การทำดัชนีวิดีโอโดยใช้เวคเตอร์ของคำ

A Word Vector Technique for Indexing

นายณัฐพงษ์ แก่นจันทร์ รหัส 48362247

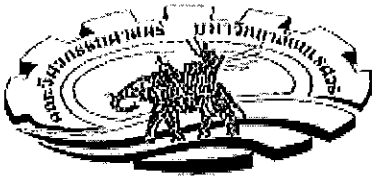
นายธรรมนูญ มีสนุ่น รหัส 48362285

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2551



ใบรับรองโครงงานวิศวกรรม

หัวข้อโครงงาน การทำดัชนีวิดีโอโดยใช้เวคเตอร์ของคำ

ผู้ดำเนินโครงงาน นายณัฐพงษ์ แก่นจันทร์ รหัส 48362247

นายธรรมนูญ มีสนุ่น รหัส 48362285

อาจารย์ที่ปรึกษา ดร.ไพศาล มุณีสว่าง

สาขาวิชา วิศวกรรมคอมพิวเตอร์

ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์

ปีการศึกษา 2551

.....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้โครงงานฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะกรรมการสอบโครงงานวิศวกรรม

.....ประธานกรรมการ

(ดร.ไพศาล มุณีสว่าง)

.....กรรมการ

(อาจารย์ศิริพร เดชะศิลารักษ์)

.....กรรมการ

(อาจารย์เศรษฐา ตั้งค้าวานิช)

หัวข้อโครงงาน	การทำดัชนีวิดีโอโดยใช้เวคเตอร์ของคำ
ผู้ดำเนินโครงงาน	นายณัฐพงษ์ แก่นจันทร์ รหัส 48362247
	นายธรรมนูญ มีสนุ่น รหัส 48362285
อาจารย์ที่ปรึกษา	ดร.ไพศาล มุณีสว่าง
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2551

บทคัดย่อ

โครงงานนี้เป็นโครงงานที่ทำเกี่ยวกับการทำดัชนีในการสืบค้นข้อมูลวิดีโอคลิปข่าวภาษาอังกฤษโดยใช้เวคเตอร์ของคำพูดที่ปรากฏในคลิปข่าว มาทำเป็นดัชนีในการสืบค้นโดยการนำไฟล์คลิปวีดีโอข่าวภาษาอังกฤษ มาทำการถอดคำพูดที่ปรากฏในคลิปให้ออกมาเป็นตัวหนังสือโดยใช้โปรแกรมแปลงเสียงพูดเป็นตัวอักษร จากนั้นจะนำข้อมูลคำที่ถอดได้มาสร้างเป็นดัชนีในรูปแบบฮีสโทแกรมของคำ การสืบค้นใช้หลักการเวคเตอร์สเปซโมเดล สามารถทำการสืบค้นวิดีโอคลิปผ่านทางเว็บแอปพลิเคชั่น

ผลที่ได้จากการทำโครงงานนี้จึงน่าจะช่วยเพิ่มรูปแบบหรือช่องทางในการสืบค้นข้อมูลวีดีโอคลิปให้สามารถเข้าถึงข้อมูลได้สะดวกและรวดเร็ว

Project Title	A Word Vector Technique for Indexing		
Name	Mr. Nathapong	Keanchan	ID. 48362247
	Mr. Thammanoon	Meesanun	ID. 48362285
Project Advisor	Paisarn Muneesawang, Ph.D.		
Major	Computer Engineering		
Department	Electrical and Computer Engineering		
Academic Year	2008		

ABSTRACT

This project is concerned with making the index for searching English news video clips. To make the index, we use vector of speech in news clips to translate from speech to text by using program to covert speech to text. Then the extracted words bring to make the index in histogram format. The searching is based on the Vector Space Model. User can search clips with database of video clips via web application.

The result shown that our project can search for video clips by using new method. Thus, the user can search video clips conveniently and rapidly.

กิตติกรรมประกาศ

โครงงานนี้สามารถสำเร็จลุล่วงไปได้ด้วยดีด้วยความช่วยเหลือจาก ดร. ไพศาล มุณีสว่าง อาจารย์ที่ปรึกษาโครงงาน ที่ได้ให้คำแนะนำและข้อคิดเห็นต่างๆของโครงงานมาโดยตลอด

อาจารย์ ศิริพร เดชะศิลารักษ์ และ อาจารย์ เศรษฐา ตั้งค้าวานิช ที่คอยให้คำแนะนำ และเป็นกรรมการในการสอบโครงงานในครั้งนี้

อาจารย์ ภานุพงศ์ สอนคม ที่คอยให้ความช่วยเหลือและให้คำแนะนำในการทำโครงงานนี้จนสำเร็จลุล่วงไปได้ด้วยดี

อาจารย์ทุกท่าน บิดา มารดา ญาติพี่น้อง และเพื่อนๆ ที่คอยให้การสนับสนุน ให้คำปรึกษาแนะนำ และติชม ผู้ดำเนินโครงงาน

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวรที่ได้อนุมัติเงินค่าใช้จ่ายในการทำโครงงานนี้

ผู้จัดทำโครงงานขอกราบขอบพระคุณเป็นอย่างสูง มา ณ โอกาสนี้

นายณัฐพงษ์ แก่นจันทร์

นายธรรมนูญ มีสนุ่น

สารบัญ

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญตาราง

สารบัญรูป

สารบัญรูป(ต่อ)

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของโครงงาน

ปัจจุบันนี้โลกได้เข้าสู่ยุคดิจิตอลหรือยุคของข้อมูลข่าวสาร การติดต่อสื่อสารจึงมีความสำคัญและจำเป็นอย่างมากในชีวิตประจำวัน อีกทั้งการเข้าถึงข้อมูลข่าวสารก็เป็นไปได้โดยสะดวกรวดเร็วและกว้างขวางจากสื่อประเภทต่างๆ เช่นโทรทัศน์ อินเตอร์เน็ต เป็นต้น โดยเฉพาะอินเตอร์เน็ตเป็นการเข้าถึงข้อมูลประเภทหนึ่งที่สามารถใช้ในการเข้าถึงข้อมูลได้รวดเร็วและกว้างขวาง การเข้าถึงข้อมูลผ่านทางอินเตอร์เน็ตได้รับความนิยมอย่างมากโดยเฉพาะการสืบค้นข้อมูลผ่านทางอินเตอร์เน็ตผ่านเว็บไซต์ที่ให้บริการการสืบค้นข้อมูล นอกจากนี้เว็บไซต์ต่างๆยังมีการสืบค้นข้อมูลจากฐานข้อมูลของเว็บไซต์เองอีกด้วย แต่จะเห็นว่าวิธีการที่ใช้ในการสืบค้นข้อมูลจากเว็บไซต์เหล่านี้ส่วนใหญ่จะใช้วิธีการเปรียบเทียบตัวอักษรที่ใช้สืบค้นกับตัวอักษรที่ใช้เป็นดัชนีของไฟล์ แต่ในปัจจุบันเริ่มมีการจัดเก็บข้อมูลในรูปแบบไฟล์เสียงหรือวีดีโอคลิปกันมากขึ้น บางครั้งการนำวิธีการดังกล่าวมาใช้ในการสืบค้นข้อมูลประเภทวีดีโอคลิปต่างๆนั้นอาจจะได้ข้อมูลที่ไม่ตรงตามความต้องการของผู้สืบค้น ดังนั้นจึงได้จัดทำโครงงานนี้ขึ้นเพื่อเป็นการรองรับการสืบค้นข้อมูลวิดีโอคลิป

โครงงานนี้เป็นโครงงานที่ทำเกี่ยวกับการทำดัชนีในการสืบค้นข้อมูลวิดีโอคลิปข่าวภาษาอังกฤษโดยใช้เวคเตอร์ของคำพูดที่ปรากฏในวิดีโอคลิปมาทำเป็นดัชนีในการสืบค้นโดยการนำไฟล์วีดีโอคลิปข่าวภาษาอังกฤษมาทำการถอดคำพูดที่ปรากฏในวิดีโอคลิปให้ออกมาเป็นตัวหนังสือด้วยโปรแกรมแปลงเสียงพูดเป็นตัวอักษร จากนั้นจะนำข้อมูลตัวหนังสือที่ถอดมาได้มาผ่านกระบวนการในการสร้างดัชนีคำเก็บไว้ในรูปแบบฮีสโทแกรมของคำลงในฐานข้อมูลของดัชนีคำเพื่อใช้ในการสืบค้นจากนั้นทำการสืบค้นวิดีโอคลิปกับฐานข้อมูลวิดีโอคลิปผ่านทางเว็บแอปพลิเคชั่น สามารถสืบค้นโดยวิธีการเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดในวิดีโอคลิปและสืบค้นโดยวิธีการใช้การเปรียบเทียบเนื้อหาคำพูดในวิดีโอคลิประหว่างวิดีโอคลิปด้วยกัน

โครงงานนี้จึงน่าจะช่วยเพิ่มรูปแบบหรือช่องทางในการสืบค้นข้อมูลวีดีโอคลิปให้สามารถเข้าถึงข้อมูลได้ถูกต้องและสะดวกรวดเร็ว

1.2 วัตถุประสงค์ของโครงงาน

1.2.1 เพื่อประยุกต์ใช้โปรแกรมแปลงเสียงพูดเป็นตัวอักษร (Speech to text)

1.2.2 พัฒนารูปแบบการสืบค้นข้อมูลประเภทวีดีโอคลิป

1.3 ขอบเขตของโครงงาน

1.3.1 สามารถค้นหาวิดีโอคลิปได้อย่างมีประสิทธิภาพ ถูกต้องแม่นยำ

1.3.2 สามารถสืบค้นข้อมูลวิดีโอคลิปข่าวได้โดยการใช้เนื้อหาคำพูดในข่าวเป็นดัชนีในการสืบค้น

1.3.3 สามารถสืบค้นข้อมูลได้เฉพาะประเภทวีดีโอคลิปข่าวที่เป็นภาษาอังกฤษเท่านั้น

1.4 ขั้นตอนการดำเนินงาน

ตารางที่ 1.1 ขั้นตอนของการดำเนินงาน

รายละเอียด	ปี 2551							ปี 2552		
	มิ.ย	ก.ค	ส.ค	ก.ย	ต.ค	พ.ย	ธ.ค	ม.ค	ก.พ	มี.ค
1. ศึกษาเนื้อหารายละเอียดต่างๆ ที่จำเป็นต้องใช้ในการสืบค้นข้อมูลประเภทไฟล์เสียง วิดีโอ										
2. ออกแบบโปรแกรมในการสืบค้นข้อมูลประเภทไฟล์เสียง วีดีโอ										
3. ทำการพัฒนาโปรแกรมในการสืบค้นข้อมูลประเภทไฟล์เสียง วีดีโอ										

ตารางที่ 1.2 ขั้นตอนของการดำเนินงาน (ต่อ)

รายละเอียด	ปี 2551							ปี 2552		
	มิ.ย	ก.ค	ส.ค	ก.ย	ต.ค	พ.ย	ธ.ค	ม.ค	ก.พ	มี.ค
4. ทดลองใช้งานโปรแกรมสืบค้นข้อมูลประเภทไฟล์เสียง วีดีโอ ประเมินผลและแก้ไขโปรแกรม								■	■	
5. สรุปผลการทำโครงงานและจัดทำรายงาน										■

1.5 ผลที่คาดว่าจะได้รับ

1.5.1 ความเข้าใจในการเข้าถึงข้อมูลประเภทวีดีโอคลิป

1.5.2 การประยุกต์ใช้งานโปรแกรมแปลงเสียงพูดเป็นตัวอักษร กับงานด้านวิศวกรรมคอมพิวเตอร์

1.5.3 การสร้างระบบการค้นหาข้อมูลประเภทวีดีโอคลิปที่มีคุณภาพสูง

1.6 งบประมาณของโครงงาน

1.6.1 ค่าหนังสือที่ใช้ประกอบการทำงาน	1,000	บาท
1.6.2 จัดทำรูปเล่ม	500	บาท
1.6.3 ค่าใช้จ่ายอื่นๆ ที่เกี่ยวข้อง	500	บาท
รวมเป็นเงิน (สองพันบาทถ้วน)	2,000	บาท

หมายเหตุ : ถัวเฉลี่ยทุกรายการ

บทที่ 2
ทฤษฎีและหลักการที่เกี่ยวข้อง

การทำดัชนีวิดีโอด้วยใช้เวคเตอร์ของคำ (A Word Vector Technique for Indexing) ได้จัดทำขึ้นเพื่อใช้ในการสืบค้นข้อมูลวิดีโอคลิปข่าวภาษาอังกฤษโดยใช้เวคเตอร์ของคำพูดที่ปรากฏในวิดีโอคลิปข่าวมาทำเป็นดัชนีในการสืบค้นโดยผ่านกระบวนการในการสร้างดัชนีคำเก็บไว้ในรูปแบบฮีสโทแกรมของคำลงในฐานข้อมูลของดัชนีคำเพื่อใช้ในการสืบค้นแล้ว จากนั้นทำการสืบค้นวิดีโอคลิปกับฐานข้อมูลวิดีโอคลิปผ่านทางเว็บแอปพลิเคชั่น

ในการทำดัชนีวิดีโอโดยใช้เวคเตอร์ของคำนั้นต้องมีความรู้และหลักการที่เกี่ยวข้องเพื่อนำมาใช้ประกอบในการทำโครงงาน ซึ่งความรู้หลักการและทฤษฎีทั้งหมดจะได้มีการอธิบายไว้ดังต่อไปนี้

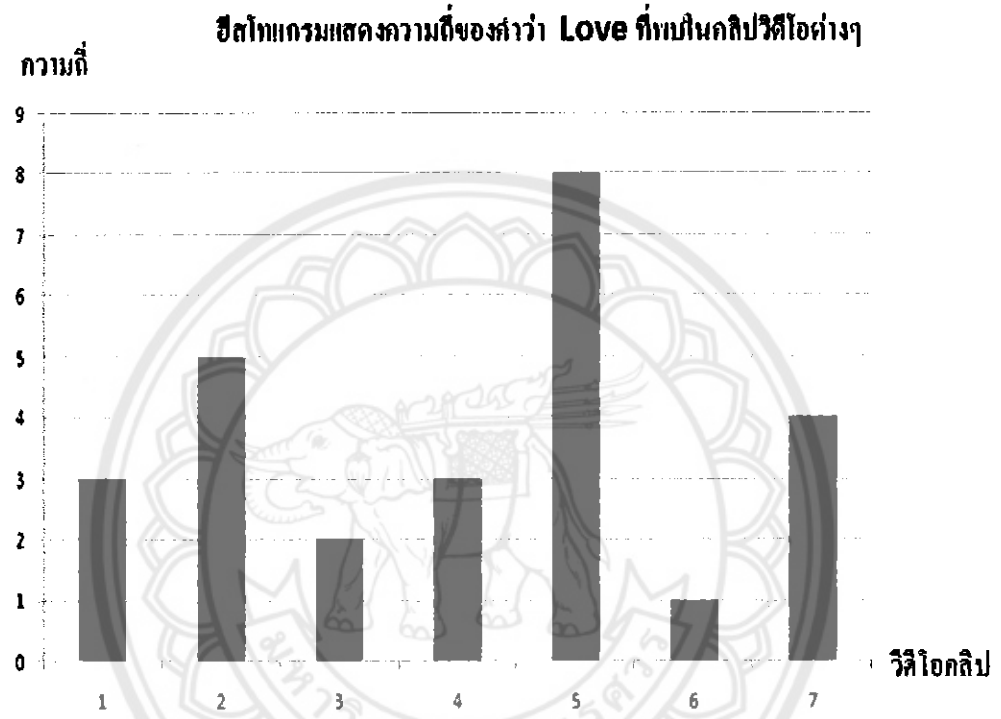
2.1 วิดีโอคลิป

วิดีโอคลิปหรือคลิปวิดีโอ (บางครั้งเรียกสั้นๆ ว่า คลิป) คือ ไฟล์คอมพิวเตอร์ที่บรรจุเนื้อหาเป็นภาพยนตร์สั้น ส่วนใหญ่จะตัดตอนมาจากภาพยนตร์ทั้งเรื่องซึ่งมีขนาดความยาวมาก โดยปกติวิดีโอคลิปมักจะเป็นส่วนที่สำคัญหรือต้องการนำมาแสดงความสนุกสนาน หรืออาจเป็นเรื่องความลับที่ต้องการนำมาเผยแพร่จากต้นฉบับเดิม แหล่งของวิดีโอคลิปได้แก่ รายการโทรทัศน์ วิดีโอข่าว ข่าวกีฬา ภาพยนตร์ เป็นต้น ปัจจุบันมีการใช้วิดีโอคลิปแพร่หลาย เนื่องจากไฟล์วิดีโอคลิปนี้มีขนาดเล็ก สามารถส่งผ่านอีเมล์หรือดาวน์โหลดจากเว็บไซต์ได้สะดวก ในประเทศตะวันตกเรียกการแพร่หลายของวิดีโอคลิปนี้ว่า วัฒนธรรมคลิป (Clip Culture)

ในปัจจุบันการใช้งานอินเทอร์เน็ตความเร็วสูง ยิ่งทำให้วิดีโอคลิปเป็นที่นิยมและแพร่หลายมากขึ้นไปอีก ประมาณว่าในแต่ละวันมีวิดีโอคลิปให้ประชาชนทั่วไปสามารถดูและโหลดออนไลน์นับล้านไฟล์ ซึ่งเว็บไซต์ยอดนิยมในการดูวิดีโอคลิปและดาวน์โหลดของคนส่วนใหญ่ได้แก่ www.ifilm.com, www.youtube.com, video.google.com

