

เครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์

Comment Analysis Tool for Online Mobile Phone Review



นายพงศ์พี วีระกุล รหัส 51371406

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

ปีการศึกษา 2555

ห้องสมุดคณะวิศวกรรมศาสตร์
วันที่รับ..... 20 ต.ค. 2558
เลขทะเบียน..... 16826799
เลขเรียกหนังสือ..... ผ.ร.
มหาวิทยาลัยเกษตรศาสตร์ W161


๑ 2555




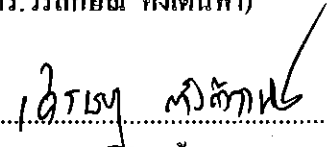
ใบรับรองโครงการวิศวกรรม

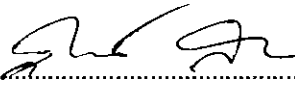
หัวข้อโครงการ เครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์
ผู้ดำเนินโครงการ นายพงษ์ศรีพี วีระกุล รหัส 51371406
อาจารย์ที่ปรึกษา อาจารย์ปรัชญภูมิ วรรณสาสน์
สาขาวิชา วิศวกรรมคอมพิวเตอร์
ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา 2555

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยบรจรัม อนุมัติให้โครงการฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะกรรมการสอบโครงการวิศวกรรม

.....ประธานกรรมการ
(อาจารย์ปรัชญภูมิ วรรณสาสน์)

.....กรรมการ
(ดร.วรลักษณ์ คงเด่นฟ้า)

.....กรรมการ
(อาจารย์เสริญฐา ตั้งคำวานิช)

.....กรรมการ
(อาจารย์ภาณุพงศ์ สอนคม)

หัวข้อโครงการ	เครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์
ผู้ดำเนินโครงการ	นายพงษ์พี วีระกุล รหัส 51371406
อาจารย์ที่ปรึกษา	อาจารย์รัฐภูมิ วรรณสาสน์
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2555

บทคัดย่อ

โทรศัพท์เคลื่อนที่ เป็นอุปกรณ์การสื่อสารที่ได้รับความนิยม เนื่องจากมีความกะทัดรัดสามารถพกพาได้สะดวกและมีฟังก์ชันการใช้งานที่สมบูรณ์ การนำข้อความการแสดงความคิดเห็นของผู้ใช้โทรศัพท์เคลื่อนที่ต่างๆ มาวิเคราะห์ข้อมูล โดยอาศัยหลักการการวิเคราะห์และสรุปผลจากข้อความแสดงความคิดเห็นเกี่ยวกับโทรศัพท์เคลื่อนที่ในระดับคุณลักษณะ ระบบนี้พัฒนาขึ้นเพื่อใช้ในการวิเคราะห์ข้อความแสดงความคิดเห็นเกี่ยวกับคุณลักษณะโทรศัพท์เคลื่อนที่รุ่นต่างๆ เพื่อง่ายต่อการตัดสินใจบริโภค โดยระบบจะรับข้อความแสดงความคิดเห็นภาษาไทย มาวิเคราะห์แล้วระบุว่า ข้อความนั้น อยู่ในกลุ่มคำเชิงบวก หรือเชิงลบ แล้วนำมาสรุปในเชิงสถิติและแสดงผลในรูปแบบกราฟิก

Project Title **Comment Analysis Tool for Online Mobile Phone Review.**
Name **Mr.Pongrapee Weerakool ID. 51371406**
Project Advisor **Mr.Rattapoom Waranusast**
Major **Computer Engineering.**
Department **Electrical and Computer Engineering.**
Academic Year **2012**

ABSTRACT

Mobile phones have become part of our daily lives due to their compactness and mobility. Opinions of mobile phone users in consumer forums can be used to help making a decision to buy a new mobile phone. This system was developed with the aim of analyzing opinions on mobile phones in discussion forums. The extracts keywords from opinion messages and labels them as positive, negative, or neutral words. Statistical scores are then computed from these categories and displayed as charts to help a phone buyer.

กิตติกรรมประกาศ

โครงการวิศวกรรมคอมพิวเตอร์สำเร็จได้ด้วยดี เนื่องจากความอนุเคราะห์จากท่านอาจารย์ที่ปรึกษาโครงการอาจารย์สิรภพ กษรัตน์ และอาจารย์รัฐภูมิ วรรณสาสน์ ที่กรุณาสละเวลาให้คำแนะนำในการทำงาน ตลอดจนตรวจสอบการทำงานพร้อมทั้งเสนอแนวทางการแก้ไขปัญหาลดระยะเวลาการทำโครงการ สุดท้ายนี้ขอขอบพระคุณคณะกรรมการ อาจารย์ทุกท่าน ครอบครัว และผู้ที่ยังไม่ได้เอ่ยนามที่คอยให้การสนับสนุนผู้ดำเนินโครงการ ให้สามารถทำโครงการนี้จนสำเร็จลุล่วงไปได้ด้วยดี

ขออำนาจคุณพระศรีรัตนตรัย และสิ่งศักดิ์สิทธิ์ทั้งหลายบันดาลให้ ท่านอาจารย์ ครอบครัว เพื่อนๆ มีสุขภาพแข็งแรง และเป็นแรงผลักดันให้คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร มีความก้าวหน้าต่อไป

ผู้ดำเนินโครงการวิศวกรรม

นายพงศ์พี วีระกุล

27 พฤษภาคม 2556



สารบัญ

หน้า

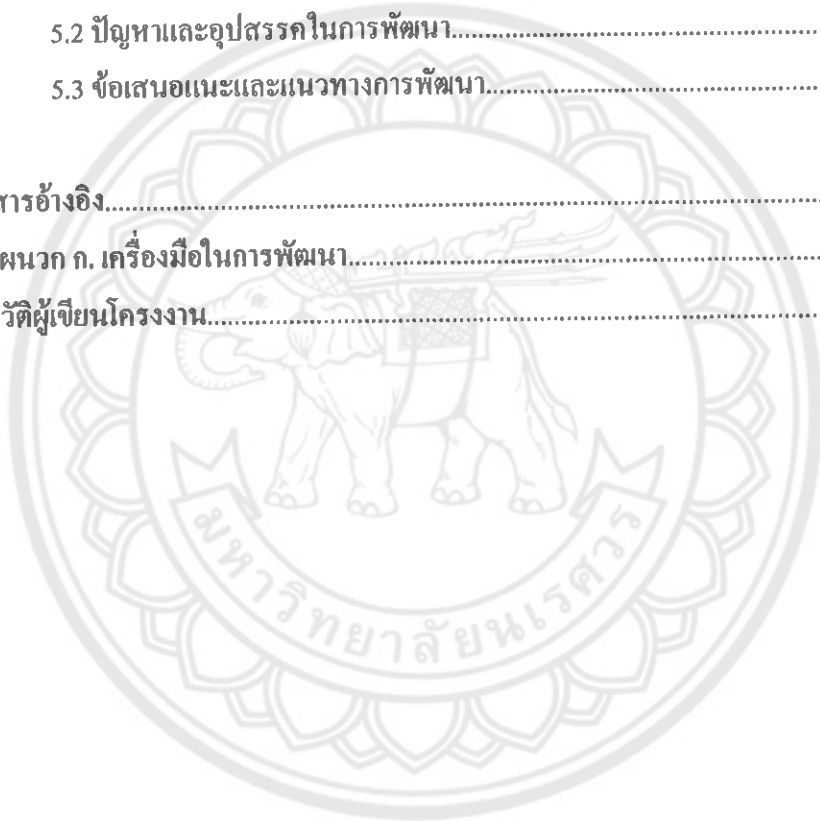
บทคัดย่อ.....	ก
ABSTRACT.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของโครงการ.....	1
1.2 วัตถุประสงค์ของโครงการ.....	2
1.3 ประโยชน์ที่คาดว่าจะได้รับ.....	2
1.4 ขอบเขตการทำโครงการ.....	2
1.5 ขั้นตอนการดำเนินงาน.....	3
1.6 แผนการดำเนินงาน.....	3
1.7 รายละเอียดงบประมาณตลอดโครงการ.....	4
บทที่ 2 หลักการและทฤษฎี.....	5
2.1 การวิเคราะห์ข้อความแสดงความคิดเห็น(Opinion Mining).....	5
2.2 การตัดคำสำหรับข้อความภาษาไทย(Thai Word Segmentation).....	5
2.3 การทำเหมืองข้อความ(Text Mining).....	9
2.3.1 Document Summarization.....	9
2.3.2 Document Classification.....	9
2.3.3 Document Clustering.....	9
2.4 เทคนิคจินตทัศน์(Information Visualization).....	10

สารบัญ(ต่อ)

	หน้า
บทที่ 3 วิธีดำเนินโครงการ.....	11
3.1 เนื้อเรื่องโดยย่อ.....	11
3.2 โครงสร้างของโปรแกรม.....	13
3.3 การวิเคราะห์คุณลักษณะ.....	13
3.3.1 หลักการวิเคราะห์ข้อความแสดงความคิดเห็นในระดับคุณลักษณะ.....	14
3.4 การทำงานของโปรแกรม.....	17
3.4.1 ส่วนต่างๆ ของโปรแกรม.....	18
3.4.1.1 การวิเคราะห์และการวิเคราะห์เปรียบเทียบคุณลักษณะ.....	19
3.4.1.2 การแสดงผลการวิเคราะห์.....	20
3.4.1.3 ดึงข้อมูลจากเว็บ.....	22
3.5 แผนภาพการทำงานของ โปรแกรม.....	23
3.6 การสร้างคลังข้อความ.....	24
3.6.1 Corpus Positive.....	24
3.6.2 Corpus Negative.....	25
3.6.3 Corpus Capability.....	25
3.6.4 Corpus Design.....	26
3.6.5 Corpus Price.....	26
3.7 หน้าที่การทำงานของแต่ละฟังก์ชันใน โปรแกรม.....	27
บทที่ 4 ผลการทดลองและวิเคราะห์.....	29
4.1 การทดสอบการวิเคราะห์คำคุณลักษณะ.....	29
4.1.1 การวิเคราะห์คุณลักษณะ.....	29
4.1.2 การสกัดความคิดเห็นโดยระบุข้อความที่มีและไม่มีในคลังข้อความ.....	30
4.1.3 การสกัดความคิดเห็นในระดับคุณลักษณะ.....	30
4.2 การประเมินผล.....	31
4.3 ตัวอย่างแบบประเมินข้อความแสดงความคิดเห็นของผู้ใช้.....	32

สารบัญ(ต่อ)

	หน้า
บทที่ 5 สรุปและข้อแนะนำ.....	33
5.1 ผลการทดลอง.....	33
5.1.1 การวิเคราะห์คุณลักษณะ.....	33
5.1.2 การสกัดความคิดเห็น โดยระบุข้อที่มีและไม่มีในคลังข้อความ.....	33
5.1.3 การสกัดความคิดเห็นระดับคุณลักษณะ.....	33
5.2 ปัญหาและอุปสรรคในการพัฒนา.....	33
5.3 ข้อเสนอแนะและแนวทางการพัฒนา.....	34
เอกสารอ้างอิง.....	35
ภาคผนวก ก. เครื่องมือในการพัฒนา.....	36
ประวัติผู้เขียนโครงการ.....	37



สารบัญตาราง

ตารางที่	หน้า
ตารางที่ 1.5 แผนการดำเนินงาน.....	3
ตารางที่ 4.1 แสดงการเปรียบเทียบความคิดเห็น โดยระบุข้อได้ ที่มีและไม่มีค่าในคลังข้อความ.....	30
ตารางที่ 4.2 การประเมินผลการสกัดข้อความแสดงความคิดเห็น.....	31
ตารางที่ 4.3 แบบประเมินความคิดเห็นของผู้ใช้โทรศัพท์เคลื่อนที่บอกว่า เป็นบวกหรือลบ.....	32



สารบัญรูป

รูปที่	หน้า
รูปที่ 2.1 องค์ประกอบของการแสดงความคิดเห็น.....	6
รูปที่ 2.2 แสดงตัวอย่างรูปแบบข้อความแสดงความคิดเห็น.....	7
รูปที่ 2.3 แสดงภาพตารางการเปรียบเทียบเทคนิคที่ประยุกต์ใช้ในการตัดคำ.....	8
รูปที่ 2.4 การทำเหมืองข้อความ (Text Mining).....	10
รูปที่ 2.5 ตัวอย่างเทคนิคจินตทัศน์ แบบ Pie Chart 2 มิติ.....	10
รูปที่ 3.1 หน้าเว็บรีวิวสินค้าในเว็บไซต์ siamphone.com.....	11
รูปที่ 3.2 ข้อความแสดงความคิดเห็นของผู้ใช้โทรศัพท์จากเว็บไซต์ siamphone.com.....	12
รูปที่ 3.3 แผนภาพการทำงานของการวิเคราะห์คำคุณลักษณะ.....	14
รูปที่ 3.4 แสดงภาพตัวอย่างการตัดคำ โดยใช้ ICUNET for CH.....	15
รูปที่ 3.5 แสดงภาพการหาคำคุณลักษณะ.....	15
รูปที่ 3.6 แสดงภาพการหาทัศนคติของคำ.....	16
รูปที่ 3.7 แสดงภาพตัวอย่างการวิเคราะห์ความคิดเห็นเชิงบวก.....	16
รูปที่ 3.8 แสดงภาพตัวอย่างการวิเคราะห์ความคิดเห็นเชิงลบ.....	16
รูปที่ 3.9 แสดงภาพตัวอย่างการสรุปการวิเคราะห์ความคิดเห็น.....	17
รูปที่ 3.10 แสดงภาพตัวอย่างการสรุปการวิเคราะห์ความคิดเห็นทั้งสองทัศนคติแบบข้อความ.....	17
รูปที่ 3.11 โปรแกรมวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์ที่เคลื่อนที่.....	18
รูปที่ 3.12 แสดงส่วนประกอบต่างๆของโปรแกรม.....	18
รูปที่ 3.13 แสดงส่วนประกอบการวิเคราะห์คุณลักษณะสินค้า.....	19
รูปที่ 3.14 แสดงส่วนประกอบการเปรียบเทียบคุณลักษณะสินค้า.....	19
รูปที่ 3.15 แสดงผลการวิเคราะห์แบบกราฟ.....	20
รูปที่ 3.16 แสดงผลการวิเคราะห์แบบข้อความ.....	21
รูปที่ 3.17 แสดงผลการวิเคราะห์ Overall แบบกราฟ.....	21
รูปที่ 3.18 แสดงผลการวิเคราะห์แบบเปรียบเทียบสินค้าแบบกราฟ.....	22
รูปที่ 3.19 การดึงข้อมูลจากเว็บ.....	22
รูปที่ 3.20 แผนภาพการทำงานของโปรแกรม.....	23
รูปที่ 3.21 แสดงตัวอย่าง Corpus Positive.....	24
รูปที่ 3.22 แสดงตัวอย่าง Corpus Negative.....	25

