิติของข้อสอบได้ทันที ทำให้ไม่เสียเวลาในการบริหารจัดการ สามารถนำข้อมูลไปใช้ในการตัดสินใจสำหรับการให้คำแนะนา ส่งเสริมและช่วยเหลือผู้เรียน นอกจากนี้ผู้เรียนยังสามารถนำเอาผลการประเมินไปปรับปรุงและพัฒนาความสามารถของตนเองให้ดียิ่งขึ้นอีกด้วย

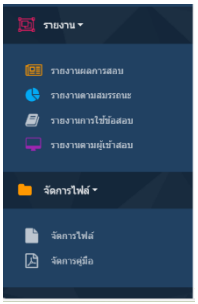
2) การจัดการของผู้ดูแลระบบ

การจัดการข้อสอบเป็นส่วนที่ให้อาจารย์ หรือผู้เกี่ยวของในการจัดการของผู้ดูแลระบบใส่ข้อสอบที่ผ่านการคัดเลือกเป็นไปตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) และคลังข้อสอบที่ใช้ในเว็บแอพพลิเคชั่นนี้เป็นข้อสอบสมรรถนะความรู้ทางวิชาชีพครู และเฉลย เป็นข้อสอบการเตรียมความพร้อมเพื่อขอรับใบประกอบวิชาชีพครูที่มาจากสภาคณบดีคณะครุศาสตร์มหาวิทยาลัยราชภัฏ (สครภ) ลักษณะข้อสอบคู่ขนานตามผังการออกข้อสอบที่ได้มีการวิเคราะห์ตามกรอบมาตรฐานวิชาชีพคุรุสภา และกรอบมาตรฐานคุณวุฒิอุดมศึกษา ที่ผ่านการวิพากษ์และแก้ไขปรับปรุงโดยผู้ทรงคุณวุฒิแล้ว จำนวน 270 ข้อ มาวิเคราะห์ตามจุดประสงค์ทางการศึกษาใหม่ตามแนวคิดของ Bloom et al. (1956); Anderson et al. (2001, p. 21) และ Krathwohl (2002, p. 213) โดยการปรับลำดับขั้นและคำศัพท์ที่ใช้ในกระบวนการพุทธิปัญญา ซึ่งมี 6 กระบวนการเหมือนเดิมแต่ 3 กระบวนการแรกเปลี่ยนชื่อเป็น จำ (Remember) เข้าใจ (Understand) และประยุกต์ (Apply) ส่วน 3 กระบวนการหลังเปลี่ยนชื่อที่มีลักษณะเป็นคำนามไปเป็นคำกริยา และสลับที่กระบวนการที่ 5 กับ 6 และสร้างสรรค์ (Create) เปลี่ยนชื่อมาจากการสังเคราะห์ (Synthesis) (Anderson et al., 2001, p. 21; Krathwohl, 2002, p. 213) และเปลี่ยนโครงสร้างจากมิติเดียวเป็นสองมิติโดยได้เพิ่มโครงสร้างในมิติด้านความรู้ (Knowledge Dimension) เข้ามาในโครงสร้างของจุดประสงค์ทางการศึกษาด้านพุทธิปัญญา ทำให้โครงสร้างใหม่มีลักษณะเป็นสองมิติประกอบด้วยมิติด้านกระบวนการพุทธิปัญญาและมิติด้านความรู้ รวมทั้งอธิบายความหมายของมิติด้านความรู้ โดยมิติด้านความรู้ แบ่งเป็น 4 ส่วน คือ ความรู้เกี่ยวกับข้อเท็จจริง (Factual Knowledge) ความรู้เกี่ยวกับมโนทัศน์ (Conceptual Knowledge) ความรู้เกี่ยวกับ

วิธีดำเนินการ (Procedural Knowledge) และความรู้เกี่ยวกับอภิปัญญา (Metacognitive Knowledge) และมิติด้านกระบวนการพุทธิปัญญา แบ่งเป็น 6 กระบวนการ คือ จำ (Remember) เข้าใจ (Understand) ประยุกต์ (Apply) วิเคราะห์ (Analyze) ประเมินค่า (Evaluate) และสร้างสรรค์ (Create) วิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ ตามแนวคิดทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (MIRT) ด้วยโมเดลปกติสะสมแบบพหุมิติ (Multidimensional normal ogive model) โดยใช้โปรแกรมสำเร็จรูป Noharm 4.0 ได้ข้อสอบที่มีคุณภาพเป็นไปตามเกณฑ์ (ค่าอำนาจจำแนกแบบพหุมิติตั้งแต่ 0.00 ขึ้นไป ค่าความยากแบบพหุมิติมีค่าอยู่ในช่วง -4.00 ถึง +4.00 และกำหนดให้ค่าการเดาของข้อสอบทุกข้อมีค่าเท่ากับ 0.20) จำแนกออกเป็น 5 มิติ ดังนี้ คือ 1) มิติวัดผลและวิจัยการเรียนรู้ จำนวน 60 ข้อ 2) มิติการประกันคุณภาพการศึกษา จำนวน 45 ข้อ 3) มิติจิตวิทยาการศึกษา จำนวน 40 ข้อ 4) มิติการเปลี่ยนแปลงบริบทโลก จำนวน 50 ข้อ และ 5) มิติหลักสูตรและการสอน จำนวน 47 ข้อ รวม 242 ข้อ

ในส่วนของการจัดการคลังข้อสอบ เพื่อทำการเพิ่มหรือทำการลบ และแก้ไขข้อสอบเดิมที่มีอยู่ในคลังข้อสอบ หลังจากผู้ใช้เลือก “เข้าสู่ระบบ” ในหน้าหลักเว็บแอพพลิเคชั่นจะเข้าสู่หน้าจอหลักของเว็บแอพพลิเคชั่นของส่วนการจัดการของผู้ดูแลระบบ จะมีเมนูย่อยหลักๆ 5 ส่วน ประกอบด้วย 1) จัดการสมาชิก 2) จัดการคลังข้อสอบ 3) จัดการแบบสอบถาม 4) รายงานผล และ 5) จัดการไฟล์ แสดงได้ ดังภาพที่ 1

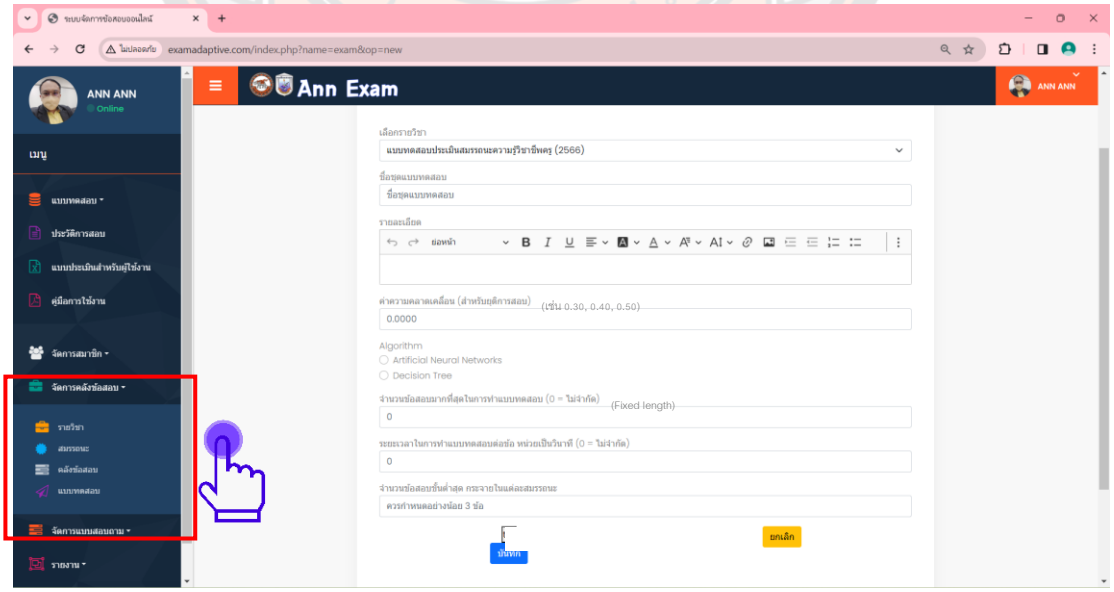




ภาพที่ 1 เมนูการจัดการของผู้ดูแลระบบ (Amin)

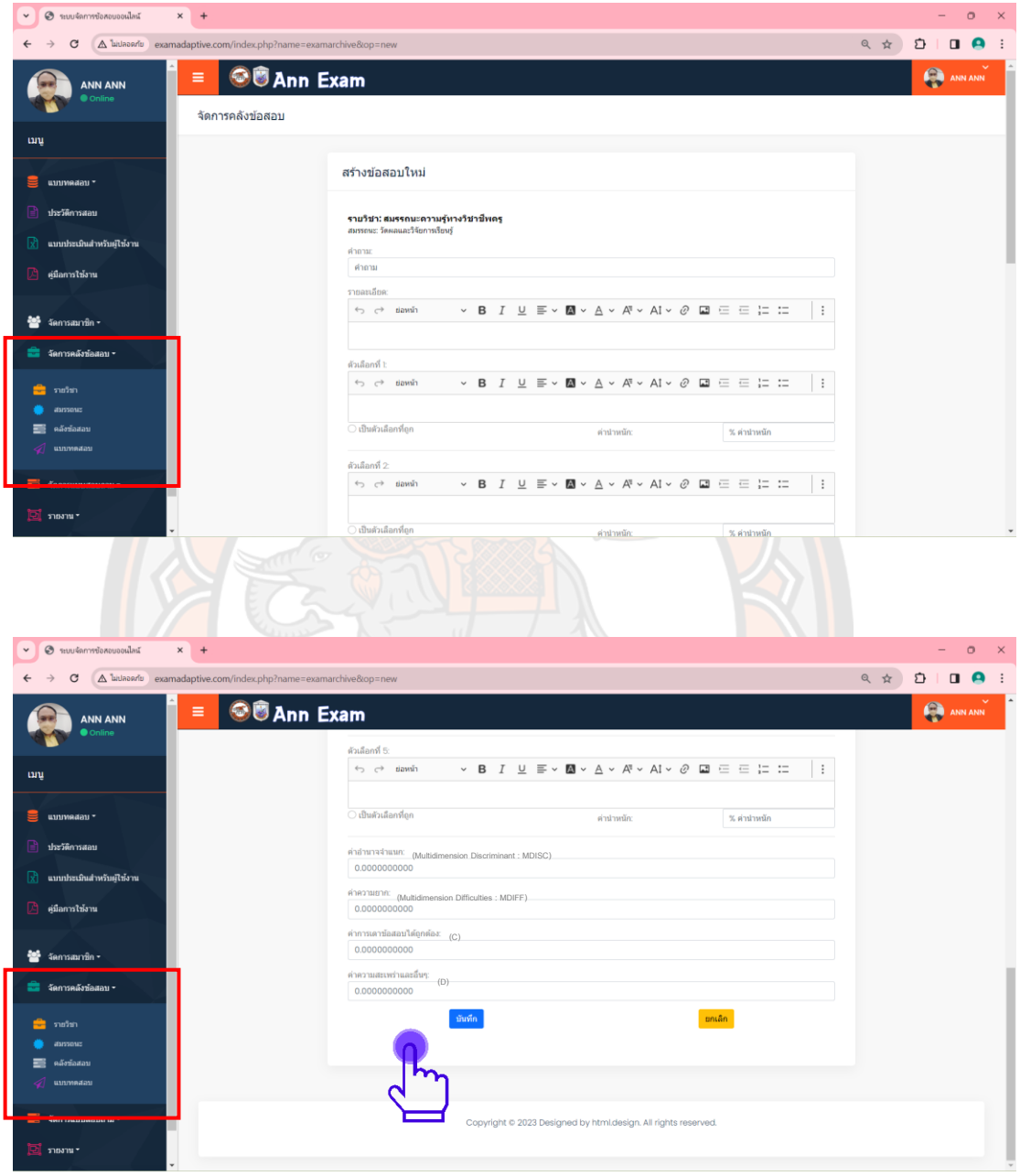
การเพิ่มข้อสอบในคลังข้อสอบ

การเพิ่มข้อสอบลงในเว็บแอพพลิเคชั่น โดยให้ผู้ใช้กดปุ่ม “จัดการคลังข้อสอบ” แล้วเว็บแอพพลิเคชั่นจะเข้าสู่ หน้าจอ และให้กดจัดการรายวิชา เพื่อที่จะสร้างรายวิชา สร้างรายวิชา โดยให้ระบุชื่อรายวิชา ชื่อชุดแบบทดสอบรายละเอียดหรือคำชี้แจงการทดสอบ การกำหนดค่าความคลาดเคลื่อน (สำหรับยุติการทดสอบ) เลือก Algorithm ในการคัดเลือกข้อสอบ เช่น Algorithm Artificial Neural Networks และให้เลือกจำนวนข้อสอบมากที่สุดในการทำแบบทดสอบ (0 = ไม่จำกัด) หรือเกณฑ์ยุติการทดสอบที่นิยมใช้กันอยู่ในขณะนี้มี 2 ลักษณะ ได้แก่ จำกัดจำนวนข้อสอบ และกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (ศิริชัย กาญจนวาสี, 2555, น. 204) หรืออาจจะใช้ทั้ง 2 เกณฑ์ร่วมกัน ในการวิจัยครั้งนี้ ใช้เกณฑ์การยุติการสอบแต่ละมิติ ด้วยการจำกัดจำนวนข้อสอบ (Fixed length) จำนวน 20 ข้อ และกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ที่ SE ≤ 0.3 ดังภาพที่ 2