2.2 การเก็บดัชนีวิดีโอคลิปในรูปแบบฮีสโทแกรมของคำ

การเก็บข้อมูลของดัชนีวิดีโอคลิปในรูปแบบฮีสโทแกรมของคำนั้นสามารถทำได้โดยการนำคำพูดในแต่ละคลิปวิดีโอมาเก็บในรูปแบบของแผนภูมิของคำซึ่งแสดงความถี่ของคำที่มีอยู่ในวิดีโอคลิปแต่ละอัน ฮีสโทแกรมของคำต่างๆจะถูกเก็บไว้ในฐานข้อมูลดัชนีคำดังในรูปที่ 2.1



รูปที่ 2.1 แสดงตัวอย่างฮีสโทแกรมของคำ

2.3 การสืบค้นวิดีโอคลิป

การสืบค้นวิดีโอคลิปนั้นหากจำนวนไฟล์หรือจำนวนวิดีโอคลิปมีมาก ก็จะส่งผลให้การค้นหามีโอกาสได้วิดีโอตามความต้องการสูง หลักการทำงานคือเมื่อผู้ใช้ป้อนคำสำคัญหรือวลีสั้นๆที่ต้องการ ผลการสืบค้นที่ได้จะไม่ได้เป็นการจับคู่คำสำคัญกับชื่อรายการวิดีโอคลิป แต่ระบบการสืบค้นวีดีโอคลิปจะนำคำเหล่านี้ไปจับคู่กับคำพูดในคลิปวิดีโอจากดัชนีวีดีโอที่มีอยู่ทำให้การสืบค้นเข้าถึงเนื้อหาข้อมูลในคลิปวิดีโอได้ ผลการสืบค้นที่ได้จึงมีความครอบคลุมมากขึ้นและตอบสนองความต้องการของผู้บริโภคมากกว่าบริการสืบค้นวีดีโอแบบเดิมๆ

2.4 การแปลงเสียงพูดเป็นตัวอักษร (speech to text or speech recognition) [1]

การแปลงเสียงพูดเป็นตัวอักษรจะเกี่ยวข้องกับการจับคลื่นเสียงและการแปลงคลื่นเสียงเป็นข้อมูลดิจิตอล (digital) การแปลงให้อยู่ในหน่วยย่อยของ (phonemes) การสร้างคำจากหน่วยเสียงย่อย และการวิเคราะห์คำรอบข้างเพื่อให้แน่ใจว่าสะกดคำได้ถูกต้องสำหรับคำที่ออกเสียงใกล้เคียงกัน เช่น write กับ right

2.4.1 หน่วยเสียงย่อย (Phonemes)

หน่วยเสียงย่อยที่สุดที่นำมารวมกันในการสร้างคำแต่ละคำออกมา (Text to (TTS) engine) มีการใช้ความรู้และพื้นฐานหลักของภาษาและพิจารณาลักษณะการออกเสียงมาเป็นองค์ประกอบในการสร้างเสียง เช่น กอ คอ บอ เป็นต้น จะบันทึกเก็บไว้ในรูปแบบไฟล์คลื่น(wave file)

2.4.2 การทำงานของการแปลงเสียงพูดเป็นตัวอักษร

การทำงานคร่าวๆของการแปลงเสียงพูดเป็นตัวอักษร ซึ่งมีลำดับดังต่อไปนี้

1. ผู้ใช้พูดผ่าน ไมโครโฟน หรือ เปิดไฟล์เสียงที่ต้องการแปลงขึ้นมา
2. ทำการจับคลื่นเสียงและสร้างแรงกระตุ้นไฟฟ้า
3. การ์ดเสียง (Sound card) จะทำการแปลงสัญญาณเสียงเป็นสัญญาณดิจิตอล
4. ตัวแปลงเสียงพูดเป็นตัวอักษร (Speech recognition engine) จะทำการแปลงสัญญาณดิจิตอล เป็นพยางค์เสียง โดยการเปรียบเทียบสัญญาณที่ได้มากับสัญญาณที่ได้ทำการบันทึกไว้ แล้วจึงทำการรวมหน่วยเสียงที่แปลงได้นั้นเป็นคำ
5. โปรแกรมจะทำงานตามคำที่แปลงได้

2.4.3 ตัวแปลงเสียงพูดเป็นตัวอักษร (Recognizer or speech recognition engine)

เป็นซอฟท์แวร์ที่แปลงสัญญาณเสียงให้กลายเป็นสัญญาณดิจิตอลแล้วส่งคำที่แปลงได้ออกมาเป็นข้อความให้โปรแกรมทำงาน โดยทั่วไปแล้วตัวแปลงเสียงพูดเป็นตัวอักษรส่วนมากจะสนับสนุนคำพูดที่มีความต่อเนื่อง ดังนั้นจึงสามารถแปลงคำพูดที่พูดอย่างเป็นธรรมชาติได้ด้วยความเร็วที่ใช้ในการสนทนากันตามปกติ

2.5 เวคเตอร์สเปซโมเดล (Vector Space Model) [2]

อาศัยเวคเตอร์ (Vector) ในการทำงานโดยจะแทนเอกสาร (Document) และเทอม (Term) ที่ผู้ใช้คิวรี่ (query) เข้ามาด้วยเวคเตอร์ จากนั้นทำการเปรียบเทียบหาเอกสารที่มีเวคเตอร์คล้ายกับเวคเตอร์ของผู้ใช้มากที่สุดเป็นผลลัพธ์ สำหรับการแทนเป็นเวคเตอร์สามารถทำได้โดยการแทนขนาดแต่ละมิติ (Dimension) ของเวคเตอร์ด้วย TF-IDF weighting และทำการเปรียบเทียบความคล้ายกันของเวคเตอร์ (Similarity Measurement) ด้วยการทำอินเนอร์โปรดักส์ (Inner product) หรือโคไซน์โปรดักส์ (Cosine product)

2.5.1 การคิดค่าน้ำหนักทีเอฟ-ไอดีเอฟ (TF-IDF weighting)

การคิดค่าน้ำหนักทีเอฟ-ไอดีเอฟ คือ การใช้ค่าทีเอฟและไอดีเอฟในการคิดค่าน้ำหนักแทนขนาดในแต่ละมิติของเวคเตอร์

1. ค่าทีเอฟหรือความถี่ของเทอม (tf or Term frequency) คือ อัตราส่วนของความถี่ของเทอมในเอกสารนั้นกับความถี่ของเทอมในเอกสารที่มีความถี่สูงสุด

ให้ f_{ij} คือ ความถี่ของเทอม i ในเอกสาร j

$$f_{ij} = \text{ความถี่ของเทอม i ในเอกสาร j} \tag{2.1}$$

ให้ tf คือ อัตราส่วนของความถี่ของเทอมในเอกสารนั้นกับความถี่ของเทอมในเอกสารที่มีความถี่สูงสุด

$$tf_{ij} = f_{ij} / \max\{f_{ij}\} \tag{2.2}$$

2. ดีเอฟ (Document frequency) คือ จำนวนเอกสารที่มีเทอมนั้นอยู่ ถ้าค่าดีเอฟน้อยแสดงถึงเทอมนั้นปรากฏอยู่ในเอกสารกลุ่มเล็กๆ ไม่ได้เป็นคำทั่วๆไป

3. ไอดีเอฟ (Inverse document frequency) หาได้จากสมการ

$$idf_i = \log_2(n/df_i) \tag{2.3}$$

โดย n คือ จำนวนเอกสารทั้งหมด

df_i คือ ค่าดีเอฟของเทอม i

หรือ $$idf_i = 1 + \log(|d| + df_i) \quad (2.4)$$

โดย |d| คือ จำนวนเอกสารทั้งหมด

กำหนดให้ ค่าน้ำหนักทีเอฟ-ไอดีเอฟของเทอม i ใน เอกสาร j แทนด้วย W_{ij}

เอกสาร j ใดๆ แทนด้วย $d_j = (W_{1j}, W_{2j}, \ldots, W_{ij})$

จะได้ว่า

$$W_{ij} = tf_{ij} \times idf_i \quad (2.5)$$

หรือ $$W_{ij} = idf_{ij} \times \log_2(n/df_i) \quad (2.6)$$

$$\begin{bmatrix} & D_1 & D_2 \ldots\ldots & D_n \\ T_1 & W_{11} & W_{21} \ldots\ldots & W_{1n} \\ T_2 & W_{21} & W_{22} \ldots\ldots & W_{2n} \end{bmatrix}$$

เอกสาร n ชุด สามารถนำเสนอเป็นเมทริกซ์ของเอกสารกับเทอม (term document matrix) ซึ่งภายในเมทริกซ์ก็จะบรรจุค่าน้ำหนัก และเทอมที่เป็น 0 จะหมายถึงเทอมที่มีข้อมูลไม่เกี่ยวกับในเอกสารนั้นๆ

กำหนดให้ความถี่ของเทอมต่างๆในเอกสารดังนี้เทอม A=3, เทอม B =2 และเทอม C=1 และมีจำนวนเอกสารทั้งหมด 10,000 เอกสาร และเทอมต่างๆปรากฏในเอกสารทั้งหมดจำนวนทั้งสิ้นดังนี้ เทอม A=50, เทอม B =1300 และเทอม C=250

แล้วสามารถคำนวณค่า ทีเอฟ-ไอดีเอฟ ได้ดังนี้

A: tf = 3/3; idf = $\log_2$ (10,000/50) = 5.3; tf-idf = 5.3

B: tf = 2/3; idf = $\log_2$ (10,000/1300) = 2.0; tf-idf = 1.3

C: tf = 1/3; idf = $\log_2$ (10,000/250) = 3.7; tf-idf = 1.2

2.5.2 อินเนอร์โปรดักส์ (Inner product)

การเปรียบเทียบเวคเตอร์ด้วยการนำเวคเตอร์มาดอท (dot) กัน ถ้ามีค่ามากแสดงว่าเวคเตอร์มีความคล้ายกันมาก สามารถคำนวณได้จาก

$$sim(d_j,q) = d_j \cdot q = \sum_{i=1}^{t} w_{ij} \cdot w_{iq} \qquad (2.7)$$

โดย d_j คือ เอกสาร j ใดๆ

q คือ คิวรี่

w_{ij} คือ ค่าน้ำหนักของเทอม i ในเอกสาร j

w_{iq} คือ ค่าน้ำหนักของเทอม i ในคิวรี่

มี 2 วิธีดังนี้

1. ไบนารีเวคเตอร์ (Binary vectors)

ตัวอย่าง อินเนอร์โปรดักส์แบบไบนารีเวคเตอร์ระหว่างเอกสาร D กับคิวรี่ Q

ตารางที่ 2.1 ตัวอย่างอินเนอร์โปรดักส์แบบไบนารีเวคเตอร์

	retrieval	Database	architecture	computer	text	management	information
D	1	1	1	0	1	1	0
Q	1	0	1	0	0	1	1

D = 1, 1, 1, 0, 1, 1, 0

Q = 1, 0, 1, 0, 0, 1, 1

sim(D, Q) = 3

สำหรับเอกสารที่แทนด้วยแบบไบนารีเวคเตอร์ ถ้ามีเทอมเหล่านั้นปรากฏอยู่ในเอกสารแทนด้วย 1, ไม่มีแทนด้วย 0 การดอทกันจะได้จำนวนของคำในคิวรี่ที่ match กับเอกสาร

2. เวทเทอมเวคเตอร์ (Weighted term vectors)

ตัวอย่าง อินเนอร์โปรดักส์แบบเวทเทอมเวคเตอร์

กำหนดให้ เอกสาร D1 มีความถี่ของเทอมต่างๆดังนี้ T_1=2, T_2=3, T_3=5

เอกสาร D2 มีความถี่ของเทอมต่างๆดังนี้ T_1=3, T_2=7, T_3=1

คิวรี่ Q มีความถี่ของเทอมต่างๆดังนี้ T_1=0, T_2=0, T_3=2

sim(D1 , Q) = (2×0) + (3×0) + (5×2) = 10

sim(D2 , Q) = (3×0) + (7×0) + (1×2) = 2

สำหรับเอกสารที่แทนด้วยแบบเวทเทอมเวคเตอร์ การดอทกันจะได้ผลรวมของค่าน้ำหนักที่คิวรี่เข้าคู่กันกับเอกสาร

2.5.3 โคไซน์โปรดักส์ (Cosine product)

การเปรียบเทียบเวคเตอร์ด้วยการวัดมุมระหว่างเวคเตอร์ ถ้ามีมุมต่างกันน้อยแสดงว่าเวคเตอร์มีความคล้ายกันมาก สามารถคำนวณได้จากสมการ

$$\text{CosSim}(d_j,q) = \frac{\vec{d}_j \cdot \vec{q}_j}{|\vec{d}_j| \cdot |\vec{q}_j|} = \frac{\sum_{i=1}^{t} (W_{ij} \cdot W_{iq})}{\sqrt{\sum_{i=1}^{t} (W_{ij})^2 \cdot \sum_{i=1}^{t} (W_{iq})^2}} \quad (2.8)$$

โดย d_j คือ เอกสาร j ใดๆ

q คือ คิวรี่

w_{ij} คือ ค่าน้ำหนักของเทอม i ในเอกสาร j

w_{iq} คือ ค่าน้ำหนักของเทอม i ในคิวรี่

ตัวอย่าง โคไซน์โปรดักส์

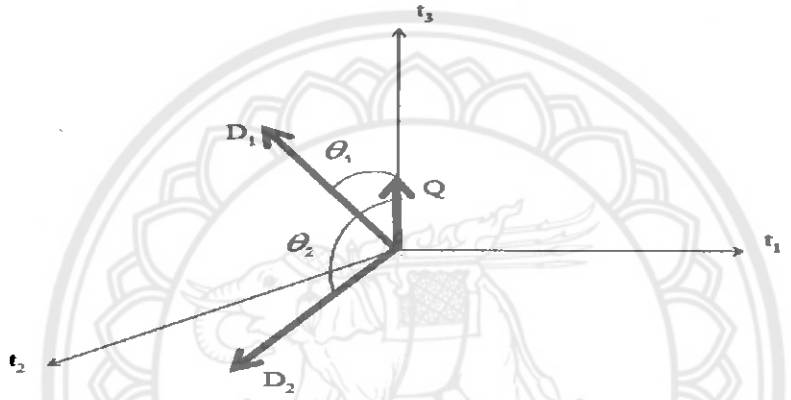
กำหนดให้ เอกสาร D1 มีความถี่ของเทอมต่างๆดังนี้ T_1=2, T_2=3, T_3=5

เอกสาร D2 มีความถี่ของเทอมต่างๆดังนี้ T_1=3, T_2=7, T_3=1

คิวรี่ Q มีความถี่ของเทอมต่างๆดังนี้ T_1=0, T_2=0, T_3=2

$$CosSim(D1 , Q) = (2\times0) + (3\times0) + (5\times2) / ((4+9+25)(0+0+4))^{1/2} = 0.81$$

$$CosSim(D2 , Q) = (3\times0) + (7\times0) + (1\times2) / ((9+49+1)(0+0+4))^{1/2} = 0.13$$



รูปที่ 2.2 แสดงการคำนวณโดยเวคเตอร์โคไซน์โปรดักส์ [3]

โดยปกติค่า CosSim ของคิวรี่จะเท่ากับ 1 เสมอ เอกสารที่มีค่า CosSim เข้าใกล้ 1 มากก็จะมีความคล้ายกับคิวรี่มาก ในรูปที่ 2.2 จะเห็นว่ามุม θ_1 ระหว่างเวคเตอร์ของเอกสาร D1 กับเวคเตอร์ของคิวรี่ Q มีขนาดเล็กกว่ามุม θ_2 ซึ่งเป็นมุมระหว่างเวคเตอร์ของเอกสาร D2 กับเวคเตอร์ของคิวรี่Q

2.5.4 ปัญหาของ เวคเตอร์สปาร์สซิเนส (Vector Sparseness)

ภาษามีคำศัพท์จำนวนมาก ทำให้เวคเตอร์แทนเอกสารอาจมีมิติมากเกินไป แต่ว่าเอกสารส่วนใหญ่ใช้คำไม่มาก ดังนั้นเวคเตอร์ส่วนใหญ่จะสปาร์ส (sparse) ซึ่งมีวิธีการจัดการสามารถสรุปได้ดังนี้

ถ้าเก็บเอกสารขนาด n คำใช้ Hash Table ใช้ คำเป็น key ของ hash table และค่าน้ำหนักของคำเป็นค่าของคีย์ (key) ใช้พื้นที่เก็บ ~1.5n ในการแก้ไขหรือดึงค่ามาดู O(1) เวลาที่ใช้ในการสร้าง Vector O(n)

2.5.5 ดัชนีอินเวอร์ส (Inverted Index)