สารบัญรูป (ต่อ)

รูปที่	หน้า
รูปที่ 3.23 แสดงตัวอย่าง Corpus Capability.....	25
รูปที่ 3.24 แสดงตัวอย่าง Corpus Design.....	26
รูปที่ 3.25 แสดงตัวอย่าง Corpus Price.....	26



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของโครงการงาน

อุตสาหกรรมเทคโนโลยีการสื่อสารทั่วโลกได้มีการพัฒนาอย่างรวดเร็ว โทรศัพท์เคลื่อนที่หรือที่เรียกกันว่าโทรศัพท์มือถือนั้น โทรศัพท์มือถือ จัดเป็นอุปกรณ์การสื่อสารที่ได้รับความนิยมมากที่สุด ในปัจจุบัน ในวันนี้ ผู้ใช้โทรศัพท์มือถือมีปริมาณมากมาย แตกต่างจากในอดีตอย่างเห็นได้ชัด อย่างไรก็ตาม แม้ว่าเราจะพบอยู่เสมอๆ ว่า โทรศัพท์มือถือรุ่นต่างๆ ทั้งจากฝั่งยุโรป หรือฝั่งเอเชีย จะถูกผลิตขึ้น และวางจำหน่ายรุ่นแล้ว รุ่นเล่า พร้อมทั้งมีการนำเอาคุณสมบัติต่างๆ เพิ่มเข้ามาบนตัวโทรศัพท์อย่างไม่หยุดยั้ง แต่ในอีกด้านหนึ่ง นั่นคือด้านของเทคโนโลยีการส่งข้อมูลบนโทรศัพท์มือถือ หลายๆ คนกลับไม่ให้ความสนใจเท่าที่ควร ซึ่งอันที่จริงแล้ว การเข้ามาสัมผัสและรู้จักกับเทคโนโลยีประเภทนี้ นับเป็นสิ่งที่น่าสนใจไม่น้อย

เนื่องจากปัจจุบันการเข้าถึงข้อมูลข่าวสารทางอินเทอร์เน็ตของผู้บริโภคเป็นไปได้ง่าย ดังนั้นเจ้าของสินค้าและบริการมีความจำเป็นอย่างยิ่ง ในการรับรู้ถึงอารมณ์ และความรู้สึกของลูกค้า ต่อสินค้าและบริการนั้นๆ ของเจ้าของสินค้าและบริการในประเภทต่างๆ ว่าลูกค้ามีการตอบสนองอย่างไรต่อสินค้าและบริการ ในทางที่เป็นความคิดเห็นด้านที่ดีต่อสินค้า หรือด้านที่ไม่ดีต่อสินค้าและบริการของเจ้าของสินค้านั้นๆ โดยโครงการนี้จะเป็นการใช้โปรแกรมวิเคราะห์ข้อความแสดงความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์ โดยผ่านทางข้อความของลูกค้าที่ได้แสดงความคิดเห็น ในที่ต่างๆบนโลกออนไลน์ เปรียบเสมือนการประเมินความพึงพอใจของลูกค้า (Customer Satisfaction) ต่อสินค้า และการให้บริการของบริษัทจากบล็อก (Blog) กระทั่ง แสดงความคิดเห็น (Web Board) รวมถึงสังคมออนไลน์ต่างๆ (Social network) ข้อความที่แสดงออกมาเหล่านั้นนั้น มักจะถ่ายทอดถึงอารมณ์ และความรู้สึกของลูกค้าอีกด้วย ซึ่งในการตลาดมักจะเรียกข้อมูลประเภทนี้ว่า เป็นข้อมูลเชิงจิตวิทยา (Psychological Data) ซึ่งข้อมูลที่ได้มาเหล่านั้นจะนำมาเพื่อไปใช้ในการปรับปรุงสินค้าและบริการของเจ้าของสินค้านั้นๆ ได้ เพราะถ้าหากนำข้อความที่ได้ใช้โปรแกรมวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์จะทำให้เกิดประโยชน์สูงสุดทั้งลูกค้าที่ใช้สินค้าและบริการ รวมถึงเจ้าของสินค้าและบริการ โดยที่ลูกค้าจะได้ประโยชน์จากการใช้สินค้าและบริการเมื่อแสดงความคิดเห็นในสื่อออนไลน์ ทำให้เจ้าของสินค้านำข้อมูลที่ได้ไปปรับปรุงสินค้าและบริการ เพื่อให้เกิด

ประโยชน์สูงสุดแก่ลูกค้า ส่วนเจ้าของสินค้าก็สามารถเพิ่มยอดขายได้ ถ้าหากปรับปรุงสินค้า และบริการของตนเองให้เป็นที่ยอมรับของบุคคลทั่วไป

****หมายเหตุ** โครงการฉบับนี้คำว่า โทรศัพท์เคลื่อนที่และโทรศัพท์มือถือ ใช้ในความหมายเดียวกัน

1.2 วัตถุประสงค์ของโครงการ

1. เพื่อเป็นเครื่องมือในการประเมินความพึงพอใจของผู้ใช้ (User) ต่อสินค้าโทรศัพท์เคลื่อนที่
2. เพื่อเป็นการลดระยะเวลา และประหยัดจำนวนคน ในการมาประเมินความพึงพอใจของผู้ใช้โทรศัพท์เคลื่อนที่ ที่แสดงต่อสินค้า ผ่านทางสื่อออนไลน์
3. เพื่อนำข้อดี และ ข้อเสียจากการแสดงความคิดเห็นของลูกค้า ผ่านข้อความทางสื่อออนไลน์มาใช้พิจารณาในการเลือกซื้อเลือกใช้นสินค้าได้

1.3 ขอบข่ายของโครงการ

1. โปรแกรมสามารถวิเคราะห์ความคิดเห็นต่อสินค้าในเชิงบวกและเชิงลบได้ โดยนำข้อความแสดงความคิดเห็นมาวิเคราะห์คุณลักษณะ
2. โปรแกรมสามารถนำเสนอผลการวิเคราะห์คุณลักษณะออกมาในลักษณะของกราฟได้
3. โปรแกรมสามารถวิเคราะห์คุณลักษณะเปรียบเทียบสินค้าได้

1.4 ขั้นตอนการดำเนินงาน

1. ศึกษา ค้นคว้า ข้อมูล และกำหนดขอบเขตของงาน
2. เขียนโปรแกรมทำเป็นซอฟต์แวร์
3. ทำการทดสอบซอฟต์แวร์ และทดลองวิเคราะห์อารมณ์ของสินค้า
4. ทดสอบประสิทธิภาพซอฟต์แวร์โดยผู้ใช้
5. ทำการสรุป รวบรวม ข้อมูลทั้งหมดของโครงการ
6. จัดทำรายงาน

1.5 แผนการดำเนินงาน

ตารางที่ 1.1 แผนการดำเนินงาน

กิจกรรม	เดือน - ปี											
	2554		2555									
	พฤศจิกายน	ธันวาคม	มกราคม	กุมภาพันธ์	มีนาคม	เมษายน	พฤษภาคม	มิถุนายน	กรกฎาคม	สิงหาคม	กันยายน	
1. ศึกษาค้นคว้าหาข้อมูลและกำหนดขอบเขตของงาน	←→											
2. เขียนโปรแกรมทำเป็นซอฟต์แวร์				←→								
3. ทำการทดสอบซอฟต์แวร์ทดลองวิเคราะห์ค่าธรรมเนียมของสินค้า						←→						
4. ทดสอบประสิทธิภาพซอฟต์แวร์โดยผู้ใช้						←→						
5. ทำการสรุป รวบรวมข้อมูลทั้งหมดของโครงการ							←→					
6. จัดทำรายงาน									←→			

1.6 ผลที่คาดว่าจะได้รับ

1. ได้โปรแกรมที่สามารถนำไปใช้วิเคราะห์สินค้า โทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์
2. สามารถนำโปรแกรมไปประยุกต์ใช้กับสินค้าชนิดอื่นๆ ได้
3. ลดขั้นตอนการสำรวจความพึงพอใจของลูกค้าต่อสินค้าโทรศัพท์เคลื่อนที่ได้

1.7 งบประมาณที่ใช้

1. ซื้อหนังสือเพื่อศึกษา	500	บาท
2. ถ่ายเอกสาร และค่าจัดพิมพ์	<u>500</u>	บาท
	รวม	<u>1,000</u> บาท



บทที่ 2

หลักการและทฤษฎีเบื้องต้น

เครื่องมือวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์ (Comment Analysis Tools for Online Mobile Phone Review) มีหลักการและทฤษฎีที่เกี่ยวข้องดังต่อไปนี้

2.1 การวิเคราะห์ข้อความแสดงความคิดเห็น (Opinion Mining)

ในปัจจุบันเว็บไซต์โดยทั่วไปมีลักษณะที่ส่งเสริมการแบ่งปันข้อมูลการพัฒนาในด้านแนวความคิด และการออกแบบ รวมถึงการร่วมสร้างข้อมูลบนโลกของอินเทอร์เน็ต แทนที่แนวคิดของเว็บไซต์ในแบบเก่าที่เป็นลักษณะการให้บริการอ่านเพียงอย่างเดียว ดังนั้นข้อมูลบนเว็บไซต์จึงแบ่งได้ออกเป็น 2 รูปแบบคือ ข้อเท็จจริง และข้อคิดเห็น [1]

1. ข้อเท็จจริง เป็นข้อมูลจริงที่สามารถพิสูจน์ได้ เช่น พระอาทิตย์เป็นศูนย์กลางระบบสุริยะ

2. ข้อคิดเห็น คือการแสดงทัศนคติที่ผู้เขียนมีต่อสิ่งๆหนึ่งหรือหัวข้อๆหนึ่ง โดยแบ่งออกเป็น

ความคิดเห็นแบบตรง (Direct opinion) สามารถบอกทัศนคติของผู้เขียนหรือผู้พูดได้อย่างชัดเจน เช่น “รูปที่ถ่ายจากกล้องตัวนี้ คุณภาพดี”

ความคิดเห็นเชิงเปรียบเทียบ (Comparisons) คือการให้ความเห็นโดยนำสิ่งที่ต้องการแสดงความคิดเห็น ไปเปรียบเทียบกับอีกสิ่งหนึ่ง เช่น “กล้อง A แพงกว่ากล้อง B”

องค์ประกอบของการแสดงความคิดเห็น ประกอบด้วย 3 ส่วนดังปรากฏในรูปที่ 2.1 คือ

1. ผู้แสดงความคิดเห็น (Opinion Holder)

2. วัตถุหรือหัวข้อที่นำมาแสดงความคิดเห็น (Object)

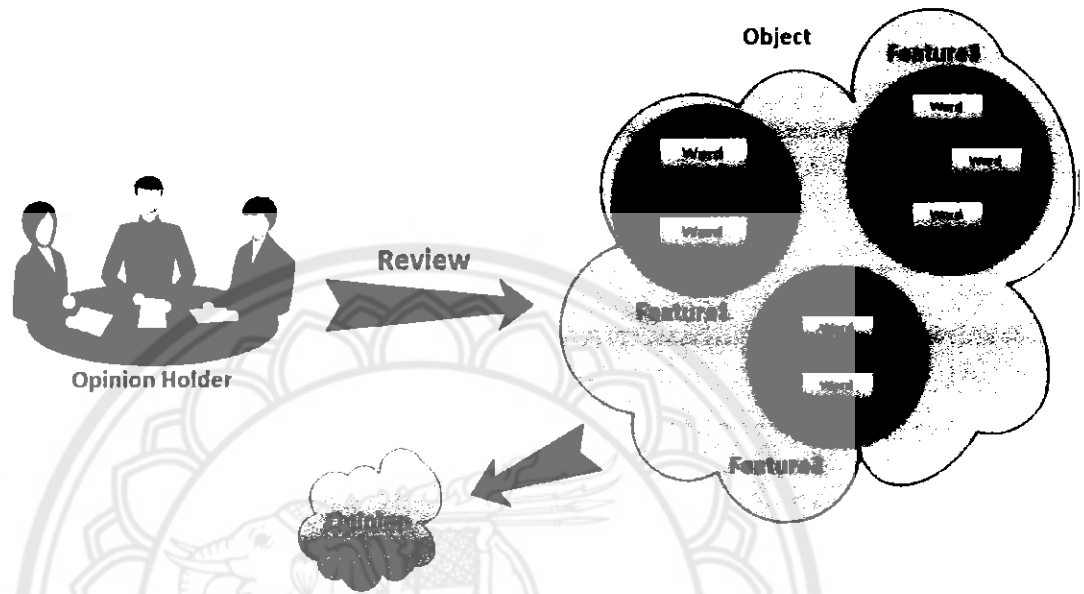
3. ความคิดเห็นหรือทัศนคติของผู้แสดงความคิดเห็นที่มีต่อหัวข้อที่นำมาแสดงความคิดเห็น (Opinion) ซึ่งแต่ละหัวข้อที่ถูกนำมาแสดงความคิดเห็นจะประกอบไปด้วยคุณลักษณะต่างๆ (feature) สามารถเขียนในรูปแบบของเซตข้อมูลได้ดังนี้

$$\text{Object} = \{F_1, F_2, F_3, \dots, F_n\}$$

โดย F_n คือ คุณลักษณะที่วัตถุหรือหัวข้อที่นำมาแสดงความคิดเห็นมี ซึ่งสามารถแทนได้ด้วยเซตของข้อมูล

$$F_n = \{W_1, W_2, W_3, \dots, W_n\}$$

โดย W_n คือ คำที่มีความหมายเหมือนกัน (word synonyms) มีความหมายที่สามารถแทนที่กับคุณลักษณะนั้นๆ ได้