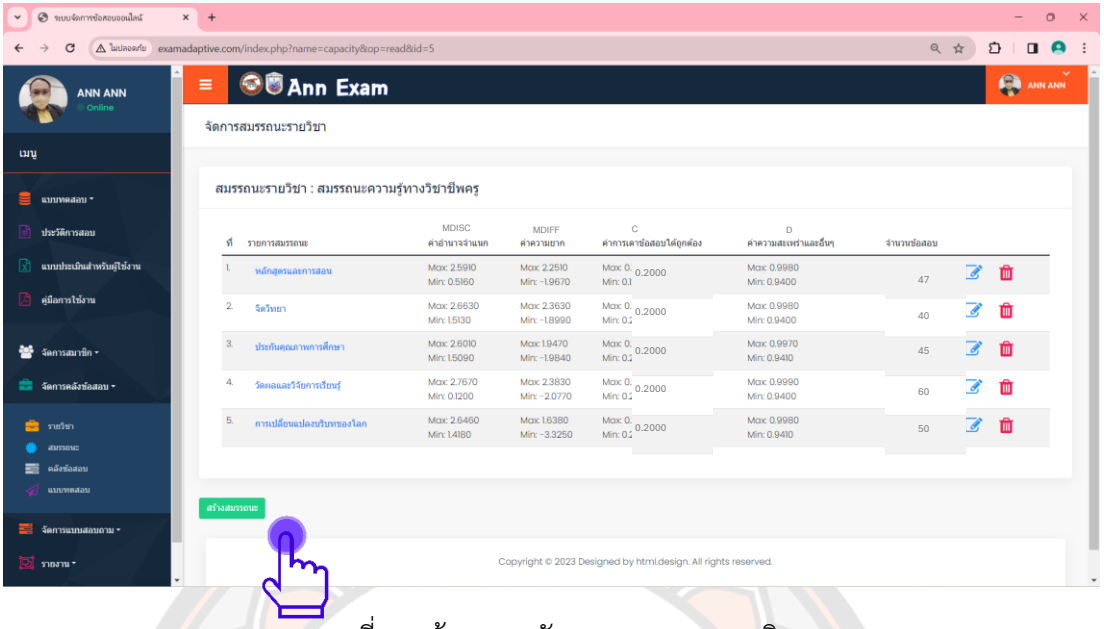
ภาพที่ 2 หน้าจอการสร้างรายวิชา

การเพิ่มสร้างข้อสอบลงในคลังข้อสอบ โดยให้ผู้ใช้กดปุ่ม แล้วเว็บแอพพลิเคชั่นแสดงได้ดังภาพที่ 3



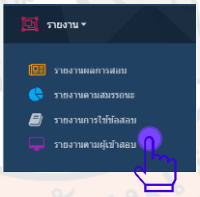
ภาพที่ 3 หน้าจอการเพิ่มข้อสอบใหม่

การจัดการสมรรถนะรายวิชาเพื่อดูค่าพารามิเตอร์สูงสุด ต่ำสุดของข้อสอบแต่ละมิติ ในคลังข้อสอบ โดยให้ผู้ใช้กดปุ่ม แล้วเว็บแอพพลิเคชั่นจะปรากฏ ดังภาพที่ 4



ภาพที่ 4 หน้าจอการจัดการสมรรถนะรายวิชา

ในส่วนของการรายงานผลการสอบต่างๆ ผู้ดูแลระบบ สามารถเข้าดูรายงานผลการสอบต่างๆ จะมีเมนูย่อยหลักๆ 4 ส่วน คือ รายงานผลสอบการสอบ 2) รายงานตามสมรรถนะ 3) รายงานการใช้ข้อสอบ และ 4) รายงานตามผู้เข้าสอบ ได้ ปรากฏดังภาพที่ 5



ภาพที่ 5 รายงานผลการสอบ

3) การจัดการสอบ

การใช้งานเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ประกอบด้วยขั้นตอนการใช้งานโปรแกรมดังนี้

- ขั้นตอนการเข้าสู่โปรแกรม
- ขั้นตอนการลงทะเบียน
- ขั้นตอนการเริ่มทำการทดสอบ
- ขั้นตอนการแสดงผลการสอบ
- ขั้นตอนการพิมพ์รายงานผลการสอบ
- ขั้นตอนการออกจากโปรแกรม

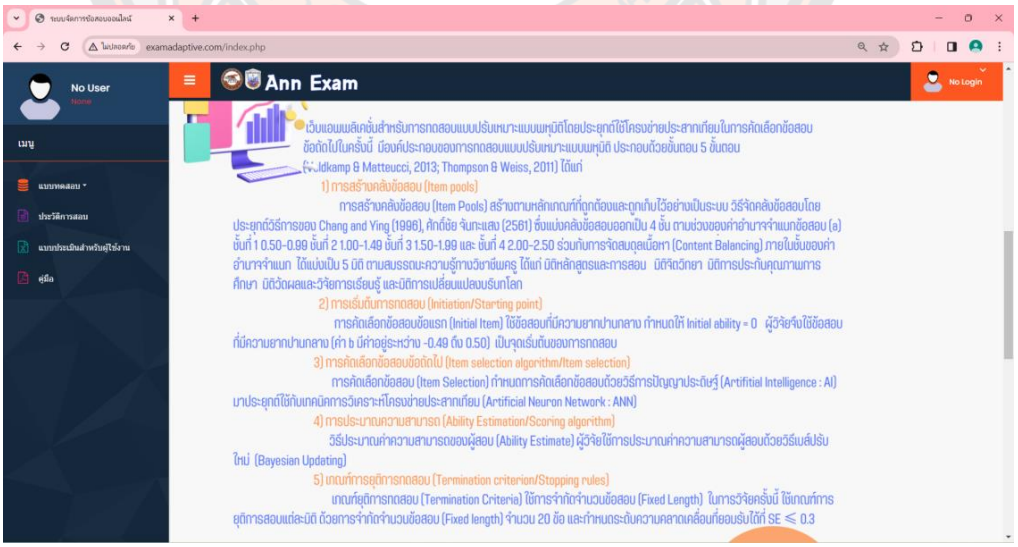
ขั้นตอนการเข้าสู่โปรแกรม

1) เข้าใช้งานโปรแกรมที่เว็บไซต์ http://examadaptive.com/

หน้าจอแรกของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ แสดงดังภาพที่ 6

ภาพที่ 6 หน้าจอแรกของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

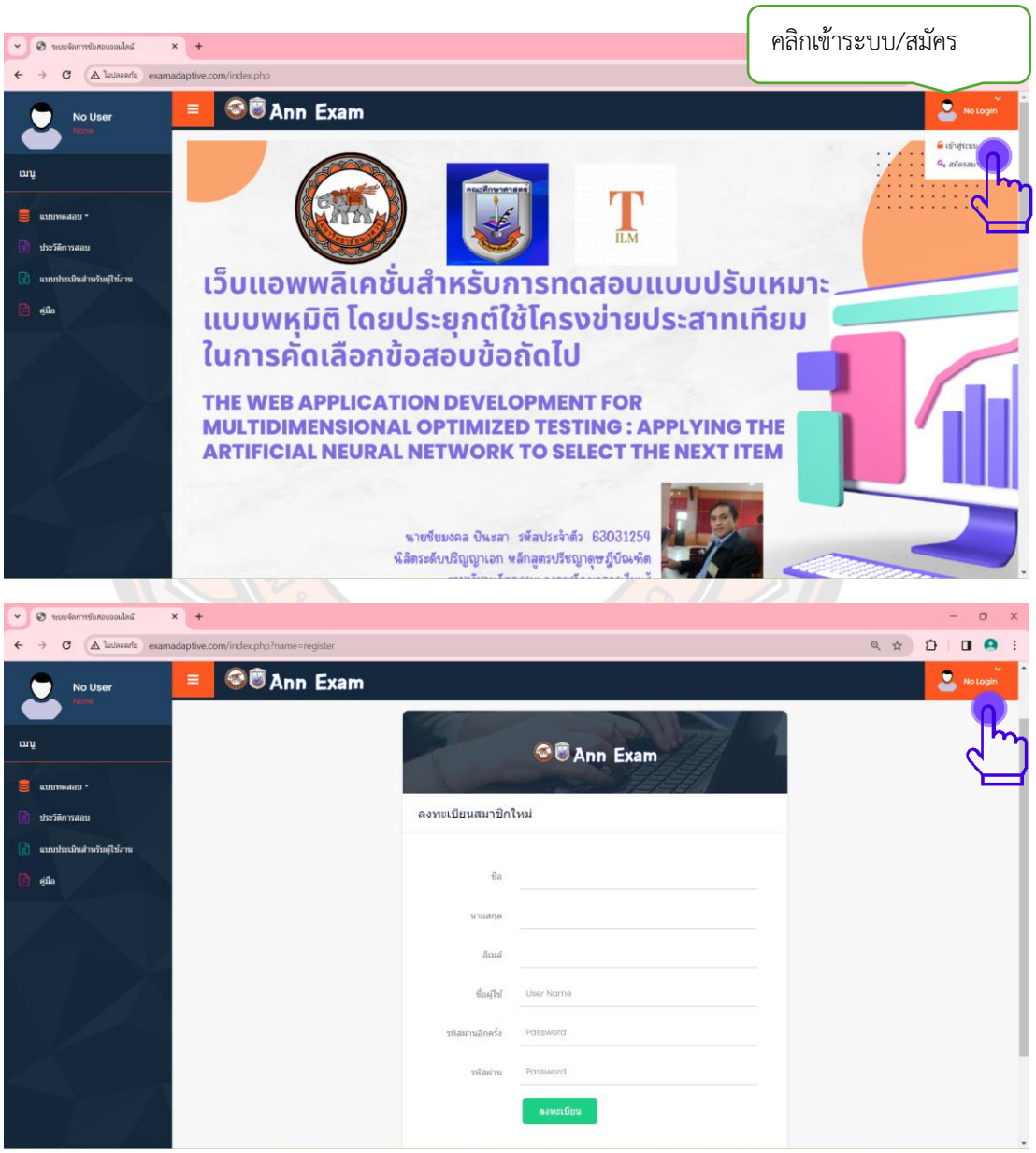
เมื่อเข้าใช้งานโปรแกรมที่เว็บไซต์ http://examadaptive.com/index.php



ภาพที่ 7 หน้าจอแสดงความหมายและขั้นตอนการทำงานของเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยสรุป