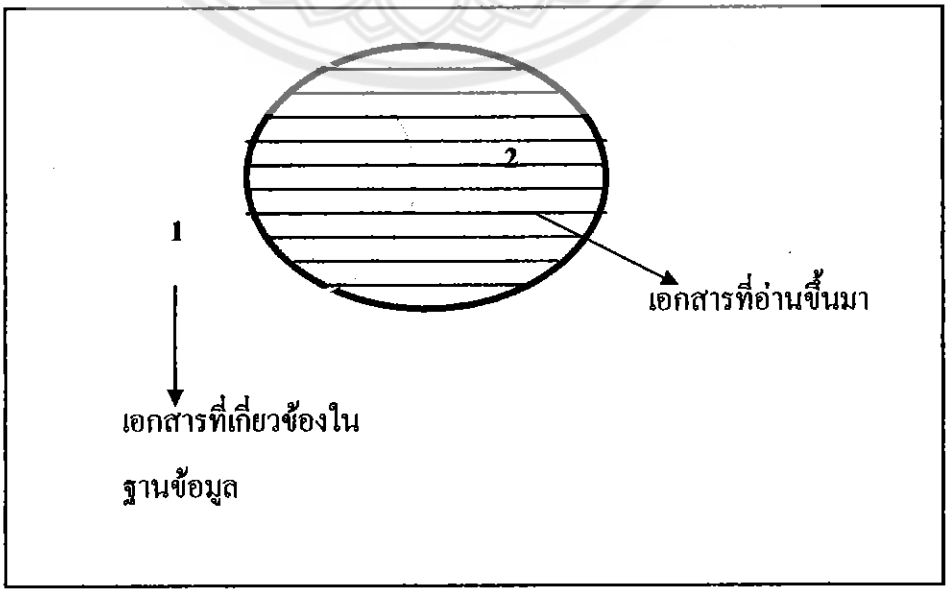
การเก็บโดยใช้เวคเตอร์เอกสาร (Document vectors) เปรียบเสมือนการเก็บว่า "เอกสารนี้มีเทอมอะไรอยู่บ้าง" เพื่อจะนำไปคำนวณตามวิธีขั้นต้น แต่โดยปกติจะไม่เก็บเวคเตอร์ของเอกสารโดยตรงแต่จะเก็บเป็นอินเวอร์สไฟล์ (Inverted file) แทนคือ " เทอมนี้มีอยู่ในเอกสารใดบ้าง" อยู่ในรูปของ Hash table, Sorted array, B-Tree หรืออื่นๆ ซึ่งเวลาในการทำงานดีกว่าวิธีที่กล่าวมาในข้างต้นมากและการสร้างมีความง่ายกว่าด้วย เพราะหากเก็บเป็นเวคเตอร์เอกสารโดยใช้ค่าน้ำหนักแบบทีเอฟ-ไอดีเอฟ ต้องทราบค่าทีเอฟและค่าดีเอฟของแต่ละเทอมก่อน แต่ค่าดีเอฟคือจำนวนเอกสารทั้งหมดที่เทอมนั้นปรากฏอยู่ ดังนั้นต้องรอให้เก็บเอกสารทั้งหมดก่อนจึงจะเริ่มสร้างดัชนีได้

2.6 การวัดประสิทธิภาพของการสืบค้น (Performance Evaluation)

2.6.1 ความแม่นยำและเรียกมาใช้ (Precision and Recall)

Precision และ Recall คือ วิธีการหนึ่งสำหรับการประเมินประสิทธิภาพการสืบค้นเอกสารซึ่งแสดงให้เห็นตามรูปที่ 2.3 ซึ่งสามารถอธิบายได้ดังนี้

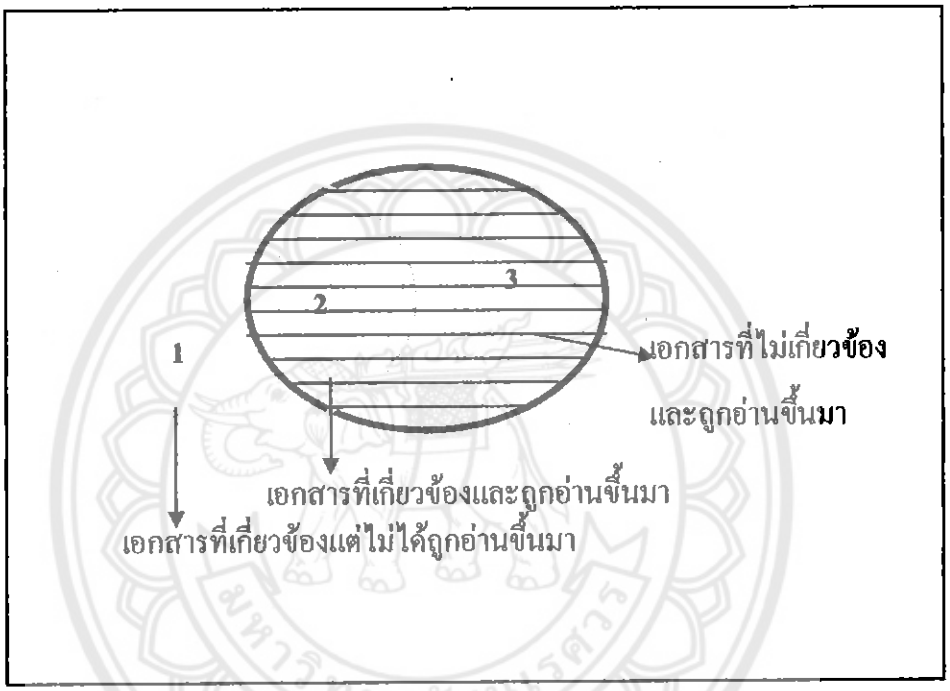
1. กลุ่มของเอกสารที่เก็บไว้ในฐานความรู้ซึ่งมีความสัมพันธ์กันกับข้อมูลของคำสืบค้น (1)
2. กลุ่มของเอกสารที่ถูกเลือกขึ้นมาเป็นผลลัพธ์ (2)



รูปที่ 2.3 แสดงความสัมพันธ์ระหว่างกลุ่มข้อมูลในฐานข้อมูลที่เกี่ยวข้องกับคำสืบค้น

จากรูปที่ 2.4 สามารถอธิบายได้ดังนี้

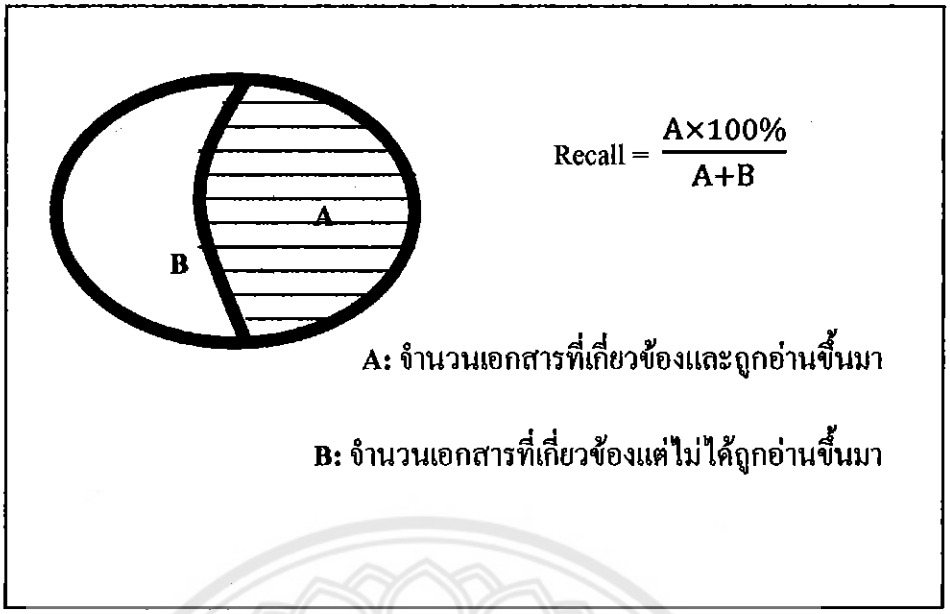
1. กลุ่มของเอกสารที่เกี่ยวข้องกับข้อมูลสืบค้นแต่ไม่ได้เลือกขึ้นมา (1)
2. กลุ่มของเอกสารที่เกี่ยวข้องกับข้อมูลสืบค้นและถูกเลือกขึ้นมา (2)
3. กลุ่มของเอกสารที่ไม่เกี่ยวข้องกับข้อมูลสืบค้นแต่ถูกเลือกขึ้นมา (3)



รูปที่ 2.4 แสดงความสัมพันธ์ระหว่างกลุ่มของข้อมูลที่ใช้สืบค้นกับคำสืบค้น

รูปที่ 2.5 คือการหาค่าของ Recall เมื่อ B คือจำนวนเอกสารที่เกี่ยวข้องหรือสัมพันธ์กับคำสืบค้นแต่ไม่ได้ถูกเลือกมาเป็นผลลัพธ์ และ A คือจำนวนของเอกสารที่เกี่ยวข้องหรือสัมพันธ์กับคำสืบค้นและถูกเลือกมาเป็นผลลัพธ์ ซึ่งสามารถคำนวณได้จากสมการ

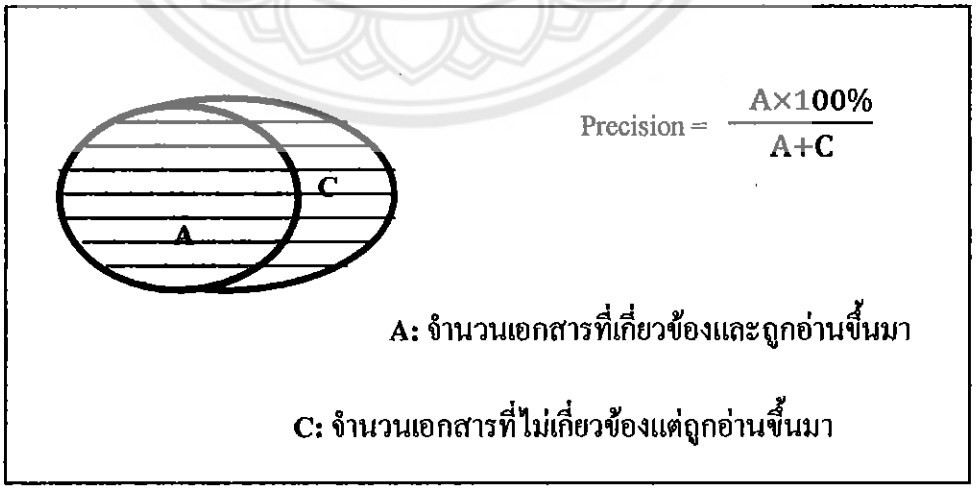
$$\text{Recall} = \frac{A \times 100\%}{A+B} \qquad (2.9)$$



รูปที่ 2.5 แสดงสมการการคำนวณในการหาค่า Recall

รูปที่ 2.6 คือการหาค่าของ Precision เมื่อ C คือจำนวนของเอกสารที่ไม่เกี่ยวข้องหรือไม่สัมพันธ์กับคำสืบค้นแต่ถูกเลือกมาเป็นผลลัพธ์และ A คือจำนวนของเอกสารที่เกี่ยวข้องหรือสัมพันธ์กับคำสืบค้นและถูกเลือกมาเป็นผลลัพธ์ ซึ่งสามารถคำนวณได้จากสมการ

$$\text{Precision} = \frac{A \times 100\%}{A+C} \qquad (2.10)$$



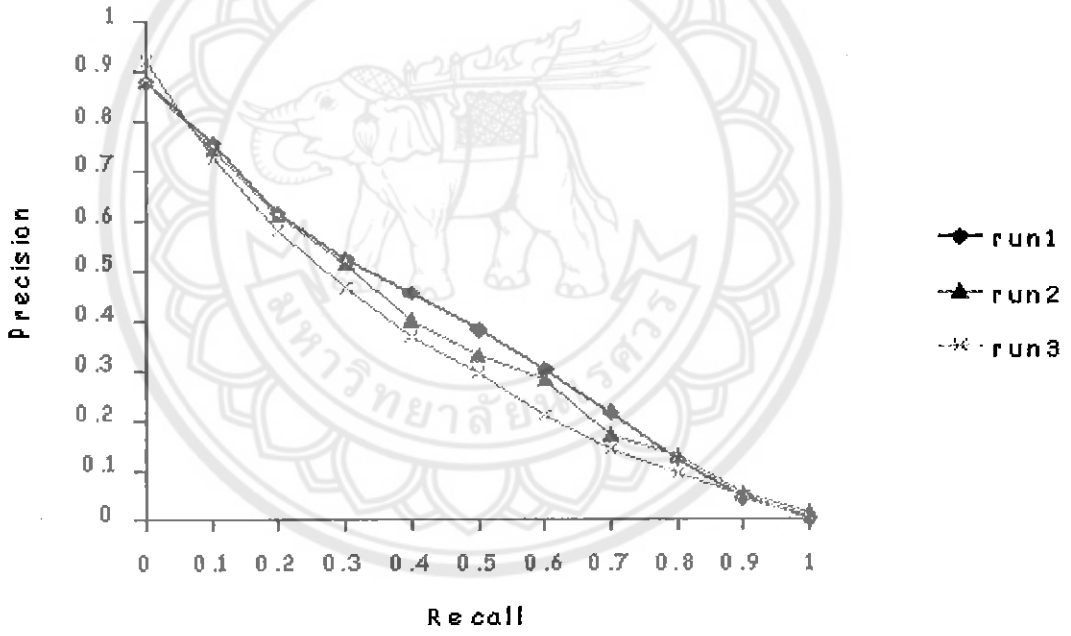
รูปที่ 2.6 แสดงสมการการคำนวณในการหาค่า Precision

แนวความคิดของผลการค้นหาจะอยู่ตรงกลางระหว่าง Precision และ Recall ถ้าเน้น Recall มากไปเอกสารที่ไม่เกี่ยวข้องก็อาจจะหลุดออกมามากส่งผลให้ Precision ต่ำหรือถ้าหากเน้น Precision มากไป เอกสารที่เกี่ยวข้องบางฉบับจะถูกคัดออกไปส่งผลให้ Recall ต่ำ

2.6.2 การใช้กราฟ R-P Curve

การใช้ R-P Curve ในการวิเคราะห์ประสิทธิภาพนั้น กำหนดให้ Recall เป็นแกนนอน และ Precision เป็นแกนตั้งเส้นกราฟที่ใกล้มุมขวาบนแสดงถึงประสิทธิภาพที่ดีกว่า

รูปที่ 2.7 ตัวอย่างกราฟ R-P Curve แสดงการเปรียบเทียบประสิทธิภาพของโปรแกรมสืบค้นจำนวน 3 โปรแกรม จากกราฟจะเห็นว่าประสิทธิภาพของโปรแกรมเรียงตามลำดับจากมากไปน้อย คือ โปรแกรม run1, โปรแกรม run2 และโปรแกรม run3



รูปที่ 2.7 แสดงกราฟตัวอย่าง R-P Curve[4]

บทที่ 3

การออกแบบและพัฒนาระบบ

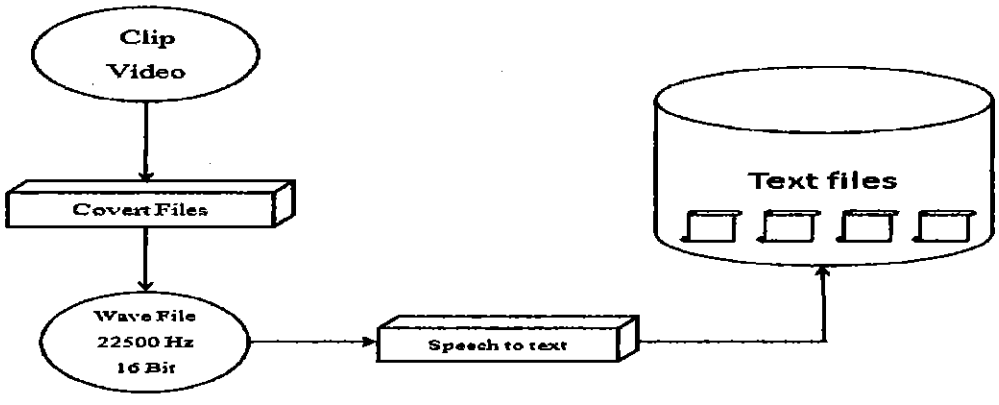
ในการออกแบบและพัฒนานั้นจะใช้ทฤษฎีที่ศึกษาในบทที่ 2 มาใช้ในการออกแบบระบบและพัฒนาระบบ

3.1 ขั้นตอนการออกแบบและพัฒนาระบบ

1. ศึกษาและรวบรวมข้อมูลเกี่ยวกับการแปลงเสียงพูดเป็นตัวอักษร, การใช้งานโปรแกรม speech recognition, ศึกษาการทำดัชนีโดยการใช้เวคเตอร์สเปซโมเดล, ศึกษาการใช้ภาษา Java, และการสร้างการสืบค้นข้อมูลผ่านเว็บแอปพลิเคชั่น
2. เก็บรวบรวมวิดีโอคลิป
3. ออกแบบกระบวนการทำงานของโครงงาน
4. พัฒนาระบบตามกระบวนการที่ได้ออกแบบไว้
5. ทดสอบการทำงานของระบบ
6. วิเคราะห์ประสิทธิภาพการทำงานของระบบโดยอ้างอิงจากการวัดประสิทธิภาพของการสืบค้น (Performance Evaluation)

3.2 กระบวนการถอดคำพูดในวิดีโอคลิปเป็นตัวอักษร

แบ่งเป็นออกเป็น 2 ขั้นตอนดังรูปที่ 3.1



รูปที่ 3.1 แสดงกระบวนการถอดคำพูดในวิดีโอคลิปเป็นตัวอักษร

3.2.1 การแปลงไฟล์วิดีโอคลิปเป็นไฟล์เสียง

นำคลิปวีดีโอที่เก็บรวบรวมจากเว็บไซต์ต่างๆมาทำการแปลงเป็นไฟล์เสียง โดยใช้โปรแกรม ImTooMPEG Encoder จากนั้นนำไฟล์เสียงที่ได้มาปรับความถี่ของคลื่นเสียงเป็น 22500 เฮิร์ต และอัตราของบิตเรตเป็น 16 บิต โดยใช้โปรแกรม WavePad เพื่อช่วยเพิ่มประสิทธิภาพในการแปลงไฟล์เสียงให้เป็นไฟล์ตัวอักษรในขั้นตอนต่อไป

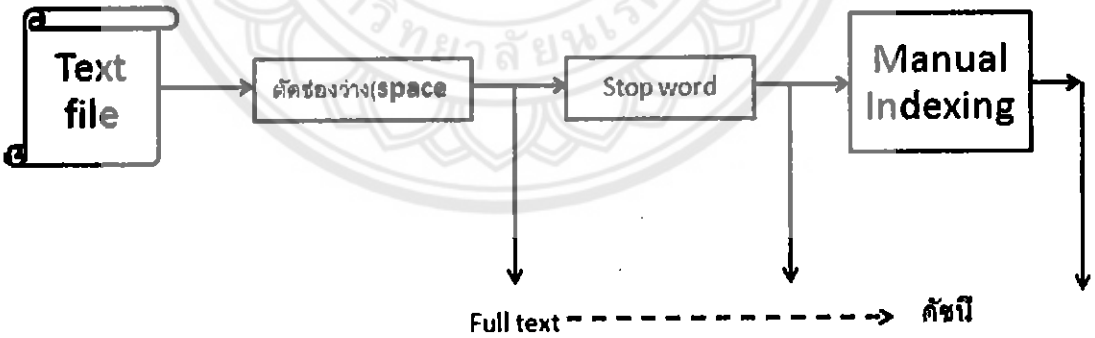
3.2.2 การแปลงไฟล์เสียงเป็นไฟล์ตัวอักษร

นำไฟล์เสียงที่ได้จากหัวข้อที่ 3.2.1 มาแปลงเป็นไฟล์ตัวอักษร (Text file) โดยใช้โปรแกรมแปลงเสียงพูดให้เป็นตัวอักษรที่ชื่อว่า Voice Recognize แล้วเก็บไฟล์ตัวอักษรที่ได้ลงในฐานข้อมูลของไฟล์ตัวอักษร