รูปที่ 2.1 องค์ประกอบของการแสดงความคิดเห็น [2]

เมื่อผู้ให้ความคิดเห็นแสดงความคิดเห็นเกี่ยวกับวัตถุหรือหัวข้อบนคุณลักษณะ ที่เป็นส่วนหนึ่งของวัตถุหรือหัวข้อนั้นๆ ก็จะถูกนำไปตรวจสอบกับคำที่มีความหมายเหมือนในแต่ละคุณลักษณะแล้วนำมาตัดสินว่าเป็นการแสดงความคิดเห็นในแง่ใด ซึ่งอาจจะเป็นจะเป็นขั้วบวก (positive) ขั้วลบ (negative) หรือเป็นกลาง (neutral)

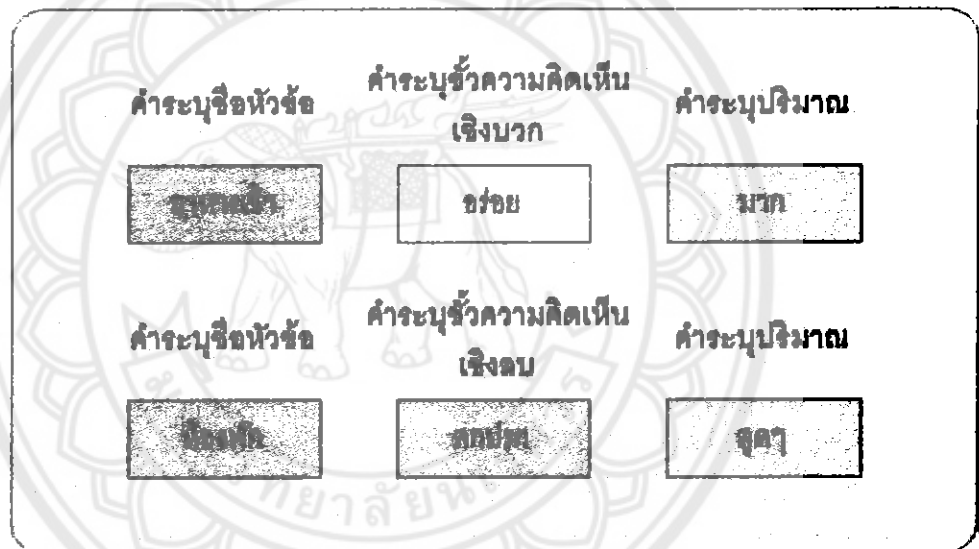
ซึ่งระบบวิเคราะห์ข้อความแสดงความคิดเห็น คือ กระบวนการอัตโนมัติเพื่อใช้ตรวจสอบทัศนคติของผู้พูดหรือผู้เขียนในหัวข้อเรื่องใดเรื่องหนึ่ง โดยอาศัยการสอนคอมพิวเตอร์ให้พิจารณาอารมณ์ความรู้สึกด้วยการประมวลผลด้วยภาษาธรรมชาติ (Natural Language Processing, NLP)

แบ่งการทำงานออกเป็น 3 ระดับคือ

1. ระดับเอกสาร (Document Level) เป็นการวิเคราะห์ข้อความแสดงความคิดเห็นในแบบหยาบ เนื่องจากเป็นการนำข้อความแสดงความคิดเห็นทั้งหมดจากเอกสาร มาสรุปแยกข้อความความคิดเห็นเป็นขั้วบวก ขั้วลบ หรือเป็นกลาง

2. ระดับประโยค (Sentence Level) เป็นการวิเคราะห์ข้อความแสดงความคิดเห็น โดยแยกข้อความที่เป็นข้อความแสดงความคิดเห็นออกจากข้อความที่เป็นข้อเท็จจริงในระดับที่เป็นประโยค แล้วนำมาแยกข้อความความคิดเห็นเป็นขั้วบวก ขั้วลบ หรือเป็นกลาง

3. ระดับคุณลักษณะ (Feature Level) เป็นการวิเคราะห์ข้อความแสดงความคิดเห็น โดยแยกลักษณะที่สนใจของวัตถุหรือหัวข้อที่ถูกแสดงความคิดเห็นออกมาก่อน แล้วจึงนำมาแบ่งข้อความความคิดเห็นเป็นขั้วบวก ขั้วลบ หรือเป็นกลาง และนำมาจัดกลุ่มเข้ากับคำที่มีความหมายเหมือนในแต่ละคุณลักษณะซึ่งระบบวิเคราะห์ข้อความ จะนำข้อความแสดงความคิดเห็นของลูกค้ำมาวิเคราะห์ข้อความแสดงความคิดเห็นในระดับคุณลักษณะ แล้วนำข้อมูลที่ได้มาประมวลผลเพื่อให้ผู้ใช้เข้าใจได้ง่ายขึ้น



รูปที่ 2.2 แสดงตัวอย่างรูปแบบข้อความแสดงความคิดเห็น [2]

2.2 การตัดคำสำหรับข้อความภาษาไทย (Thai Word Segmentation)

การตัดคำเป็นพื้นฐานที่สำคัญอย่างยิ่งในการวิเคราะห์และประมวลผลภาษารวมทั้งการพัฒนา ระบบที่เกี่ยวข้องต่างๆ เช่น การกำหนดหน้าที่ของคำ (Part-of-Speech Tagging) การแปลภาษาโดย เครื่อง (Machine Translation) การจดจำและสังเคราะห์เสียงพูด (Speech Recognition/Synthesis) , การ ค้นคืนสารสนเทศและเสิร์จเอนจิน (Information Retrieval & Search Engine) การกรองสารสนเทศ (Information Filtering) การทำเหมืองข้อความ (Text Mining) การสร้างความรู้และโครงสร้างความรู้ เชิงความหมาย (Knowledge Base and Ontology)

ในการประมวลผลภาษาธรรมชาติ ภาษาไทยถูกจัดอยู่ในประเภทของภาษาที่ไม่ตัดคำ (Unsegmented language)

- ไม่มีการใช้ตัวอักษรใดๆ ในการบ่งบอกขอบเขตของคำอย่างชัดเจน
- ต้องอาศัยเทคนิคการตัดคำในการบอกขอบเขตของคำ ซึ่งมีข้อจำกัดคือ ขึ้นอยู่กับภาษานั้นๆ : พจนานุกรมและไวยากรณ์ของภาษา ยังไม่มีเทคนิคที่ให้ความถูกต้องได้ 100% ซึ่งเกิดจาก 2 ปัญหาหลักคือ

คำไม่รู้จัก (Unknow words) เช่น Context-dependent : มากกว่า, ปีกว่า, ตากลม Context-independent : มากล้น, การสอบ

คำกำกวม (Ambiguity) เช่น Explicit : ไมโครซอฟท์, ไอโฟน Hidden : สมชาย, การบินไทย, แม่สาย Mixed : สีนามิ, นาช่า, ดีแทค

เทคนิคที่ประยุกต์ใช้สามารถแบ่งออกเป็น 3 วิธีหลัก ได้แก่

- การใช้กฎไวยากรณ์ทางภาษา (Rule-based)
- การอ้างอิงคำจากพจนานุกรม (Dictionary-based)
- การสร้าง โมเดลเรียนรู้จากฐานข้อความขนาดใหญ่ (Machine Learning or Corpus based)

Approach	Advantage	Disadvantage
Rule-based	<ul style="list-style-type: none"> - Relatively fast - No need to store dictionary in memory 	<ul style="list-style-type: none"> - Difficult to form a "perfect" rule set grammatically - Only segment syllable level - Cannot effectively handle unknown words, especially transliterated words
Dictionary-based	<ul style="list-style-type: none"> - Yields relatively high accuracy - Solve unknown word problem by updating dictionary - Allow segmentation of compound words (more semantic focused) 	<ul style="list-style-type: none"> - Require large memory to store dictionary - Difficult to maintain dictionary for the case of new unknown words
Corpus-based	<ul style="list-style-type: none"> - No need to store dictionary in memory - Partially solve unknown word problem by learning from a large text corpus - Automatically form a rule set by learning from a large text corpus - Can have different set of learning rules for each particular subject 	<ul style="list-style-type: none"> - Only segment syllable level - Require a large corpus and learning time plus a good learning algorithm

รูปที่ 2.3 แสดงภาพตารางการเปรียบเทียบเทคนิคที่ประยุกต์ใช้ในการตัดคำ [3]

2.3 การทำเหมืองข้อความ (Text Mining)

Text Mining [4] คือขบวนการทำงานที่เรียกว่า Process ที่สกัดข้อมูล (Extract data) จากฐานข้อมูลขนาดใหญ่ (Large Textual Information) เพื่อให้ได้สารสนเทศ (Usefull Textual Information) โดยข้อมูลที่ถูกลำมา Mining เป็นข้อมูลที่มีลักษณะเป็น Text data sets การนำ Text Mining เข้ามาช่วยงานทางด้านเอกสาร สามารถแบ่งเป็นรูปแบบต่างๆ ได้แก่

2.3.1 Document Summarization

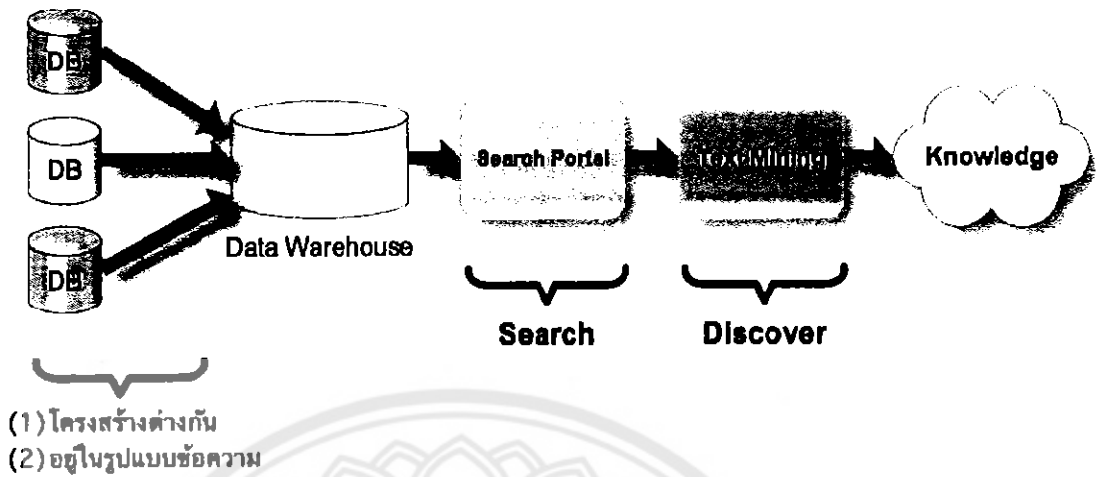
ในบางครั้งเราอาจต้องการทราบเพียงข้อมูลที่เป็นสาระสำคัญของเอกสารเท่านั้น การนำเทคนิคนี้มาใช้ ก็เพื่อทำการลดรายละเอียดของข้อมูลที่ไม่จำเป็น เพื่อให้คงไว้เฉพาะส่วนที่เป็นสาระสำคัญซึ่งยังคงทำให้เอกสารนั้นมีความสมบูรณ์ในเชิงเนื้อหาเป็นต้น ตัวอย่างงานที่เห็นได้ชัดเจนคือ Google นั้นเอง เคยสังเกตไหมว่า เมื่อเรา search ข้อมูล Google จะแสดงบางส่วนของเนื้อหาของแต่ละผลลัพธ์ เพื่อให้เราเห็นภาพรวมของ website นั้นๆ ก่อนที่เราจะคลิกเข้าไปดู

2.3.2 Document Classification

เป็นเทคนิคช่วยในจำแนกประเภทเอกสาร ทั้งนี้เราต้องทราบก่อนแล้วว่าต้องการจำแนกเอกสารออกเป็นกี่ประเภท (Class) ดังนั้นการใช้เทคนิคนี้ จำเป็นต้องทำการสอนระบบ (train model) ให้รู้จำรูปแบบของเอกสารในแต่ละ class ก่อน ถ้าเราเคยสังเกตตอนสมัคร E-mail ตาม free e-mail ต่างๆ นั้น จะมีหน้าต่างเงื่อนไขการใช้บริการ ถ้าเราอ่านเงื่อนไขทั้งหมดจะพบว่า 1 ในหลายๆ ข้อนั้น จะมีเงื่อนไขของการยินยอมให้ทางผู้ให้บริการ E-mail สามารถอ่านเนื้อหาภายในได้ ทั้งนี้ส่วนหนึ่งก็เพื่อใช้ในการกรอง พวก Spam mail ออกจาก E-mail ปกตินั่นเอง อีกตัวอย่างหนึ่งของการทำเทคนิค Document Classification ไปใช้ คือใช้ในการจำแนกข้อมูลที่มีการ Post อยู่ใน Social network เพื่อใช้ในการวิเคราะห์หรือดูแนวโน้มในเรื่องต่างๆ ได้อีกด้วย