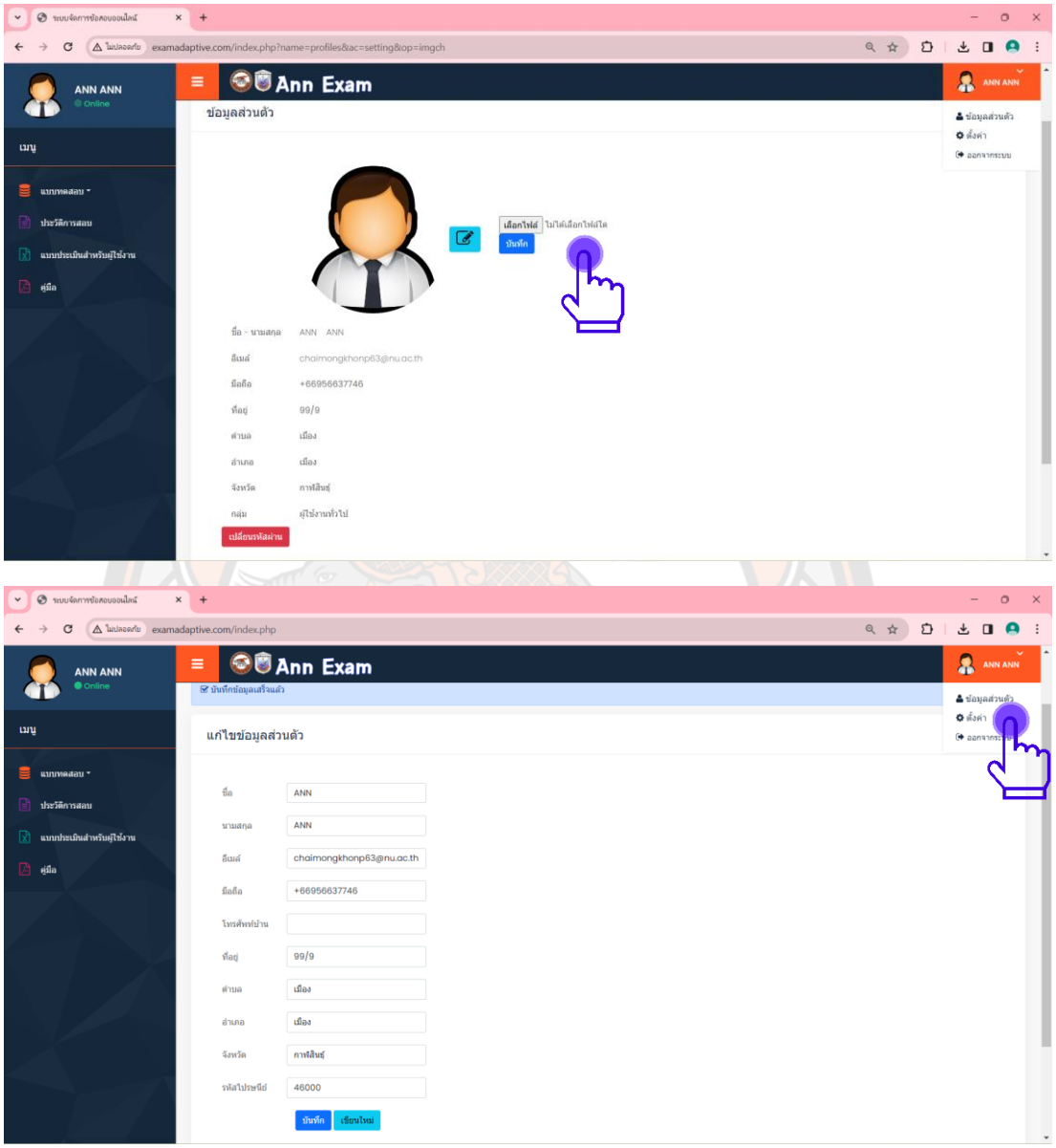
ขั้นตอนการลงทะเบียน

1) หน้าจอการเข้าระบบและลงทะเบียนสมัครสมาชิก มุมบนขวา มีลักษณะเป็นแบบฟอร์มให้ผู้สอบพิมพ์และเลือกข้อมูลของตนเองตามความเป็นจริงลงในช่องว่างตามหัวข้อที่กำหนด ประกอบด้วย 1) ชื่อ 2) นามสกุล 3) อีเมล์ 4) ชื่อผู้ใช้ และ 5) รหัสผ่าน/รหัสผ่านอีกครั้ง ดังภาพที่ 91



ภาพที่ 8 หน้าจอแสดงแบบฟอร์มให้ผู้สอบพิมพ์ข้อมูลและเลือกข้อมูลของตนเองตามความเป็นจริง เมื่อกดปุ่ม ลงทะเบียน

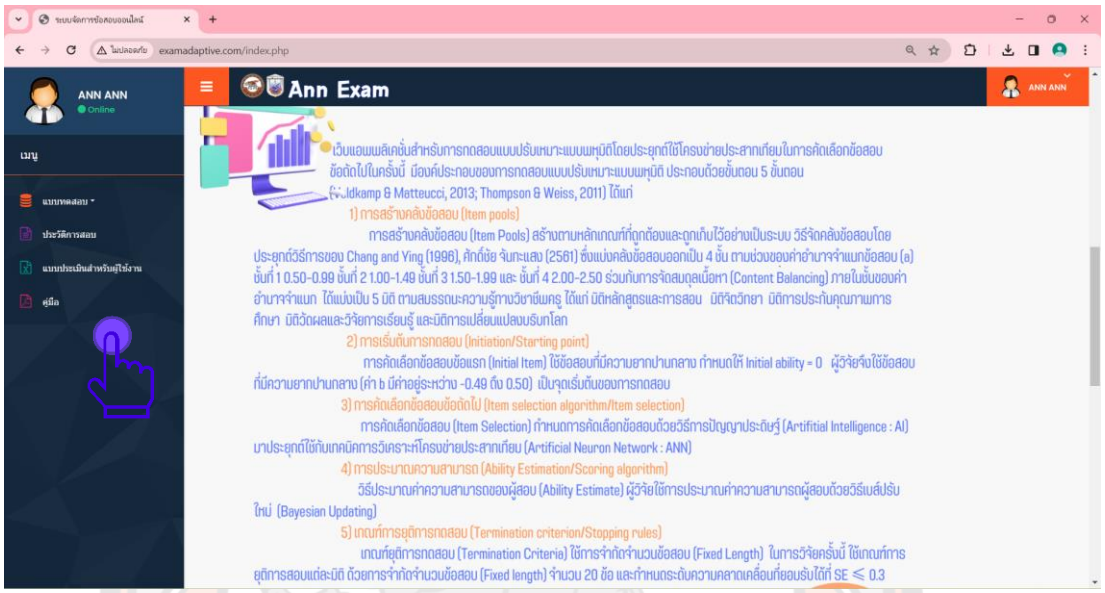
3) หน้าจอแรกของการเข้าใช้งานโปรแกรมโดยกดปุ่ม **เข้าระบบ** ที่แถบเมนูด้านบนเพื่อเข้าใช้งานแอพพลิเคชั่น แสดงดังภาพที่ 9



ภาพที่ 9 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้ สามารถตั้งค่าการจัดการข้อมูลผู้ใช้และแก้ไขข้อมูลส่วนตัว

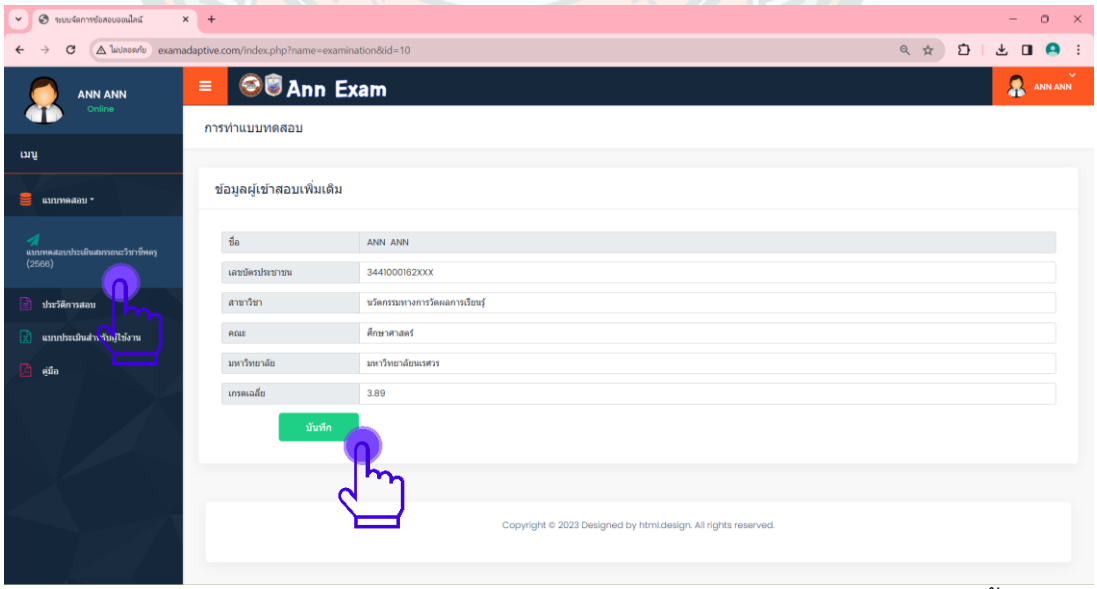
ขั้นตอนการเริ่มทำแบบทดสอบ

1) หน้าจอเริ่มทำข้อสอบ มุมซ้ายมือ ประกอบด้วย 1) แบบทดสอบ 2) ประวัติการสอบ 3) แบบประเมินสำหรับผู้ทดสอบ และ 4) คู่มือการใช้แอพพลิเคชั่น แสดงดังภาพที่ 10



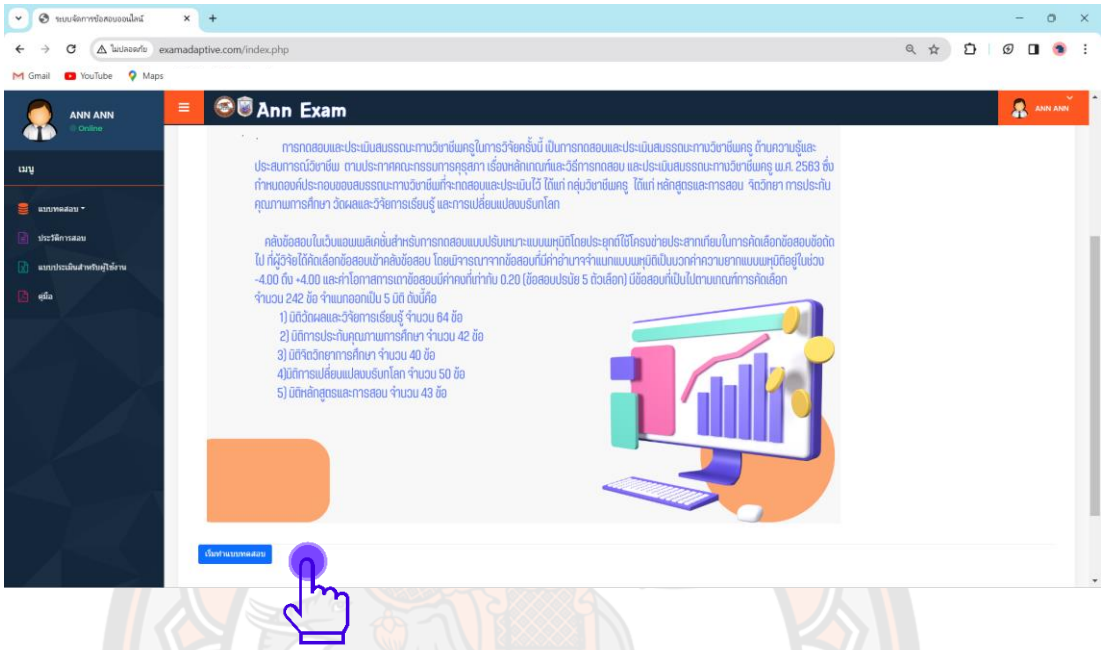
ภาพที่ 10 หน้าจอเริ่มทำข้อสอบ มุมซ้ายมือ

2) หน้าจอเริ่มทำข้อสอบมุมซ้ายมือก่อนจะทำแบบทดสอบ ระบบจะให้กรอกข้อมูลเบื้องต้นของผู้สอบเพิ่มเติม แล้วกด **บันทึก** แสดงดังภาพที่ 11

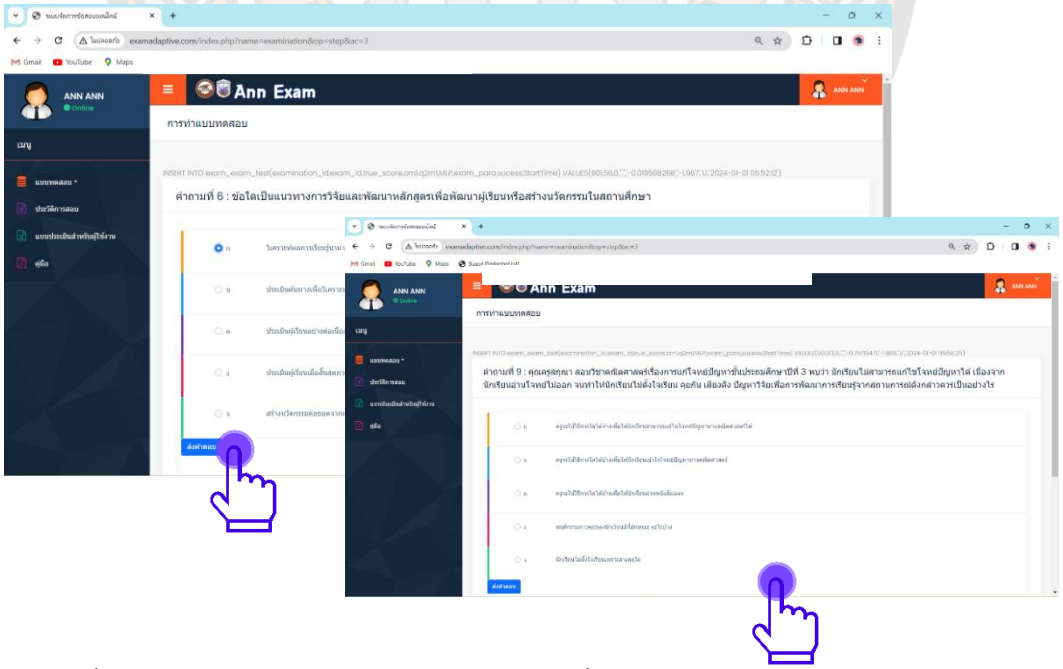


ภาพที่ 11 หน้าจอแสดงข้อมูลของผู้สอบตามที่ได้ลงทะเบียนไว้ ผู้สอบตรวจสอบข้อมูลเบื้องต้นของผู้สอบแล้วกดปุ่ม บันทึก

3) หลังจากผู้สอบตรวจสอบข้อมูลเบื้องต้นของผู้สอบแล้ว จะปรากฏหน้าจอข้อมูลรายวิชาและรายละเอียดแบบทดสอบ ให้ผู้สอบอ่านคำชี้แจงให้ละเอียด ก่อนกด **เริ่มทำแบบทดสอบ**



