

การประเมินคุณภาพการใช้งานซอฟต์แวร์โดยใช้การทำเหมืองบทวิจารณ์ผู้ใช้งาน

นายวฤต เลี้ยวไพโรจน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2555

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)

are the thesis authors' files submitted through the Graduate School.

EVALUATING SOFTWARE QUALITY IN USE USING USER REVIEWS MINING

Mr.Warit Leopairote

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Computer Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2012

Copyright of Chulalongkorn University

วฤช เลี้ยวโพโรจน์: การประเมินคุณภาพการใช้งานซอฟต์แวร์โดยใช้การทำเหมืองบทวิจารณ์
ผู้ใช้งาน (EVALUATING SOFTWARE QUALITY IN USE USING USER REVIEWS MINING)

อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. นครทิพย์ พร้อมพูล, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม :
ผศ.ดร.อรรถสิทธิ์ สุรฤกษ์, 103 หน้า.

ข้อมูลบทวิจารณ์ที่อยู่ในรูปภาษาธรรมชาติและให้ความคิดเห็นเกี่ยวกับคุณภาพของซอฟต์แวร์
จากผู้ใช้ที่มีประสบการณ์ในการใช้งานที่ปรากฏบนอินเทอร์เน็ตนั้นมีจำนวนมาก ซึ่งข้อมูลนี้มีประโยชน์
อย่างยิ่งเพื่อใช้ประกอบการตัดสินใจในการเลือกใช้ซอฟต์แวร์ที่มีความสามารถและคุณภาพตามความ
ต้องการของผู้ใช้ จึงเป็นสิ่งที่ท้าทายในการรวบรวม สกัด และสรุปข้อมูลบทวิจารณ์เพื่อให้คะแนน
คุณภาพในการใช้งานให้อยู่ในรูปแบบที่ง่ายต่อการนำไปใช้ดังกล่าว

วิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อหาวิธีการในการประเมินคุณภาพซอฟต์แวร์จากบทวิจารณ์ให้
อยู่ในรูปของคะแนน วิธีการที่นำเสนอมีสี่ขั้นตอน ขั้นตอนที่หนึ่ง การรวบรวมและการเตรียมข้อมูลให้
เหมาะสมสำหรับวิธีการประเมินที่นำเสนอ ขั้นตอนที่สอง การหาความเกี่ยวข้องของประโยคจากบท
วิจารณ์กับคุณลักษณะของคุณภาพซอฟต์แวร์จากมุมมองของผู้ใช้งานตามมาตรฐานไอเอสโอ 9126 ที่
ได้จำแนกคุณภาพในการใช้งานเป็นสี่คุณลักษณะ คือ ประสิทธิภาพ ผลผลิต ความปลอดภัย และความพึง
พอใจ ขั้นตอนที่สาม การจำแนกความคิดเห็นของประโยค ประยุกต์ใช้รายการคำศัพท์ที่เกี่ยวข้องกับ
ความรู้สึกกับกฎที่สร้างขึ้นจากประโยคในบทวิจารณ์ ขั้นตอนที่สี่ การคำนวณคะแนนคุณภาพการใช้งาน
ทั้งในระดับประโยค ระดับบทวิจารณ์ และระดับซอฟต์แวร์ จากผลการทดลองสรุปได้ว่า ประการแรก
คือการขยายออนโทโลยีของคุณลักษณะของคุณภาพซอฟต์แวร์ด้วยอรรถาภิธานไม่ได้ช่วยเพิ่ม
ประสิทธิภาพของการหาความเกี่ยวข้องของประโยค ประการที่สองการจำแนกความคิดเห็นของประโยค
ด้วยวิธีที่นำเสนอให้ประสิทธิภาพด้านความถูกต้องดีกว่าด้วยวิธีการใช้คำศัพท์เกี่ยวกับความรู้สึกเพียง
อย่างเดียวคิดเป็นร้อยละ 4 ประการที่สามคือการคำนวณคะแนนคุณภาพการใช้งานด้วยวิธีที่นำเสนอ
ให้ประสิทธิภาพด้านความถูกต้องดีกว่าด้วยวิธีการใช้คำศัพท์เกี่ยวกับความรู้สึกเพียงอย่างเดียว คิดเป็น
ร้อยละ 1.7 ประการที่สี่คือคะแนนคุณภาพการใช้งานมีความสอดคล้องกับความคิดเห็นของบทวิจารณ์
มากกว่าคะแนนที่คิดจากความคิดเห็นของประโยคเพียงอย่างเดียวคิดเป็นร้อยละ 0.5 วิธีการที่นำเสนอ
นี้สามารถนำไปประยุกต์ใช้กับคุณภาพซอฟต์แวร์ด้านอื่นๆ หรือสินค้าประเภทอื่นๆ ได้

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต

สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

ปีการศึกษา 2555 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม

5470361821 : MAJOR COMPUTER ENGINEERING

KEYWORDS : OPINION MINING/ SENTIMENT ANALYSIS/ SOFTWARE QUALITY/ POLARITY CLASSIFICATION/ ONTOLOGY

WARIT LEOPAIROTE : EVALUATING SOFTWARE QUALITY IN USE USING USER REVIEWS MINING. ADVISOR : ASST. PROF. NAKORNTHIP, CO-ADVISOR : ASST. PROF. ATHASIT SURARERKS, Ph.D., 103 pp.

On the internet, mass of review information is in natural language and provides experience users' comments that involved software quality. This information is useful for making a decision for software product acquirement in order that software product capability and quality meet user's requirements. It is challenging to aggregate, extract and summarize the review information for quality in use score in a form that it is easy to apply.

This thesis aims to propose a methodology for an evaluating software quality in a form of score based on reviews. This proposed methodology is consisted of 4 steps. 1) Gathering and preparing the data for the proposed evaluating methodology. 2) Mapping the relevant review sentences onto the characteristics of software quality model assessed from user perspective. According to ISO 9126 standard, the software quality in use is classified into 4 characteristics: effectiveness, productivity, safety and satisfaction. 3) Classifying the polarity of sentence opinion. This approach is applied from lists of sentiment words combining with rules made up of review sentences. 4) Calculating quality in use score in sentence level, review level and software level. According to the experiment result, the conclusions are 1) the ontology expansion using thesaurus does not improve the effectiveness of the sentence mapping 2) the effectiveness of proposed sentence polarity classification yields 4 percent of the accuracy improvement from the classification using only sentiment words 3) the effectiveness of quality in use score calculation using proposed approach yields 1.7 percent of the accuracy improvement from the calculation using only sentiment words 4) while comparing review polarity relation, quality in use score is more 0.5 present related than the score calculated based on only sentence polarity. This proposed methodology can be applied to other aspects of software quality and other types of products.

Department:Computer Engineering..... Student's Signature

Field of Study: Computer Engineering.... Advisor's Signature

Academic Year:2012..... Co-advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ได้รับคำแนะนำ ข้อคิดเห็น และแนวทางในการทำวิจัยที่ดีต่างๆ จากอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ นครทิพย์ พร้อมพูล และอาจารย์ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์ ที่คอยช่วยเหลือให้ความรู้ทางด้านที่จำเป็นต่อวิทยานิพนธ์ และคำปรึกษาที่สามารถแก้ไขปัญหาดังกล่าวที่เกิดขึ้นได้ นอกจากนี้ยังถ่ายทอดประสบการณ์ต่างๆ ให้เป็นข้อคิด และแนวทางที่ดีต่อการทำงานต่อไป ผู้วิจัยจึงขอกราบขอบพระคุณอาจารย์ที่ปรึกษาทั้งสอง

กราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.สมชาย ประสิทธิ์จตุระกุล ผู้ช่วยศาสตราจารย์ ดร.ธนารัตน์ ชลิตาพงศ์และ ผู้ช่วยศาสตราจารย์ ดร.เบญจพร ลิ้มธรรมมาภรณ์ ที่สละเวลามมาให้ข้อเสนอแนะ และแนวคิดต่างๆ ที่เป็นประโยชน์ต่อการพัฒนาวิทยานิพนธ์ฉบับนี้