2.3.3 Document Clustering

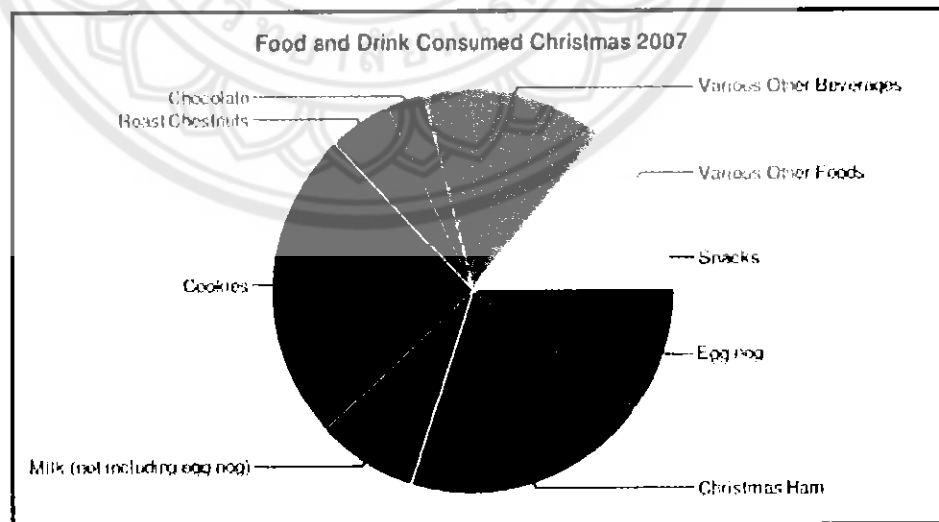
เป็นเทคนิคที่ช่วยในการจัดกลุ่มเอกสารต่างๆ ไป ที่เราไม่ทราบว่าเป็นเอกสารประเภทใด โดยพิจารณาจากเนื้อหาของเอกสารที่คล้ายคลึงกันให้อยู่กลุ่มเดียวกัน ทั้งนี้จะเห็นว่า เราสามารถนำไปใช้งานด้าน Search engine เพื่อทำการจัดกลุ่มข้อมูลที่มีอยู่มากมาย ออกเป็นกลุ่มย่อยๆ หรือ Categories เมื่อ user ระบุ key word หรือ คำค้น เข้ามา ระบบ Search engine จะทำการค้นข้อมูลใน Category เป้าหมายก่อน เพื่อลดเวลาในการ Search แทนที่จะต้องทำการค้นหาข้อมูลจากฐานข้อมูลทั้งก้อน



รูปที่ 2.4 การทำเหมืองข้อความ (Text Mining) [4]

2.4 เทคนิคจินตทัศน์ (Information Visualization)

เทคนิคจินตทัศน์ [5] คือ การนำเสนอข้อมูลที่เข้าใจยาก เช่นตัวเลข หรือ ตัวหนังสือ ที่มีปริมาณ มากๆ มาแสดงให้อยู่ในรูปแบบที่ผู้อ่านเข้าใจได้ง่ายกว่าเดิม เช่น กราฟ แผนภาพแบบต่างๆ โดยเลือกใช้ วิธีที่จะนำเสนอให้ผู้อ่านเข้าใจได้ง่ายที่สุด ซึ่งขึ้นกับข้อมูลที่จะนำมานำเสนอว่าจะจะเป็นข้อมูลที่อยู่ใน ลักษณะใด การนำเสนอข้อมูลผ่านเทคนิคจินตทัศน์ทำได้หลายรูปแบบ แต่ที่นิยมใช้กันมากนั้นคือการ นำเสนอในรูปแบบที่เป็น 2 มิติ เช่น Bar Chart, Pie Chart เป็นต้น



รูปที่ 2.5 ตัวอย่างเทคนิคจินตทัศน์ แบบ Pie Chart 2 มิติ [5]

บทที่ 3 วิธีดำเนินโครงการ

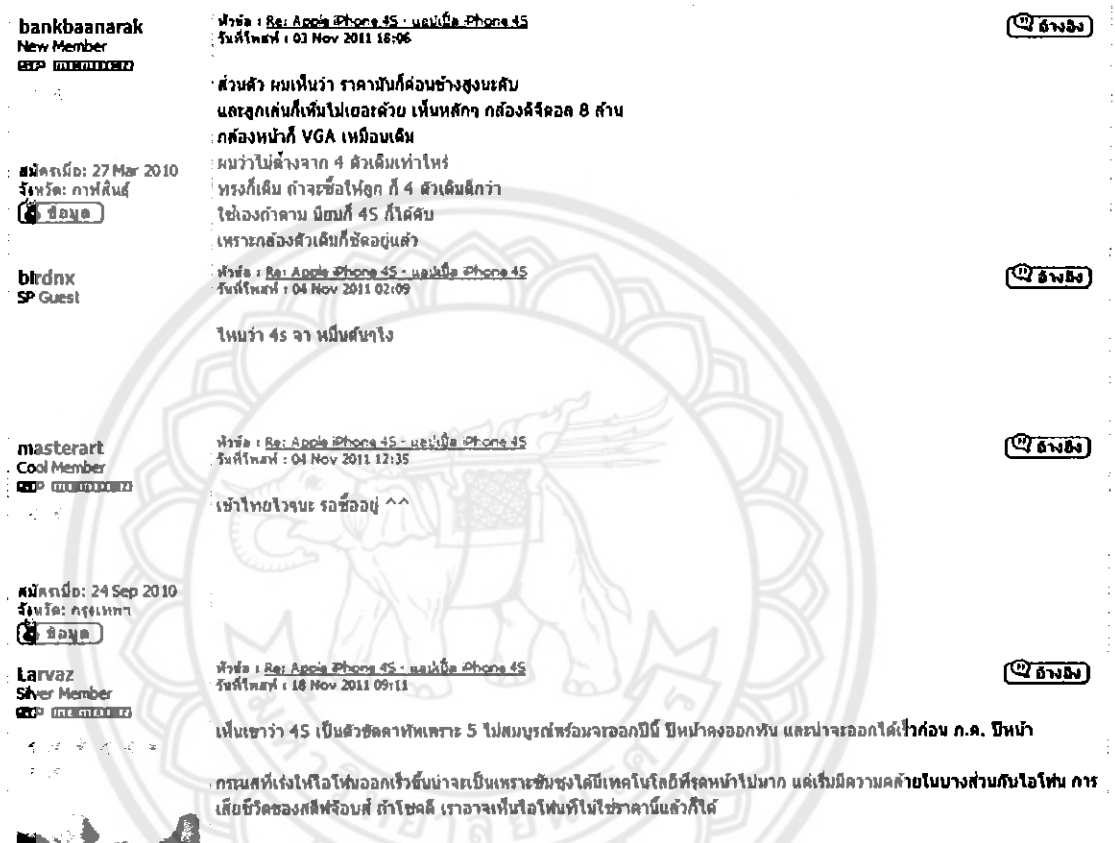
3.1 เนื้อเรื่องโดยย่อ

เมื่อนายเอ ต้องการซื้อโทรศัพท์มือถือเพื่อให้ทันกับเทคโนโลยีในสมัยนี้ จึงต้องการข้อมูลใน ส่วนของ Capability (ความสามารถและประสิทธิภาพทั่วไป) Design (ความสวยงาม) และ Price (ราคา) ซึ่งนายเอต้องการ โทรศัพท์มือถือที่มีความสามารถที่ดี ดีไซน์หรู และราคาไม่แพงมากด้วย นายเอจึงเข้าไปดูข้อความแสดงความคิดเห็นภายในเว็บไซต์ siamphone.com ดังแสดงในรูปที่ 3.1



รูปที่ 3.1 หน้าเว็บรีวิวสินค้าในเว็บไซต์ siamphone.com

ภายในเว็บไซต์มีข้อมูลของโทรศัพท์มือถือที่นำมาเสนอมากมาย และข้อความแสดงความคิดเห็นต่างๆ ออกเป็นจำนวนมากดังแสดงในรูปที่ 3.2 ซึ่งในรูปนั้นก็เป็นส่วนหนึ่งของข้อความแสดงความคิดเห็นเท่านั้น



รูปที่ 3.2 ข้อความแสดงความคิดเห็นของผู้ใช้โทรศัพท์จากเว็บไซต์ siamphone.com

เนื่องจากภายในเว็บไซต์ siamphone.com มีการรีวิวโทรศัพท์มือถือหลายรุ่นหลายยี่ห้อและการแสดงความคิดเห็นมีความคิดเห็นเป็นจำนวนมาก ดังนั้นเครื่องมือวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์มือถือ สามารถสรุปข้อมูลความเห็นเกี่ยวกับโทรศัพท์มือถือให้นายเอได้ ช่วยให้ประหยัดเวลาในการหาข้อมูลและช่วยนายเอสามารถตัดสินใจซื้อได้ง่ายขึ้น

3.2 โครงสร้างของโปรแกรม

โปรแกรมเครื่องมือวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์ได้ออกแบบไว้ประกอบด้วย 3 ส่วนใหญ่ๆ ดังนี้

1. คึงข้อมูลจากเว็บ ไซด์ เป็นส่วนแรกที่น่ามาใช้ในการวิเคราะห์คือจะต้องทำการคึงข้อมูลจากเว็บ ไซด์ (www.siamphone.com) ซึ่งเป็นข้อมูลจริงๆ ที่มีผู้ใช้และผู้ที่มีความสนใจโทรศัพท์มือถือ มาแสดงความคิดเห็นเกี่ยวกับโทรศัพท์เคลื่อนที่เป็นอย่างมาก

2. การวิเคราะห์คุณลักษณะในความเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ ในส่วนนี้เมื่อได้รับข้อมูลที่นำมาจากเว็บ ไซด์แล้ว ก็จะมีการวิเคราะห์ความเห็นซึ่งจะแบ่งออกเป็น 2 ลักษณะคือ ความคิดเห็นเชิงบวก และความคิดเห็นเชิงลบ

3. การแสดงผลลัพธ์ออกมาเป็นกราฟ ในส่วนนี้เมื่อมีการวิเคราะห์คุณลักษณะ เสร็จแล้ว โปรแกรมก็จะนำผลลัพธ์ที่ได้มาสร้างกราฟในลักษณะของ กราฟวงกลม 2 มิติ (Pie chart) เพื่อผู้ใช้เข้าใจได้ง่ายขึ้น

3.3 การวิเคราะห์คุณลักษณะ

การวิเคราะห์ข้อความแสดงความความคิดเห็น (Opinion Mining) คือ กระบวนการอัตโนมัติเพื่อใช้ตรวจสอบทัศนคติของผู้พูดหรือผู้เขียนในหัวข้อเรื่องใดเรื่องหนึ่ง โดยอาศัยการสอนคอมพิวเตอร์ให้พิจารณาอารมณ์ความรู้สึกด้วยการประมวลผลด้วยภาษารธรรมชาติ (Natural Language Processing, NLP) แบ่งการทำงานออกเป็น 3 ระดับดังนี้

1. ระดับเอกสาร (Document Level)
2. ระดับประโยค (Sentence Level)
3. ระดับคุณลักษณะ (Feature Level)

ซึ่งเครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่นี้ จะนำข้อความแสดงความความคิดเห็นของผู้ใช้งาน โทรศัพท์เคลื่อนที่มาวิเคราะห์ข้อความแสดงความความคิดเห็นในระดับคุณลักษณะ แล้วนำข้อมูลที่ได้มาประมวลผลเพื่อให้ผู้ใช้เข้าใจง่ายขึ้น

3.3.1 หลักการวิเคราะห์ข้อความแสดงความคิดเห็นในระดับคุณลักษณะ



รูปที่ 3.3 แผนภาพการทำงานของกรวิเคราะห์คำคุณลักษณะ

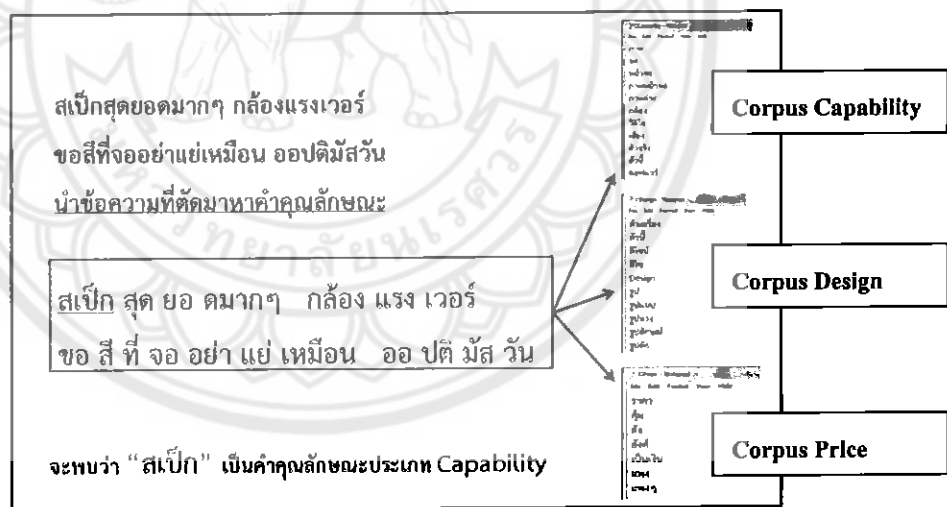
1. ทำการตัดคำให้กับแต่ละข้อความ การตัดคำในโปรแกรมนี้เราได้ใช้ทางด้าน Library ของ ICU4NET ซึ่งเป็นเครื่องมือที่พัฒนามาเพื่อใช้สำหรับ C#

Input: แพคเกจคอมพิวเตอร์จาก apple ทำให้พวกเราได้ใช้เทคโนโลยีใหม่ๆ
แล้วจะมีใครรุ่นใหม่มาใช้ IPHONE4S
ช่วงในเหมือนเดิมก็ได้ แต่ปรับภายนอกให้ดูว่า ดีและรุ่นใหม่
ถือไปไหน ดู IPHONE 4S น้อยค

Output: แพคเกจคอมพิวเตอร์จาก apple ทำให้พวกเราได้ใช้เทคโนโลยีใหม่ๆ
แล้วจะมีใครรุ่นใหม่มาใช้ IPHONE4S
ช่วงในเหมือนเดิมก็ได้ แต่ปรับภายนอกให้ดูว่า ดีและรุ่นใหม่
ถือไปไหน ดู IPHONE 4S น้อยค