3.3 การทำดัชนี (Indexing)

เมื่อมีการเปลี่ยนแปลงเอกสารหรือการเพิ่มเอกสารที่สัมพันธ์กับวิดีโอคลิปใดๆแล้ว ระบบจะต้องมีการสร้างดัชนีใหม่เพื่อให้สอดคล้องกับเอกสารใหม่

3.3.1 การประมวลผลข้อความ (Text Processing)



รูปที่ 3.2 แสดงขั้นตอนการประมวลผลข้อความสำหรับนำไปสร้างดัชนีคำ

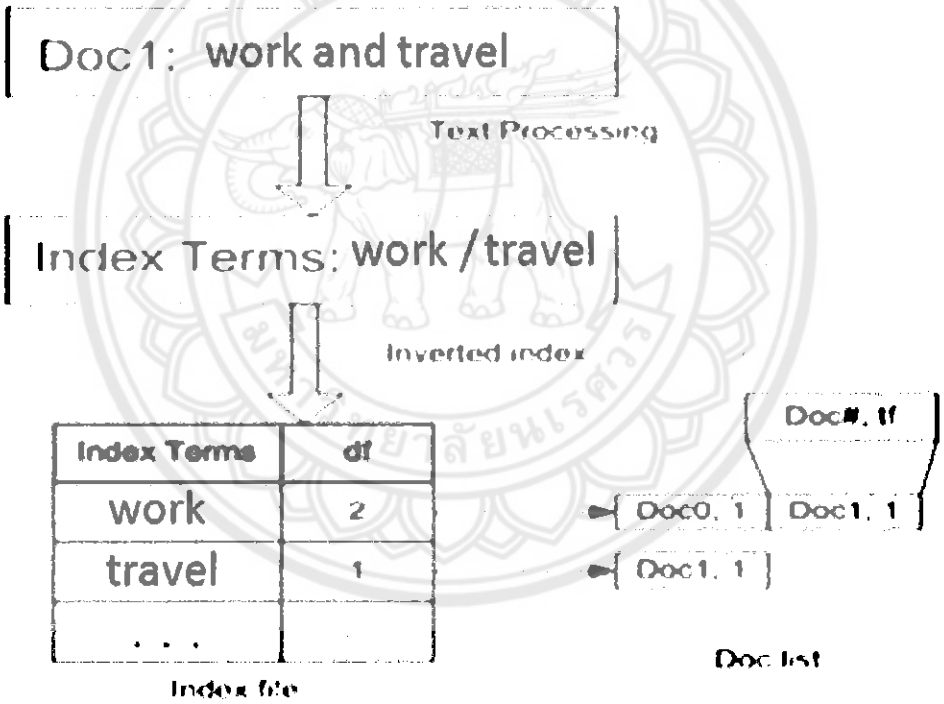
การสร้างดัชนีให้กับวิดีโอคลิปนั้นเอกสารที่สัมพันธ์กับวิดีโอคลิปจำเป็นที่จะต้องผ่านการประมวลผลข้อความก่อน เนื่องจากการประมวลผลข้อความบนพื้นฐานของระบบการค้นคืนสารสนเทศ (Information Retrieval) สิ่งที่เป็นพื้นฐานที่จำเป็นอย่างยิ่งคือ "หน่วยคำ" ดังนั้นการหาขอบเขตของแต่ละคำจึงเป็นสิ่งแรกที่ต้องคำนึงถึง เพราะหากเลือกการหาขอบเขตคำไม่เหมาะสม

อาจนำมาสู่ระบบการประมวลผลข้อความที่ไม่ถูกต้อง ซึ่งในส่วนประมวลผลความนี้สำหรับนำไปใช้สร้างดัชนีคำ (Index Terms) ซึ่งในโครงงานนี้จะใช้ Full Text Search ในการสร้างดัชนี

จากรูปที่ 3.2 แบ่งขั้นตอนการประมวลผลข้อความได้ดังนี้

1. ตัดช่องว่าง (Space)
2. ทำการตัดคำหยุด (Stopword) ซึ่งเป็นคำที่ไม่จำเป็นหรือคำที่ใช้บ่อย ๆ เช่น "and" "the" "a" "on" "from" เป็นต้น
3. นำคำที่ได้ทั้งหมด มาเก็บไว้ในฐานข้อมูล เพื่อนำไปทำดัชนีคำ (Index)

3.3.2 การสร้างดัชนีแบบอินเวิร์สไฟล์ดัชนี



รูปที่ 3.3 แสดงการสร้างดัชนีแบบ Inverted Index[5]

โครงงานนี้ได้ใช้การสร้างดัชนีแบบอินเวอร์สไฟล์ดัชนี(Inverted File Index) โดยเมื่อเอกสารได้ผ่านการประมวลผลข้อความแล้ว หลังจากนั้นจึงนำเอกสารเหล่านั้นเข้าสู่กระบวนการสร้างดัชนีโดยดัชนีที่ได้จะถูกจัดเก็บเป็นไฟล์ดัชนีซึ่งภายในจะประกอบด้วยคำต่าง ๆ พร้อมทั้ง

จำนวนเอกสารที่คำเหล่านั้นปรากฏอยู่ และแต่ละคำก็จะมีข้อมูลซึ่งระบุรายการหมายเลขของเอกสารพร้อมทั้งจำนวนคำที่ปรากฏอยู่ในเอกสารนั้น ดังรูปที่ 3.3

วิธีการในการจัดทำดัชนีของคำหลักที่พบภายในเอกสาร โดยการกำหนดความสำคัญของคำเมื่อเทียบกับเอกสารทั้งระบบโดยใช้ค่า TF/IDF (Term Frequency/Inverse Document Frequency) มีสูตร ดังนี้

$$W_{ij} = tf_{ij} \times \log_2 N / n_j \qquad (3.1)$$

โดยที่ W_{ij} คือค่าน้ำหนักของคำที่ t_j ในเอกสาร D_i

tf_{ij} คือความถี่ของคำที่ t_j ในเอกสาร D_i

N คือจำนวนเอกสารทั้งหมด

n_j คือจำนวนเอกสารที่มีคำที่ t_j ปรากฏอยู่อย่างน้อย 1 ครั้ง

พารามิเตอร์ n_j ในสมการนี้จะหาค่าได้เมื่อระบบได้รู้ทุกเอกสารที่มีในระบบ ในกรณีนี้จะต้องมีฐานข้อมูลของคำไว้ให้ระบบ โดยคำในฐานข้อมูลนี้ได้มีการประมวลผลข้อความแล้ว เพื่อกรองคำที่ไม่สื่อถึงความสำคัญของเอกสาร ดังนั้นคำที่มีค่าความสำคัญในขอบเขตที่กำหนดจะถูกเลือกเป็นคำสำคัญ (Significant word) และจะถูกจัดเก็บลงในฐานข้อมูล

3.4 การค้นหา (Searching)

เป็นส่วนที่นำอินพุตที่ต้องการค้นหาจากผู้ใช้ผ่านทาง User Interface มาค้นหาในฐานข้อมูล และส่งผลการค้นหากลับไปแสดงบนหน้าจอ ซึ่งโมเดล (Model) ที่นำมาใช้คือ เวคเตอร์สเปซโมเดล โดยการค้นคืนแบบเวคเตอร์สเปซโมเดลนั้นจะทำการค้นคืนได้วิดีโอคลิปที่มีแม้แต่ความเหมือนแค่เพียงบางส่วน โดยที่ในวิดีโอคลิปไม่ต้องมีคำพูดที่ตรงกันทั้งหมด ซึ่งค่าความเหมือนระหว่างเอกสารที่สัมพันธ์กับวิดีโอคลิปกับอินพุต โดยคุณสมบัติของเวคเตอร์ทำให้สามารคำนวณค่าความคล้ายคลึงจากการคำนวณโดยวิธีอินเนอร์โปรดักส์และคำนวณค่าของมุมโคไซน์โปรดักส์ซึ่งพิจารณาจากความถี่ของคำและเอกสารที่ค้นคืนได้สามารถนำไปจัดอันดับ (Ranking) ได้

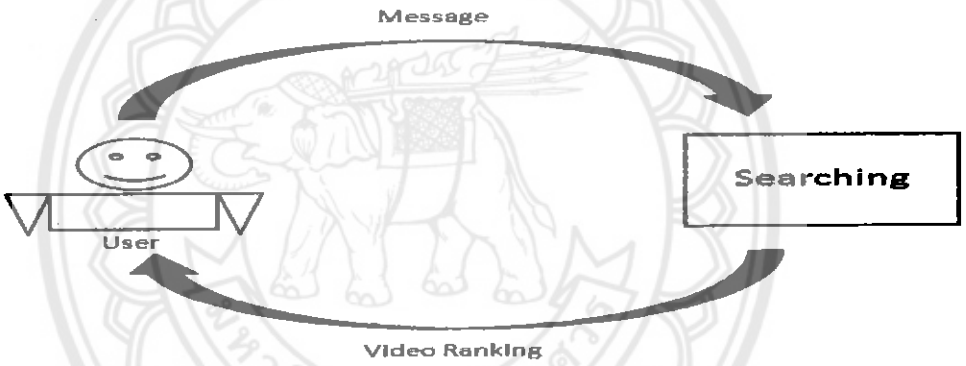
3.4.1 การสืบค้นโดยการเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดในวิดีโอคลิป

การสืบค้นโดยการเปรียบเทียบข้อความหรือคำกับคำพูดในวิดีโอคลิป จะทำการค้นหาโดยการนำคำศัพท์แต่ละคำที่ได้จากอินพุตมาเปิดหาในฐานข้อมูลดัชนีคำและนำผลลัพธ์ที่ได้จากการ

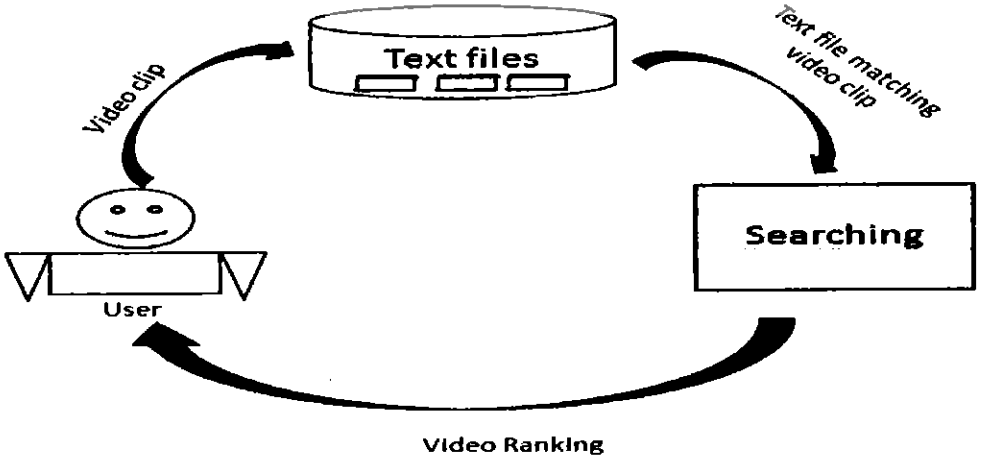
ค้นหาของแต่ละครั้งจัดอันดับหาเอกสารที่มีความเหมือนกับอินพุตมากที่สุดตามลำดับ แล้วจึงแสดงชื่อและลิงค์ของไฟล์วีดีโอคลิปที่สัมพันธ์กับเอกสารเหล่านั้นกลับไปแสดงทางหน้าจอของผู้ใช้งาน

3.4.2 การสืบค้นโดยการเปรียบเทียบเนื้อหาคำพูดในวิดีโอคลิประหว่างวิดีโอคลิปกับคลิปวิดีโอ

การเปรียบเทียบเนื้อหาคำพูดในวิดีโอคลิประหว่างวิดีโอคลิปด้วยกัน จะทำการค้นหาโดยการนำชื่อของวิดีโอคลิปที่รับเป็นอินพุตเข้ามาไปทำการติดต่อกับฐานข้อมูลของเอกสารแล้วนำเอกสารที่สัมพันธ์กับวิดีโอคลิปอินพุตมาใช้ในการค้นหาโดยการนำคำศัพท์แต่ละคำที่อยู่ในเอกสารมาเปิดหาในฐานข้อมูลดัชนีคำและนำผลลัพธ์ที่ได้จากการค้นหาของแต่ละครั้งมาจัดอันดับหาเอกสารที่มีความเหมือนกับเอกสารที่สัมพันธ์กับวิดีโอคลิปอินพุตมากที่สุดตามลำดับ แล้วจึงแสดงชื่อและลิงค์ของไฟล์วีดีโอคลิปที่สัมพันธ์กับเอกสารเหล่านั้นกลับไปแสดงทางหน้าจอของผู้ใช้งาน



รูปที่ 3.4 แสดงการค้นหาแบบอินพุตเป็นคำหรือข้อความ



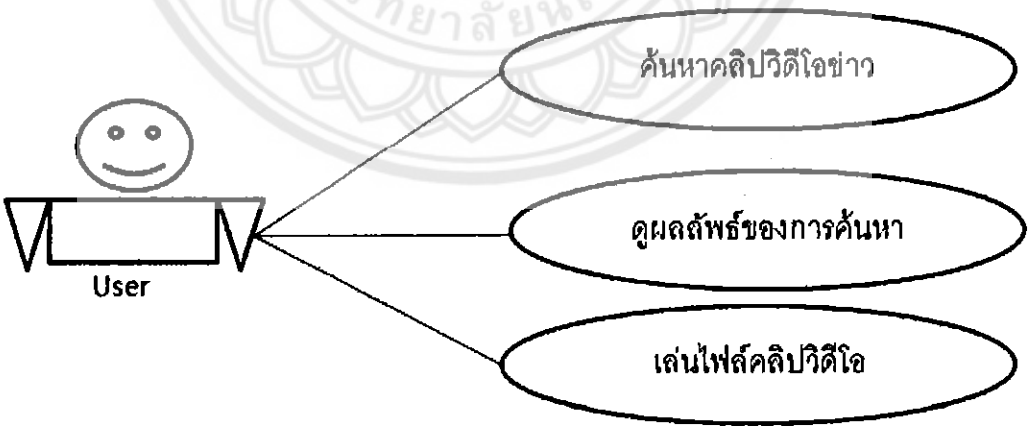
รูปที่ 3.5 แสดงการค้นหาแบบอินพุตเป็นวิดีโอคลิป

3.4.3 การจัดอันดับ (Ranking)

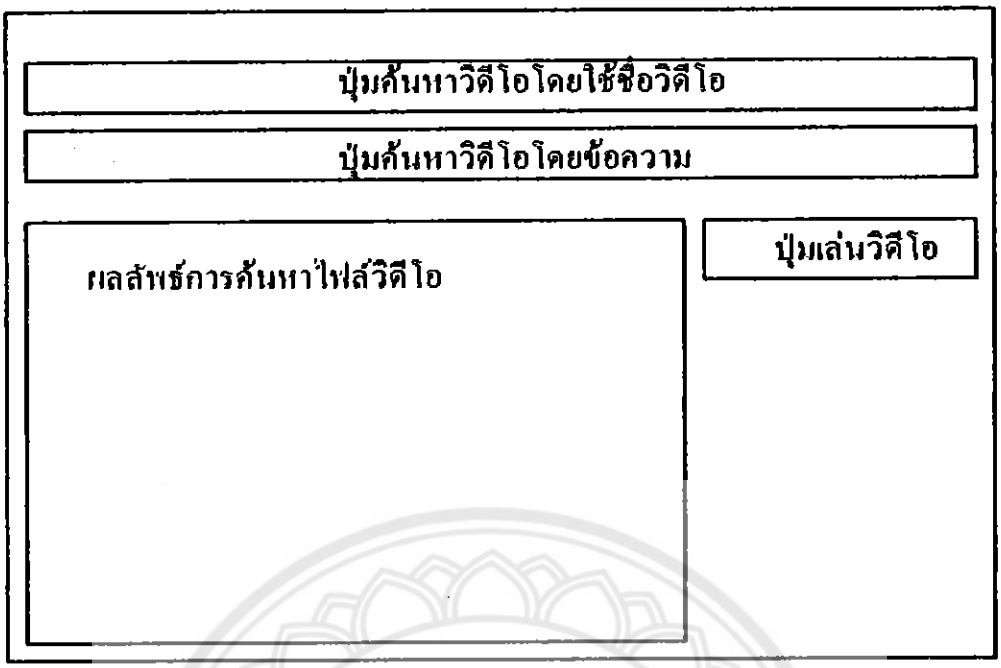
เมื่อทำการค้นหาข้อมูล (Searching) เพื่อนำมาแสดงให้กับผู้ใช้ จะมีการจัดทำลำดับของวิดีโอคลิป (Ranking) ตามลำดับความเหมือนของเอกสารที่สัมพันธ์กับวิดีโอคลิป กับข้อความที่ต้องการค้นหา (Similarity) ซึ่งวิดีโอคลิปที่เกี่ยวข้องกับผู้ใช้มากที่สุด (High Relevance) จะถูกแสดงอยู่ในลำดับต้น ๆ ซึ่งการจัดลำดับจะดูจากความถี่ของคำในไฟล์วิดีโอคลิป