ขอขอบคุณจุฬาลงกรณ์มหาวิทยาลัยที่ให้ความรู้ และประสบการณ์ที่ดีต่างๆ ที่เป็นส่วนหนึ่งในการใช้ชีวิตในมหาวิทยาลัย รวมทั้งขอขอบคุณอาจารย์ทุกๆ ท่านที่ได้อบรม สั่งสอนให้ผู้วิจัยได้รับความรู้ในวิชาต่างๆ ที่เป็นประโยชน์ในการทำงาน อีกทั้งเพื่อนๆ พี่ๆ น้องๆ ที่ดีในแลป SE ที่คอยช่วยเหลือ และเป็นกำลังใจให้ในช่วงเวลาต่างๆ

และสุดท้าย ขอขอบคุณทุกคนในครอบครัวทั้ง คุณพ่อ คุณแม่ น้องชาย และญาติๆ ที่คอยสนับสนุนให้ผู้ทำวิจัยได้เรียนรู้ และทำในสิ่งที่ดี คอยเป็นกำลังใจให้กันและกันตลอดมา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฎ
สารบัญรูป.....	ฐ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 ปัญหาของงานวิจัย.....	4
1.3 วัตถุประสงค์ของงานวิจัย.....	5
1.4 ขอบเขตของงานวิจัย.....	5
1.5 ประโยชน์ของงานวิจัย.....	5
1.6 ขั้นตอนและวิธีดำเนินการวิจัย.....	6
1.7 โครงสร้างของเนื้อหาในวิทยานิพนธ์.....	6
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ความรู้และทฤษฎีที่เกี่ยวข้อง.....	7
2.1.1 บทวิจารณ์หรือความคิดเห็น.....	7
2.1.2 การทำเหมืองข้อความ (Text mining).....	10
2.1.3 การทำเหมืองความคิดเห็น (Opinion mining).....	12
2.1.4 แบบจำลองคุณภาพซอฟต์แวร์ (Software quality model).....	13
2.2 งานวิจัยที่เกี่ยวข้อง.....	17
2.2.1 ระบบแนะนำซึ่งมีความเชี่ยวชาญโดยการแนะนำนั้นพิจารณาจากบทวิจารณ์สินค้าของ ลูกค้า (Informed Recommender: Basing Recommendations on Consumer Product Reviews).....	17
2.2.2 การทำเหมืองเชิงขั้วโดยมีออนโทโลยีสนับสนุน (Ontology-Supported Polarity Mining).....	18
2.2.3 การทำเหมืองความคิดเห็นที่ใช้ออนโทโลยีสำหรับบทวิจารณ์เกี่ยวกับภาพยนตร์ (Ontology Based Opinion Mining for Movie Reviews).....	19

2.2.4 การทำเหมืองความคิดเห็นที่ใช้ลักษณะเฉพาะโดยกระบวนการที่แนะนำโดยออนโทโลยี (Ontology-Guided Approach to Feature-Based Opinion Mining).....	20
2.2.5 การทำเหมืองและสรุปบทวิจารณ์ของลูกค้า (Mining and Summarizing Customer Reviews).....	21
2.2.6 สรุปและวิเคราะห์แนวคิดในงานวิจัยต่างๆ	23
บทที่ 3 แนวคิดที่นำเสนอ	26
3.1 ภาพรวมกระบวนการทำเหมืองบทวิจารณ์ด้านคุณภาพซอฟต์แวร์	26
3.2 การรวบรวมและจัดเตรียมข้อมูล	28
3.3 กระบวนการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน	29
3.4 กระบวนการจำแนกความคิดเห็นของประโยค	31
3.4.1 กระบวนการจำแนกความคิดเห็นของประโยคแบบพื้นฐาน	32
3.4.2 กระบวนการจำแนกความคิดเห็นของประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ ผู้ใช้งาน	32
3.5 การคำนวณคะแนนคุณภาพในการใช้งาน	38
3.5.1 การคำนวณคะแนนคุณภาพในการใช้งานในระดับประโยค	39
3.5.2 การคำนวณคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์.....	39
3.5.3 การคำนวณคะแนนคุณภาพการใช้งานในระดับซอฟต์แวร์	40
บทที่ 4 การทดลองและการวิเคราะห์ผล	42
4.1 การวางแผนการทดลอง	42
4.1.1 จุดประสงค์การทดลอง.....	42
4.1.2 ออกแบบการทดลอง	42
4.1.3 มาตรการวัด	44
4.2 การเตรียมการทดลองและข้อมูลทดลอง	46
4.3 การวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใ้ งาน	48
4.3.1 ขั้นตอนการทดลอง.....	48
4.3.2 ผลการทดลองและการวิเคราะห์ผล	50
4.3.3 สรุปผลการทดลอง.....	51
4.3.4 ข้อมูลสนับสนุนผลการทดลอง	52

4.4 การจำแนกประโยชน์ว่าเป็นประโยชน์ที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง	52
4.4.1 ขั้นตอนการทดลอง	53
4.4.2 ผลการทดลองและการวิเคราะห์ผล	56
4.4.3 สรุปผลการทดลอง.....	58
4.4.4 ข้อมูลสนับสนุนผลการทดลอง	59
4.5 การวัดประสิทธิผลส่วนการคำนวณคะแนนคุณภาพการใช้งาน	61
4.5.1 ขั้นตอนการทดลอง	62
4.5.2 ผลการทดลองและการวิเคราะห์ผล	63
4.5.3 สรุปผลการทดลอง.....	64
4.6 การวัดประสิทธิผลของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง.....	64
4.6.1 ขั้นตอนการทดลอง	66
4.6.2 ผลการทดลองและการวิเคราะห์ผล	68
4.6.3 สรุปผลการทดลอง.....	69
4.6.4 ข้อมูลสนับสนุนผลการทดลอง	69
4.7 สรุปผลทั้งหมดของทุกการทดลอง	72
บทที่ 5 การพัฒนาเครื่องมือ	76
5.1 ความต้องการเชิงหน้าที่.....	76
5.2 การออกแบบเครื่องมือ	77
5.2.1 การออกแบบสถาปัตยกรรมของเครื่องมือ.....	77
5.2.2 ความต้องการของระบบ	78
5.2.3 แบบจำลองข้อมูล	78
5.2.4 เครื่องมือสนับสนุนที่ใช้.....	80
5.2.5 รายละเอียดในการออกแบบ	81
5.3 การทำงานและส่วนต่อประสานผู้ใช้งาน.....	82
5.3.1 ป้อนข้อมูลบทวิจารณ์.....	82
5.3.2 ดูคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์และบทวิจารณ์.....	84
5.3.3 ดูคะแนนคุณภาพการใช้งานระดับบทวิจารณ์.....	85
5.3.4 สร้างกฎ.....	85

5.3.5 ทำการทดลอง.....	87
บทที่ 6 สรุปผลการวิจัย	88
6.1 สรุปผลการวิจัย	88
6.2 ข้อจำกัดของงานวิจัย	90
6.3 งานวิจัยในอนาคต.....	90
6.4 ผลงานตีพิมพ์จากวิทยานิพนธ์	91
รายการอ้างอิง	92
ภาคผนวก.....	96
ภาคผนวก ก.....	97
ภาคผนวก ข.....	100
ภาคผนวก ค.....	102
ประวัติผู้เขียนวิทยานิพนธ์.....	103