รูปที่ 3.4 แสดงภาพตัวอย่างการตัดคำโดยใช้ ICU4NET for C#

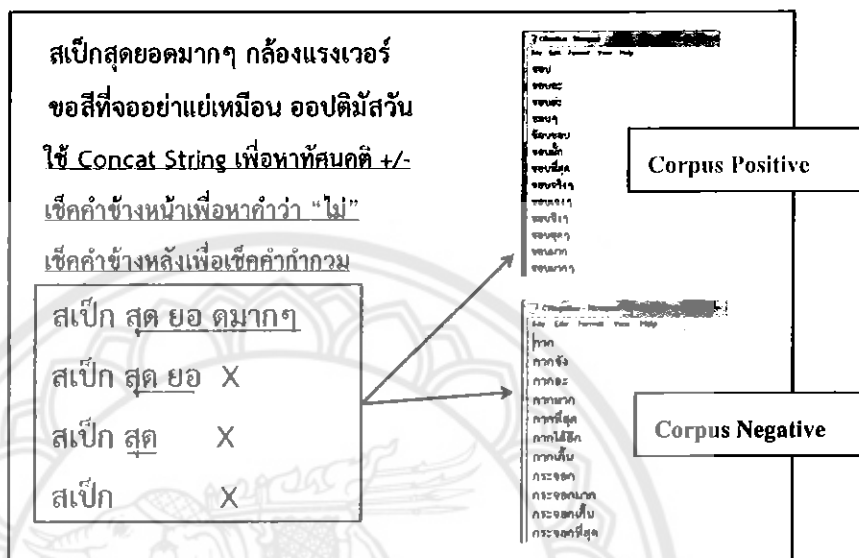
2. หากำคุณลักษณะที่ตรงกับคุณลักษณะที่ต้องการ 3 คุณลักษณะ เมื่อทำการตัดคำเสร็จแล้ว โปรแกรมจะทำการหาคำคุณลักษณะ เมื่อเจอคำคุณลักษณะที่ตรงกับคลังข้อความ ให้เก็บคำคุณลักษณะ นั้นไว้



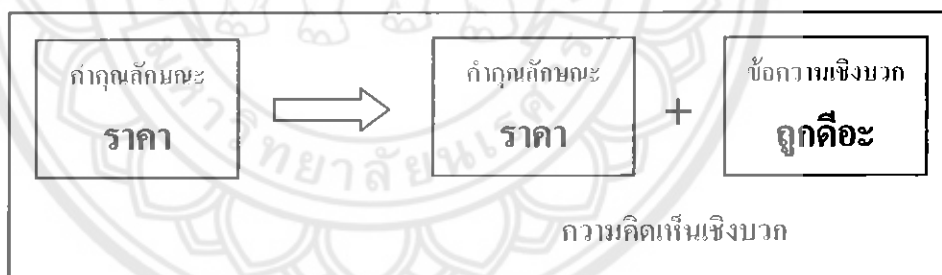
รูปที่ 3.5 แสดงภาพการหาคำคุณลักษณะ

3. นำคำคุณลักษณะที่ได้ในแต่ละข้อความแสดงความคิดเห็นมาตรวจสอบว่ามีที่สนคตัวอย่างไร (เชิงบวกและเชิงลบ)

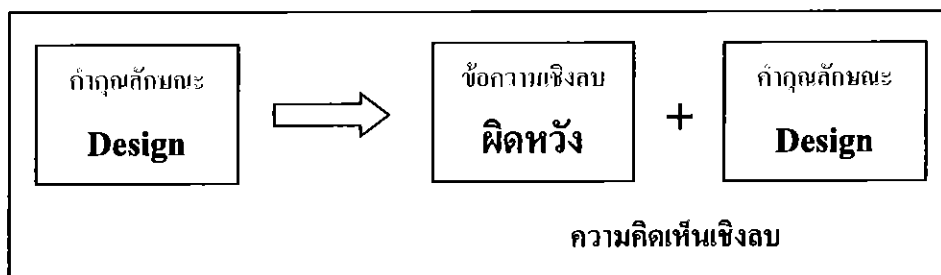
การวิเคราะห์ข้อความแสดงความคิดเห็น โดยแยกลักษณะที่สนใจของวัตถุหรือหัวข้อที่ถูกแสดงความคิดเห็นออกมาก่อน แล้วจึงนำมาแบ่งข่าว และนำมาจัดกลุ่มเข้ากับความหมายเหมือนในแต่ละคุณลักษณะเช่น



รูปที่ 3.6 แสดงภาพการหาที่สนคติของคำ

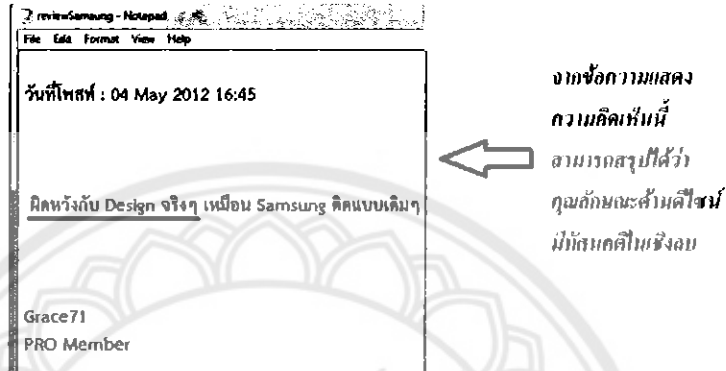


รูปที่ 3.7 แสดงภาพตัวอย่างการวิเคราะห์ความคิดเห็นเชิงบวก



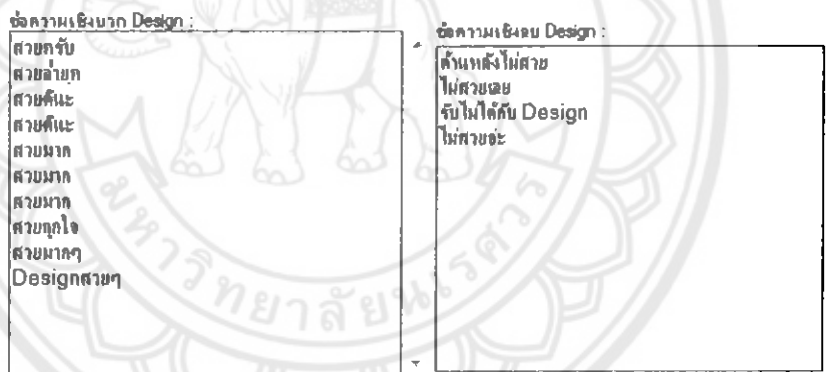
รูปที่ 3.8 แสดงภาพตัวอย่างการวิเคราะห์ความคิดเห็นเชิงลบ

4. สรุปทัศนคติให้กับข้อความแสดงความคิดเห็นนั้นซึ่งหากข้อความนั้นตรงกับคลังข้อความความคิดเห็นเชิงบวก ข้อความแสดงความคิดเห็นนั้นก็จะเป็นบวก ในทางกลับกันหากข้อความนั้นตรงกับคลังข้อความความคิดเห็นเชิงลบ ข้อความแสดงความคิดเห็นนั้นก็จะกลายเป็นลบ



จากข้อความแสดง
ความคิดเห็น
สามารถสรุปได้ว่า
คุณลักษณะส่วนดีใน
มีทัศนคติในเชิงลบ

รูปที่ 3.9 แสดงภาพตัวอย่างการสรุปการวิเคราะห์ความคิดเห็น



รูปที่ 3.10 แสดงภาพตัวอย่างการสรุปการวิเคราะห์ความคิดเห็นทั้งสองทัศนคติแบบข้อความ

3.4 การทำงานของโปรแกรม

หลังจากที่เราได้ศึกษาค้นคว้าหาทฤษฎีและหลักการต่างๆ มาใช้ในการวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ ขั้นตอนต่อไปคือการเขียนโปรแกรมรองรับการทำงานตามแผนที่ได้กำหนดไว้ โดยหน้าตาของโปรแกรมเป็นดังรูปที่ 3.11

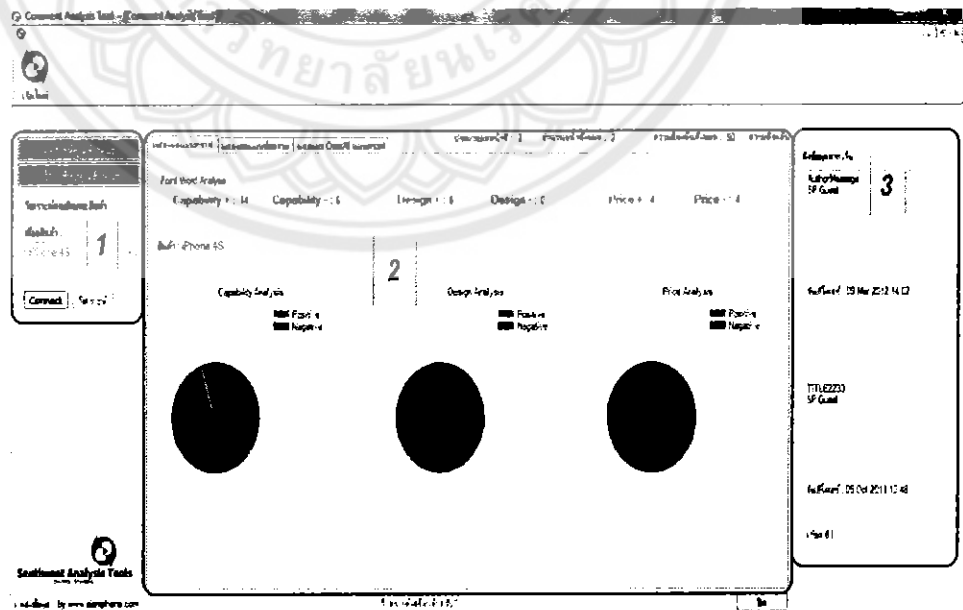


รูปที่ 3.11 โปรแกรมวิเคราะห์ความเห็นต่อสินค้าโทรศัพท์ที่เคลื่อนที่

3.4.1 ส่วนต่างๆของโปรแกรม

โปรแกรมนี้มีองค์ประกอบทั้งหมด 3 ส่วนใหญ่ๆ ดังนี้

1. การวิเคราะห์และการวิเคราะห์เปรียบเทียบคุณลักษณะ
2. การแสดงผลการวิเคราะห์
3. คึงข้อมูลจากเว็บ



รูปที่ 3.12 แสดงส่วนประกอบต่างๆของโปรแกรม

3.4.1.1 การวิเคราะห์และการวิเคราะห์เปรียบเทียบคุณลักษณะ

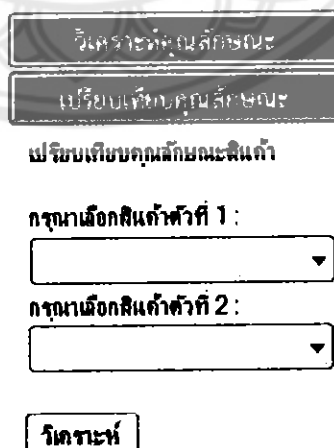
ในส่วนนี้การวิเคราะห์และการวิเคราะห์เปรียบเทียบคุณลักษณะจะแบ่งเป็น 2 ส่วนคือ

1. วิเคราะห์คุณลักษณะ มีหน้าที่วิเคราะห์ข้อความแสดงความคิดเห็น โดยในส่วนนี้จะสามารถวิเคราะห์เพียงหนึ่งข้อเท่านั้น ประกอบด้วยปุ่ม Connect และปุ่มวิเคราะห์ ปุ่ม Connect ทำหน้าที่ดึงข้อมูลจากเว็บไซต์ siamphone.com เมื่อทำการ Connect เสร็จแล้วปุ่มวิเคราะห์ถึงจะสามารถกดได้ เพื่อวิเคราะห์ความคิดเห็นของสินค้าตามที่ใช้เลือก ดังรูปที่ 3.13

2. เปรียบเทียบคุณลักษณะ มีหน้าที่วิเคราะห์เปรียบเทียบความคิดเห็น โดยสามารถเลือกสินค้าที่จะเปรียบเทียบได้สองข้อเท่านั้น ดังรูปที่ 3.14



รูปที่ 3.13 แสดงส่วนประกอบกรวิเคราะห์คุณลักษณะสินค้า



รูปที่ 3.14 แสดงส่วนประกอบกรเปรียบเทียบคุณลักษณะสินค้า

3.4.1.2 การแสดงผลการวิเคราะห์

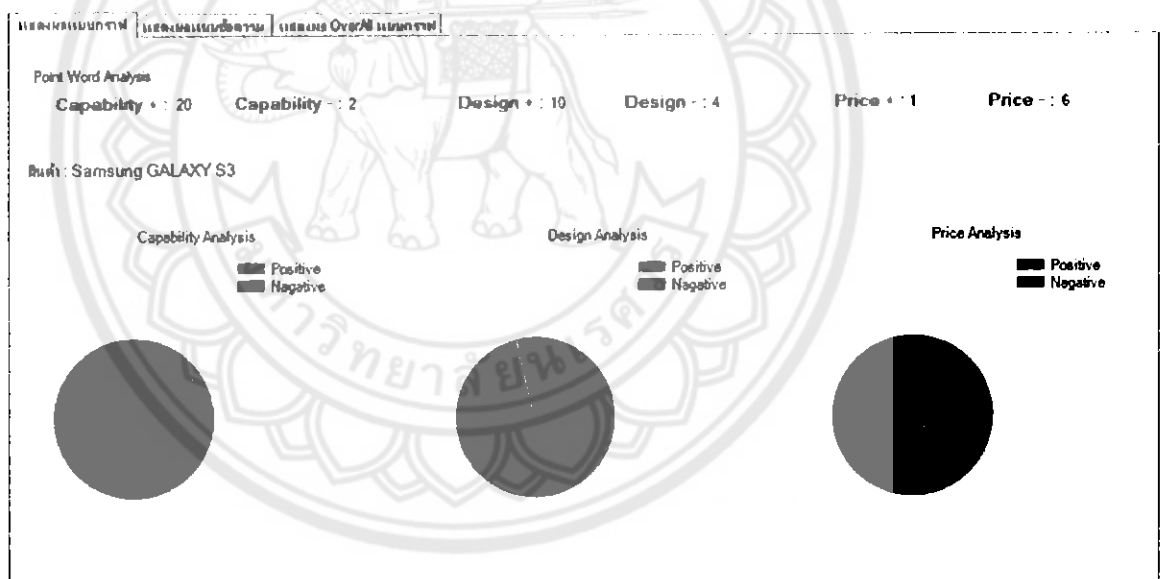
ส่วนนี้เป็นส่วนของการแสดงผลโดยจะแบ่งออกเป็น 3 ส่วนคือ