3.5 การออกแบบการใช้งานการสืบค้นผ่านเว็บแอปพลิเคชั่น(User Interface)

ในส่วนที่ติดต่อกับผู้ใช้มีหน้าที่ในการรับอินพุตที่เป็นข้อความหรือคำ (Query) และรับอินพุตที่เป็นวิดีโอคลิปจากผู้ใช้ โดยผู้ใช้สามารถป้อนข้อความหรือคำที่ต้องการสืบค้นและสามารถป้อนชื่อของวิดีโอคลิปที่ใช้ในการสืบค้น เพื่อนำไปสู่กระบวนการค้นหาต่อไป และเป็นส่วนที่แสดงผลจากการค้นคืนให้กับผู้ใช้ นอกจากนี้ยังเป็นส่วนที่ผู้ใช้สามารถเข้าไปชมวิดีโอคลิปที่เป็นผลลัพธ์จากการค้นหาได้อีกด้วย โดยจะแสดงเป็นแผนภาพยูสเคสไดอะแกรม (Use case Diagram) ดังรูปที่ 3.6 เพื่อให้ง่ายต่อการทำความเข้าใจ ส่วนรูปแบบหน้าเว็บแอปพลิเคชั่นที่ใช้การค้นหานั้นก็แสดงดังรูปที่ 3.7



รูปที่ 3.6 แสดงแผนภาพ Use Case Diagram การใช้งานเว็บแอปพลิเคชั่นของ User



รูปที่ 3.7 แสดงรูปแบบเว็บแอปพลิเคชั่น

บทที่ 4

ผลการทดลอง

ในบทนี้จะเป็นการทดสอบว่าโครงงานที่ออกแบบมานั้นมีประสิทธิภาพในระดับใด โดยแบ่งการทดลองออกเป็น 3 แบบดังตารางที่ 4.1

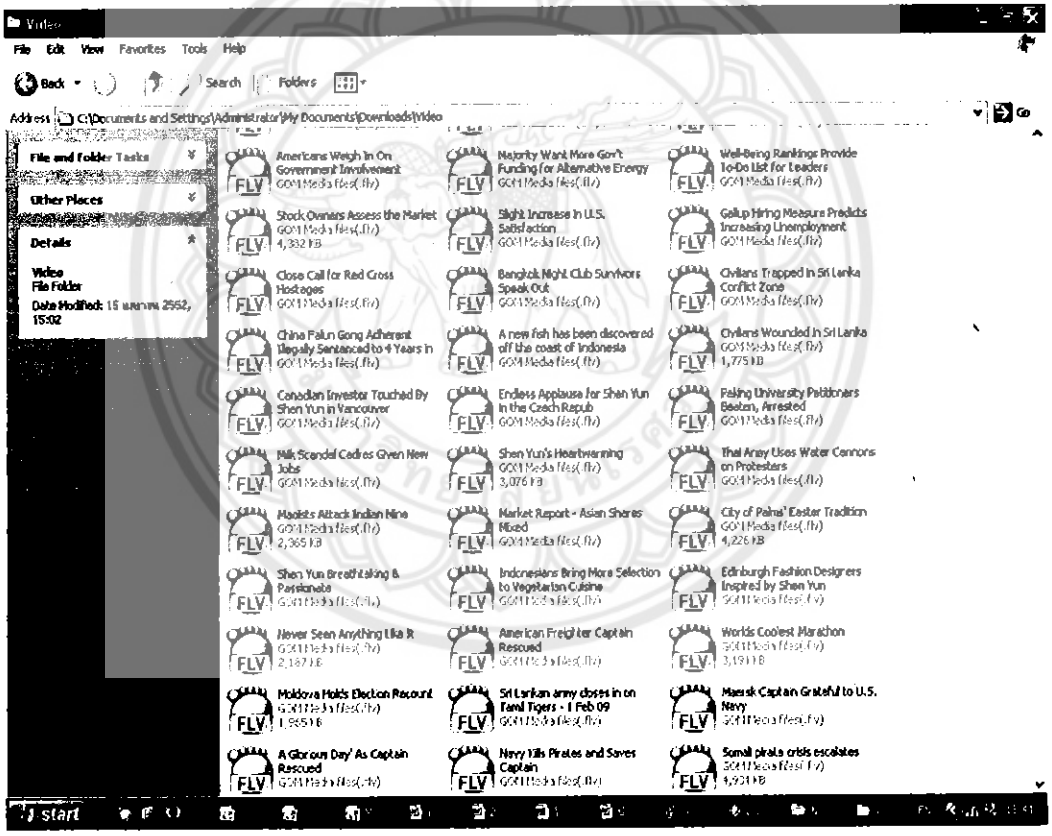
ตารางที่ 4.1 ตารางแสดงแผนการทดลอง

ลำดับ	การทดลอง
1	การทดลองกระบวนการสร้างดัชนี ● การทดสอบการใช้โปรแกรมแปลงวิดีโอคลิปเป็นไฟล์เสียงและปรับอัตราบิตเรทกับความถี่ของไฟเสียง ● การทดสอบการแปลงไฟล์เสียงเป็น text file ● การทดสอบการสร้างฐานข้อมูลดัชนีคำ
2	การทดลองการค้นหาและการแสดงผลเกี่ยวกับวิดีโอคลิป ● การทดสอบการค้นหา - ทดสอบการค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป - ทดสอบการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป ● การทดสอบการเล่นไฟล์วีดีโอคลิป
3	การทดลองหาประสิทธิภาพการค้นหา ● การวัดประสิทธิภาพในการค้นหาโดยการเปรียบเทียบเนื้อหาคำพูดในวิดีโอคลิประหว่างวิดีโอคลิป ● การวัดประสิทธิภาพในการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป

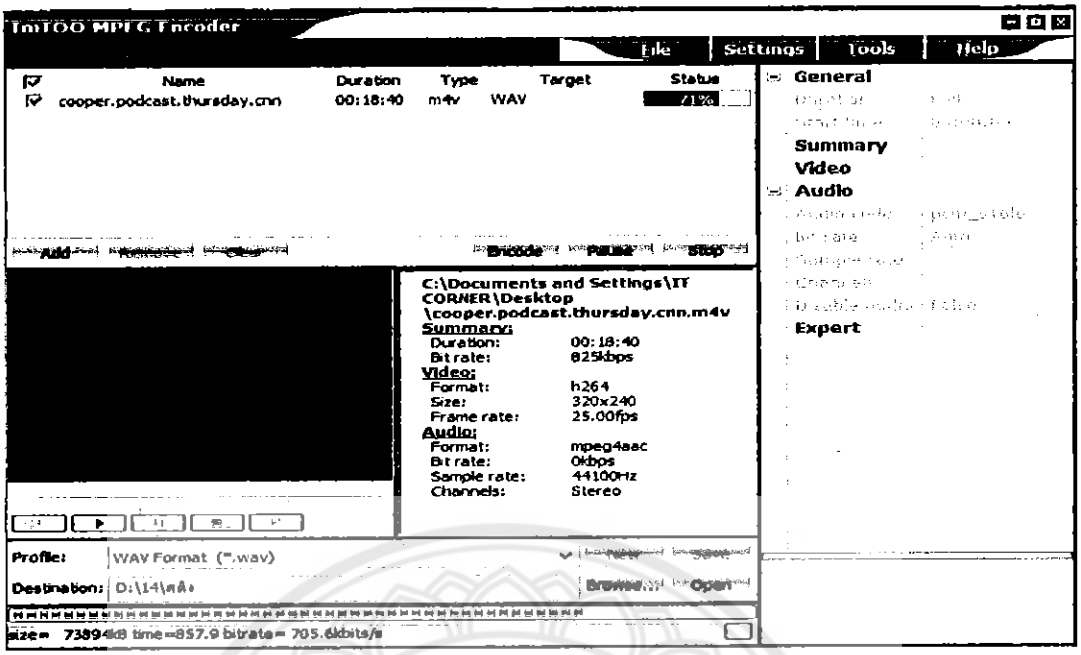
4.1 การทดลองกระบวนการสร้างดัชนี

4.1.1 การทดสอบการใช้โปรแกรมแปลงวิดีโอคลิปเป็นไฟล์เสียงและปรับอัตราบิตเรตกับความถี่ของไฟล์เสียง

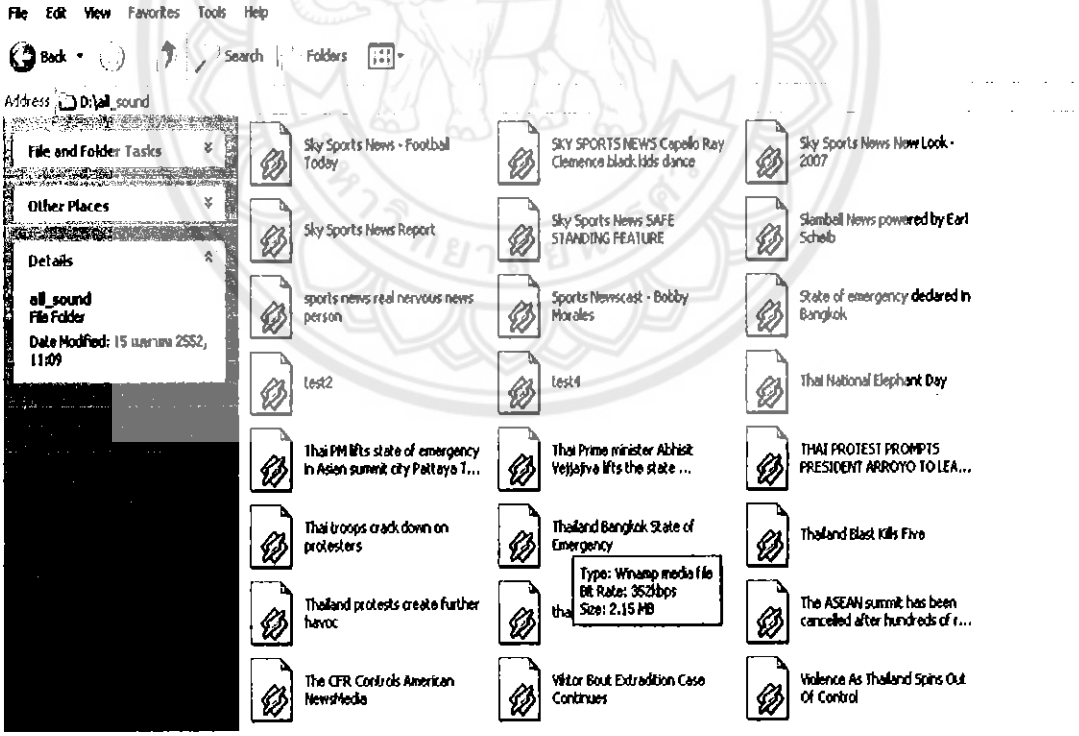
จะทดสอบโดยการนำวิดีโอคลิปข่าวจำนวน 100 คลิป ที่มีอยู่ดังในรูปที่ 4.1 มาแปลงเป็นไฟล์เสียงโดยใช้โปรแกรมสำเร็จรูป ImTOOMPEG Encoder ดังในรูปที่ 4.2 เมื่อแปลงไฟล์เสร็จเรียบร้อยจะได้ไฟล์เสียงออกมาดังในรูป 4.3 เมื่อได้ไฟล์เสียงมาแล้วนำมาปรับอัตราบิตกับความถี่ของเสียงโดยใช้โปรแกรมสำเร็จรูป WavePad โดยแปลงความถี่ของคลื่นเสียงเป็น 22500 เฮิร์ต และอัตราของบิตเรตเป็น 16 บิต ดังในรูปที่ 4.4



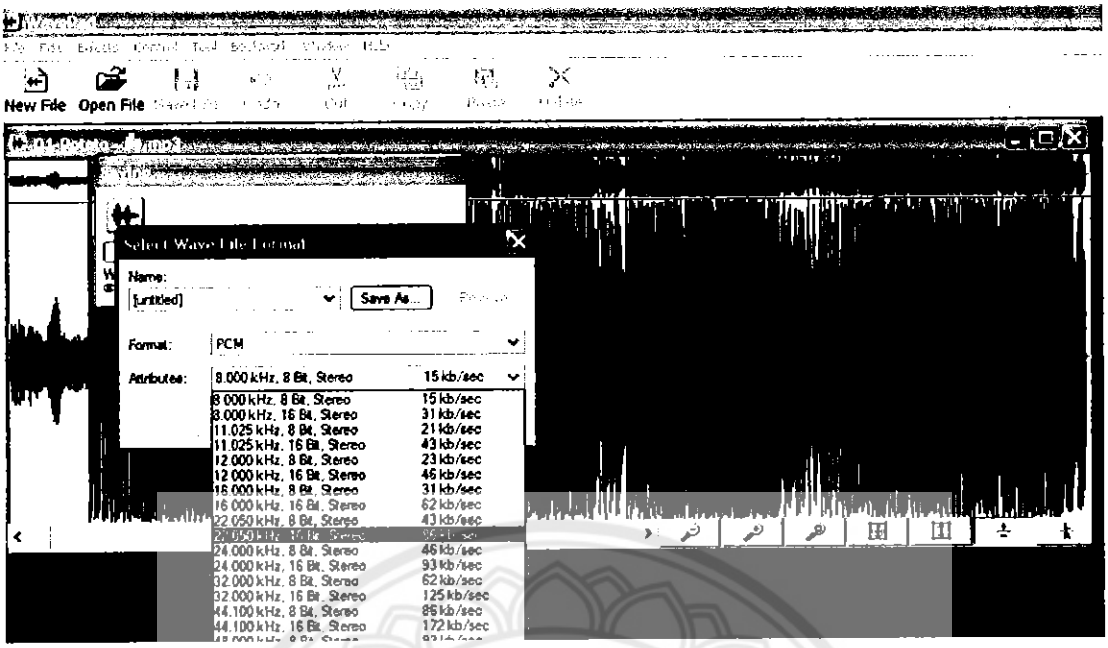
รูปที่ 4.1 รูปวิดีโอคลิปตัวอย่างก่อนการแปลงเป็นไฟล์เสียง



รูปที่ 4.2 แสดงการทำงานของโปรแกรม ImTOOMPEG Encoder



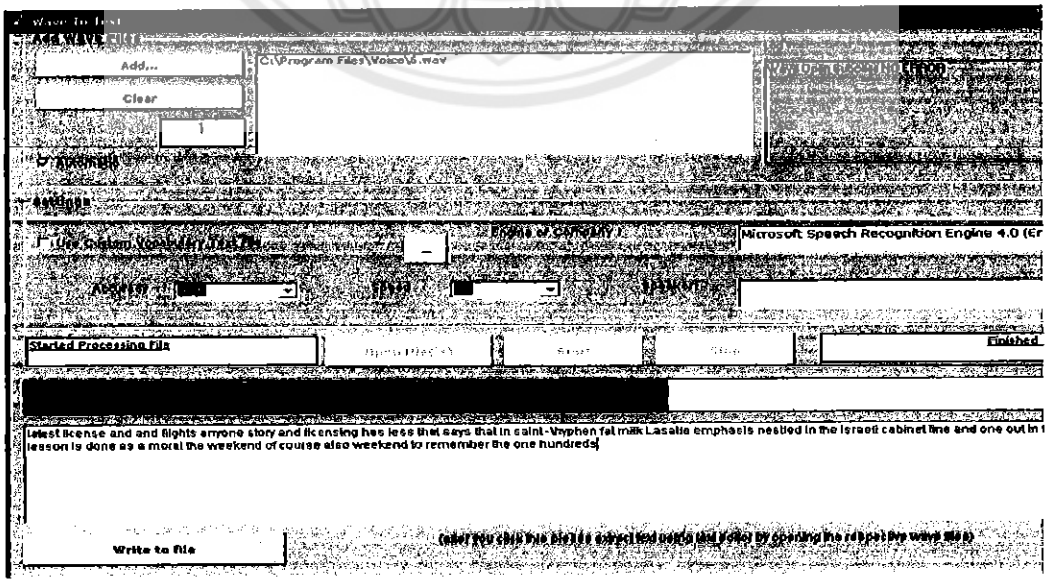
รูปที่ 4.3 แสดงไฟล์เสียงหลังจากการแปลงไฟล์โดยใช้โปรแกรม ImTOOMPEG Encoder



รูปที่ 4.4 แสดงการที่แปลงความถี่ของคลื่นเสียงเป็น 22500 เฮิร์ต และอัตราของบิตเรตเป็น 16 บิต

4.1.2 การทดสอบการแปลงไฟล์เสียงเป็นไฟล์ตัวอักษร

จะทดสอบโดยการนำไฟล์เสียงที่ผ่านการปรับค่าความถี่และอัตราบิตเสร็จเรียบร้อยมาทำการแปลงเป็นไฟล์ตัวอักษรโดยใช้โปรแกรมแปลงคำพูดเป็นตัวอักษรที่ชื่อว่า Voice Recognize ดังในรูปที่ 4.5 ซึ่งเมื่อแปลงไฟล์เสร็จแล้วจะได้ไฟล์ตัวอักษรออกมาดังรูปที่ 4.6 ซึ่งไฟล์ตัวอักษรทั้งหมดจะถูกเก็บให้อยู่ในโฟลเดอร์ที่ชื่อว่า all_text ดังรูปที่ 4.7