สารบัญตาราง

	หน้า
ตารางที่ 1.1 คำศัพท์และคำอธิบาย.....	4
ตารางที่ 2.1 ปัจจัยสำคัญของแนวคิดในงานวิจัยต่างๆ	24
ตารางที่ 3.1 ตัวอย่างรายการคำศัพท์หลังการตัดคำ.....	29
ตารางที่ 3.2 ผลของการจำแนกประโยคด้วยสองเครื่องมือจำแนก	32
ตารางที่ 3.3 ตัวอย่างการสร้างเครื่องมือจากสี่ชุดกฎ.....	35
ตารางที่ 3.4 ผลการจำแนกประโยคโดยวิธีการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ ผู้ใช้งาน	36
ตารางที่ 4.1 จำนวนของสิ่งที่ถูกจำแนกได้สองประเภทโดยระบบเปรียบเทียบกับผู้เชี่ยวชาญ.....	45
ตารางที่ 4.2 จำนวนของสิ่งที่ถูกจำแนกได้สามประเภทโดยระบบเปรียบเทียบกับผู้เชี่ยวชาญ.....	45
ตารางที่ 4.3 ข้อมูลบทวิจารณ์.....	47
ตารางที่ 4.4 ความถูกต้องของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้าน ต่างๆ.....	50
ตารางที่ 4.5 ประสิทธิภาพของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้านต่างๆ	51
ตารางที่ 4.6 วิธีในการทดลองส่วนการจำแนกประโยค.....	53
ตารางที่ 4.7 ประสิทธิภาพของการจำแนกประโยคสี่วิธีเมื่อใช้ข้อมูลทดสอบชุดที่หนึ่ง.....	57
ตารางที่ 4.8 ประสิทธิภาพของการจำแนกประโยคสี่วิธีเมื่อใช้ข้อมูลทดสอบชุดที่สอง	57
ตารางที่ 4.9 จำนวนประโยคที่มีความคิดเห็นเชิงบวก เชิงลบและเป็นกลางในข้อดีและข้อเสีย	59
ตารางที่ 4.10 ความถูกต้องของสมมติฐานในกรณีต่างๆ.....	60
ตารางที่ 4.11 ความถูกต้องของการจำแนกประโยคที่ประกอบด้วยคำนิเสธ.....	61
ตารางที่ 4.12 ความถูกต้องของการคำนวณคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์	63
ตารางที่ 4.13 ความถูกต้องของการวิเคราะห์ความคิดเห็นของบทวิจารณ์	68
ตารางที่ 4.14 สัมประสิทธิ์ของสมการและค่าเฉลี่ยของความผิดพลาดของสมการ 4.11.....	70
ตารางที่ 4.15 สัมประสิทธิ์ของสมการและค่าเฉลี่ยของความผิดพลาดของสมการ 4.12.....	71
ตารางที่ 4.16 ความสัมพันธ์ระหว่างคุณภาพด้านแต่ละด้านกับจำนวนดาวในซอฟต์แวร์แต่ละประเภท	71
ตารางที่ 4.17 สรุปการทดลองในหัวข้อ 4.3.....	72
ตารางที่ 4.18 สรุปการทดลองในหัวข้อ 4.4.....	73
ตารางที่ 4.19 สรุปการทดลองในหัวข้อ 4.5.....	74

ตารางที่ 4.20 สรุปการทดลองในหัวข้อ 4.6.....	75
ตารางที่ 5.1 โครงสร้างตารางข้อมูล Software.....	79
ตารางที่ 5.2 โครงสร้างตารางข้อมูล Review.....	79
ตารางที่ 5.3 โครงสร้างตารางข้อมูล Sentence.....	80
ตารางที่ 5.4 โครงสร้างตารางข้อมูล Rule.....	80
ตารางที่ ก.1 คำศัพท์ของออนโทโลยีที่ไม่ได้ถูกขยาย.....	97
ตารางที่ ก.2 คำศัพท์ของออนโทโลยีที่ถูกขยายหนึ่งรอบ.....	98

สารบัญรูป

หน้า

รูปที่ 2.1 ตัวอย่างหน้าต่างสำหรับกรอกบทวิจารณ์บนเว็บไซต์ www.cnet.com	9
รูปที่ 2.2 ตัวอย่างของบทวิจารณ์ในเว็บไซต์ www.cnet.com	9
รูปที่ 2.3 คุณภาพของซอฟต์แวร์ในวงจรการพัฒนาซอฟต์แวร์	14
รูปที่ 3.1 แผนภาพกิจกรรมของภาพรวมของแนวคิดที่นำเสนอ	27
รูปที่ 3.2 แผนภาพกิจกรรมของการรวบรวมและจัดเตรียมข้อมูล	28
รูปที่ 3.3 แผนภาพกิจกรรมของกระบวนการหาความเกี่ยวข้องเกี่ยวกับคุณลักษณะของคุณภาพในระดับ ประโยค	29
รูปที่ 3.4 ออนโทโลยี	30
รูปที่ 3.5 แผนภาพกิจกรรมของกระบวนการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน	33
รูปที่ 3.6 ขั้นตอนการหาค่าความคิดเห็นของประโยค	36
รูปที่ 3.7 แผนภาพกิจกรรมของการคำนวณคะแนนคุณภาพในการใช้งานผู้ใช้งาน	38
รูปที่ 4.1 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับ คุณลักษณะของคุณภาพการใช้งาน	49
รูปที่ 4.2 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิผลของการจำแนกประโยค	55
รูปที่ 4.3 แผนภาพกิจกรรมของภาพรวมการวัดความถูกต้องของการคำนวณคะแนนคุณภาพการใ้ งาน	62
รูปที่ 4.4 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิผลการวิเคราะห์ความคิดเห็นของบทวิจารณ์	67
รูปที่ 5.1 แผนภาพยูสเคสของเครื่องมือ	77
รูปที่ 5.2 สถาปัตยกรรมของเครื่องมือ	78
รูปที่ 5.3 แผนภาพแสดงความสัมพันธ์ระหว่างเอนทิตี	79
รูปที่ 5.4 หน้าจอหลักของเครื่องมือ	83
รูปที่ 5.5 หน้าจอส่วนการกรอกข้อมูลซอฟต์แวร์	83
รูปที่ 5.6 หน้าจอส่วนการกรอกข้อมูลบทวิจารณ์	84
รูปที่ 5.7 หน้าจอส่วนการกรอกข้อมูลโดยใช้ที่อยู่ของเว็บไซต์	84
รูปที่ 5.8 หน้าจอส่วนการดูคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์	85
รูปที่ 5.9 หน้าจอส่วนการเลือกซอฟต์แวร์เพื่อดูคะแนนคุณภาพการใช้งานของบทวิจารณ์	85
รูปที่ 5.10 หน้าจอส่วนการดูคะแนนคุณภาพการใช้งานระดับบทวิจารณ์	86
รูปที่ 5.11 หน้าจอส่วนการสร้างกฎ	86

รูปที่ 5.12 หน้าจอส่วนการเตรียมข้อมูลทดลองโดยผู้เชี่ยวชาญ.....	87
รูปที่ ข.1 ตัวอย่างเวกเตอร์ที่สร้างจากเครื่องมือ TMSK.....	101

บทที่ 1

บทนำ

ในบทนี้อธิบายถึงแนวคิดหลักของงานวิจัย ซึ่งประกอบด้วย ที่มาและความสำคัญของปัญหา วัตถุประสงค์ ขอบเขต ประโยชน์ ขั้นตอนและวิธีดำเนินการ

1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบันอินเทอร์เน็ตเข้ามาเป็นส่วนหนึ่งของชีวิตประจำวัน อินเทอร์เน็ตถูกนำเข้ามาช่วยในการทำงานรวมถึงการซื้อขายของบนอินเทอร์เน็ต พาณิश्यอิเล็กทรอนิกส์หรืออีคอมเมิร์ซ (e-commerce) จึงเกิดขึ้นจำนวนมากที่มีการซื้อขายและให้บริการ เช่น อีเบย์ (eBay) อะเมซอน (Amazon) ซีเน็ต (Cnet) เป็นต้น เว็บไซต์เหล่านี้เป็นที่นิยมมากขึ้นเนื่องจากมีข้อมูลสำคัญที่มีส่วนช่วยในการตัดสินใจคือ บทวิจารณ์ (review) หรือความคิดเห็น (comment) ของผู้ที่มีประสบการณ์ซื้อสินค้านั้น โดยทั่วไปบทวิจารณ์จะถูกเขียนขึ้นเพื่อแบ่งบอกข้อมูลเกี่ยวกับคุณลักษณะและคุณภาพของสินค้าหรือบริการในมุมมองผู้ใช้งาน