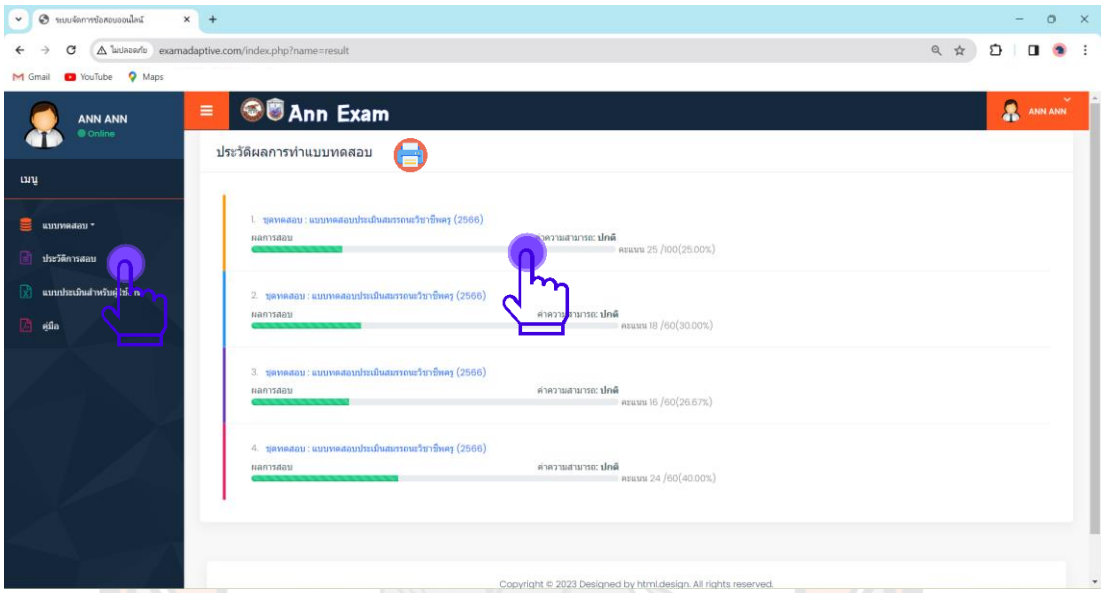
4) หน้าจอแสดงการทำแบบทดสอบ ส่วนนี้หน้าจอจะแสดงคำถาม และตัวเลือก 5 ตัวเลือก



ภาพที่ 12 หน้าจอแสดงตัวเลือกตอบ 5 ตัวเลือก เมื่อผู้สอบอ่านคำถามจนเข้าใจแล้วผู้สอบจะต้องกดเลือกตอบเพียง 1 ตัวเลือก และกดปุ่มส่งคำตอบ

ขั้นตอนการแสดงผลการสอบ

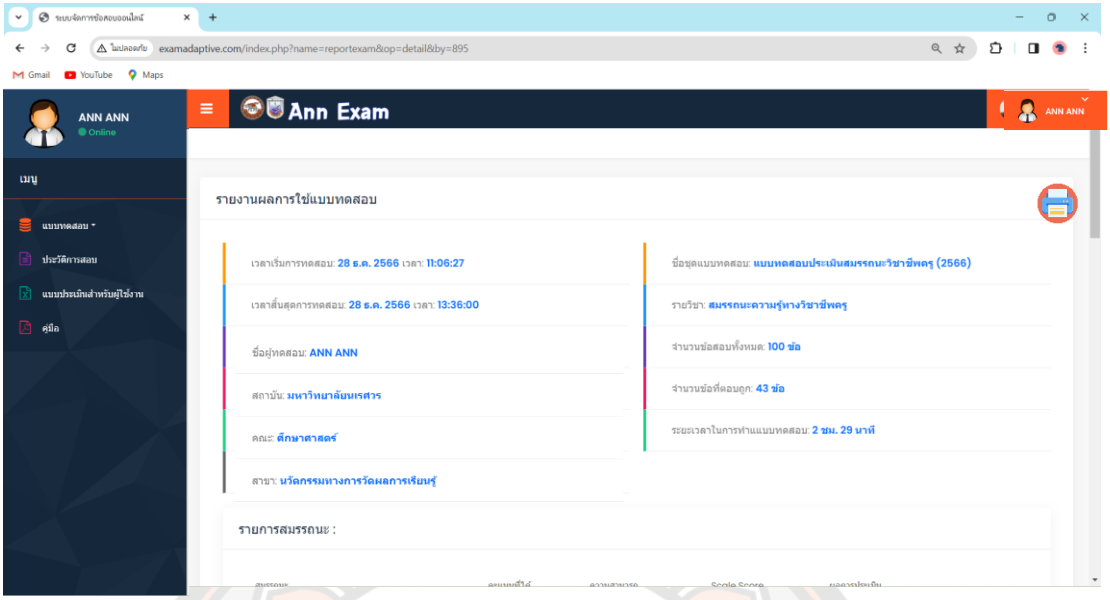
1) หน้าจอแสดงผลการทำแบบทดสอบ เมื่อผู้สอบทำข้อสอบครบทั้ง 5 มิติ แล้วให้ผู้สอบกดปุ่ม **ประวัติการสอบ** ที่แถบเมนูด้านซ้าย เพื่อดูผลการสอบ โปรแกรมจะแสดงข้อมูลตามที่ผู้สอบลงทะเบียนไว้ และผลการทำแบบทดสอบ ดังภาพที่ 13



ภาพที่ 13 หน้าจอแสดงข้อมูลตามที่ผู้สอบลงทะเบียนไว้ และประวัติผลการทำแบบทดสอบ

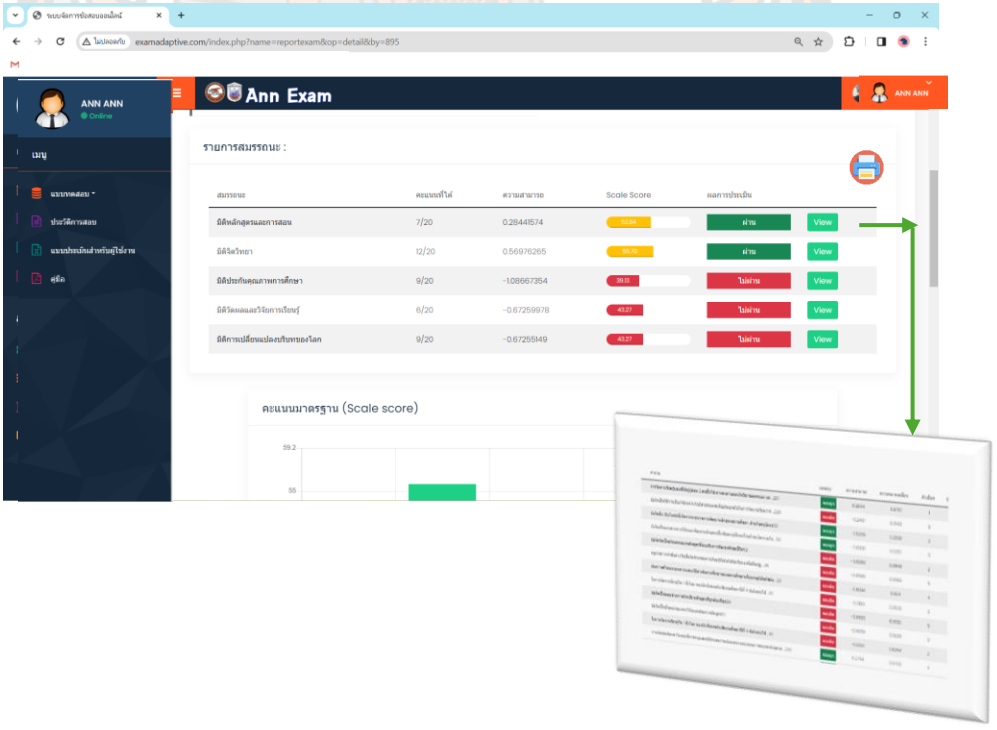
2) หน้าจอแสดงผลประวัติผลการทำแบบทดสอบ สามารถกดดูรายงานผลการใช้แบบทดสอบ ผลการวิเคราะห์ความสามารถของผู้สอบแยกตามมิติของข้อสอบ และสรุปผลการตอบแบบทดสอบแยกตามมิติของข้อสอบ

3) หน้าจอผลการวิเคราะห์ความสามารถของผู้สอบแยกตามมิติของข้อสอบหน้าจอนี้จะแสดงค่าความสามารถและผลการประเมินระดับความสามารถของผู้สอบแยกตามมิติของข้อสอบ คือ มิติหลักสูตรและการสอน มิติจิตวิทยา มิติการประกันคุณภาพการศึกษา มิติวัดผลและวิจัยการเรียนรู้ และมิติการเปลี่ยนแปลงบริบทโลกและความสามารถโดยรวม ดังภาพที่ 14



ภาพที่ 14 หน้าจอแสดงผลรายงานผลการใช้แบบทดสอบ

4) หน้าจอผลการวิเคราะห์ความสามารถ Theta (θ) ของผู้สอบแยกตามมิติของข้อสอบ



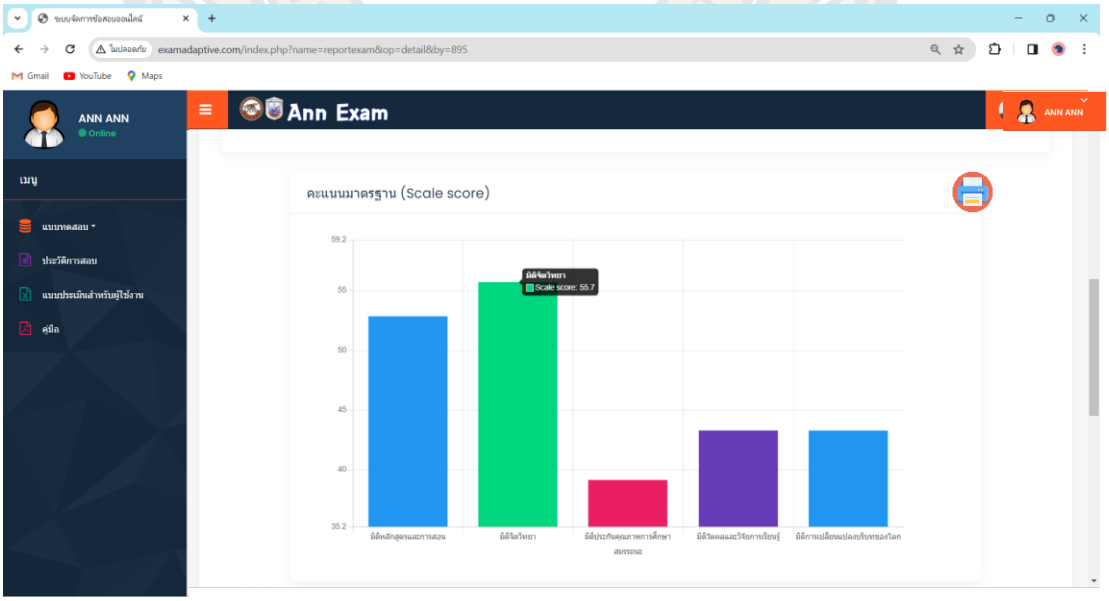
ภาพที่ 15 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบ Theta (θ) , ค่า Scale Score และผลการประเมินแยกตามมิติ

จากภาพ 15 การรายงานผลการทดสอบจะเป็นการรายงานผลคะแนนความสามารถในแต่ละมิติ ผลการทดสอบรายงานในรูปของคะแนนความสามารถ คือค่า Theta (θ) ซึ่งได้จากการประมาณค่าตามแนวทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ โดยค่า Theta มีช่วงของคะแนนอยู่ระหว่าง -2.00 ถึง +2.00 และเพื่อให้การแปลความหมายคะแนนง่ายต่อการรายงานผลการทดสอบ จึงได้ปรับคะแนน Theta ให้อยู่ในรูปของคะแนนมาตรฐาน (scale score) ซึ่งเป็นการปรับค่าคะแนนให้มีค่าเฉลี่ยเท่ากับ 50 และมีค่าส่วนเบี่ยงมาตรฐานเท่ากับ 10 โดยคะแนนมาตรฐานนี้มีช่วงคะแนนอยู่ระหว่าง 20 ถึง 80 คะแนน ผู้วิจัยได้เลือกใช้เกณฑ์การแปลความหมาย ดังนี้

ตาราง 1 คะแนนความสามารถTheta ที่ปรับเป็นคะแนนมาตรฐาน (scale score)