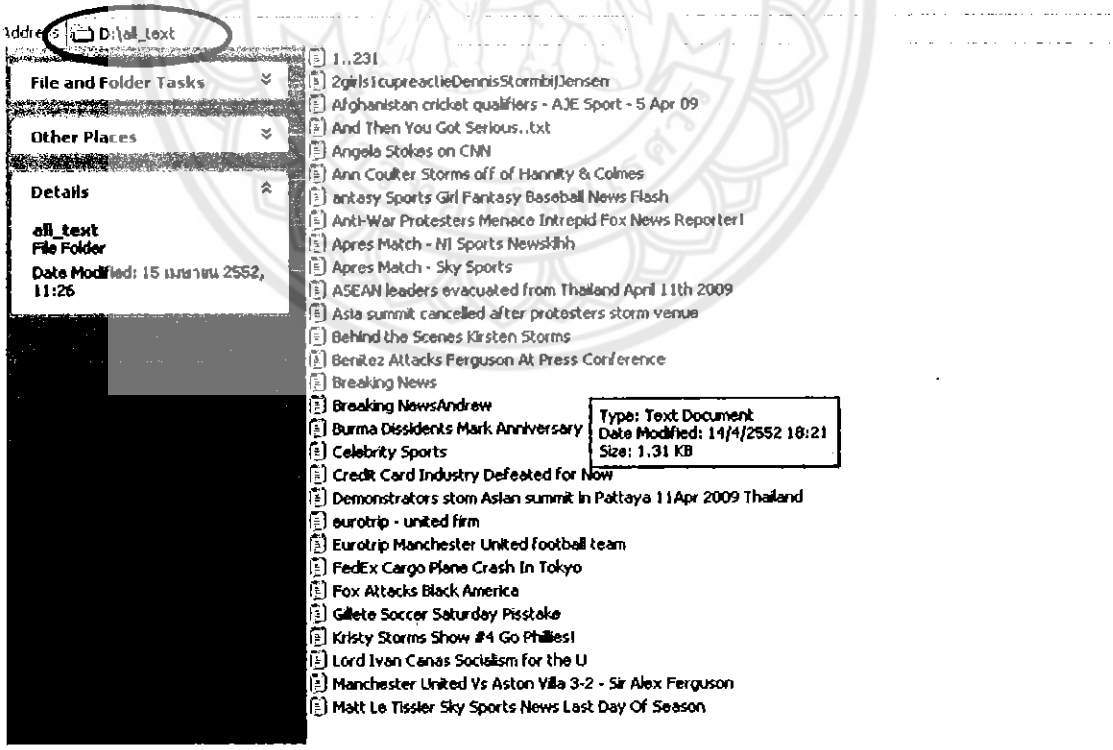
รูปที่ 4.5 รูปการแปลงไฟล์เสียงเป็นตัวอักษรโดยใช้โปรแกรม Voice Recognize

6 - Notepad

File Edit Format View Help

some if they line million the Zavala the members will is visiting Elizabeth Glaser remain well nine sixty and with the fees in Lincoln high on the left and then this guy with women ingestion ninety fan club from Ohio humans is the noise was saying the man's his view new-line so in nuns \comma much of this is not be a dime law and in a bomb in the law and and who stay at home Tahoe the see you and can do them the commitment and Indian who would go away valley and visitors and home rates are living human they view the lighting and only the died :\colon so what the hell hath name is now off San offering Louisville and with John is a new monument growth in the high AI - \hyphen museum in new-line the following the Phoenix a Vietnam has only to Ers he is in the early going in one allow liaison for Zen nothing his wife Lance journalist the only

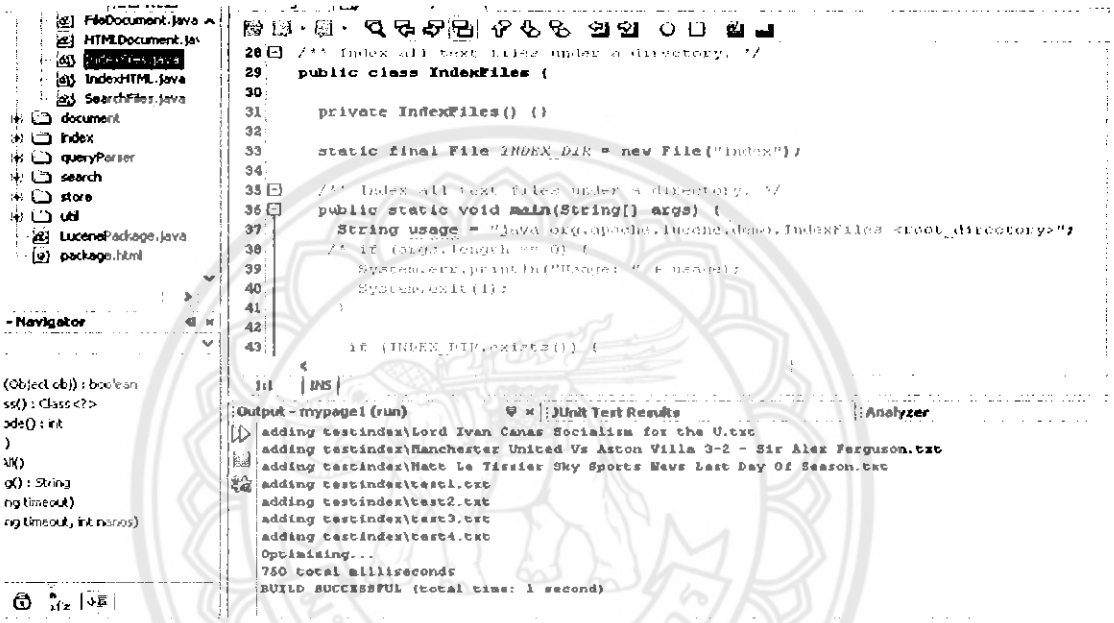
รูปที่ 4.6 แสดง Text file ที่ได้จากการแปลงจากโปรแกรม



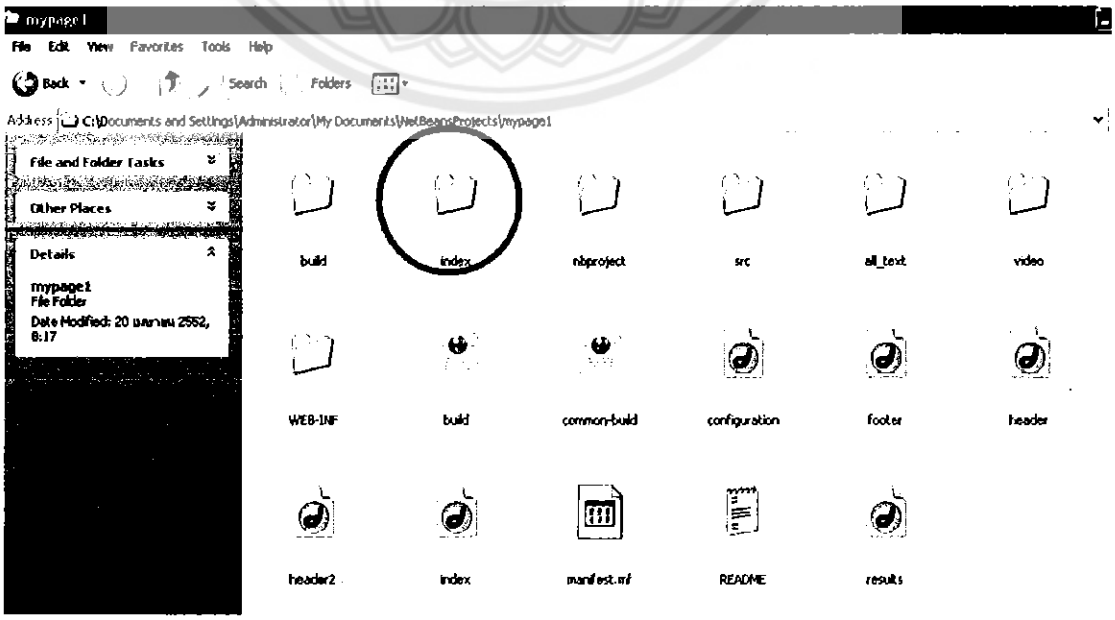
รูปที่ 4.7 แสดง Text file ที่ผ่านการแปลงไฟล์เสร็จเรียบร้อย

4.1.3 การทดสอบการสร้างฐานข้อมูลดัชนีคำ

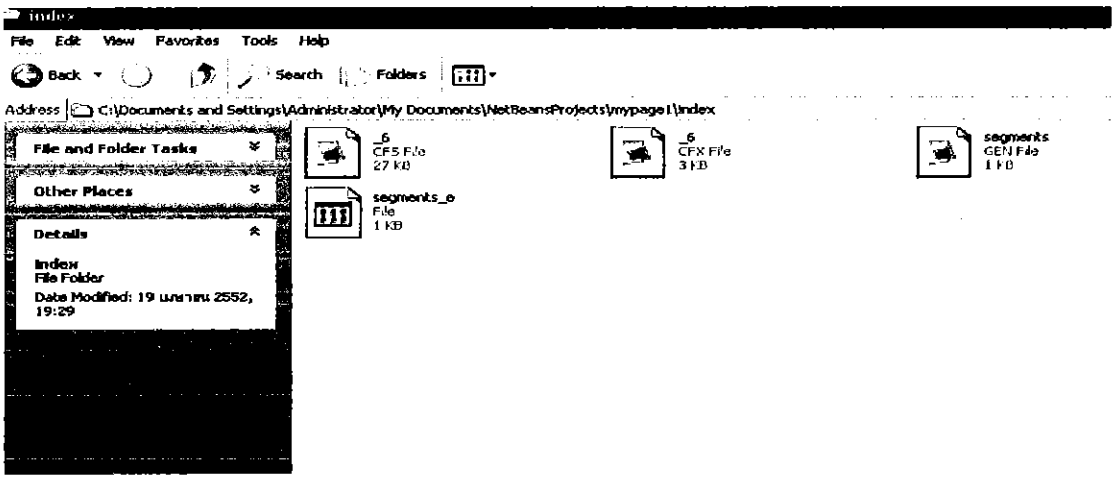
จะทดสอบโดยการนำไฟล์ตัวอักษรที่ได้จากการแปลงไฟล์ในทุกขั้นตอนแล้วมาทำดัชนี โดยการสร้างดัชนีนั้นจะทำการรันโปรแกรมในรูปที่ 4.8 ผลการทำงานของโปรแกรมจะได้โฟลเดอร์ที่ชื่อ index ดังรูปที่ 4.9 โดยข้างในโฟลเดอร์ index นั้นจะเป็นไฟล์เก็บดัชนีคำที่ได้จากการสร้างดัชนีโดยโปรแกรมดังในรูปที่ 4.10



รูปที่ 4.8 แสดงการรันโปรแกรมการทำ indexing



รูปที่ 4.9 แสดงโฟลเดอร์ชื่อว่า index ซึ่งเก็บไฟล์เก็บดัชนีคำที่ได้จากการทำ indexing



รูปที่ 4.10 แสดงไฟล์เก็บดัชนีคำที่อยู่ในโฟลเดอร์ index

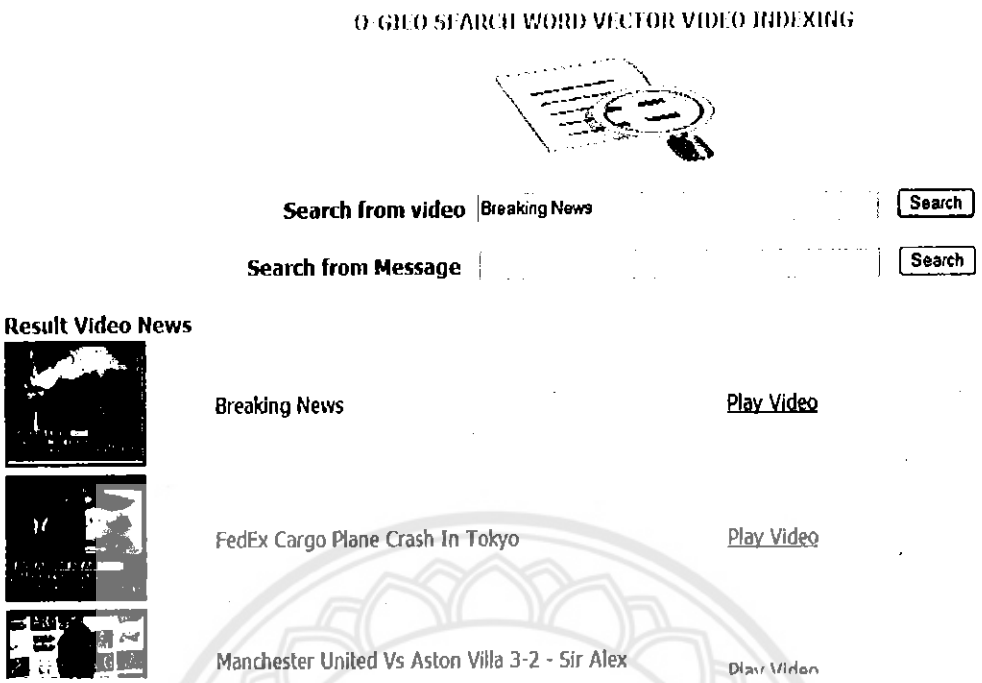
4.2 การทดลองการสืบค้นและการแสดงผลเกี่ยวกับวิดีโอคลิป

4.2.1 การทดสอบการสืบค้น

การทดสอบการค้นหาเพื่อตรวจสอบการทำงานของระบบการสืบค้นสามารถทำได้ดังต่อไปนี้

1. เริ่มจากการเปิดหน้าเว็บแอปพลิเคชั่นที่ใช้ในการสืบค้นขึ้นมาดังรูปที่ 4.11
2. จากนั้นใส่คำหรือข้อความที่ต้องการสืบค้นในช่องสำหรับใช้สืบค้น
3. คลิก search ผลลัพธ์ของการสืบค้นที่ได้จะถูกแสดงดังรูปที่ 4.12 และ 4.13

รูปที่ 4.11 แสดงหน้าเว็บแอปพลิเคชั่นที่ใช้ในการสืบค้น

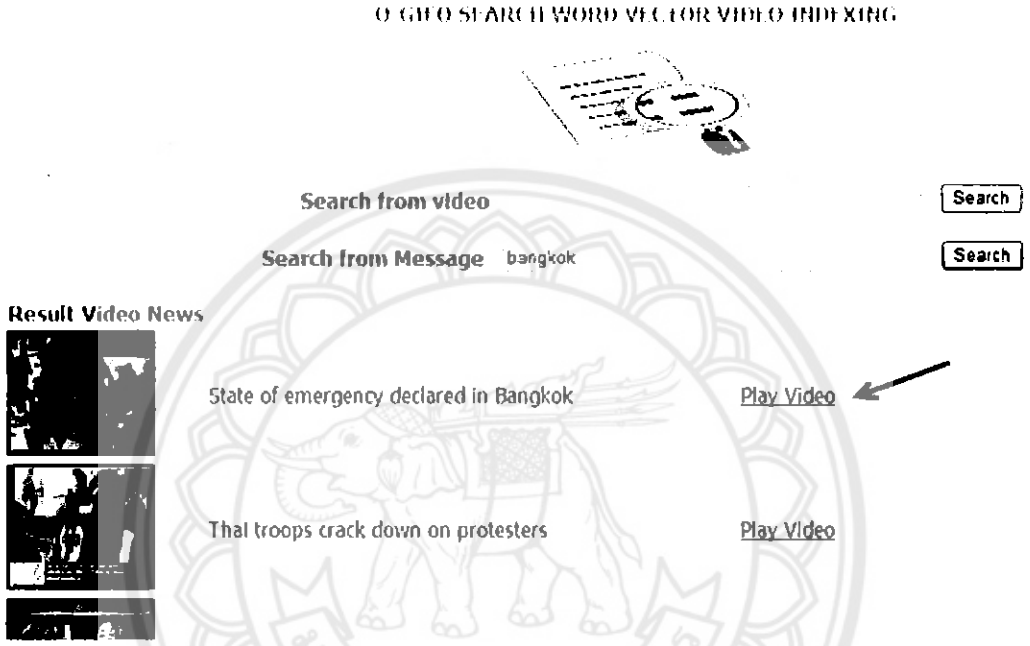


รูปที่ 4.12 แสดงผลการค้นหาโดยใช้การเปรียบเทียบเนื้อหาคำพูดระหว่างวิดีโอคลิปกับคลิปวิดีโอ

รูปที่ 4.13 แสดงผลการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป

4.2.4 การทดสอบการเล่นไฟล์วีดีโอ

การทดสอบการเล่นไฟล์วีดิโอเพื่อตรวจสอบการทำงานของการแสดงการเล่นไฟล์วีดิโอจากวีดีโอคลิปที่เป็นผลลัพธ์จากการสืบค้น โดยสามารถเข้าดูวีดีโอคลิปด้วยการคลิกที่ลิงค์ของวีดีโอคลิปดังรูปที่ 4.14 และผลลัพธ์จากการเล่นวีดีโอจะแสดงอยู่บนหน้าเว็บดังรูปที่ 4.15



รูปที่ 4.14 แสดงการเข้าดูวีดีโอคลิปด้วยการคลิกที่ลิงค์ของวีดีโอคลิป

รูปที่ 4.15 แสดงวีดีโอคลิปที่เลือกดู

4.3 การทดลองหาประสิทธิภาพการค้นหา

เป็นการวัดประสิทธิภาพในการทำงานของการสืบค้น โดยใช้ผลการทดลองจากการสืบค้นวีดิโอคลิปโดยการใช้หลักการเปรียบเทียบผลของค่าความถูกต้องได้แก่ ค่า Recall และค่า Precision มาวิเคราะห์หาประสิทธิภาพซึ่งเป็นค่าที่ได้จากการคำนวณสมการต่อไปนี้

$$\text{Recall} = \frac{\text{จำนวนวิดีโอคลิปที่เกี่ยวข้องและถูกนำมาแสดงเป็นผลลัพธ์}}{\text{จำนวนคลิปวิดีโอที่เกี่ยวข้องทั้งหมด}} \quad (4.1)$$

$$\text{Precision} = \frac{\text{จำนวนวิดีโอคลิปที่เกี่ยวข้องที่ถูกนำมาแสดงเป็นผลลัพธ์}}{\text{จำนวนคลิปวิดีโอที่เป็นผลลัพธ์ทั้งหมด}} \quad (4.2)$$

$$\text{Precision top 10} = \frac{\text{จำนวนวิดีโอคลิปที่เกี่ยวข้องและเป็นผลลัพธ์ใน10 อันดับแรก}}{10} \quad (4.3)$$