ในเว็บไซต์อีคอมเมิร์ซส่วนมากจะมีสินค้าและบริการจำนวนมากและหลากหลายประเภทจนผู้ซื้ออาจไม่สามารถเลือกดูได้ทั้งหมดจากข้อจำกัดด้านเวลา ผู้ซื้อส่วนมากใช้เวลาในการหาสินค้าที่มีคุณสมบัติและคุณภาพตามที่ต้องการด้วยการหาสินค้าที่คาดว่าจะตรงตามความต้องการ และอ่านข้อมูลสินค้านั้นทั้งหมดสำหรับเป็นข้อมูลประกอบการตัดสินใจ แนวคิดที่ช่วยให้ใช้เวลาเฉลียวจึงถูกเสนอขึ้น ซึ่งแนวคิดที่เป็นที่นิยมและน่าสนใจมีสองแนวคิด แนวคิดแรกคือการจัดทำระบบแนะนำ (recommender system) ระบบนี้จะช่วยหาและแนะนำสินค้าที่น่าสนใจอีกทั้งระบุว่าสินค้าใดที่ไม่ควรนำมาแสดง [1] วิธีที่ใช้ในระบบแนะนำที่นิยมมีสองอย่างคือ คอลลาบอราทีฟฟิวเตอร์ริง (collaborative filtering) และคอนเทนต์เบสฟิวเตอร์ริง (content-based filtering) [2] ระบบแนะนำจะทำเพียงแค่เสนอรายการของสินค้าที่ระบบประมวลผลแล้วให้ผลลัพธ์ที่ตรงกับความต้องการของผู้ซื้อซึ่งช่วยลดเวลาในส่วนของการหาสินค้าที่คาดว่าจะตรงตามความต้องการ แต่อย่างไรก็ตามผู้ซื้อยังต้องเสียเวลาในการอ่านข้อมูลเกี่ยวกับสินค้าประกอบการตัดสินใจ

เพื่อช่วยลดเวลาในการอ่านข้อมูลสินค้า อีกแนวคิดที่ถูกนำเสนอคือการทำเหมืองความคิดเห็น (opinion mining) แนวคิดนี้เสนอวิธีการสกัดข้อมูล (information extraction) การวิเคราะห์ข้อมูล (information analysis) และการเสนอกลุ่มของข้อมูล (information aggregation) จากบทวิจารณ์ ข้อมูลส่วนใหญ่ที่ผู้วิจารณ์เขียนในบทวิจารณ์จะประกอบด้วยระดับความพึงพอใจที่แสดงได้ด้วยจำนวนดาว และข้อความที่เขียนบรรยายว่าคุณลักษณะหรือคุณภาพของสินค้านั้นดีหรือไม่อย่างไร ข้อมูลของระดับความพึงพอใจหรือจำนวนดาวนี้บอกคุณภาพของสินค้าในมุมมองของผู้วิจารณ์ ข้อมูลส่วนนี้ในเว็บไซต์ส่วนมากจะสรุปอยู่ในภาพรวมของผู้วิจารณ์ทั้งหมดเพื่อเป็นข้อมูลเบื้องต้นในการตัดสินใจ แต่เป็นข้อมูลที่ไม่มีการละเอียด ข้อมูลจากข้อความที่เขียนบรรยายจึงมีความสำคัญที่จะบอกได้ว่าสินค้าหรือบริการนี้ดีหรือไม่ดีอย่างไร ดังนั้นการตีความหรือสกัดข้อมูลจะมุ่งเน้นไปในการค้นหาคุณลักษณะหรือคุณภาพของสินค้าหรือบริการ และพิจารณาตีความว่าข้อความที่เกี่ยวกับคุณลักษณะนั้นดีหรือไม่ดี

การค้นหาคุณลักษณะเฉพาะหรือคุณภาพของสินค้าบริการว่ามีอะไรบ้างที่ถูกเขียนในบทวิจารณ์ ในงานวิจัยที่เคยเสนอมามากจะมุ่งเน้นในการพิจารณาหาคุณลักษณะเฉพาะของสินค้าที่เป็นคำนามที่ถูกกล่าวถึงในบทวิจารณ์เป็นหลัก [3, 4] หรือในบางงานวิจัย จะเสนอออนโทโลยีหรือโครงสร้างข้อมูลที่เกี่ยวข้องกับสินค้านั้น [5, 6] เพื่อให้รู้ว่าสินค้านั้นประกอบด้วยลักษณะเฉพาะอะไรบ้าง หลังจากนั้นพิจารณาความเกี่ยวข้องกันระหว่างออนโทโลยีกับส่วนของข้อความหรือประโยคในบทวิจารณ์ ทั้งหมดนี้ข้อความในบทวิจารณ์จะถูกแบ่งเป็นส่วนๆ ที่รู้ว่าแต่ละส่วนกล่าวถึงลักษณะเฉพาะใด หลังจากนั้นจึงมีการตีความว่าข้อความนั้นมีความหมายหรือความรู้สึกอย่างไรนั้นเป็นประเด็นปัญหาที่หลายงานวิจัยเสนอแนวคิดเพื่อแก้ไขประเด็นปัญหานี้ บางงานวิจัยจะสร้างรายการคำจากคำที่รู้ความหมายอยู่แล้วไปเทียบกับคำในข้อความหรือประโยคเพื่อพิจารณา [7-9] แต่รายการคำที่กำหนดนั้นไม่สามารถใช้ได้ในทุกประเภทของสินค้าหรือบริการ ตัวอย่างเช่น “It is unpredictable” จะให้ความรู้สึกเชิงบวกในความคิดเห็นเกี่ยวกับภาพยนตร์ แต่จะให้ความรู้สึกเชิงลบในความคิดเห็นเกี่ยวกับเครื่องใช้ไฟฟ้า อีกหนึ่งวิธีการแก้ปัญหาที่ช่วยแก้ปัญหานี้คือกระบวนการเรียนรู้ของเครื่อง (machine learning) [6, 10, 11] กระบวนการนี้มุ่งเน้นการวิเคราะห์ความรู้สึกของประโยคโดยการจำแนกหรือประเมินผลข้อความหรือประโยคว่าแสดงความรู้สึกในเชิงบวกหรือเชิงลบ

เมื่อพิจารณาสินค้าหรือผลิตภัณฑ์ประเภทซอฟต์แวร์ บทวิจารณ์และความคิดเห็นที่สั้นนั้นจะให้ความคิดเห็นในด้านคุณภาพที่ไม่ได้ระบุไปในลักษณะเฉพาะใดลักษณะหนึ่งของซอฟต์แวร์ ดังนั้นการ