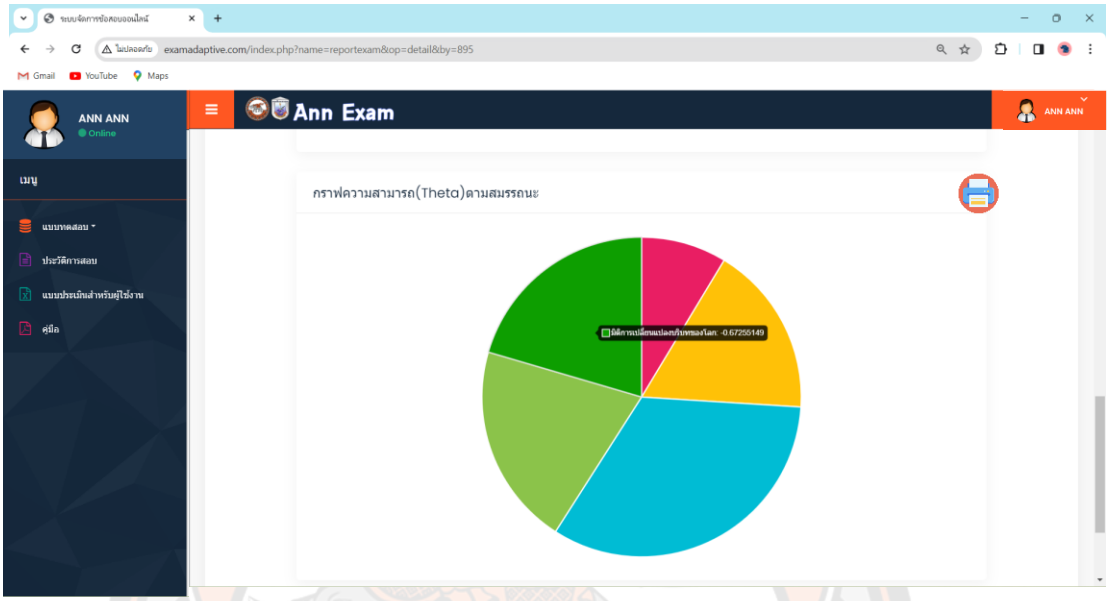
คะแนนที่ปรับสเกลมาตรฐาน (scale score) =50+(10* Theta (θ))	ผลการประเมิน
น้อยกว่า 50	ไม่ผ่าน
50 – 64	ผ่าน
ตั้งแต่ 65 ขึ้นไป	ดี

5) หน้าจอแสดงผลคะแนนที่ปรับในสเกลมาตรฐาน (scale score) ของผู้สอบแยกตามมิติของข้อสอบ ดังภาพที่ 16



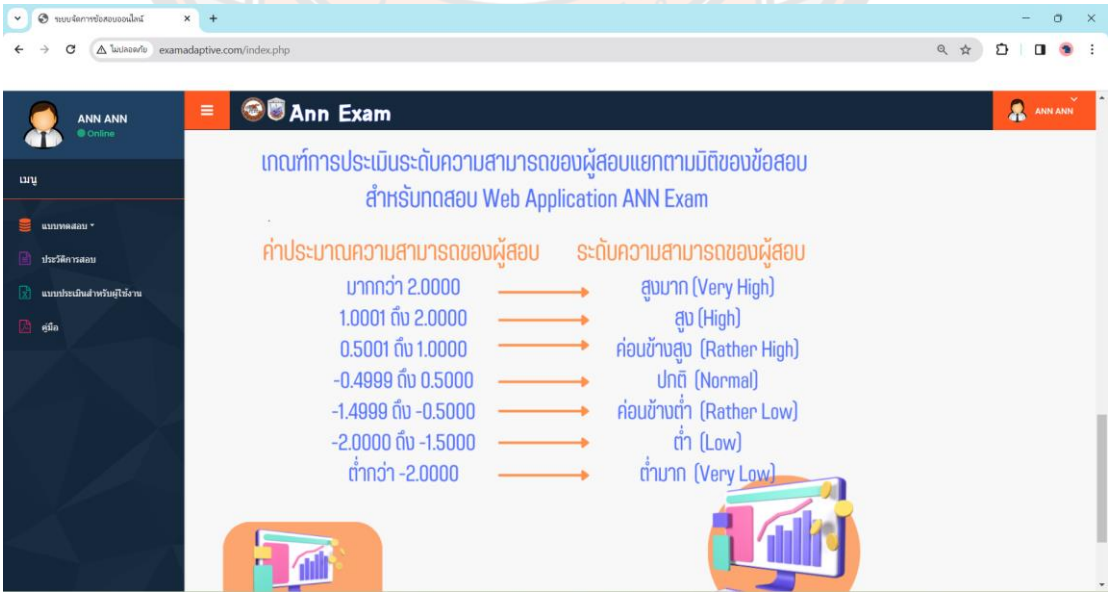
ภาพที่ 16 หน้าจอแสดงผลคะแนนที่ปรับในสเกลมาตรฐาน (scale score) ของผู้สอบแยกตามมิติของข้อสอบ

19) หน้าจอแสดงผลกราฟความสามารถ Theta (θ) ของผู้สอบแยกตามมิติของข้อสอบ ดังภาพที่ 17



จากภาพที่ 17 หน้าจอแสดงผลการวิเคราะห์ความสามารถของผู้สอบและผลการประเมินระดับความสามารถของผู้สอบแยกตามมิติของข้อสอบ และสรุปความสามารถโดยรวม ซึ่งผู้สอบสามารถเทียบระดับความสามารถของตนเองได้โดยนำค่าความสามารถที่ได้จากการทดสอบไปเทียบกับเกณฑ์ ดังภาพที่ 18

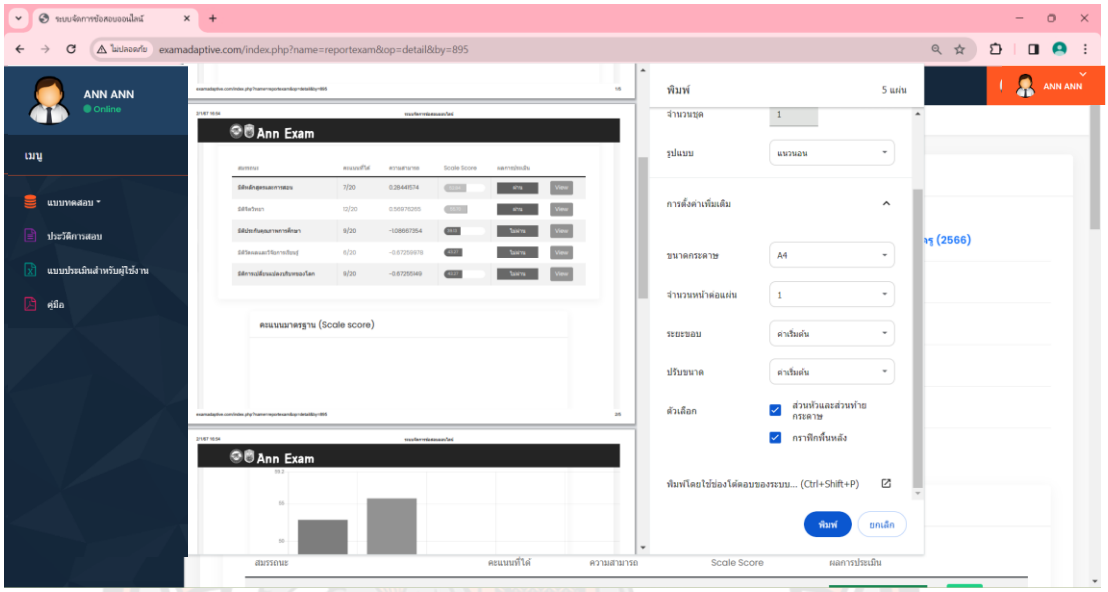
19) หน้าจอแสดงเกณฑ์การประเมินความสามารถของผู้สอบแยกตามมิติของข้อสอบ



ภาพที่ 18 หน้าจอเกณฑ์การประเมินความสามารถของผู้สอบแยกตามมิติของข้อสอบ

ขั้นตอนการพิมพ์รายงานผลการสอบ

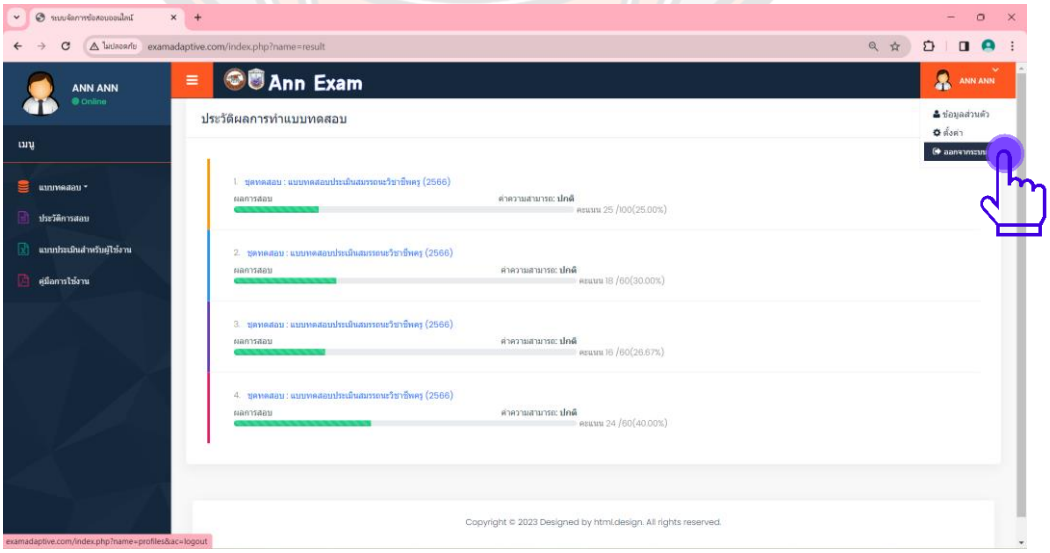
1) เมื่อผู้สอบต้องการพิมพ์สรุปผลการทดสอบความสามารถ ให้กดปุ่ม Ctel+P ที่ คีย์บอร์ด Keyboard จะแสดงหน้าจอที่มีเมนูให้เลือกเครื่องพิมพ์ และปุ่มสั่งพิมพ์ แสดงดังภาพที่ 19



ภาพที่ 19 หน้าจอแสดงการกดปุ่ม Print จะแสดงหน้าจอที่มีเมนูให้เลือกเครื่องพิมพ์และปุ่มสั่งพิมพ์

ขั้นตอนการออกจากโปรแกรม

1) เมื่อผู้สอบต้องการออกจากโปรแกรม ให้กดปุ่ม LOGOUT ที่แถบเมนูด้านบนขวามือของเว็บแอพพลิเคชั่น การออกจากระบบ แสดงดังภาพที่ 20



ภาพที่ 20 หน้าจอแสดงการกดปุ่ม ที่แถบเมนูด้านบนขวา เพื่อออกจากระบบ

การพัฒนาวิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียม

1. วิธีการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียมซึ่งเป็นเลือกใช้โครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น (Multi-Layer Perceptron : MLP) การเรียนรู้แบบมีผู้สอน (supervised learning) แบบการแพร่ย้อนกลับ (back propagation Learning) ซึ่งประกอบด้วยเซลล์ประสาท 3 ชั้น ได้แก่ชั้นชั้นข้อมูลป้อนเข้า (n_{input}) เท่ากับ 25 โหนด และจำนวนโหนดในชั้นซ่อน $(n_{hidden1})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden2})$ จำนวน 10 โหนด โหนดในชั้นซ่อน $(n_{hidden3})$ จำนวน 10 โหนด และจำนวนโหนดชั้นข้อมูลส่งออก (n_{output}) จำนวน 2 โหนด (รูปแบบข้อมูล 25-10-10-10-2) ทำการออกแบบโมเดลโดยใช้เทคนิคดาต้าไมน์นิง (Data mining) หรือการทำเหมืองข้อมูล ในการจำแนกประเภทข้อมูล (Classification) และใช้โปรแกรม RapidMiner Studio Educational 10.1.003 วิเคราะห์โดยใช้อัลกอริธึม Neural Net เพื่อเป็นการเพิ่มประสิทธิภาพความแม่นยำในการคัดเลือกข้อสอบข้อถัดไปให้มีความถูกต้องมากที่สุด และผู้วิจัยได้กำหนดค่าพารามิเตอร์ในกระบวนการของโครงข่ายประสาทเทียม ประกอบด้วย อัตราการเรียนรู้เท่ากับ 0.95 โมเมนตัม (Momentum) เท่ากับ 0.05 ในแต่ละสถานการณ์จะมีการทำซ้ำ 500 รอบ