4.3.1 การวัดประสิทธิภาพการค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิป

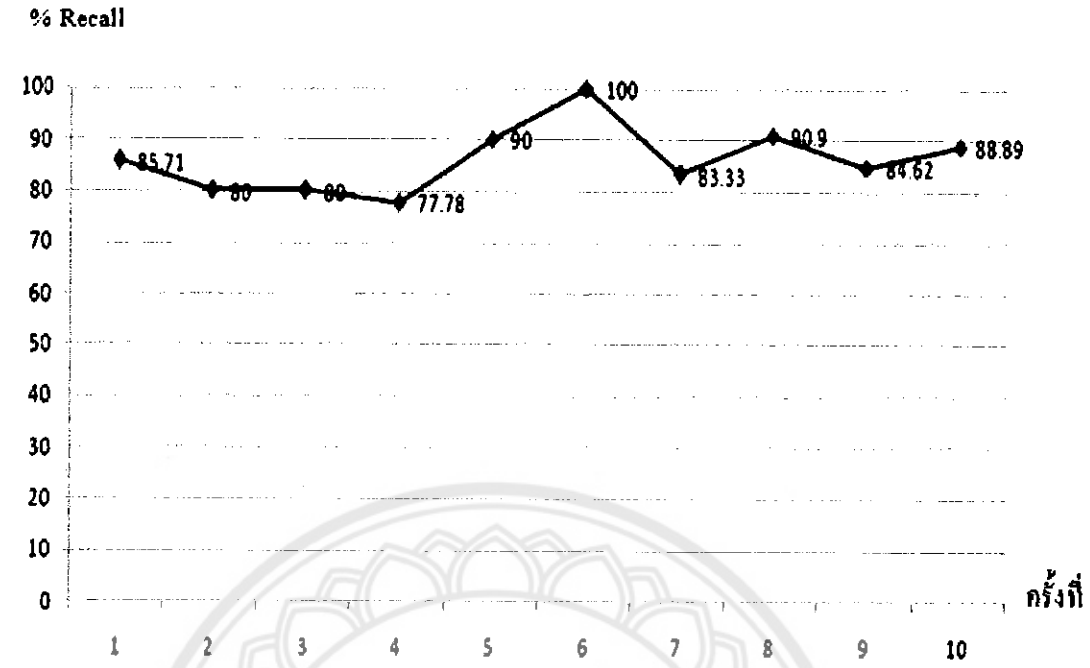
การทดลองสืบค้นโดยการเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปจำนวน 10 ครั้ง รายชื่อวิดีโอคลิปที่ใช้ในการสืบค้นได้แก่

- And Then You Got Serious
- Singapore's sexual revolution
- Eurotrip Manchester United football team
- Lord Ivan Canas Socialism
- Burma Dissidents Mark Anniversary
- Asia summit cancelled after protesters storm venue
- Sports Newscast - Bobby Morales
- Apres Match - NI Sports Newsklhh
- Thai troops crack down on protesters
- Behind the Scenes Kirsten Storms

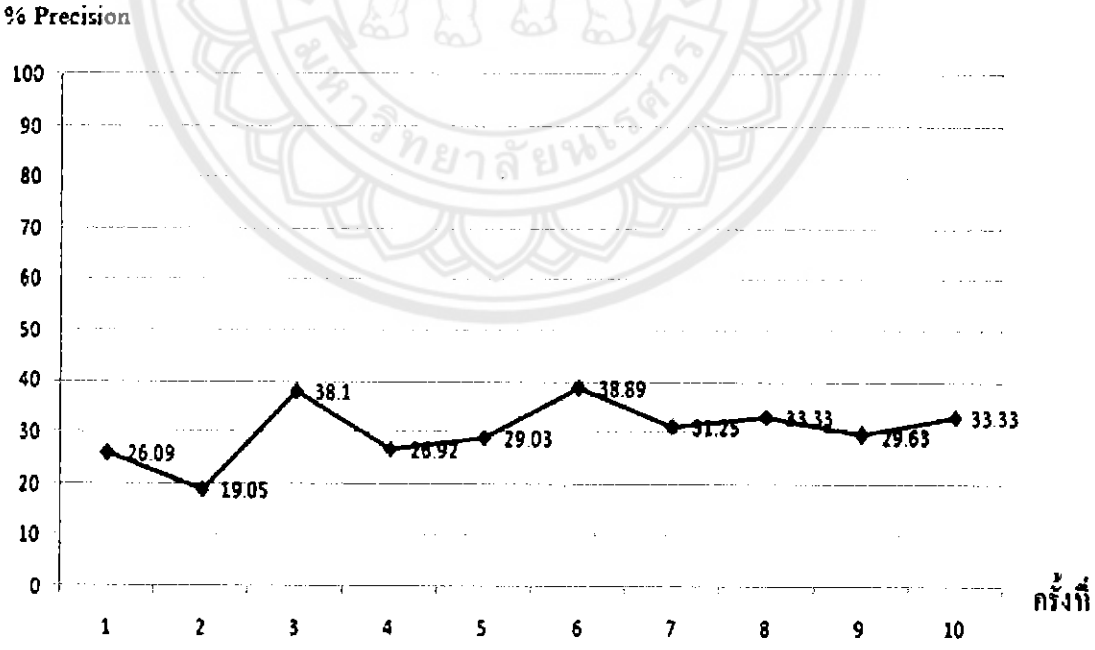
ปรากฏผลการทดลองดังตารางที่ 4.2

ตารางที่ 4.2 ตารางแสดงผลการทดลองการค้นหาการเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป

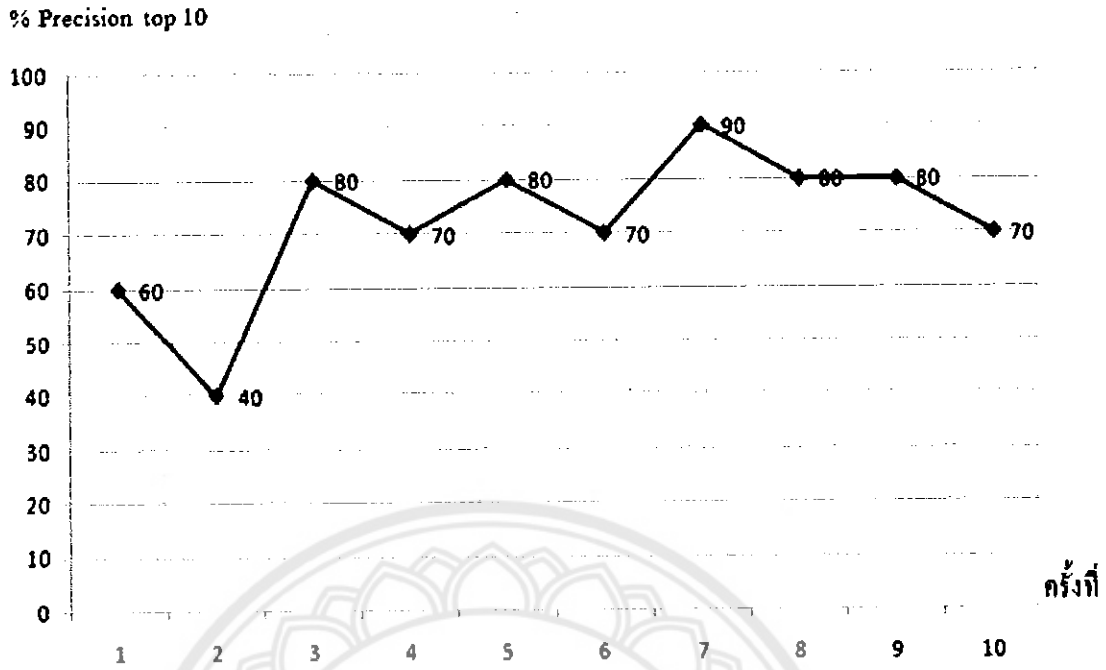
ครั้งที่	จำนวนวิดีโอคลิปที่เกี่ยวข้องและถูกนำมาแสดงเป็นผลลัพธ์	จำนวนวิดีโอคลิปที่เกี่ยวข้องและถูกนำมาแสดงเป็นผลลัพธ์ใน 10 อันดับแรก	จำนวนวิดีโอคลิปที่เกี่ยวข้องทั้งหมด	จำนวนวิดีโอคลิปที่เป็นผลลัพธ์ทั้งหมด	% Recall	% Precision	% Precision top 10
1	6	6	7	23	85.71	26.09	60.00
2	4	4	5	21	80.00	19.05	40.00
3	8	8	10	21	80.00	38.10	80.00
4	7	7	9	26	77.78	26.92	70.00
5	9	8	10	31	90.00	29.03	80.00
6	7	7	7	18	100.00	38.89	70.00
7	10	9	12	32	83.33	31.25	90.00
8	10	8	11	30	90.90	33.33	80.00
9	11	8	13	27	84.62	29.63	80.00
10	8	7	9	24	88.89	33.33	70.00
		ค่าเฉลี่ย			86.12	30.56	72.00



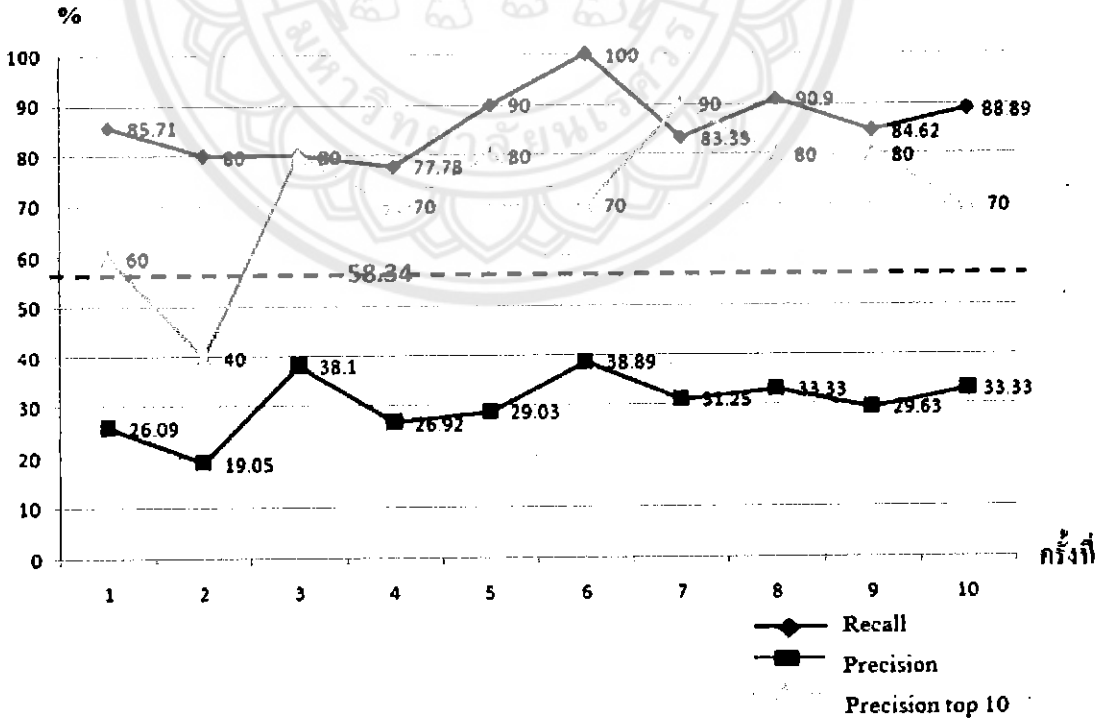
รูปที่ 4.16 แสดงกราฟค่า Recall ของการทดลองค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป



รูปที่ 4.17 แสดงกราฟค่า Precision ของการทดลองค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป



รูปที่ 4.18 แสดงกราฟค่า Precision top 10 ของการทดลองค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป



รูปที่ 4.19 แสดงกราฟเปรียบเทียบค่าRecall, ค่า Precision และค่า Precision top 10 ของการทดลองค้นหาโดยใช้การเปรียบเทียบในเนื้อหาคำพูดระหว่างวิดีโอคลิปกับวิดีโอคลิป

นำข้อมูลที่ได้จากการทดลองสืบค้นมาเขียนกราฟแสดงค่าเปอร์เซ็นต์ Recall กราฟแสดงเปอร์เซ็นต์ค่า Precision กราฟแสดงเปอร์เซ็นต์ค่าPrecision top 10 และกราฟเปรียบเทียบค่า Recall ค่า Precision และ Precision top 10 ได้ดังรูปที่ 4.16 รูปที่ 4.17 รูปที่ 4.19 และรูปที่ 4.19 ตามลำดับ จากกราฟทั้งหมดข้างต้นจะเห็นว่าค่าเปอร์เซ็นต์ Recall มีค่าอยู่ในช่วง 75-100 และเปอร์เซ็นต์ค่า Precision มีค่าอยู่ในช่วง 19-40

จากการสังเกตกราฟในรูปที่ 4.19 พื้นที่ตรงกลางระหว่างเส้นกราฟ Recall และPrecision ที่ถูกแบ่งโดยเส้นประเป็นสองฝั่งตรงจุดกึ่งกลางระหว่างค่าเฉลี่ยของเส้นกราฟทั้งสองนั้นค่อนข้างมีความสมดุลกัน นอกจากนี้สังเกตเส้นกราฟ Precision top 10 จะเห็นว่าวิดีโอคลิปผลลัพธ์ที่มีเนื้อหาเกี่ยวข้องกันกับวิดีโอที่ใช้สืบค้นจะมีอัตราการถูกจัดให้อยู่ใน 10 อันดับแรกโดยเฉลี่ยสูงถึงร้อยละ 72 อีกด้วย แสดงถึงประสิทธิภาพในการค้นหาค้นที่ดี แต่ก็ยังมีความผิดพลาดเกิดขึ้นแต่เป็นเพียงส่วนน้อย เนื่องมาจากคำพูดในวิดีโอคลิปที่ใช้สืบค้นกับคำพูดในบางวิดีโอคลิปเป้าหมายที่เกี่ยวข้องไม่ตรงกันเลย เมื่อทำการการค้นหาและจึงไม่พบคลิปวิดีโอที่เกี่ยวข้องดังกล่าว

4.3.2 การค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป

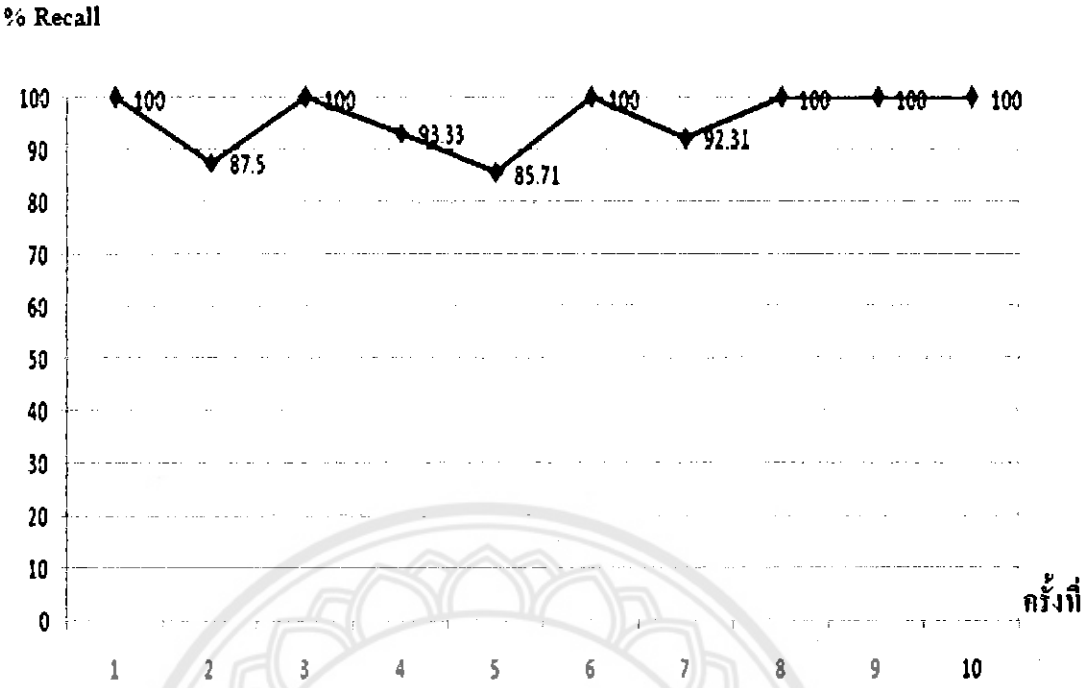
การทดลองสืบค้นด้วยข้อความหรือคำสำคัญจำนวน 10 ครั้ง ได้แก่

- Thailand
- Bangkok
- soccer
- red shirt
- play football
- tsunami
- elephant
- Manchester United football team
- Sports
- Korea