สกัดข้อมูลจะสกัดข้อมูลด้านคุณภาพของซอฟต์แวร์จากบทวิจารณ์หลายบทวิจารณ์แทนการสกัดข้อมูลด้านลักษณะเฉพาะ ในงานวิจัยนี้จึงเสนอกระบวนการประมวลผลบทวิจารณ์เพื่อให้ได้ข้อมูลเกี่ยวกับคุณภาพของซอฟต์แวร์ (software product quality) คุณภาพของซอฟต์แวร์ถูกวัดด้วยการใช้แบบจำลองต่างๆ ซึ่งแบบจำลองหนึ่งของแบบจำลองคุณภาพซอฟต์แวร์ (software quality model) ที่เป็นที่ยอมรับในด้านวิศวกรรมซอฟต์แวร์ (software engineer) คือ ไอเอสโอ 9126 (ISO 9126) [12] ในแบบจำลองนี้มีส่วนหนึ่งที่วัดคุณภาพในมุมมองของผู้ใช้คือ ส่วนคุณภาพในการใช้งาน (quality in use) ในงานวิจัยนี้จึงใช้แบบจำลองนี้ในการแสดงคุณภาพของสินค้าซึ่งประกอบด้วยคุณลักษณะประสิทธิผล (effectiveness) ผลิตภาพ (productivity) ความปลอดภัย (safety) และความพึงพอใจ (satisfaction) ในแนวคิดที่ได้นำเสนอ วิธีการประมวลผลขั้นแรกเริ่มจากการรวบรวมและจัดเตรียมข้อมูลเพื่อให้เหมาะสมกับวิธีที่นำเสนอ หลังจากนั้น วิธีการประมวลผลขั้นที่สองคือหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพซอฟต์แวร์ด้วยการจัดเข้ากลุ่มออนโทโลยี ออนโทโลยีประกอบด้วยคำศัพท์ (terminology) ที่ถูกกำหนดโดยการสกัดคำสำคัญที่เกี่ยวข้องกับแต่ละคุณลักษณะ แล้วถูกขยายเพิ่มคำศัพท์มากขึ้นด้วยบรรดาภิธานตามขั้นตอนวิธีที่ได้นำเสนอ วิธีการประมวลผลขั้นที่สามคือการจำแนกในการหาความเป็นเชิงบวกหรือเชิงลบของประโยคด้วยการประยุกต์คำศัพท์ที่เกี่ยวข้องกับความรู้สึก (sentiment words) และวิธีการจำแนกด้วยกฎ (rule-based classification) ซึ่งเป็นหนึ่งในวิธีการเรียนรู้ของเครื่อง วิธีการประมวลผลขั้นสุดท้ายเป็นการคำนวณข้อมูลบทวิจารณ์จากที่ถูกประมวลผลให้เป็นคะแนนคุณภาพการใช้งาน (quality in use score) ของซอฟต์แวร์ โดยรายละเอียดวิธีการคำนวณจะถูกนำเสนอในงานวิจัยนี้

เพื่อช่วยให้เข้าใจแนวคิดที่จะนำเสนอไปในทางเดียวกัน คำศัพท์แต่ละคำจะถูกอธิบายในตารางที่ 1

ตารางที่ 1.1 คำศัพท์และคำอธิบาย

คำศัพท์	คำอธิบาย
บทวิจารณ์ (review)	ข้อความที่ถูกเขียนขึ้นจากผู้ที่มีประสบการณ์โดยจะมีเนื้อหาไปในทางแสดงความคิดเห็นที่เกี่ยวข้องสิ่งที่ถูกวิจารณ์
การทำเหมืองความคิดเห็น (opinion mining)	กระบวนการที่ประมวลผลค้นหาสำหรับสิ่งที่ให้มา สร้างรายการของคุณลักษณะของสินค้า (คุณภาพ ลักษณะเฉพาะอื่นๆ) และรวบรวมหรือประมวลความคิดเห็นที่เกี่ยวข้องกับคุณลักษณะของสินค้านั้น (ไม่ดี กลาง ดี) [13]
การทำเหมืองเชิงขั้ว (polarity mining)	การจำแนกประเภทของข้อความโดยรวมทั้งหมดที่เป็นความคิดเห็นว่าข้อความนั้นแสดงความคิดเห็นไปในเชิงบวกหรือเชิงลบ [14]
ออนโทโลยี (ontology)	รายการของแนวคิดและหมวดหมู่ในหัวข้อใกล้เคียงกันที่แสดงความสัมพันธ์ระหว่างกัน
ความคิดเห็นเชิงบวก (pros)	ข้อความแสดงความคิดเห็นไปในแง่ดีหรือมีทัศนคติที่ดี
ความคิดเห็นเชิงลบ (cons)	ข้อความแสดงความคิดเห็นไปในแง่ร้ายหรือมีทัศนคติที่ไม่ดี
ประโยคเชิงบวก (positive sentence)	ประโยคที่แสดงความคิดเห็นไปในแง่ดี
ประโยคเชิงลบ (negative sentence)	ประโยคที่แสดงความคิดเห็นไปในแง่ร้ายหรือไม่ดี
แก่นคำ (stemmed word)	รากศัพท์หรือส่วนสำคัญของคำที่ยังให้ความหมายไปในแนวทางเดียวกันเมื่อถูกเพิ่มเติมตัวอักษรเข้าไปด้านท้าย

1.2 ปัญหาของงานวิจัย

คุณภาพของซอฟต์แวร์สามารถหาได้จากบทวิจารณ์ของซอฟต์แวร์นั้นได้อย่างไร

1.3 วัตถุประสงค์ของงานวิจัย

- 1) เพื่อออกแบบวิธีการประเมินข้อมูลคุณภาพการใช้งานซอฟต์แวร์โดยใช้การทำเหมืองบทวิจารณ์ผู้ใช้งาน และนำเสนอข้อมูลที่ได้จากการประเมินข้อมูลคุณภาพการใช้งานในรูปแบบที่เข้าใจง่าย
- 2) เพื่อสร้างเครื่องมือสำหรับประมวลผลบทวิจารณ์ของซอฟต์แวร์เพื่อให้ได้ข้อมูลเกี่ยวกับคุณภาพการใช้งานของซอฟต์แวร์ และนำเสนอข้อมูลในรูปแบบที่เข้าใจง่าย

1.4 ขอบเขตของงานวิจัย

- 1) งานวิจัยนี้จะพิจารณาเฉพาะบทวิจารณ์ที่เป็นภาษาอังกฤษเท่านั้น
- 2) งานวิจัยนี้จะพิจารณาเฉพาะบทวิจารณ์ที่เกี่ยวกับซอฟต์แวร์ที่อยู่ในเว็บไซต์ www.cnet.com เท่านั้น เนื่องจากเว็บไซต์นี้เป็นเว็บไซต์ที่มีโครงสร้างบทวิจารณ์ที่มีส่วนแสดงความคิดเห็นเชิงบวกกับความความคิดเห็นเชิงลบชัดเจน โดยแนวคิดจะนำส่วนนี้สามารถนำข้อมูลนี้มาเป็นข้อมูลสอนและข้อมูลทดสอบได้
- 3) งานวิจัยนี้มีข้อมูลทดสอบอย่างน้อย 30 ซอฟต์แวร์และในแต่ละซอฟต์แวร์มีอย่างน้อย 100 บทวิจารณ์
- 4) เครื่องมือที่ใช้ในการหาคำศัพท์ที่พ้องความหมายและความหมายตรงข้ามคือเวิร์ดเน็ต
- 5) เครื่องมือที่ใช้ในการสร้างเครื่องมือจำแนกประเภทด้วยกฎคือเครื่องมือริคเท็ค (RIKTEXT)

1.5 ประโยชน์ของงานวิจัย

- 1) ได้วิธีการประเมินข้อมูลคุณภาพการใช้งานซอฟต์แวร์โดยใช้การทำเหมืองบทวิจารณ์ผู้ใช้งาน และนำเสนอข้อมูลคุณภาพการใช้งานซอฟต์แวร์ในรูปแบบที่เข้าใจง่าย
- 2) ได้เครื่องมือสำหรับประมวลผลบทวิจารณ์ของซอฟต์แวร์เพื่อให้ได้ข้อมูลเกี่ยวกับคุณภาพการใช้งานของซอฟต์แวร์ และนำเสนอข้อมูลในรูปแบบที่เข้าใจง่าย

1.6 ขั้นตอนและวิธีดำเนินการวิจัย

- 1) ศึกษาความรู้เกี่ยวกับการทำเหมืองข้อความ (text mining)
- 2) ศึกษาความรู้เกี่ยวกับการทำเหมืองเชิงขั้ว (polarity mining)
- 3) ศึกษาความรู้เกี่ยวกับแบบจำลองคุณภาพซอฟต์แวร์ (software quality model)
- 4) ศึกษางานวิจัยที่เกี่ยวข้อง
- 5) ออกแบบแนวคิดในภาพรวมและแนวคิดในรายละเอียดส่วนต่างๆเพื่อแก้ปัญหา
- 6) ออกแบบวิธีการตรวจสอบแนวคิดที่นำเสนอกำหนดมาตรฐานวัดที่ใช้ในการทดลอง
- 7) ศึกษาและพัฒนาเครื่องมือเพื่อใช้ในการตรวจสอบแนวคิด
- 8) ตรวจสอบความถูกต้องของเครื่องมือและแนวคิดที่นำเสนอ
- 9) สรุปผลการทดลองและจัดทำเล่มวิทยานิพนธ์