1. การแสดงผลแบบกราฟ จะประกอบด้วยการแสดงผลของ Point Word Analysis ซึ่งเป็นคะแนนและกราฟวงกลม 2 มิติของผลการวิเคราะห์ความเห็นของทั้ง 3 คุณลักษณะ ดังรูปที่ 3.15

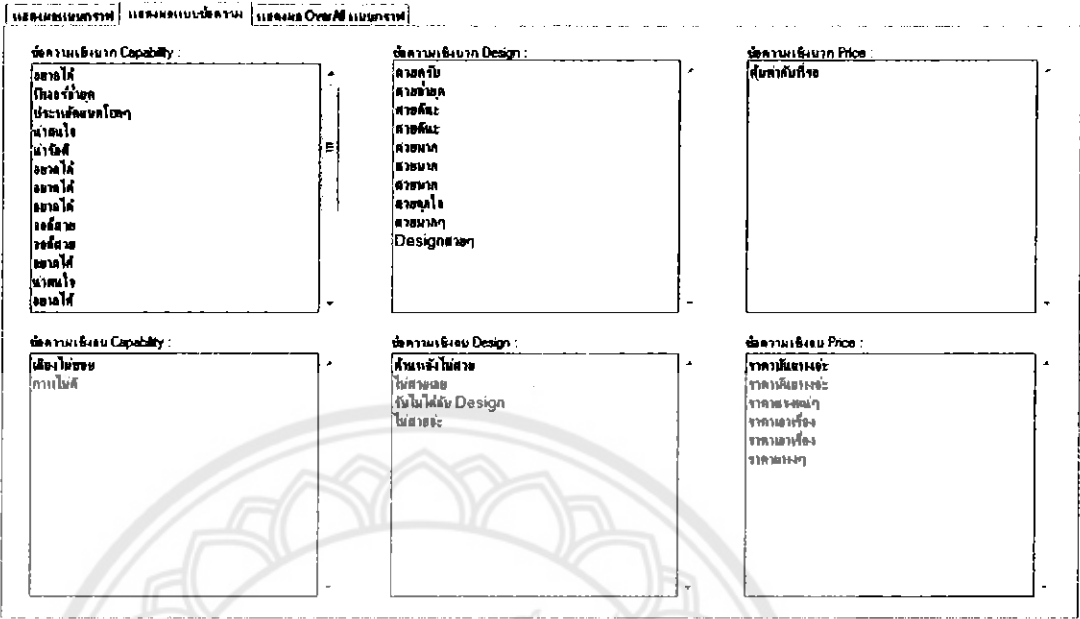
2. การแสดงผลแบบข้อความ จะประกอบด้วยข้อความแสดงความคิดเห็นด้านบวกและด้านลบของทั้ง 3 คุณลักษณะ ดังรูปที่ 3.16

3. การแสดงผลแบบกราฟ Overall จะประกอบด้วย Point Comment Overall และกราฟวงกลมแบบ 2 มิติ ซึ่งแบบเป็น 3 ข้อคือ ความคิดเห็นด้านบวก, ความคิดเห็นด้านลบ, ความคิดเห็นเป็นกลาง ดังรูปที่ 3.17

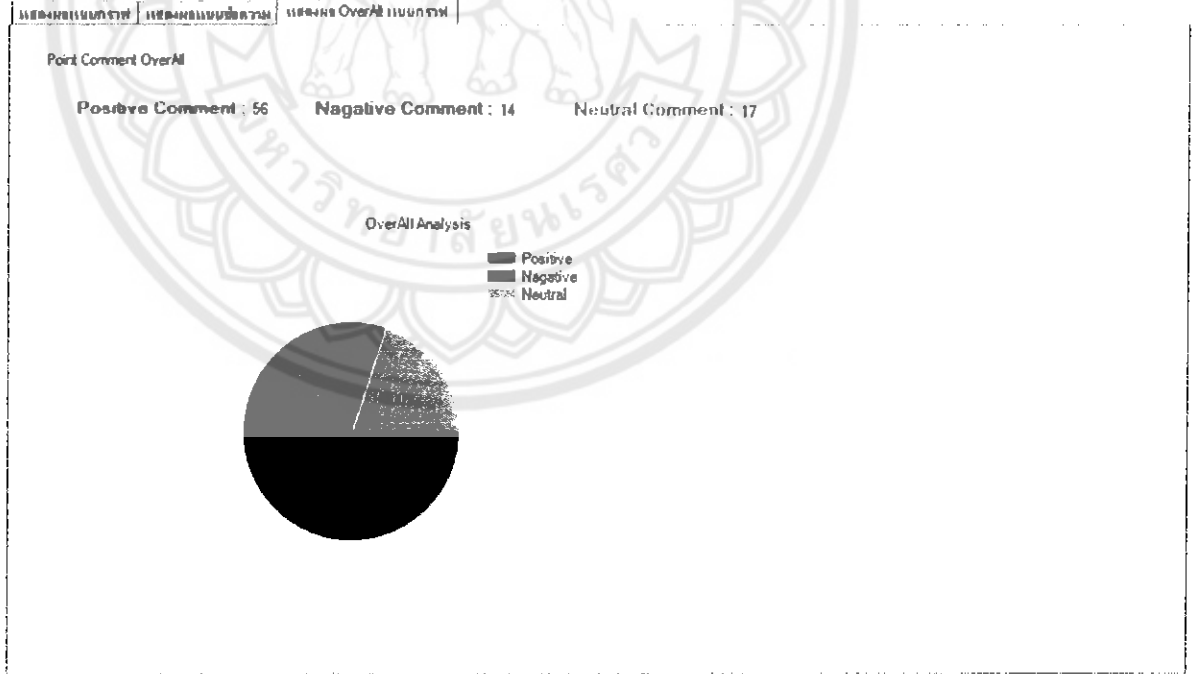
4. การแสดงผลกราฟแบบเปรียบเทียบ แสดงถึงกราฟวงกลม 2 มิติ ของสินค้า 2 ยี่ห้อ ดังรูปที่ 3.18



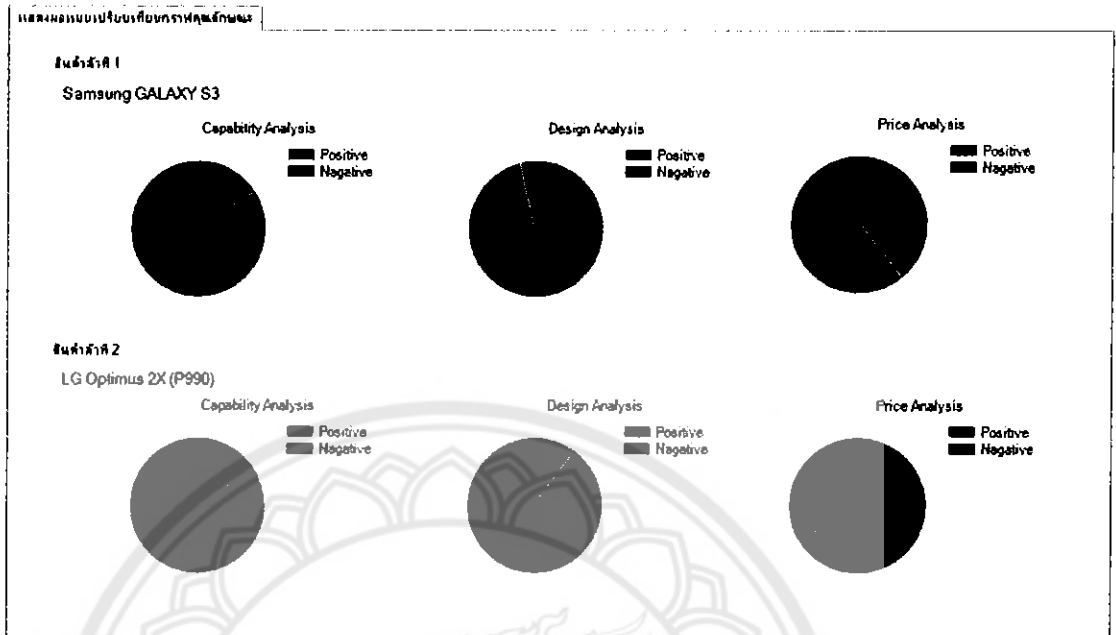
รูปที่ 3.15 แสดงผลการวิเคราะห์แบบกราฟ



รูปที่ 3.16 แสดงผลการวิเคราะห์แบบข้อความ



รูปที่ 3.17 แสดงผลการวิเคราะห์ Overall แบบกราฟ



รูปที่ 3.18 แสดงผลการวิเคราะห์แบบเปรียบเทียบสินค้าแบบกราฟ

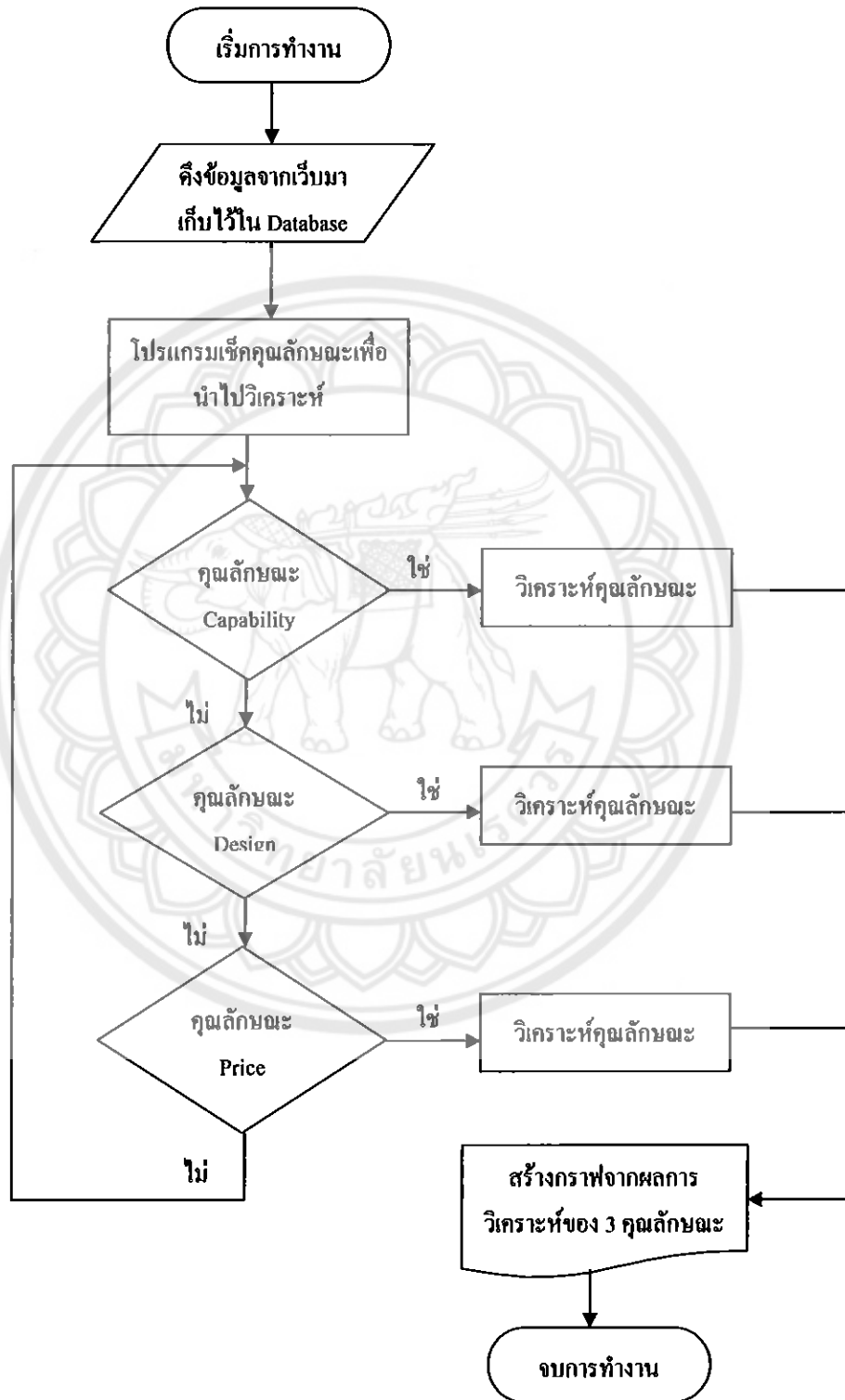
3.4.1.3 ดึงข้อมูลจากเว็บ

ในส่วนของการดึงข้อมูลจากเว็บ มีหน้าที่ดึงข้อมูลจากเว็บไซต์โดยโปรแกรมจะทำการนำข้อความแสดงความคิดเห็นจากสินค้าที่ผู้ใช้เลือกทั้งหมดมาเก็บไว้ใน Database (Text file) เพื่อนำไปวิเคราะห์ต่อไป ดังรูปที่ 3.19