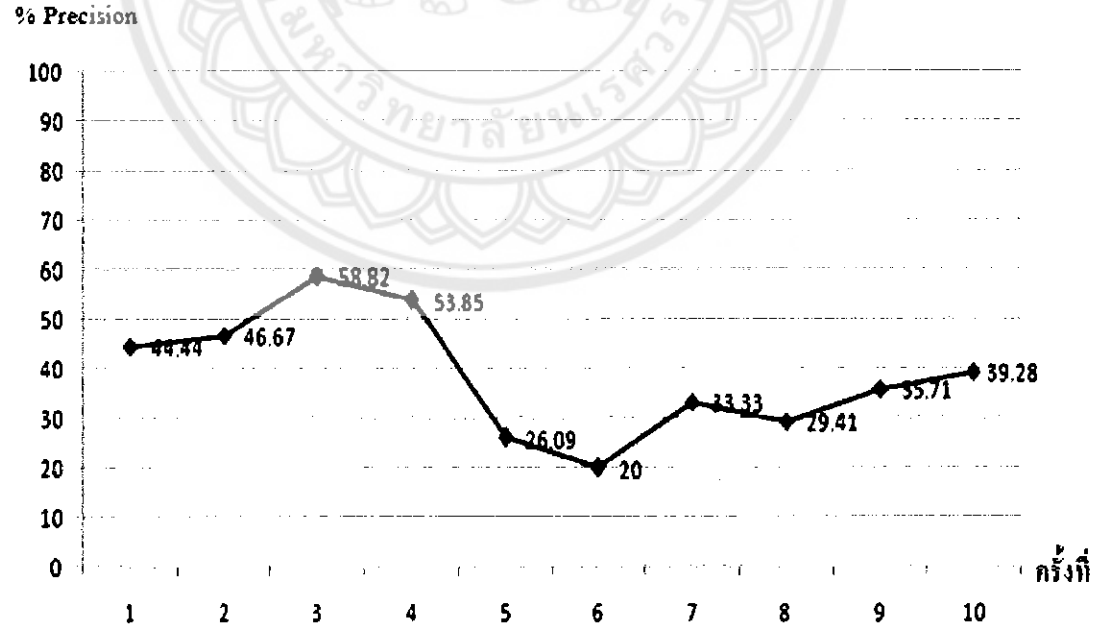
ปรากฏผลการทดลองดังตารางที่ 4.3

ตารางที่ 4.3 ตารางแสดงผลการทดลองการค้นหาการเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป

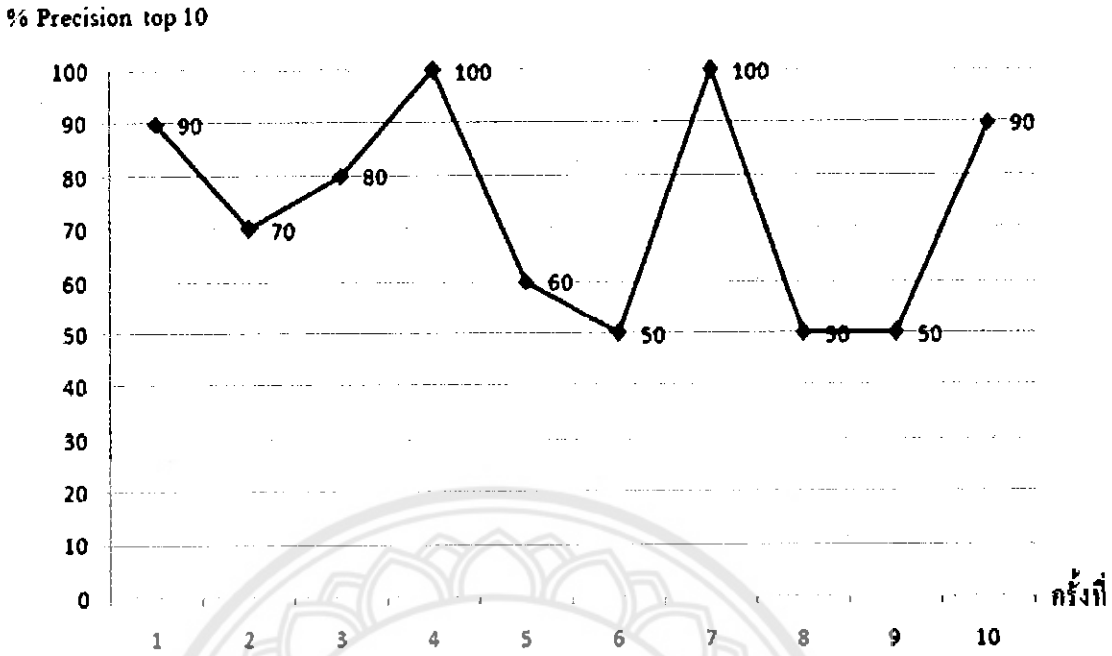
ครั้งที่	จำนวนวิดีโอคลิปผลลัพธ์ที่เกี่ยวข้องและถูกสืบค้นมาเป็นผลลัพธ์	จำนวนวิดีโอคลิปผลลัพธ์ที่เกี่ยวข้องและถูกสืบค้นมาเป็นผลลัพธ์ใน 10 อันดับแรก	จำนวนวิดีโอคลิปที่เกี่ยวข้องทั้งหมด	จำนวนวิดีโอคลิปที่เป็นผลลัพธ์ทั้งหมด	% Recall	% Precision	% Precision top 10
1	12	9	12	27	100	44.44	90.00
2	7	7	8	15	87.5	46.67	70.00
3	10	8	10	17	100	58.82	80.00
4	14	10	15	26	93.33	53.85	100.00
5	6	6	7	23	85.71	26.09	60.00
6	5	5	5	25	100	20.00	50.00
7	12	10	13	36	92.31	33.33	100.00
8	5	5	5	17	100	29.41	50.00
9	5	5	6	14	100	35.71	50.00
10	11	9	11	28	100	39.28	90.00
		ค่าเฉลี่ย			95.89	38.76	74.00



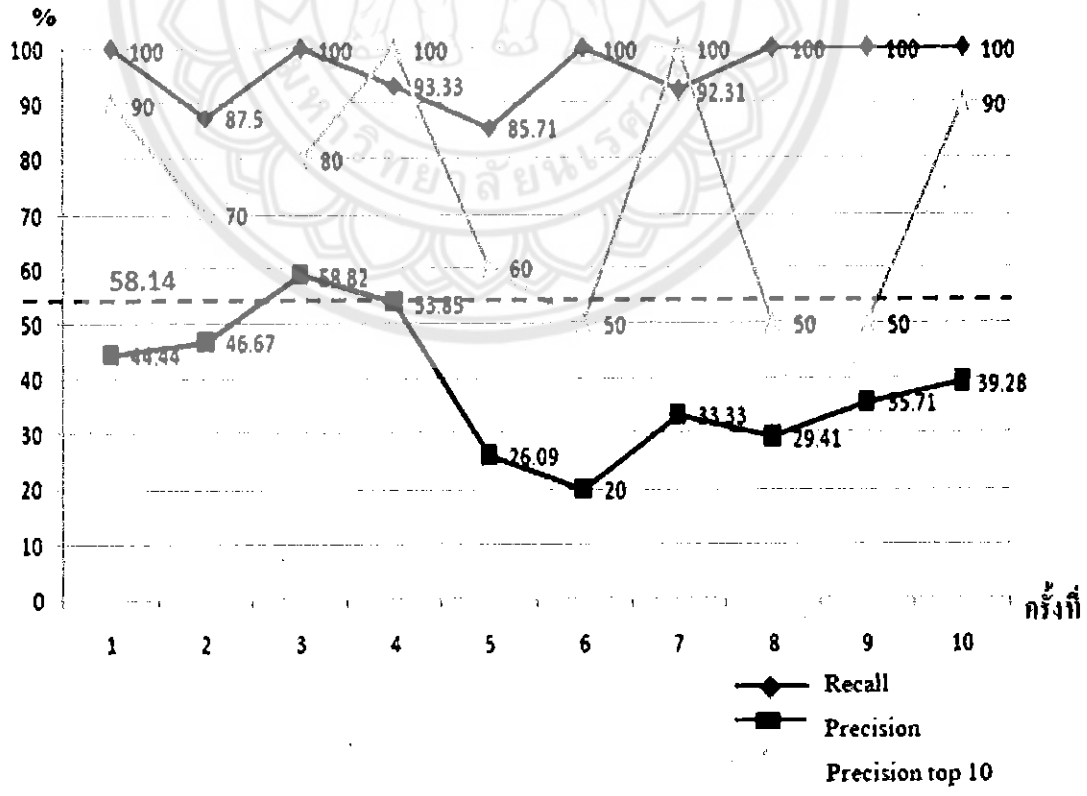
รูปที่ 4.20 แสดงกราฟค่า Recall ของการทดลองการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป



รูปที่ 4.21 แสดงกราฟค่า Precision ของการทดลองการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป



รูปที่ 4.22 แสดงกราฟค่า Precision Top 10 ของการทดลองการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญกับคำพูดที่มีในวิดีโอคลิป



รูปที่ 4.23 แสดงกราฟเปรียบเทียบค่าRecall และ ค่า Precision และ Precision Top 10 ของการทดลองการค้นหาโดยใช้การเปรียบเทียบข้อความหรือคำสำคัญที่มีในคลิปวิดีโอ

นำข้อมูลที่ได้จากการทดลองสืบค้นมาเขียนกราฟแสดงค่าเปอร์เซ็นต์ Recall กราฟแสดงเปอร์เซ็นต์ค่า Precision กราฟแสดงเปอร์เซ็นต์ค่า Precision top 10 และกราฟเปรียบเทียบค่า Recall ค่า Precision top 10 และ ค่า Precision ได้ดังรูปที่ 4.20 รูปที่ 4.21 รูปที่ 4.22 และรูปที่ 4.23 ตามลำดับ จากกราฟทั้งหมดข้างต้นจะเห็นว่าค่าเปอร์เซ็นต์ Recall มีค่าอยู่ในช่วง 85-100 และเปอร์เซ็นต์ค่า Precision มีค่าอยู่ในช่วง 20-60

จากกราฟในรูปที่ 4.23 จะเห็นว่าในส่วนของพื้นที่ตรงกลางระหว่างเส้นกราฟ Recall และ Precision ที่ถูกแบ่งโดยเส้นประเป็นสองส่วนตรงจุดกึ่งกลางระหว่างค่าเฉลี่ยของเส้นกราฟทั้งสองนั้นไม่ค่อยมีความสมดุลกันค่อนข้างโน้มเอียงไปทางเส้นกราฟ Recall มากกว่าเล็กน้อย นอกจากนี้จะเห็นว่าเส้นกราฟ Precision top 10 จะเห็นว่าวิดีโอคลิปผลลัพธ์ที่มีเนื้อหาเกี่ยวข้องกันกับวิดีโอที่ใช้สืบค้นจะมีอัตราการถูกจัดให้อยู่ใน 10 อันดับแรกโดยเฉลี่ยสูงถึงร้อยละ 74 อีกด้วย บ่งบอกถึงความสามารถในการเข้าถึงข้อมูลคลิปวิดีโอที่เกี่ยวข้องกับข้อความหรือคำสำคัญที่ใช้ในการสืบค้นได้ในระดับที่ค่อนข้างดี

บทที่ 5
สรุปผลการดำเนินงาน

โครงงานนี้เป็นโครงงานที่ทำเกี่ยวกับการทำดัชนีในการสืบค้นข้อมูลคลิปข่าวภาษาอังกฤษ โดยใช้เวคเตอร์ของคำพูดที่ปรากฏในคลิปข่าว มาทำเป็นดัชนีในการสืบค้นโดยการนำไฟล์คลิปวีดีโอข่าวภาษาอังกฤษ มาทำการถอดคำพูดที่ปรากฏในคลิปให้ออกมาเป็นตัวหนังสือโดยใช้โปรแกรมแปลงเสียงเป็นตัวอักษร จากนั้นจะนำข้อมูลคำที่ถอดมาทำเป็นดัชนีหรือฮีสโทแกรมเพื่อใช้ในการสืบค้นแล้ว จากนั้นทำการสืบค้นวิดีโอคลิปกับฐานข้อมูลดัชนีวิดีโอคลิปผ่านทางเว็บแอปพลิเคชั่น โมเดลที่นำมาใช้คือ เวคเตอร์สเปซโมเดล โดยการค้นคืนในรูปแบบนี้สามารถทำการค้นคืนได้วิดีโอคลิปที่มีเนื้อหาคำพูดที่มีแม้แต่ความเหมือนแค่เพียงบางส่วน โดยที่ไม่ต้องเป็นคำที่ตรงกันทั้งหมด

5.1 ผลการดำเนินงาน

1. สามารถสืบค้นวีดิโอคลิปข่าวภาษาอังกฤษผ่านทางเว็บแอปพลิเคชั่นได้และวิดีโอคลิปที่ได้จากการค้นหานั้นมีคำพูดปรากฏในวิดีโอคลิปตรงกับคำหรือข้อความที่ใช้ในการสืบค้น

2. สามารถสืบค้นวีดิโอคลิปข่าวภาษาอังกฤษผ่านทางเว็บแอปพลิเคชั่นได้ โดยการเปรียบเทียบเนื้อหาคำพูดระหว่างคลิปวิดีโอ

3. สามารถแสดงผลลัพธ์ในการจัดลำดับวีดิโอคลิปที่ค้นหาได้ตามลำดับความเหมือนระหว่างเนื้อหาคำพูดในวิดีโอคลิปกับคำหรือข้อความที่ใช้ในการสืบค้นค้นหา และเนื้อหาคำพูดระหว่างคลิปวิดีโอ

4. สามารถเข้ารับชมวิดีโอคลิปที่ค้นหาได้ผ่านเว็บแอปพลิเคชั่น

5.2 ปัญหาที่พบในการทำโครงงาน

1. กระบวนการในการถอดคำพูดในวิดีโอคลิปออกมาเป็นตัวหนังสือนั้นค่อนข้างที่ช้าเนื่องจากมีหลายขั้นตอนและความสามรถทางด้านฮาร์ดแวร์ของอุปกรณ์ที่ใช้ไม่สามารถตอบสนองต่อกระบวนการดังกล่าวได้รวดเร็วเพียงพอ

2. แหล่งข้อมูลความรู้ ตำรา ที่จะนำมาใช้ในการสร้างพัฒนาโครงงานค่อนข้างมีน้อยหาได้ลำบาก

5.3 ข้อเสนอแนะ

1. ควรใช้อุปกรณ์ทางด้านฮาร์ดแวร์ที่มีสมรรถนะที่สูงมากในกระบวนการถอดคำพูดในวิดีโอคลิปออกมาเป็นตัวหนังสือซึ่งจะช่วยในกระบวนการส่วนนี้ได้มาก

2. บางครั้งเนื้อหาคำพูดในวิดีโอคลิปอาจไม่ได้มีการกล่าวระบุถึง ชื่อบุคคล สถานที่ ทั้งที่เนื้อหาสาระในวิดีโอคลิปเกี่ยวข้องกับสิ่งเหล่านั้น เมื่อทำการสืบค้นด้วยคีย์เวิร์ดเหล่านั้นผลของการค้นหาจึงไม่พบวิดีโอคลิปเหล่านั้น การสร้างดัชนีให้นำหัวข้อชื่อวิดีโอคลิปมาทำเป็นดัชนีด้วยเพื่อเพิ่มประสิทธิภาพในการสืบค้น

3. ควรปรับปรุงเว็บแอปพลิเคชั่นให้สามารถบอกผลลัพธ์ที่แสดงออกมาว่าสามารถแสดงผลลัพธ์ออกมาได้กี่วิดีโอคลิปจากวิดีโอคลิปทั้งหมด จะทำให้สามารถคำนวณRecall และ Precision ได้ง่ายขึ้นกว่าเดิม

4. ในการสืบค้นควรปรับปรุงให้สามารถตรวจสอบความถูกต้องของคำหรือข้อความที่ใช้สืบค้นด้วย

5.4 แนวทางในการพัฒนาโครงงานวิจัย

1. พัฒนาเว็บแอปพลิเคชั่นให้ผู้ใช้สามารถอัพโหลดวิดีโอคลิปเข้าไปในฐานข้อมูลวิดีโอคลิปของเว็บได้ และสามารถนำวิดีโอคลิปไปสร้างดัชนีแล้วเก็บลงในฐานข้อมูลดัชนีได้โดยอัตโนมัติ

2. พัฒนากระบวนการถอดคำพูดออกมาเป็นตัวหนังสือให้มีเพียงขั้นตอนเดียว

3. พัฒนาการสร้างดัชนีให้นำหัวข้อชื่อวิดีโอคลิปมาทำเป็นดัชนีด้วย

เอกสารอ้างอิง

[1] "Speech Processing" [Online]. Available: 202.28.94.55/web/320491/2547/seminar/g31/doc/present.ppt

[2] "The Classic Vector Space Model Description, Advantages and Limitations of the Classic Vector Space Model" [Online]. Available: http://www.miislita.com/term-vector/term-vector-3.html

[3] "IR Models" [Online]. Available: http://www.csie.ndhu.edu.tw/~showyang /WebInfoSys2005f/02IRmodels.pdf

[4] "Vector Space Model (VSM) " [Online]. Available: http://trec.nist.gov/presentations/TREC8/intro/sld033.htm

[5] "Search Engine" [Online]. Available: http://catadmin.cattelecom.com/km/blog/kittichonm/category/search-engine/ทำ-search-engine-เองกันเถอะ/

[6] "Lucene" [Online]. Available: http://www.thaihealth.net/mirror/apache/lucene/java/

[7] ดวงพร เกี๋ยงคำ. อินไซท์ **Dreamweaver 8.** กรุงเทพฯ : โปรวิชั่น. 2549

ประวัติผู้เขียนโครงงาน

ชื่อ นายณัฐพงษ์ แก่นจันทร์

ภูมิลำเนา 112 ม.7 ต.ศรีถ้อย อ.แม่ใจ จ.พะเยา 56130

ประวัติการศึกษา

- จบการศึกษาระดับมัธยมศึกษาจากโรงเรียนพะเยาพิทยาคม
- ปัจจุบันกำลังศึกษาอยู่ในระดับปริญญาตรีชั้นปีที่ 4

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยนเรศวร

E-mail : prince_of_samurai@hotmail.com

ชื่อ นายธรรมนูญ มีสนุ่น

ภูมิลำเนา 83 ม.1 ต.นาสนุ่น อ.ศรีเทพ จ.เพชรบูรณ์ 67170

ประวัติการศึกษา

- จบการศึกษาระดับมัธยมศึกษาจากโรงเรียนนาสนุ่นวิทยาคม
- ปัจจุบันกำลังศึกษาอยู่ในระดับปริญญาตรีชั้นปีที่ 4

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยนเรศวร

E-mail : gleomuztozy@hotmail.com