1.7 โครงสร้างของเนื้อหาในวิทยานิพนธ์

โครงสร้างเนื้อหาของวิทยานิพนธ์ฉบับนี้ถูกแบ่งออกเป็นหกบท คือ บทที่ 1 เป็นบทนำ บทที่ 2 เป็นบทที่กล่าวถึงทฤษฎีและงานวิจัยต่างๆ ที่เกี่ยวข้อง บทที่ 3 เป็นบทที่อธิบายถึงขั้นตอนแนวคิดการทำเหมืองบทวิจารณ์ผู้ใช้งาน ซึ่งประกอบด้วยกระบวนการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน กระบวนการจำแนกความคิดเห็นของประโยคและการคำนวณคะแนนคุณภาพการใช้งาน บทที่ 4 เป็นบทที่อธิบายการทดลองและผลการทดลองของแนวคิดที่นำเสนอ รวมถึงการวิเคราะห์ผลการทดลอง บทที่ 5 เป็นบทที่อธิบายการออกแบบและพัฒนาเครื่องมือตามแนวคิดที่นำเสนอ และบทที่ 6 เป็นบทสุดท้ายโดยบทนี้จะเป็นบทสรุปและข้อจำกัดของงานวิจัยรวมทั้งงานวิจัยในอนาคต และบทความวิชาการที่ตีพิมพ์

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้ ทฤษฎี ความรู้และงานวิจัยที่เกี่ยวข้องกับงานวิจัยจะถูกอธิบาย โดยเริ่มที่ความรู้และทฤษฎีที่เกี่ยวข้องเพื่อเป็นความรู้พื้นฐาน แล้วจึงอธิบายงานวิจัยที่เกี่ยวข้องตามมาเพื่อเห็นภาพโดยรวมเกี่ยวกับงานวิจัยอื่นๆ ตามลำดับ

2.1 ความรู้และทฤษฎีที่เกี่ยวข้อง

ในหัวข้อนี้อธิบายเกี่ยวกับความรู้พื้นฐานที่เกี่ยวข้องกับแนวคิดที่จะนำเสนอ เริ่มจากอธิบายภาพรวมของบทวิจารณ์หรือความคิดเห็นที่อยู่บนเว็บไซต์ การทำเหมืองข้อความ การทำเหมืองความคิดเห็นและแบบจำลองคุณภาพซอฟต์แวร์ตามลำดับ

2.1.1 บทวิจารณ์หรือความคิดเห็น

ในการซื้อขายของหรือบริการผ่านทางเว็บไซต์ ข้อมูลที่จัดเตรียมสำหรับผู้ซื้อจึงเป็นส่วนสำคัญมากต่อการตัดสินใจ นอกจากข้อมูลเกี่ยวกับสินค้าที่ผู้ผลิตหรือผู้ขายให้เป็นข้อมูลที่ถูกต้องอ่านสนใจเป็นอันดับแรก ข้อมูลที่เกี่ยวกับความคิดเห็นหรือบทวิจารณ์นั้นมีความสำคัญในด้านสะท้อนความรู้สึกหรือมุมมองของผู้บริโภคที่มีประสบการณ์การใช้แล้ว เนื่องจากข้อมูลเหล่านั้นเปรียบเสมือนข้อมูลของการประเมินสินค้าหรือบริการโดยผู้ที่มีประสบการณ์ การให้บทวิจารณ์หรือความคิดเห็นมีหลากหลายโครงสร้างขึ้นอยู่กับแต่ละเว็บไซต์ จากการศึกษามากมายเว็บไซต์ เช่น <http://www.amazon.com> <http://www.cnet.com> <https://play.google.com/store> เป็นต้น ลักษณะของข้อมูลที่เกี่ยวข้องกับการให้ความคิดเห็นแบ่งเป็นสี่ส่วน

- 1) ข้อความหรือบทวิจารณ์ ข้อมูลที่เป็นข้อความตัวอักษรถูกเขียนขึ้นโดยผู้ที่มีประสบการณ์ บทวิจารณ์หรือความคิดเห็นอย่างละเอียดจะถูกเขียนในส่วนนี้ซึ่งเป็นส่วนที่สำคัญที่สุด เพื่อความชัดเจนในการกรอกข้อมูล เว็บไซต์อาจจะเตรียมช่องสำหรับกรอกข้อความหลายช่อง โดยแต่ละช่องมีจุดประสงค์ในการกรอกแตกต่างกันออกไป เช่น ช่องกรอกข้อความสำหรับใส่ข้อดีของสินค้า ช่องกรอกข้อความสำหรับใส่ข้อเสียของสินค้า ช่องกรอกข้อความสำหรับใส่ความคิดเห็นโดยสรุป เป็นต้น





2) ระดับความพึงพอใจ โดยทั่วไปจะแสดงเป็นการให้ดาวเพื่อแสดงความพึงพอใจ จำนวนดาวมาก แสดงถึงความพึงพอใจมาก ในทางกลับกันจำนวนดาวน้อยแสดงถึงความพึงพอใจน้อย ช่วงของจำนวนของดาวที่จะสามารถให้ได้จะขึ้นอยู่กับแต่ละเว็บไซต์เป็นผู้กำหนด จำนวนดาวที่แสดงความพึงพอใจส่วนมากจะมีการแสดงอยู่ควบคู่กับบทวิจารณ์หรือความเห็น เพราะเป็นข้อมูลที่แสดงให้ผู้ที่มาอ่านเข้าใจง่ายว่าผู้เขียนบทวิจารณ์หรือความคิดเห็นมีทัศนคติต่อสิ่งนั้นดีหรือไม่ดี

3) ข้อมูลเกี่ยวกับบทวิจารณ์ ข้อมูลทั้งหมดที่เกี่ยวกับการให้บทวิจารณ์ เช่น ชื่อของคนเขียนบทวิจารณ์ วันที่เขียนบทวิจารณ์ เป็นต้น ส่วนนี้ไม่ได้เกี่ยวกับการเขียนบทวิจารณ์โดยตรง แต่ในส่วนนี้จะส่งผลต่อความน่าเชื่อถือของบทวิจารณ์ ถ้าชื่อของผู้เขียนบทวิจารณ์เหมือนสแปมเมอร์ (spammer) บทวิจารณ์นี้อาจจะให้ข้อมูลไม่ตรงกับความเป็นจริง หรือถ้าบทวิจารณ์นี้ถูกเขียนเมื่อเวลานานมากแล้ว บทวิจารณ์นั้นอาจจะไม่น่าเชื่อถือ เพราะสินค้าหรือบริการนั้นมีโอกาสปรับปรุงแก้ไขในส่วนที่ไม่ดีที่บทวิจารณ์นี้กล่าวถึงเรียบร้อยแล้ว ในเว็บไซต์ต่างๆจะมีข้อมูลส่วนนี้เพื่อให้ผู้อ่านรู้ว่าใครเป็นผู้เขียนและบทวิจารณ์นี้เขียนขึ้นเวลาใด

4) อัตราส่วนความมีประโยชน์ (helpful ratio) อัตราส่วนความมีประโยชน์คือ อัตราส่วนของคนที่มาอ่านบทวิจารณ์นั้นเป็นผู้ให้คะแนนว่าบทวิจารณ์นั้นมีประโยชน์หรือไม่มีประโยชน์ ถ้าอัตราส่วนความมีประโยชน์สูงแสดงว่าบทวิจารณ์นั้นน่าจะมีความน่าเชื่อถือ ในทางกลับกันถ้าอัตราส่วนความมีประโยชน์ต่ำแสดงว่าบทวิจารณ์นั้นอาจจะไม่น่าเชื่อถือ เว็บไซต์ที่จัดเตรียมระบบข้อมูลลักษณะนี้ไม่ได้มีทุกเว็บไซต์ ตัวอย่างเว็บไซต์ที่จัดเตรียมระบบข้อมูลลักษณะนี้ ได้แก่ www.cnet.com www.amazon.com เป็นต้น