รูปที่ 3.19 การดึงข้อมูลจากเว็บ

3.5 แผนภาพการทำงานของโปรแกรม



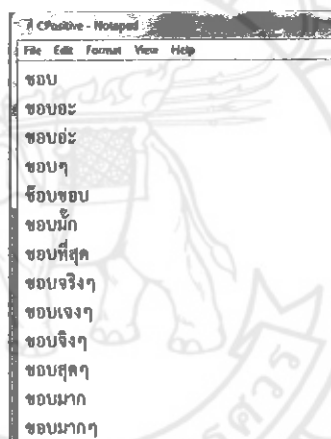
รูปที่ 3.20 แผนภาพการทำงานของโปรแกรม

3.6 การสร้างคลังข้อความ

การสร้างคลังข้อความมีความสำคัญต่อการวิเคราะห์ข้อความแสดงความคิดเห็นอย่างมากเพื่อนำมาใช้วิเคราะห์ความเห็น เนื่องจากต้องเก็บรวบรวมความคิดเห็น หรือคำเฉพาะในการแสดงความคิดเห็นต่อโทรศัพท์มือถือซึ่งปัจจุบันการแสดงความคิดเห็นนั้นมีการใช้คำทับศัพท์ ศัพท์ภาษาใหม่ที่เรียกกันในหมู่วัยรุ่น หรือที่เรียกกันว่า ศัพท์แสลง ของผู้แสดงความคิดเห็น ซึ่งโปรแกรมนี้จะประกอบด้วยคลังข้อความดังต่อไปนี้

3.6.1 Corpus Positive

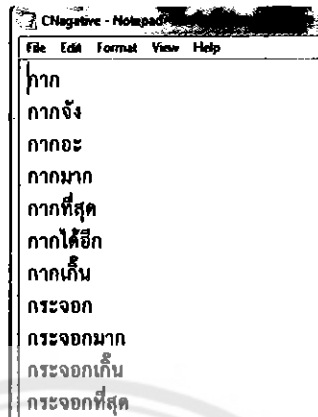
คลังข้อความนี้จะเก็บในส่วนของ ข้อความแสดงความคิดเห็นเชิงบวก ซึ่งมีการรวบรวมคำที่มีความหมายสื่อถึงสินค้าในด้านที่ดี ตัวอย่าง Corpus Positive ดังรูปที่ 3.21



รูปที่ 3.21 แสดงตัวอย่าง Corpus Positive

3.6.2 Corpus Negative

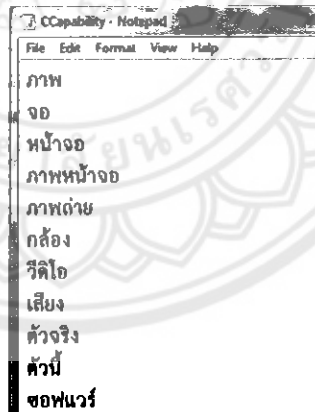
คลังข้อความนี้จะเก็บในส่วนของ ข้อความแสดงความคิดเห็นเชิงลบ ซึ่งมีการรวบรวมคำที่มีความหมายสื่อถึงสินค้าในด้านที่ไม่ดี ตัวอย่าง Corpus Negative ดังรูปที่ 3.22



รูปที่ 3.22 แสดงตัวอย่าง Corpus Negative

3.6.3 Corpus Capability

คลังข้อความนี้จะเก็บในส่วนของ คำที่มีความเกี่ยวข้องกับคุณลักษณะของ Capability (ความสามารถทั่วไป) ซึ่งเป็นหนึ่งในคุณลักษณะหนึ่งที่โปรแกรมใช้วิเคราะห์ ตัวอย่าง Corpus Capability ดังรูปที่ 3.23



รูปที่ 3.23 แสดงตัวอย่าง Corpus Capability

3.6.4 Corpus Design

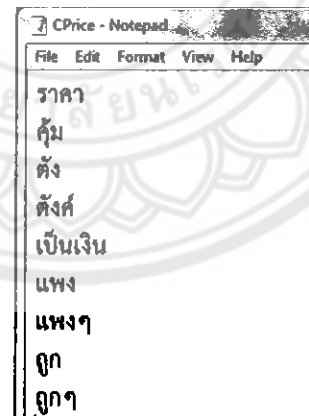
คลังข้อความนี้จะเก็บในส่วนของ คำที่มีความเกี่ยวข้องกับคุณลักษณะของ Design (ด้านความสวยงาม) ซึ่งเป็นหนึ่งในคุณลักษณะหนึ่งที่โปรแกรมใช้วิเคราะห์ ตัวอย่าง Corpus Design ดังรูปที่ 3.24



รูปที่ 3.24 แสดงตัวอย่าง Corpus Design

3.6.5 Corpus Price

คลังข้อความนี้จะเก็บในส่วนของ คำที่มีความเกี่ยวข้องกับคุณลักษณะของ Price (ด้านราคา) ซึ่งเป็นหนึ่งในคุณลักษณะหนึ่งที่โปรแกรมใช้วิเคราะห์ ตัวอย่าง Corpus Price ดังรูปที่ 3.25



รูปที่ 3.25 แสดงตัวอย่าง Corpus Price

3.7 หน้าี่การทำงานของแต่ละฟังก์ชันในโปรแกรม

โปรแกรมเครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์นี้มี Class 1 Class ชื่อ Form1 และ FormWordBreak โดยมี Method ต่างๆ ดังนี้

1. FormWordBreak ทำงานเมื่อรันโปรแกรมขึ้นมาครั้งแรก ทำหน้าที่เรียกใช้งาน Method พื้นฐาน 1 Method คือ InitializeComponent()
2. InitializeComponent ทำหน้าที่เรียก Component ของระบบที่ต้องใช้เมื่อรันโปรแกรม
3. จาก Form1 toolStripButton2_Click ทำหน้าที่เรียก FormWordBreak เมื่อคลิกปุ่มที่เมนูบาร์
4. btnF1_Click ทำหน้าที่เรียก groupBox ของฟังก์ชันวิเคราะห์คุณลักษณะ ซึ่งประกอบด้วย Combobox เพื่อเลือกชื่อของโทรศัพท์มือถือ, ปุ่มConnect, ปุ่มวิเคราะห์
5. btnF2_Click ทำหน้าที่เรียก groupBox ของฟังก์ชันเปรียบเทียบคุณลักษณะ ซึ่งประกอบด้วย 2 Combobox เพื่อเลือกชื่อของโทรศัพท์มือถือได้ 2 ชื่อ และปุ่มวิเคราะห์
6. button3_Click ทำหน้าที่ Connect กับเว็บไซต์เพื่อทำการดึงข้อมูลจากเว็บ โดยการเรียก ฟังก์ชัน WebBrowser1_DocumentCompleted, btnsplit_Click, button5_Click มาใช้งาน
7. WebBrowser1_DocumentCompleted ทำหน้าที่เปิดเว็บไซต์เพื่อทำการเก็บข้อมูลในส่วนของ Document ทั้งหมดไว้
8. btnsplit_Click ทำหน้าที่รับค่าจาก WebBrowser1_DocumentCompleted เพื่อนำมาคัดแยกเอา เฉพาะเนื้อหาในส่วนของข้อความแสดงความคิดเห็น
9. button5_Click ทำหน้าที่บันทึกข้อมูลที่คัดแยกแล้วจากฟังก์ชัน btnsplit_Click เก็บไว้ใน Textfile
10. btnUpdate_Click ทำหน้าที่เมื่อคลิกปุ่มวิเคราะห์แล้ว ฟังก์ชันนี้จะทำการวิเคราะห์ข้อความ แสดงความคิดเห็น โดยแบ่งออกเป็น 3 คุณลักษณะได้แก่ Capability, Design, Price
11. using BreakIterator ทำหน้าที่ตัดคำภาษาไทยแล้วบันทึกไว้ใน Text file โดยใช้ library ของ ICU4NET ในการตัดคำ
12. DicCapability.Contains() ทำหน้าที่ ค้นหา keyword ของคำคุณลักษณะของ Capability
13. DicDesign.Contains() ทำหน้าที่ ค้นหา keyword ของคำคุณลักษณะของ Design
14. DicPrice.Contains() ทำหน้าที่ ค้นหา keyword ของคำคุณลักษณะของ Price
15. btnAnalysis1_Click ทำหน้าที่ นำไปใช้ในฟังก์ชันเปรียบเทียบเป็นสินค้าตัวที่ 1
16. btnAnalysis2_Click ทำหน้าที่ นำไปใช้ในฟังก์ชันเปรียบเทียบเป็นสินค้าตัวที่ 2

17. btnChoose_Click ทำหน้าที่ เรียกฟังก์ชัน btnAnalysis1_Click และ btnAnalysis2_Click เพื่อทำการเปรียบเทียบสินค้าทั้ง 2 ชนิด
18. CalculatePointOverAll() ทำหน้าที่ คำนวณคะแนนของแต่ละคุณลักษณะโดยรับค่ามาทั้งในเชิงบวกและเชิงลบในหนึ่งความคิดเห็น
19. CalPointWordCapability() ทำหน้าที่รับค่าคะแนนของคุณลักษณะ Capability ทั้งเชิงบวกและเชิงลบมาคำนวณหาค่าเฉลี่ยเพื่อนำไปสร้างกราฟ
20. CalPointWordDesign() ทำหน้าที่รับค่าคะแนนของคุณลักษณะ Design ทั้งเชิงบวกและเชิงลบมาคำนวณหาค่าเฉลี่ยเพื่อนำไปสร้างกราฟ
21. CalPointWordPrice() ทำหน้าที่รับค่าคะแนนของคุณลักษณะ Price ทั้งเชิงบวกและเชิงลบมาคำนวณหาค่าเฉลี่ยเพื่อนำไปสร้างกราฟ
22. Piechart() – Piechart10() ทำหน้าที่รับค่าของฟังก์ชัน CalPointWordCapability(), CalPointWordDesign(), CalPointWordPrice() เพื่อมาสร้างเป็นกราฟวงกลม 2 มิติ
23. ShowWordAnalysis() ทำหน้าที่อ่านค่าของความเห็นที่ผ่านการวิเคราะห์แล้วทั้ง 3 คุณลักษณะมาแสดงใน textbox
24. button2_Click ทำหน้าที่ปิดหน้าต่าง โปรแกรมในส่วนของการวิเคราะห์ทั้งหมด
25. HashSet<string> ทำหน้าที่รับค่า string มาเก็บไว้มีลักษณะเหมือน (Set)
26. ArrayList() ทำหน้าที่รับค่าข้อมูลมาเก็บไว้
27. StreamReader() ทำหน้าที่อ่านข้อมูลจากไฟล์
28. StreamWriter() ทำหน้าที่เขียนข้อมูลลงไฟล์

บทที่ 4

ผลการทดลองและวิเคราะห์

จากการประเมินผลการวิเคราะห์ด้วยชุดข้อความแสดงความคิดเห็นเกี่ยวกับสินค้าโทรศัพท์เคลื่อนที่ทั้งหมด 5 ยี่ห้อ คือ iPhone 4s, LG Optimus 2x, Nokia X7, Samsung GALAXY S3, Sony Xperia S แบ่งเป็น 3 คุณลักษณะ (Capability, Design, Price) และผลการประเมินแบบ Overall จากเว็บไซต์ <http://www.siamphone.com> สรุปได้ดังนี้

4.1 การทดสอบการวิเคราะห์คำคุณลักษณะ

4.1.1 การวิเคราะห์คำคุณลักษณะ

ตัวอย่างข้อความที่ 1 สวยมากเลยครับ สวยมาไว้นั้

จะรอผู้รู้มาให้ข้อมูลอยู่ครับ

ระบบสรุปได้ว่า : สวยมากเลยครับ เป็นข้อความคุณลักษณะด้าน Design ใน “เชิงบวก”

ตัวอย่างข้อความที่ 2 สเป็กสุดยอดมากๆ กล้องแรงเวอร์

ขอสื่อที่จ่อย่าแ่เหมือน ออปคิมมีส่วน

ระบบสรุปได้ว่า : สเป็กสุดยอดมากๆ เป็นข้อความคุณลักษณะด้าน Capability ใน “เชิงบวก”

ตัวอย่างข้อความที่ 3 ไม่สวยอะ TT

ระบบสรุปได้ว่า : ไม่สวยอะ เป็นข้อความคุณลักษณะด้าน Design ใน “เชิงลบ”

ตัวอย่างข้อความที่ 4 แต่ยังไม่ชอบดีไซน์เท่าไร รอรุ่นต่อไป

ระบบสรุปได้ว่า : ไม่ชอบดีไซน์ เป็นข้อความคุณลักษณะด้าน Design ใน “เชิงลบ”

ตัวอย่างข้อความที่ 5 ราคาแรงเอาเรื่องเหมือนกันนะเนี่ย

ระบบสรุปได้ว่า : ราคาแรงเอาเรื่อง เป็นข้อความคุณลักษณะด้าน Price ใน “เชิงลบ”

ตัวอย่างข้อความที่ 6 แบตหมดไวมากอะ แอบเซ็ง

ระบบสรุปได้ว่า : แบตหมดไว เป็นข้อความคุณลักษณะด้าน Capability ใน “เชิงลบ”

ตัวอย่างข้อความที่ 7 แต่รับไม่ได้กับ Design จริงๆเห็นบ่นกันทั่วโลกเลย

ระบบสรุปได้ว่า : รับไม่ได้กับ Design เป็นข้อความคุณลักษณะด้าน Design ใน “เชิงลบ”

ตัวอย่างข้อความที่ 8 ถ้าราคาไม่แรงมากอาจเจอกัน

ระบบสรุปได้ว่า : ราคาไม่แรงมาก เป็นข้อความคุณลักษณะด้าน Price ใน “เชิงบวก”

4.1.2 การสกัดความคิดเห็นโดยระบุขั้วที่มีและไม่มีในคลังข้อความ

สรุปผลการวิเคราะห์ข้อความแสดงความคิดเห็นทั้งหมด โดยระบุขั้ว แบ่งออกเป็น 3 ขั้วคือ
ความคิดเห็นในเชิงบวก ความคิดเห็น ในเชิงลบ ความคิดเห็นที่เป็นกลาง

ตารางที่ 4.1 แสดงการเปรียบเทียบความคิดเห็นโดยระบุขั้วได้ ที่มีและไม่มีค่าในคลังข้อความ