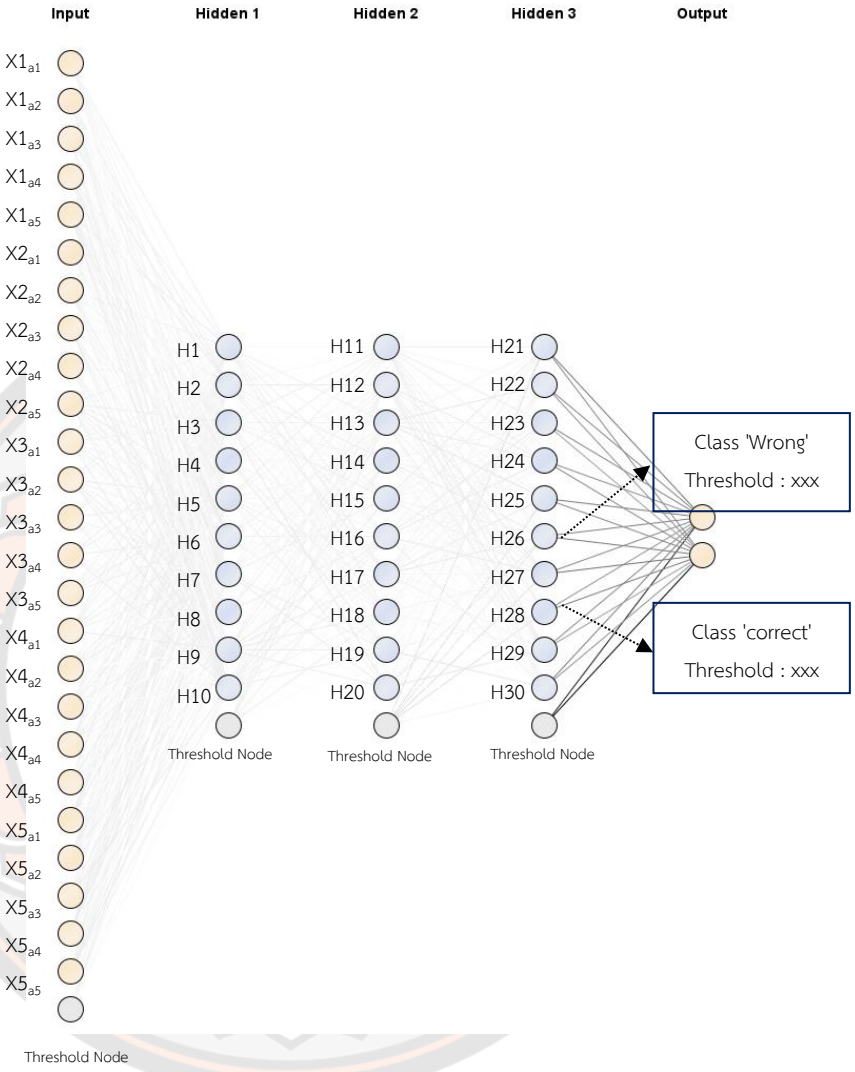
ตารางที่ 2 รายละเอียดข้อมูลสำหรับทำเหมืองข้อมูล ข้อสอบ 5 ข้อ

Variable/ Node	Detail	Key code
$X1_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 1	Parameter value
$X1_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 2	Parameter value
$X1_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 3	Parameter value
$X1_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 4	Parameter value
$X1_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 1 ในมิติที่ 5	Parameter value
$X2_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 1	Parameter value
$X2_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 2	Parameter value
$X2_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 3	Parameter value
$X2_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 4	Parameter value
$X2_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 2 ในมิติที่ 5	Parameter value
$X3_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 1	Parameter value
$X3_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 2	Parameter value

Variable/ Node	Detail	Key code
$X3_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 3	Parameter value
$X3_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 4	Parameter value
$X3_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 3 ในมิติที่ 5	Parameter value
$X4_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 1	Parameter value
$X4_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 2	Parameter value
$X4_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 3	Parameter value
$X4_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 4	Parameter value
$X4_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 4 ในมิติที่ 5	Parameter value
$X5_{a1}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 1	Parameter value
$X5_{a2}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 2	Parameter value
$X5_{a3}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 3	Parameter value
$X5_{a4}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 4	Parameter value
$X5_{a5}$	ค่าอำนาจจำแนกของข้อสอบข้อที่ 5 ในมิติที่ 5	Parameter value

จากตารางที่ 2 นำค่าพารามิเตอร์ของค่าอำนาจจำแนกแบบพหุมิติของข้อสอบ (MDISC) ในคลังข้อสอบ จำนวน 5 ข้อ ที่มีค่าระดับปานกลาง (Moderate) อยู่ระหว่าง $0.65 \leq MDISC < 1.35$ (Baker, 2001; Hasmy, 2014) มาทำการทดสอบประสิทธิภาพในการจำแนกข้อมูลโดยการวัดความถูกต้องของอัลกอริธึม Neural Net ด้วยวิธีการ 10-fold cross-validation ซึ่งจะแบ่งข้อมูลออกเป็น 10 ชุดเท่าๆ กันแล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ (Test set) 20% ส่วนที่เหลือ 9 ชุด นำมาใช้เป็นกลุ่มเรียนรู้ (Training Set) 80% แล้วทำการ วนทำเป็นจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อยๆ จนครบ ในการวัดประสิทธิภาพของการจำแนกประเภทข้อมูลได้แก่ การหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถ่วงดุล (F-measure) และค่าความถูกต้อง (Accuracy)

โมเดลการคัดเลือกข้อสอบข้อถัดไปใช้โปรแกรม RapidMiner Studio Educational 10.1 วิเคราะห์โดยใช้อัลกอริธึม Neural Net ผลการวิเคราะห์ปรากฏดังภาพ 21



ภาพที่ 21 โมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) กรณีตอบถูก และตอบผิด

การนำโมเดลการคัดเลือกข้อสอบข้อถัดไปโดยใช้วิธีโครงข่ายประสาทเทียมไปใช้จริงได้นั้น จำเป็นต้องทราบประสิทธิภาพของโมเดลก่อน โดยทั่วไปแล้วจะมีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่างๆ อยู่ 4 ค่า ดังนี้ 1) ค่าความแม่นยำ (Precision) คือค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์ 2) ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล 3) ค่าความถ่วงดุล (F-measure) คือค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก และ 4)

ค่าความถูกต้อง (Accuracy) คือ จำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส โดยต้องมีค่าสูงที่สุด

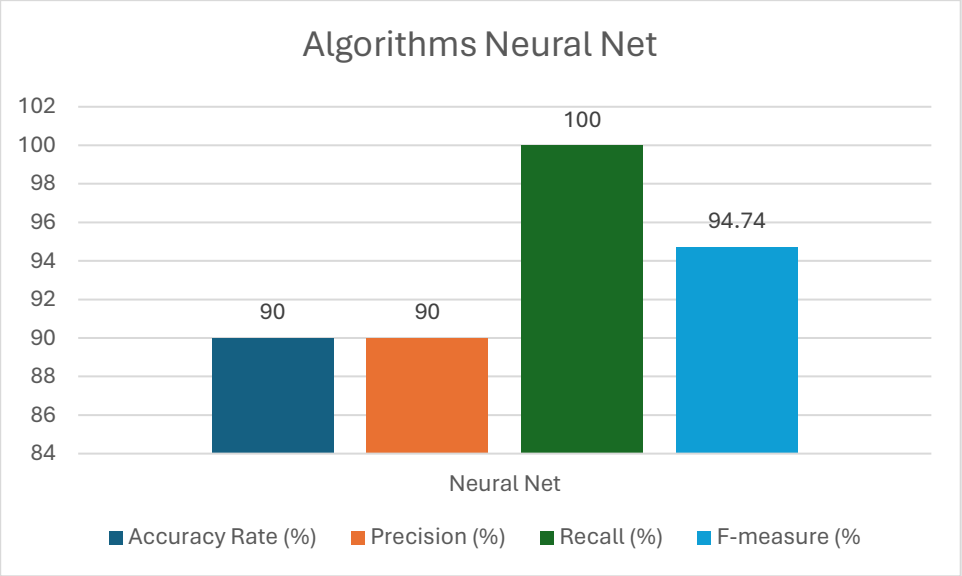
จากการทดลองโมเดล สามารถอธิบายผลการทดลองได้ ดังนี้

1. การทดสอบประสิทธิภาพ ด้วยวิธี แบบวิธี 10 Fold Cross-Validation เพื่อหาประสิทธิภาพของโมเดลโครงข่ายประสาทเทียม ดังตารางที่ 3

ตารางที่ 3 ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบ

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%
Neural Net	90	90	100	94.74

จากตารางที่ 3 พบว่าจากการประเมินประสิทธิภาพด้วยวิธีการ 10-fold cross-validation ผลลัพธ์การวัดประสิทธิภาพของอัลกอริธึม Neural Network สำหรับคัดเลือกข้อสอบข้อถัดไป ค่าความระลึก (Recall) คิดเป็นร้อยละ 100 สูงที่สุด รองลงมาคือ ค่าความถ่วงดุล (F-measure) คิดเป็นร้อยละ 94.70, ค่าความถูกต้อง (Accuracy) คิดเป็นร้อยละ 90 และการวัดค่าความแม่นยำ (Precision) คิดเป็นร้อยละ 90 ตามลำดับ



ภาพที่ 22 กราฟเปรียบเทียบค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึกและค่าความถ่วงดุลของการคัดเลือกข้อสอบข้อถัดไป

นำโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) มาเขียนสมการทางคณิตศาสตร์ ได้ดังนี้ :

Hidden Layer 1 : สำหรับ $Node_i$ (โดย i=1,2,...,10):

$$a_1^{(1)} = Sigmoid(\sum_{j=1}^{5} parameter1_a_j^{(1)} \cdot X_j + Bias)$$

Hidden Layer 2 : สำหรับ $Node_i$ (โดย i=1,2,...,10):

$$a_1^{(2)} = Sigmoid(\sum_{j=1}^{10} parameter2_a_j^{(2)} \cdot a_j^{(1)} + Bias)$$

Hidden Layer 3 : สำหรับ $Node_i$ (โดย i=1,2,...,10):

$$a_1^{(3)} = Sigmoid(\sum_{j=1}^{10} parameter3_a_j^{(3)} \cdot a_j^{(2)} + Bias)$$

Output Layer[1] : สำหรับ Class 'wrong' : กรณีตอบผิด

$$o_{wrong} = Sigmoid(\sum_{j=1}^{10} parameter_{Wrong,i}_a_j^{(3)} \cdot a_j^{(2)} + Threshold_{wrong})$$

Output Layer[2] : สำหรับ Class 'correct' : กรณีตอบถูก

$$o_{Correct} = Sigmoid(\sum_{j=1}^{10} parameter_{correct,i}_a_j^{(3)} \cdot a_j^{(2)} + Threshold_{correct})$$

โดยที่:

x_j คือ ข้อมูลนำเข้า (input) สำหรับ $Node_j$ ใน Hidden Layer 1.

$a_i^{(l)}$ คือ ผลลัพธ์ที่ได้จาก $Node_i$ ใน $Hidden\ Layer_l$.

$parameter_{a,j}^{(l)}$ และ $paramete_{class,i}$ คือ พารามิเตอร์ (parameters) ที่ถูกเรียนรู้ของโมเดลสำหรับ $Node_j$ ใน $Layer_l$ และ $Class_i$ ใน Output Layer ตามลำดับ

Bias คือ พารามิเตอร์ที่เกี่ยวข้องกับการเทียบค่าเบื้องต้น (constant term) ในทุกๆ Node.

Sigmoid (x) คือ ฟังก์ชัน Sigmoid ที่นิยมใช้ใน Neural Network และมีสูตร คือ

$$\frac{1}{1+e^{-x}}$$

ส่วนสมการคำนวณ Output และการตัดสินใจในโมเดลการคัดเลือกข้อสอบข้อถัดไปใช้การตัดสินใจ (decision-making) โดยนำการใช้ฟังก์ชัน sigmoid เพื่อแปลงค่าผลลัพธ์ในแต่ละ

โหนดของ Neural Network ให้อยู่ในช่วงระหว่าง 0 ถึง 1 ซึ่งสามารถใช้ในการตัดสินใจว่า output Layer นั้นๆ ควรจะถูกจัดอยู่ใน Class ไหน ดังสมการ ได้ดังนี้

1. สมการของ Hidden Layer 1 :

คำนวณผลลัพธ์ใน Hidden Layer 1 ได้ด้วยฟังก์ชัน Sigmoid ตามสมการ :

$$Node_i = sigmoid(\sum_{j=1}^{5} parameter_{i,a_j} . input_{a_j} + Bias_i)$$

โดยที่

i คือ หมายเลขโหนดใน Hidden Layer 1 (เช่น Node 1, Node 2, .., Node 10)

j คือ ค่าพารามิเตอร์ (parameter) ในแต่ละโหนด (จาก a_1 ถึง a_5)

ยกตัวอย่าง เช่น สำหรับ Node 1 ใน Hidden Layer 1:

$$Node_1 = sigmoid(\sum_{j=1}^{5} parameter_{1,a_j} . input_{a_j} + Bias_1)$$

และในที่นี้ a_j คือ a_1, a_2, a_3, a_4, a_5 ซึ่งแทนค่า input ที่เป็นข้อมูลนำเข้าของ Hidden Layer 1 สำหรับ Hidden Layer 2 และ Hidden Layer 3 ก็สามารถทำการคำนวณได้ตามลำดับเดียวกันโดยนำผลลัพธ์จากชั้นก่อนหน้ามาเป็น input ของชั้นถัดไป

2. สมการของ Output Layer :

ใน Output Layer มีสอง Class คือ "ผิด (wrong)" และ "ถูก (correct)" โดยใช้ฟังก์ชัน Sigmoid เช่นกันสมการสำหรับการคำนวณใน Output Layer จะเป็นดังนี้ :

$$Class\ 'wrong' = sigmoid(\sum_{i=1}^{10} Node_i . input_{a_j} + outputParameter_i + outputBias)$$

$$Class\ 'correct' = sigmoid(\sum_{i=1}^{10} Node_i . input_{a_j} + outputParameter_i + outputBias)$$