เว็บไซต์อีคอมเมิร์ซ (e-commerce website) หลายเว็บไซต์นั้นมีข้อมูลในส่วนของบทวิจารณ์ โดยงานวิจัยนี้จะให้ความสนใจในเว็บไซต์ www.cnet.com ซึ่งเป็นเว็บไซต์อีคอมเมิร์ซที่มีการขายซอฟต์แวร์และมีรูปแบบของบทวิจารณ์ที่มีประโยชน์โดยเฉพาะในส่วนของข้อความ ข้อความในการเขียนบทวิจารณ์ในเว็บไซต์นี้ถูกแบ่งออกเป็นสี่ส่วน คือ สรุปสั้น (one-line summary) ความคิดเห็นเชิงบวก (pros) ความคิดเห็นเชิงลบ (cons) และสรุป (summary) งานวิจัยนี้ใช้ข้อมูลในเว็บไซต์นี้เป็นชุดข้อมูลในการทดลอง ตัวอย่างของบทวิจารณ์ในเว็บไซต์ www.cnet.com แสดงในรูปที่ 2.1 และ 2.2

Add Your Review

Quickly sign in with:    Or  Log in or create an account to post a review.

1. Rate this product:
★★★★★

2. One-line summary: (10 characters minimum) 0 of 55 characters

3. Pros: (10 characters minimum) 0 of 1,000 characters

4. Cons: (10 characters minimum) 0 of 1,000 characters

5. Summary: (optional) 0 of 15,000 characters

Add Your Review The posting of advertisements, profanity, or personal attacks is prohibited. [Click here to review our site terms of use.](#)

รูปที่ 2.1 ตัวอย่างหน้าต่างสำหรับกรอกบทวิจารณ์บนเว็บไซต์ www.cnet.com [15]

★★★★★
"Easy to use"
December 14, 2011 | By aa885
Version: YouTube Downloader 3.4

Pros
Handy, works every time. Almost perfect. Thanks

Cons
Upgrade bother me a little bit, that's all.

Summary
What more can I say.. it's just nice.

[Reply to this review](#)

Was this review helpful? (1)  (0) 

รูปที่ 2.2 ตัวอย่างของบทวิจารณ์ในเว็บไซต์ www.cnet.com [15]

2.1.2 การทำเหมืองข้อความ (Text mining)

การทำเหมืองข้อความนั้นมีความแตกต่างจากการทำเหมืองข้อมูล (data mining) โดยการทำเหมืองข้อมูล คือการค้นหาคำหรือความรู้ รูปแบบ แนวทาง หรือความสัมพันธ์ในข้อมูล โดยข้อมูลจะมีโครงสร้างลักษณะที่ชัดเจน [16] เช่น รูปแบบข้อมูลที่เป็นตาราง เป็นต้น แต่การทำเหมืองข้อความคือ การค้นหาคำหรือความรู้ รูปแบบ แนวทางหรือความสัมพันธ์ในชุดข้อความนั้น ซึ่งชุดข้อความนั้นมีลักษณะโครงสร้างของข้อมูลไม่ชัดเจน การทำเหมืองข้อความถูกใช้ในการแก้ปัญหาที่หลากหลาย เช่น การจำแนกเอกสาร (document classification) การค้นคืนสารสนเทศ (information retrieval) การจัดกลุ่มและจัดระบบเอกสาร (clustering and organizing documents) การสกัดข้อมูล (information extraction) การพยากรณ์และประเมินผล (prediction and evaluation) เป็นต้น ความยากในการประมวลผลชุดข้อความคือ ความไร้โครงสร้างของข้อมูล เพื่อความสะดวกในการประมวลผล จึงต้องจัดการข้อความให้มีโครงสร้างที่สามารถนำไปคำนวณได้ เทคนิควิธีการประมวลผลนั้นสามารถดำเนินการได้หลายวิธีเพื่อช่วยจัดการข้อความให้เป็นข้อมูลที่เหมาะสมกับการนำไปประมวลผลให้เป็นข้อมูลที่มีโครงสร้าง เช่น

1) การกำหนดมาตรฐานของเอกสาร (document standardization) เป็นการกำหนดเอกสารให้มีรูปแบบที่ชัดเจนให้เหมือนกันในทุกเอกสารเพื่อสะดวกในการประมวลผล เช่น เอกสารหรือบทความหนึ่งจะประกอบด้วยชื่อผู้แต่ง หัวข้อ เนื้อหา เป็นต้น

2) การแยกคำศัพท์ (tokenization) เป็นการแยกคำศัพท์ในข้อความในเอกสารเพื่อทำให้เกิดความชัดเจนว่าข้อความนั้นประกอบด้วยคำศัพท์คำใดบ้าง โดยทั่วไปในภาษาอังกฤษ จะแยกคำศัพท์ด้วยการเว้นวรรค และตัดอักขระที่ไม่ใช่ตัวอักษร เช่น () < > ! ? “ เป็นต้น

3) การทำให้คำเป็นแก่นคำ (stemming) เป็นการลดรูปคำศัพท์ให้อยู่ในรูปแก่นคำ (stemmed word) โดยคำศัพท์ที่มีแก่นคำเดียวกันจะมีความสัมพันธ์หรือความหมายคล้ายกัน เช่น “stemmer” “stemming” “stemmed” มีแก่นคำคือ “stem” เป็นต้น วิธีการลดรูปมีหลากหลายวิธี โดยวิธีที่เป็นที่นิยมคือพอร์เตอร์อัลกอริทึม (porter stemming algorithm) [17] ที่ถูกใช้ในเครื่องมือลูซีน (Lucene) [18] ที่ใช้ในการจัดเก็บและค้นคืนข้อมูล

4) การแบ่งประโยค ใช้วิธีการหาจุดสิ้นสุดของประโยคในข้อความเพื่อเพิ่มประสิทธิภาพของการประมวลผลขั้นต่อไป โดยทั่วไปอักขระที่สามารถบ่งบอกจุดสิ้นสุดของประโยค ได้แก่ “.” “?” “!” เป็นต้น

5) การระบุหน้าที่ของคำ (part-of-speech tagging) เป็นส่วนที่ทำให้รู้ข้อมูลของคำศัพท์มากขึ้น ว่าคำศัพท์ที่อยู่ในประโยคนั้นมีหน้าที่เป็นคำนาม (noun) คำกริยา รูปพื้นฐาน (verb, base form) คำกริยาในรูปอดีต (verb, past tense) หรืออื่นๆ ส่วนนี้ นำความรู้เกี่ยวกับการประมวลผลภาษาธรรมชาติ (natural language processing) มาใช้ เครื่องมือที่ใช้ในส่วนนี้ ได้แก่ สแตนฟอร์ด ล็อกลิเนียร์พาร์ทออฟสปีชแท็กเกอร์ (Stanford log-linear part-of-speech tagger) [19] เป็นต้น

จากตัวอย่างเทคนิควิธีที่ช่วยในการเตรียมข้อความ การจัดการข้อมูลของข้อความให้เป็นโครงสร้างแนวทางหนึ่งที่ยอมรับใช้การจัดการข้อความคือ การใช้เวกเตอร์ การใช้เวกเตอร์เป็นตัวแทนข้อความจะมีการสร้างพจนานุกรม (dictionary) ก่อน ซึ่งโดยทั่วไปจะเป็นรายการคำศัพท์หรือเป็นส่วนที่บ่งบอกว่าความหมายในโครงสร้างเวกเตอร์ หลังจากนั้นจึงแปลงข้อความในเอกสารไปเป็นเวกเตอร์ตามพจนานุกรมที่สร้างขึ้น ในการสร้างพจนานุกรมหรือการแปลงเอกสารไปเป็นเวกเตอร์อาจมีการตัดคำศัพท์ที่น่าจะมีความสำคัญน้อยในเชิงการให้ความหมายออกด้วย เช่น คำที่ปรากฏบ่อยมากหรือน้อยมาก คำที่อยู่ในสตอปลิส (stop list) เช่น “a” “an” “the” เป็นต้น หรือบางครั้งอาจจะมีการถ่วงน้ำหนักของคำศัพท์ด้วย ตัวอย่างแบบจำลองรูปแบบหนึ่งของเวกเตอร์ที่นิยมใช้กันในการค้นคืนสารสนเทศคือ แบบจำลองปริภูมิเวกเตอร์ (vector space model) [20]