ยี่ห้อ โทรศัพท์	จำนวนข้อความทั้งหมด	ความคิดเห็นที่มีค่าในคลังข้อความ	ความคิดเห็นที่ไม่มีค่าในคลังข้อความ
iPhone 4S	90	32.22 %	67.78 %
LG Optimus 2x	306	17.32 %	82.68 %
Nokia X7	197	18.27 %	81.72 %
Samsung GALAXY S3	224	14.28 %	85.71 %
Sony Xperia S	238	15.54 %	84.45 %

จากตารางจะเห็นได้ว่า ข้อความแสดงความคิดเห็นที่เราสามารถสกัดได้นั้นมีเปอร์เซ็นต์ที่น้อยกว่าความคิดเห็นที่ไม่สามารถสกัดได้เนื่องจากคำศัพท์บางคำอาจจะไม่มีในฐานข้อมูลของตัวตัดคำ และคลังข้อความแสดงความคิดเห็นเกี่ยวกับโทรศัพท์มือถือ

4.1.3 การสกัดความคิดเห็นระดับคุณลักษณะ

ทำการวิเคราะห์ข้อความแสดงความคิดเห็นในแต่ละความคิดเห็น โดยระบุขั้ว แบ่งออกเป็น 2 ขั้วคือ ความคิดเห็นในเชิงบวก ความคิดเห็นในเชิงลบ ที่สกัดได้

ตารางที่ 4.2 การประเมินผลการสกัดข้อความแสดงความคิดเห็นระดับคุณลักษณะ

ยี่ห้อโทรศัพท์	ความครบถ้วน (Recall)	ความแม่นยำ (Precision)	F-Measure
iPhone 4S	68.42%	76.47%	72.22%
LG Optimus 2x	82.69%	67.18%	74.13%
Nokia X7	69.23%	55.10%	61.36%
Samsung GALAXY S3	77.08%	72.55%	74.75%
Sony Xperia S	72.72%	72.72%	72.72%

จากตารางแสดงการวิเคราะห์ข้อความแสดงความคิดเห็นข้างต้นสรุปได้ว่า เครื่องมือวิเคราะห์ความคิดเห็นต่อสินค้าโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์สามารถวิเคราะห์ข้อความแสดงความคิดเห็นได้ จากนั้นสามารถนำข้อความที่วิเคราะห์มาสรุปในรูปแบบกราฟ

4.2 การประเมินผล

4.2.1 การคำนวณ Recall และ Precision

Recall นั้นถูกกำหนดให้เป็นอัตราส่วนของข้อความแสดงความคิดเห็นที่ระบบ บอกว่าเป็นบวกหรือลบทั้งหมด ในขณะที่ Precision เป็นอัตราส่วนของข้อความแสดงความคิดเห็นที่ผู้ใช้โทรศัพท์เคลื่อนที่ บอกว่าเป็นบวกหรือลบ และวัดประสิทธิภาพโดยรวมด้วยค่า F-Measure (เป็นผลการเฉลี่ยของค่ารวม Recall และ Precision) สามารถให้คำนิยามของค่าต่างๆ ได้ดังนี้

ให้ $A = \text{Set}$ ของข้อความแสดงความคิดเห็นที่ระบบบอกว่าเป็นบวกหรือลบ

$B = \text{Set}$ ของข้อความแสดงความคิดเห็นที่ผู้ใช้โทรศัพท์เคลื่อนที่บอกว่าเป็นบวกหรือลบ

$$\text{Recall} = \frac{A \cap B}{A}$$

$$\text{Precision} = \frac{A \cap B}{B}$$

$$F - \text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

โดยเกณฑ์การประเมินประสิทธิภาพของโปรแกรมคือ ค่า F-Measure จะมีค่าประมาณร้อยละ 71.07 จึงถือว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพค่อนข้างดี โดยการวิเคราะห์ข้อความแสดงความคิดเห็น โดยผู้ใช้

4.3 ตัวอย่างแบบประเมินข้อความแสดงความคิดเห็นของผู้ใช้

ตารางที่ 4.3 แบบประเมินความคิดเห็นของผู้ใช้โทรศัพท์เคลื่อนที่บอกว่าเป็นบวกหรือลบ

ข้อความแสดงความคิดเห็น	ความ คิดเห็นเชิง บวก	ความ คิดเห็นเชิง ลบ
เก๋อะ ที!	x	
ยังกากได้อีก 555+		x
แต่ต้องยอมรับว่ากระแส apple ทำให้พวกเราได้ใช้เทคโนโลยีใหม่ๆแล้ว จะมีใครรู้ใหม่ว่าผมใช้ IPHONE4S ช่างในเหมือนเดิมก็ได้ แต่เปลี่ยน ภายนอกให้ดูรู้ว่า นี่แหละรุ่นใหม่ ถูไปไหน อู๊ย IPHONE 4S หน่อยดิ	x	
บันทึกเสียงระหว่างสนทนาได้ เป็นฟังก์ชันที่หามาแล้ว โทรศัพท์รุ่น เก่าๆ มีฟังก์ชันนี้ แต่ในช่วงหลังๆ พวกสมาร์ท โฟนไม่เห็นจะมีเลย ไม่รู้ว่า ฟังก์ชันนี้มันมีต้นทุนในการผลิตสูงหรือเปล่าที่ตัดมันออก เพิ่งจะเห็นใน iphone 4S อีกครั้ง เผลอๆ ได้เสียตั้งค์ซื้อเน้ ^^"	x	
ส่วนตัว ผมเห็นว่า ราคามันก็ค่อนข้างสูงนะคับ และลูกเล่นก็เพิ่มไม่เยอะ ด้วย เห็นหลักๆ กล้องดิจิตอล 8 ล้าน กล้องหน้าก็ VGA เหมือนเดิมผมว่า ไม่ต่างจาก 4 ตัวเดิมเท่าไร ทรงก็เดิม ถ้าจะซื้อให้ลูก ก็ 4 ตัวเดิมดีกว่า ใช้ เองถ้าตาม นิยมก็ 4S ก็ได้คับ เพราะกล้องตัวเดิมก็ซ้คอยู่แล้ว		x
สวยดีนะ จอก็น่าจะสวย ถ้าราคาไม่แรงมากอาจเจอกัน	x	
สเปกเทพดีครับ ขอให้ราคาไม่แรงจะดีมากๆ	x	
มารอดูราคากันเนอะ แต่ 23000 แพงไปนะ		x

บทที่ 5

บทสรุปและข้อเสนอแนะ

5.1 สรุปผลการทดลอง

เครื่องมือวิเคราะห์ความคิดเห็นต่อโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์สามารถวิเคราะห์ข้อความแสดงความคิดเห็นเกี่ยวกับโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์ได้ แต่เนื่องจากข้อความแสดงความคิดเห็นภาษาไทย ไม่มีรูปแบบที่แน่นอนจึงทำให้ยังมีข้อผิดพลาดอยู่ในการสกัดความคิดเห็น โดยระบุข้อ และการสกัดความคิดเห็นระดับคุณลักษณะ หากมีการปรับปรุงแก้ไขเพิ่มเติมก็จะสามารถทำให้เครื่องมือวิเคราะห์ความคิดเห็นต่อโทรศัพท์เคลื่อนที่ในเนื้อหาออนไลน์ได้อย่างแม่นยำยิ่งขึ้น และสามารถนำไปปรับใช้กับการวิเคราะห์ข้อความแสดงความคิดเห็นในหัวข้ออื่นๆที่น่าสนใจได้อย่างมีประสิทธิภาพ

5.1.1 การวิเคราะห์คำคุณลักษณะ

จากการวิเคราะห์คำคุณลักษณะของระบบ ระบบสามารถวิเคราะห์หาข้อความคุณลักษณะที่มีทัศนคติในเชิงบวกหรือลบได้โดยใช้ F-Measure ในการวัด ในส่วนเปอร์เซ็นต์ที่คลาดเคลื่อนเกิดจากปัญหาการตัดคำซึ่งยังมีข้อผิดพลาดอยู่

5.1.2 การสกัดความคิดเห็นโดยระบุข้อที่มีและไม่มีในคลังข้อความ

จากการทดลองพบว่า การสกัดความคิดเห็น โดยระบุข้อที่มีและไม่มีในคลังข้อความนั้น ความคิดเห็นที่มีในคลังข้อความจะมีเปอร์เซ็นต์ที่น้อยกว่าความคิดเห็นที่ไม่มีในคลังข้อความ เนื่องจากปัญหาที่เกิดขึ้นได้ดังนี้ การพิมพ์คำไม่ถูกต้องต้องตามพจนานุกรม สารเนื้อหาไม่เกี่ยวข้องกับหัวข้อ ฯลฯ

5.1.3 การสกัดความคิดเห็นระดับคุณลักษณะ

จากการทดลองจะเห็นได้ว่าการสกัดความคิดเห็นระดับคุณลักษณะโดยใช้ F-Measure ในการช่วยวัดผล จากค่าเฉลี่ยที่ได้ของ F-Measure ของโทรศัพท์เคลื่อนที่ทั้ง 5 รุ่น คือ iPhone 4s, LG Optimus 2x, Nokia X7, Samsung GALAXY S3, Sony Xperia S มีค่าเฉลี่ย 71.07% เห็นได้ว่าโปรแกรมสามารถสกัดความคิดเห็นระดับคุณลักษณะได้

5.2 ปัญหาและอุปสรรคในการพัฒนา

1. การตัดคำและการสกัดข้อความแสดงความคิดเห็นยังมีข้อผิดพลาด เนื่องจากคำศัพท์บางคำอาจจะไม่มีในฐานข้อมูลของตัวตัดคำและคลังข้อความแสดงความคิดเห็นสำหรับโทรศัพท์เคลื่อนที่

วิธีการแก้ไขปัญหาคือ นำคำที่ยังไม่มีในฐานข้อมูลหรือคลังข้อความเพิ่มเข้าไป และพัฒนาตัวตัดคำขึ้นมาใช้เอง เพื่อให้ตรงตามความต้องการของผู้ใช้

2. การแสดงความคิดเห็นของผู้ใช้นเว็บไซต์ ที่ไม่เกี่ยวข้องกับหัวข้อ ทำให้ยากต่อการวิเคราะห์ข้อความแสดงความคิดเห็น วิธีการแก้ไขปัญหาคือ เลือกรหาเว็บไซต์ที่มีเนื้อหาเกี่ยวข้องกับการแสดงความคิดเห็นเกี่ยวกับสินค้าให้ได้เยอะที่สุด

3. ข้อมูลโทรศัพท์เคลื่อนที่จากเว็บไซต์ www.siamphone.com มีเป็นจำนวนมาก ทำให้การดึงข้อมูลในส่วนของข้อความแสดงความคิดเห็นมาวิเคราะห์มีความล่าช้า ทำให้ต้องเสียเวลารอนาน วิธีการแก้ไขปัญหาคือ พัฒนาอัลกอริทึมในการดึงและเก็บข้อมูลจากเว็บไซต์ เพื่อรองรับข้อมูลจำนวนมากที่นำมาจากเว็บไซต์และมีการอัปเดตข้อมูลได้โดยไม่ต้องโหลดใหม่ทั้งหมด

5.3 ข้อเสนอแนะและแนวทางการพัฒนา

1. พัฒนาอัลกอริทึมในการวิเคราะห์ข้อความแสดงความคิดเห็นในระดับคุณลักษณะ
2. ควรมีการสร้างตัวตัดคำขึ้นมาใช้เอง เพื่อที่จะสามารถเพิ่มคำศัพท์หรือคำเฉพาะอื่นๆ ได้ทำให้ลดข้อผิดพลาดที่เกิดขึ้น ในการวิเคราะห์ให้ได้
3. ข้อความแสดงความคิดเห็นที่นำมาวิเคราะห์ควรมาจากหลายเว็บไซต์
4. ปรับปรุงแก้ไขในส่วนการดึงข้อมูลจากเว็บไซต์ให้สามารถโหลดต่อจากของเดิมได้ เพื่อการวิเคราะห์ครั้งต่อไปจะทำได้ทำไ้รวดเร็วขึ้น
5. สามารถพัฒนาให้นำไปใช้กับสินค้าชนิดอื่นที่สนใจได้

เอกสารอ้างอิง

- [1] Choochart Haruechaiyasak, et al. (2010). Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews. Human Language Technology Laboratory (HLT), National Electronics and Computer Technology Center (NECTEC), Thailand
- [2] Opinion Mining (เทคโนโลยีเหมืองข้อความแสดงความคิดเห็น). Available from: http://wiki.nectec.or.th/ru-newwiki/bin/view/IT630_11_Assignment/Gr11_OpinionMining
- [3] Natural Language Processing : [online] [cited 2011 Oct 09] Available from: http://vclass.mgt.psu.ac.th/~parinya/.../natural_language.html
- [4] Basic NLP Tools and Text Mining Applications : [online] [cited 2012 Jan 18] Available from: suanpalm3.kmutnb.ac.th/teacher/.../choochart32255411155.pdf
- [5] Google. Visualization: Pie Chart: [online] [cited 2012 Jan 18] Available from: <http://code.google.com/intl/th-TH/apis/visualization/documentation/gallery/picchart.html>
- [6] ICU4NET provides binding for .NET language: [online] [cited 2012 Apr 27] Available from: <http://code.google.com/p/icu4net/>
- [7] Microsoft Chart Controls for Microsoft .NET Framework 3.5: [online] [cited 2012 Apr 30] Available from: <http://msdn.microsoft.com/en-us/library/dd456632.aspx>

ภาคผนวก ก.
เครื่องมือในการพัฒนา

- Microsoft Visual C# Studio 2010
เป็นเครื่องมือที่ช่วยนักพัฒนาซอฟต์แวร์พัฒนาโปรแกรมคอมพิวเตอร์
- ICU4NET for C#
คือ ไลบรารีที่ใช้ในการตัดคำภาษาไทย โดย ICU4NET ตัวนี้ใช้สำหรับ C# เท่านั้น
- Microsoft Chart Controls for Microsoft .NET Framework 3.5
คือ ไลบรารีที่ใช้ในการสร้างกราฟของ Microsoft Visual Studio
- Microsoft Office Excel 2010
- Microsoft Power point 2010
- Calculate
- Paint