โดยที่ i คือ หมายเลข Node ใน Hidden Layer3

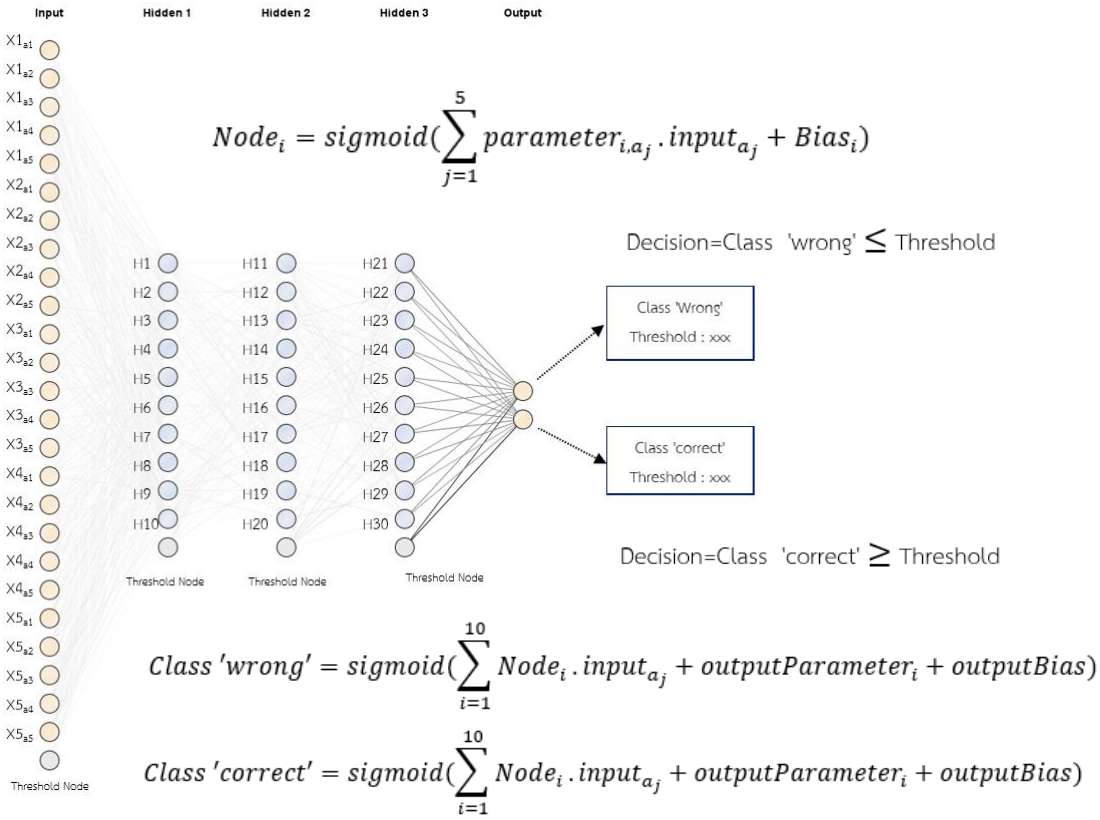
3. สมการของตัดสินใจ (decision-making) คัดเลือกข้อสอบข้อถัดไปโดยใช้ Threshold : โดยนำ Output $Class_{'correct', 'wrong'}$ มาเปรียบเทียบกับ Threshold ได้สมการได้ดังนี้

Decision=Class 'correct' $\geq$ Threshold ; กรณีตอบถูก

Decision=Class 'wrong' $\leq$ Threshold ; กรณีตอบผิด

การคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) จะทำการคัดเลือกข้อสอบที่มีการประมาณค่าพารามิเตอร์ความยากแบบพหุมิติ (Threshold) ของข้อสอบ ที่ใกล้เคียงกับค่าประมาณความสามารถ (θ) ขณะนั้นที่สุด ข้อสอบข้อไหนที่มีค่าพารามิเตอร์ความยาก (Threshold) ของข้อสอบ ใกล้เคียงกับการประมาณค่าของโหนด Output (Output Node) มากที่สุดข้อสอบข้อนั้นจะถูกเลือกเป็นข้อสอบข้อถัดไป อยู่ 2 กรณี 1) กรณีตอบถูก (correct) ค่า Threshold ข้อถัดไป ต้องสูงกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after answering correctly) 2) กรณีตอบผิด (wrong) ค่า Threshold ข้อถัดไป ต้องน้อยกว่าหรือใกล้เคียงค่า Threshold ข้อก่อนหน้า (Next question after answering incorrectly)

A model for selecting the next exam using a neural network method



ภาพที่ 23 สมการโมเดลการคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN) (25-10-10-10-2)

จากนั้นนำสมการของโมเดลการคัดเลือกข้อสอบข้อถัดไปนี้มาเขียนเป็น Code PHP เนื่องจากการเขียน Code PHP Hypertext Preprocessor มีความซับซ้อนมาก ผู้วิจัยจึงกำหนดฟังก์ชันคำนวณผลลัพธ์ โดยใช้ Forward Propagation ของแต่ละ Node และ Layer ของ Neural Network ด้วยค่าพารามิเตอร์ที่กำหนดฟังก์ชัน Sigmoid มาทำการคำนวณผลลัพธ์ของ Neural Network ดังนี้

```php
<?php
function sigmoid($x) {
return 1 / (1 + exp(-$x));
}
function makeDecision($output, $threshold) {
return $output >= $threshold;
}
function forwardPropagation($inputValues, $parameters, $biases) {
$A = $inputValues;
foreach ($parameters as $W) {
$Z = dotProduct($W, $A) + array_pop($biases);
$A = sigmoid($Z);
}
return $A;
}
// Example usage
$inputValues = [/* parameters_a1_a5.txt */];  ##(your input values)
$outputThreshold = 0.5;
$hidden1Params = [/* parameters for hidden layer 1 */];
$hidden2Params = [/* parameters for hidden layer 2 */];
$hidden3Params = [/* parameters for hidden layer 3 */];
$outputParams = [/* parameters for output layer */];
$hiddenBiases = [
$hidden1Biases = [/* biases for hidden layer 1 */],
$hidden2Biases = [/* biases for hidden layer 2 */],
$hidden3Biases = [/* biases for hidden layer 3 */],
$outputBiases = [/* biases for output layer */]
];
$parameters = [$hidden1Params, $hidden2Params, $hidden3Params, $outputParams];
$biases = $hiddenBiases;
$output = forwardPropagation($inputValues, $parameters, $biases);
$decision = makeDecision($output, $outputThreshold);
if ($decision) {
$nextQuestion = "Next question after answering correctly";
} else {
$nextQuestion = "Next question after answering incorrectly";
}
echo $nextQuestion;
function dotProduct($array1, $array2) {
$result = 0;
$length = min(count($array1), count($array2));
for ($i = 0; $i < $length; $i++) {
$result += $array1[$i] * $array2[$i];
}
return $result;
}
?>
```

ภาพที่ 24 Code PHP สำหรับคัดเลือกข้อสอบข้อถัดไปด้วยวิธีการใช้โครงข่ายประสาทเทียม (ANN)

ภาคผนวก ฉ

แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานของแอพพลิเคชั่น

สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ

ที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

(สำหรับผู้เชี่ยวชาญ)

Ann Exam

แบบประเมินสำหรับผู้เชี่ยวชาญ

แบบประเมินความถูกต้องเหมาะสมของแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

คำชี้แจง

1. แบบประเมินความเหมาะสมของแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ฉบับนี้มีจุดมุ่งหมายเพื่อทราบความคิดเห็นที่มีต่อการใช้งานแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติการทดสอบ รวมทั้งคู่มือการใช้แอพพลิเคชั่นเพื่อนำผลที่ได้มาเป็นข้อมูลในการพัฒนาแอพพลิเคชั่นต่อไป

2. แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ มีลักษณะเป็นมาตราส่วนประเมินค่า (Rating Scale) 5 ระดับ โดยแบ่งเป็น 2 ตอน ได้แก่

ตอนที่ 1 ความคิดเห็นของผู้เชี่ยวชาญ

ให้ทำเครื่องหมาย ☐ลงในช่องที่ตรงกับความคิดเห็นของท่านมากที่สุด มีเกณฑ์การประเมินระดับความคิดเห็น ดังนี้

ระดับ 5 หมายถึง ความเหมาะสมมากที่สุด
ระดับ 4 หมายถึง ความเหมาะสมมาก
ระดับ 3 หมายถึง ความเหมาะสมปานกลาง
ระดับ 2 หมายถึง ความเหมาะสมน้อย
ระดับ 1 หมายถึง ความเหมาะสมน้อยที่สุด

ตอนที่ 1 การประเมินความเหมาะสมในการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ มี 2 ตอนดังนี้

1. ด้านความสามารถของผู้ใช้โปรแกรม (Functional Requirement) จำนวน 5 ข้อ
2. ด้านการทำงานของโปรแกรม (Functional) จำนวน 5 ข้อ
3. ความสะดวกในการใช้งานโปรแกรม (Usability) จำนวน 7 ข้อ
4. ด้านความชัดเจนของคู่มือการใช้โปรแกรม (Program Manual) จำนวน 5 ข้อ

ตอนที่ 2 ข้อเสนอแนะอื่น ๆ

ในกรณีที่ท่านมีข้อเสนอแนะอื่น ๆ เพิ่มเติม เพื่อประโยชน์ในการปรับปรุงและพัฒนาแอพพลิเคชั่นต่อไปขอให้ท่านเขียนข้อเสนอแนะลงในช่องว่าง

ขอขอบพระคุณเป็นอย่างสูง
นายชัยมงคล ปีนะสา
ผู้วิจัย

รายการ	ระดับความพึงพอใจ				
	5	4	3	2	1
1.1 เว็บแอพพลิเคชั่นสามารถจัดการทดสอบได้ตรงตาม	○	○	○	○	○

Ann Exam

และแสดงผลข้อมูล	○	○	○	○	○
1.3 เว็บแอพพลิเคชั่นสามารถค่านวณเค่าความสามารถของผู้ใช้งานได้	○	○	○	○	○
1.4 มีการตรวจสอบป้องกันข้อผิดพลาดที่อาจจะเกิดขึ้นจากการใช้เว็บแอพพลิเคชั่น	○	○	○	○	○
1.5 มีระบบป้องกันความลับของผู้ใช้งาน	○	○	○	○	○
2.1 ความถูกต้องในการจัดเก็บข้อมูลนำเข้า	○	○	○	○	○
2.2 ความถูกต้องในการค้นหาข้อมูล	○	○	○	○	○
2.3 ความถูกต้องในการปรับปรุงแก้ไขข้อมูล	○	○	○	○	○
2.4 ความถูกต้องในการลบข้อมูล	○	○	○	○	○
2.5 ความถูกต้องของผลลัพธ์ที่ได้จากการประมวลผลในโปรแกรม	○	○	○	○	○
2.6 ความถูกต้องของการผลลัพธ์ในรูปแบบรายงาน	○	○	○	○	○
2.7 ความรวดเร็วในการประมวลผลของระบบ	○	○	○	○	○
2.8 ความน่าเชื่อถือได้ของระบบ	○	○	○	○	○
2.9 ความครอบคลุมของโปรแกรมที่พัฒนากับระบบงานจริง	○	○	○	○	○
2.10 การป้องกันข้อผิดพลาดที่อาจเกิดขึ้น	○	○	○	○	○
3.1 ความสะดวกในการเรียกใช้งานโปรแกรม	○	○	○	○	○
3.2 ความสะดวกในการเข้าและออกจากโปรแกรม	○	○	○	○	○
3.3 ความสะดวกในการลงทะเบียนเข้าใช้งานโปรแกรม	○	○	○	○	○
3.4 ความสะดวกในการป้อนข้อมูล	○	○	○	○	○
3.5 ความสะดวกในการกดปุ่มเลือกค่าตอบ หรือเปลี่ยนค่าตอบ	○	○	○	○	○
3.6 เมื่อทดสอบเสร็จโปรแกรมจะรายงานผลการทดสอบทันที	○	○	○	○	○

Ann Exam

3.7 ผู้ใช้งานสามารถบันทึกผลการทดสอบ และสั่งพิมพ์ได้อย่างสะดวกรวดเร็ว	○	○	○	○	○
4.1 คู่มือการใช้งานโปรแกรมสามารถอธิบายวัตถุประสงค์ของโปรแกรมได้อย่างชัดเจน	○	○	○	○	○
4.2 คู่มือการใช้โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน	○	○	○	○	○
4.3 ภาษาที่ใช้ในคู่มือการใช้งานโปรแกรมอ่านเข้าใจง่าย และกระชับ	○	○	○	○	○
4.4 คู่มือการใช้งานโปรแกรมมีภาพประกอบการอธิบายให้เห็นชัดเจน	○	○	○	○	○
4.5 หลังจากอ่านคู่มือแล้ว ผู้ใช้งานสามารถ ดูคำชี้แจงจากคู่มือการใช้งานโปรแกรมและสามารถ ปฏิบัติตามได้	○	○	○	○	○

ตอนที่ 2 ข้อเสนอแนะเพื่อการพัฒนาโปรแกรม

ส่งแบบสอบถาม

Copyright © 2023 Designed by html.design. All rights reserved.

ภาคผนวก ช
แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานของแอพพลิเคชั่น สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป
(สำหรับผู้ใช้งาน)

Ann Exam

แบบประเมินสำหรับผู้ใช้งาน

แบบประเมินความถูกต้องเหมาะสมของแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป

คำชี้แจง

1. แบบประเมินความเหมาะสมของแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติที่ประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป ฉบับนี้มีจุดมุ่งหมายเพื่อทราบความคิดเห็นที่มีต่อการใช้งานแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติการทดสอบ รวมทั้งคู่มือการใช้แอพพลิเคชั่นเพื่อนำผลที่ได้มาเป็นข้อมูลในการพัฒนาแอพพลิเคชั่นต่อไป

2. แบบประเมินความถูกต้องเหมาะสมของโปรแกรมและคู่มือการใช้งานแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ มีลักษณะเป็นมาตราส่วนประเมินค่า (Rating Scale) 5 ระดับ โดยแบ่งเป็น 2 ตอน ได้แก่

ตอนที่ 1 ความคิดเห็นของผู้ใช้งาน

ให้ทำเครื่องหมาย ☐ลงในช่องที่ตรงกับความคิดเห็นของท่านมากที่สุด มีเกณฑ์การประเมินระดับความคิดเห็น ดังนี้

ระดับ 5 หมายถึง ความเหมาะสมมากที่สุด
ระดับ 4 หมายถึง ความเหมาะสมมาก
ระดับ 3 หมายถึง ความเหมาะสมปานกลาง
ระดับ 2 หมายถึง ความเหมาะสมน้อย
ระดับ 1 หมายถึง ความเหมาะสมน้อยที่สุด

ตอนที่ 1 การประเมินความเหมาะสมในการใช้เว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ มี 2 ตอนดังนี้

1) ด้านการทำงานของเว็บแอพพลิเคชั่น(Functional) จำนวน 10 ข้อ
2) ความสะดวกในการใช้งานเว็บแอพพลิเคชั่น(Usability) จำนวน 7 ข้อ
3)ด้านความชัดเจนของคู่มือการใช้เว็บแอพพลิเคชั่น (Web application Manual) จำนวน 5 ข้อ

ตอนที่ 2 ข้อเสนอแนะอื่น ๆ

ในกรณีที่ท่านมีข้อเสนอแนะอื่น ๆ เพิ่มเติม เพื่อประโยชน์ในการปรับปรุงและพัฒนาแอพพลิเคชั่นต่อไปขอให้ท่านเขียนข้อเสนอแนะลงในช่องว่าง

ขอขอบพระคุณเป็นอย่างสูง
นายชัยมงคล ปีนะสา
ผู้วิจัย

รายการ	ระดับความพึงพอใจ				
	5	4	3	2	1

Ann Exam

2.2 ความถูกตองในการคนหาขอมูล	○	○	○	○	○
2.3 ความถูกต้องในการปรับปรุงแก้ไขข้อมูล	○	○	○	○	○
2.4 ความถูกต้องในการลบข้อมูล	○	○	○	○	○
2.5 ความถูกต้องของผลลัพธ์ที่ได้จากการประมวลผลในโปรแกรม	○	○	○	○	○
2.6 ความถูกต้องของการผลลัพธ์ในรูปแบบรายงาน	○	○	○	○	○
2.7 ความรวดเร็วในการประมวลผลของระบบ	○	○	○	○	○
2.8 ความน่าเชื่อถือได้ของระบบ	○	○	○	○	○
2.9 ความครอบคลุมของโปรแกรมที่พัฒนากับระบบงานจริง	○	○	○	○	○
2.10 การป้องกันข้อผิดพลาดที่อาจเกิดขึ้น	○	○	○	○	○
3.1 ความสะดวกในการเรียกใช้งานโปรแกรม	○	○	○	○	○
3.2 ความสะดวกในการเข้าและออกจากโปรแกรม	○	○	○	○	○
3.3 ความสะดวกในการลงทะเบียนเข้าใช้งานโปรแกรม	○	○	○	○	○
3.4 ความสะดวกในการป้อนข้อมูล	○	○	○	○	○
3.5 ความสะดวกในการกดปุ่มเลือกคำตอบ หรือเปลี่ยนคำตอบ	○	○	○	○	○
3.6 เมื่อทดสอบเสร็จโปรแกรมจะรายงานผลการทดสอบทันที	○	○	○	○	○
3.7 ผู้ใช้งานสามารถบันทึกผลการทดสอบ และสั่งพิมพ์ได้อย่างสะดวกรวดเร็ว	○	○	○	○	○
4.1 คู่มือการใช้งานโปรแกรมสามารถอธิบายวัตถุประสงค์ของโปรแกรมได้อย่างชัดเจน	○	○	○	○	○
4.2 คู่มือการใช่โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน	○	○	○	○	○
4.3 ภาษาที่ใช้ในคู่มือการใช้งานโปรแกรมอ่านเข้าใจง่าย และกระชับ	○	○	○	○	○
4.4 คู่มือการใช้งานโปรแกรมมีภาพประกอบการอธิบาย	○	○	○	○	○

Ann Exam

ให้เห็นชัดเจน					
4.5 หลังจากอ่านคู่มือแล้ว ผู้ใช้งานสามารถ ดูค่าชี้แจงจากคู่มือการใช้งานโปรแกรมและสามารถ ปฏิบัติตามได้	○	○	○	○	○

ตอนที่ 2 ข้อเสนอแนะเพื่อการพัฒนาโปรแกรม

ส่งแบบสอบถาม

Copyright © 2023 Designed by html.design. All rights reserved.

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๕๓

บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
อำเภอเมืองฯ จังหวัดพิษณุโลก ๖๕๐๐๐

๖ มกราคม ๒๕๖๖

เรื่อง ขอความร่วมมือเก็บข้อมูลเพื่อการวิจัย

เรียน

สิ่งที่ส่งมาด้วย แบบสอบถาม จำนวน.....ฉบับ

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ นิสิตมีความจำเป็นต้องขอเก็บข้อมูลจากหน่วยงานของท่าน บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร จึงใคร่ขอความอนุเคราะห์จากท่านโปรดอนุญาตให้นิสิตดำเนินการเก็บรวบรวมข้อมูลในการวิจัย ซึ่งจะเป็นประโยชน์ทางวิชาการต่อไป บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดีและขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

๑. งานวิชาการ บัณฑิตวิทยาลัย
โทร ๐-๕๕๙๖-๘๘๒๘
โทรสาร ๐-๕๕๙๖-๘๘๒๖
๒. นายชัยมงคล ปินะสา
โทร ๐๙-๕๖๖๓-๗๗๔๖

ภาคผนวก ซ

หนังสือขอความร่วมมือในการวิจัย

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๓๓

บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
อำเภอเมืองฯ จังหวัดพิษณุโลก ๖๕๐๐๐

๕ มกราคม ๒๕๖๖

เรื่อง ขอความอนุเคราะห์ตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย

เรียน ดร.อำนาจ เกษศรีโพร

สิ่งที่ส่งมาด้วย ๑. โครงร่างวิทยานิพนธ์ จำนวน ๑ ฉบับ
๒. เครื่องมือที่ใช้ในการวิจัย จำนวน ๑ ฉบับ

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร พิจารณาแล้วเห็นว่าท่านเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเนื้อหาสาระของวิทยานิพนธ์เรื่องนี้เป็นอย่างดียิ่ง จึงใคร่ขอเรียนเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัยดังแนบมาพร้อมนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

๑. งานวิชาการ บัณฑิตวิทยาลัย
โทร ๐-๕๕๙๖-๘๘๒๘
โทรสาร ๐-๕๕๙๖-๘๘๒๖
๒. นายชัยมงคล ปินะสา
โทร ๐๙-๕๖๖๓-๗๗๔๖

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๓๓ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
อำเภอเมืองฯ จังหวัดพิษณุโลก ๖๕๐๐๐

๕ มกราคม ๒๕๖๖

เรื่อง ขอความอนุเคราะห์ตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย

เรียน ผู้ช่วยศาสตราจารย์ ดร.มานิตย์ อาษานอก

สิ่งที่ส่งมาด้วย ๑. โครงร่างวิทยานิพนธ์ จำนวน ๑ ฉบับ
๒. เครื่องมือที่ใช้ในการวิจัย จำนวน ๑ ฉบับ

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร พิจารณาแล้วเห็นว่าท่านเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเนื้อหาสาระของวิทยานิพนธ์เรื่องนี้เป็นอย่างดียิ่ง จึงใคร่ขอเรียนเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัยดังแนบมาพร้อมนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

๑. งานวิชาการ บัณฑิตวิทยาลัย
โทร ๐-๕๕๙๖-๘๘๒๘
โทรสาร ๐-๕๕๙๖-๘๘๒๖
๒. นายชัยมงคล ปินะสา
โทร ๐๙-๕๖๖๓-๗๗๔๖

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๓๓

บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
อำเภอเมืองฯ จังหวัดพิษณุโลก ๖๕๐๐๐

๕ มกราคม ๒๕๖๖

เรื่อง ขอความอนุเคราะห์ตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย

เรียน รองศาสตราจารย์ ดร.ชัยวิชิต เชียรชนะ

สิ่งที่ส่งมาด้วย ๑. โครงร่างวิทยานิพนธ์ จำนวน ๑ ฉบับ
๒. เครื่องมือที่ใช้ในการวิจัย จำนวน ๑ ฉบับ

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร พิจารณาแล้วเห็นว่าท่านเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเนื้อหาสาระของวิทยานิพนธ์เรื่องนี้เป็นอย่างดียิ่ง จึงใคร่ขอเรียนเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัยดังแนบมาพร้อมนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

๑. งานวิชาการ บัณฑิตวิทยาลัย
โทร ๐-๕๕๙๖-๘๘๒๘
โทรสาร ๐-๕๕๙๖-๘๘๒๖
๒. นายชัยมงคล ปินะสา
โทร ๐๙-๕๖๖๓-๗๗๔๖

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๓๓

บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
อำเภอเมืองฯ จังหวัดพิษณุโลก ๖๕๐๐๐

๕ มกราคม ๒๕๖๖

เรื่อง ขอความอนุเคราะห์ตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย

เรียน ดร.บูรพา วิถีปัญญา

สิ่งที่ส่งมาด้วย ๑. โครงร่างวิทยานิพนธ์ จำนวน ๑ ฉบับ
๒. เครื่องมือที่ใช้ในการวิจัย จำนวน ๑ ฉบับ

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร พิจารณาแล้วเห็นว่าท่านเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเนื้อหาสาระของวิทยานิพนธ์เรื่องนี้เป็นอย่างดียิ่ง จึงใคร่ขอเรียนเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัยดังแนบมาพร้อมนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร

๑. งานวิชาการ บัณฑิตวิทยาลัย
โทร ๐-๕๕๙๖-๘๘๒๘
โทรสาร ๐-๕๕๙๖-๘๘๒๖
๒. นายชัยมงคล ปินะสา
โทร ๐๙-๕๖๖๓-๗๗๔๖

บันทึกข้อความ

ส่วนราชการ บัณฑิตวิทยาลัย งานวิชาการ โทร. ๘๘๒๗

ที่ อว ๐๖๐๓.๐๒/ว ๐๐๓๓ **วันที่** ๕ มกราคม ๒๕๖๖

เรื่อง ขอความอนุเคราะห์ตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย

เรียน รองศาสตราจารย์ ดร.ไกรศักดิ์ เกษร

ด้วย **นายชัยมงคล ปินะสา** รหัสประจำตัว ๖๓๐๓๑๒๕๔ นิสิตระดับปริญญาเอก สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ สังกัดบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร ได้รับอนุมัติให้ดำเนินการทำวิทยานิพนธ์ เรื่อง **"การพัฒนาเว็บแอพพลิเคชั่นสำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติโดยประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป"** เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต โดยมี รองศาสตราจารย์ ดร.สำราญ มีแจ้ง เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์

ในการทำวิทยานิพนธ์เรื่องนี้ บัณฑิตวิทยาลัย พิจารณาแล้วเห็นว่าท่านเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเนื้อหาสาระของวิทยานิพนธ์เรื่องนี้เป็นอย่างดียิ่ง จึงใคร่ขอเรียนเชิญท่านเป็นผู้ทรงคุณวุฒิตรวจแก้ไขเครื่องมือที่ใช้ในการวิจัย ดังที่แนบมาพร้อมนี้ บัณฑิตวิทยาลัย หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดีและขอขอบคุณอย่างสูงมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์

(รองศาสตราจารย์ ดร.อนามัย นาอุดม)
รองคณบดีฝ่ายวิชาการ ปฏิบัติราชการแทน
คณบดีบัณฑิตวิทยาลัย

ประวัติผู้วิจัย

ชื่อ-นามสกุล นายชัยมงคล ปินะสา

วัน เดือน ปี เกิด 24 มกราคม 2526

ที่อยู่ปัจจุบัน 195 หมู่ 2 บ้านสิมลี ต.เหนือ อ.เมืองกาฬสินธุ์ จ.กาฬสินธุ์

ที่ทำงานปัจจุบัน คณะครุศาสตร์ มหาวิทยาลัยราชภัฏเลย

ตำแหน่งหน้าที่ปัจจุบัน อาจารย์

ประสบการณ์การทำงาน 13 ปี