เวกเตอร์ของเอกสารและพจนานุกรมสามารถนำไปประยุกต์เพื่อสร้างตัวจำแนกเอกสาร วิธีการจำแนกเอกสารด้วยกฎ (rule-based classification) เป็นวิธีการจำแนกเอกสารวิธีหนึ่งที่เหมาะสมกับเอกสารที่มีจำนวนค่าน้อยหรือประโยคหนึ่งประโยค เป็นวิธีที่มีการสร้างกฎของการตัดสินใจ (decision rule) โดยเงื่อนไขของกฎคือคำศัพท์ ถ้าเอกสารใดมีคำศัพท์ตามกฎ เอกสารนั้นจะถูกจำแนกไปตามกฎนั้น ซึ่งกฎนี้จะถูกสร้างขึ้นจากการเรียนรู้เวกเตอร์ของเอกสารที่รู้แล้ว ตัวอย่างเครื่องมือที่ช่วยในการสร้างกฎ คือ เท็กซ์ไมเนอร์ซอฟต์แวร์คิต (Text-Miner Software Kit, TMSK) และ รูลอินดักชันคิตฟอร์เท็กซ์ (Rule Induction Kit for Text, RIKTEXT) ในส่วนของทีเอ็มเอสเค (TMSK) จะทำหน้าที่เตรียมข้อมูลจากข้อความ คือสร้างเวกเตอร์ของข้อความและพจนานุกรม และส่วนของริกเท็กซ์ (RIKTEXT) ทำหน้าที่สร้างกฎจากข้อมูลที่ได้จากทีเอ็มเอสเค [16]

2.1.3 การทำเหมืองความคิดเห็น (Opinion mining)

การทำเหมืองความคิดเห็นคือการค้นหาความคิดเห็น ความรู้สึกหรือจุดมุ่งหมายของบทวิจารณ์ เนื่องจากเป็นแนวความคิดที่ใหม่จึงยังไม่มีข้อกำหนดวิธีการหรือกระบวนการแนวคิดชัดเจน ส่วนใหญ่ผลจากการทำเหมืองความคิดเห็นนั้นเป็นข้อสรุปความคิดเห็นทั้งหมดทั่วไปในทางบวกหรือทางลบหรือข้อสรุปว่าความคิดเห็นเกี่ยวข้องกับอะไรบ้างและมีความคิดเห็นต่อสิ่งเหล่านั้นอย่างไร ปัจจัยในบทวิจารณ์ที่นำมาใช้ในการทำเหมืองความคิดเห็นมีสามปัจจัย [14] ปัจจัยแรกคือข้อมูลของบทวิจารณ์ที่เป็นตัวอักษร เป็นส่วนที่เป็นข้อความทั้งหมดของบทวิจารณ์ ปัจจัยที่สองคือข้อมูลของบทวิจารณ์ที่ไม่เป็นตัวอักษร เช่น จำนวนดาว ระดับความพึงพอใจ เป็นต้น และปัจจัยสุดท้ายคือคุณภาพของบทวิจารณ์ที่เป็นข้อมูลจากผู้อ่านบทวิจารณ์ให้ ทั้งสามปัจจัยส่งผลต่อรูปแบบของแนวคิดต่างๆที่นำเสนอในการทำเหมืองความคิดเห็น

ภาพรวมของแนวคิดในการทำเหมืองความคิดเห็น คือ การประมวลผลค้นหาสำหรับสิ่งที่ให้มาสร้างรายการของคุณลักษณะของสินค้า (คุณภาพ ลักษณะเฉพาะ อื่นๆ) และรวบรวมหรือประมวลความคิดเห็นที่เกี่ยวกับคุณลักษณะของสินค้านั้น (ไม่ดี กลาง ดี) [13] จากงานวิจัย แนวคิดพื้นฐานหลักของการทำเหมืองความคิดเห็น คือ การทำเหมืองเชิงขั้ว (polarity mining) เป็นแนวคิดที่พิจารณาส่วนของข้อความ ข้อความหรือบทวิจารณ์ว่ามีการแสดงความคิดเห็นไปในทางบวกหรือลบ ซึ่งเปรียบเสมือนการจำแนกเป็นสองกลุ่ม มีงานวิจัยหลายชิ้นเสนอวิธีคิดในการทำเหมืองเชิงขั้วที่ใช้หลากหลายปัจจัยในการพิจารณาความคิดเห็น แต่ในด้านกระบวนการสามารถแบ่งออกเป็นสองแบบคือ แบบใช้ข้อมูลสอน (supervised approach) และแบบไม่ใช้ข้อมูลสอน (unsupervised approach) ในแบบใช้ข้อมูลสอนจะเป็นการใช้วิธีการทำเหมืองข้อความ และการเรียนรู้ด้วยเครื่อง (learning machine) เช่น วิธีในงานวิจัย [6, 10, 11] และในส่วนของแบบไม่ใช้ข้อมูลสอนเป็นการใช้องค์ความรู้เกี่ยวกับการประมวลผลภาษาธรรมชาติ (natural language processing) และเครื่องมือที่เกี่ยวกับภาษา เช่น เวิร์ดเน็ต (WordNet) [21] เป็นต้น ซึ่งส่วนนี้จะมีตัวอย่างเครื่องมือที่พัฒนาขึ้นมาให้ใช้ เช่น เซนติเวิร์ดเน็ต (SentiWordNet) [22, 23] ที่พัฒนาเพิ่มเติมจากเวิร์ดเน็ตและเจนเนอรอลอินไควเรอร์ (General Inquirer) [24] เป็นต้น นอกจากนี้ในงานวิจัยจะกำหนดรายการของคำขึ้นจากการสังเกตโดยผู้เชี่ยวชาญและมีวิธีการขยายคำเพิ่มเติมด้วยวิธีการต่างๆ [7-9]

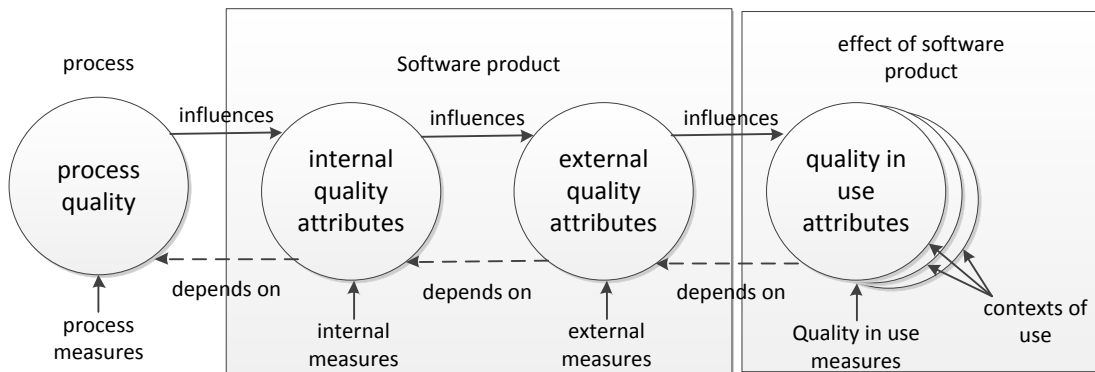
นอกจากแนวคิดในการทำเหมืองเชิงขั้วที่เป็นส่วนหนึ่งของการทำเหมืองความคิดเห็นแล้ว อีกแนวคิดที่น่าสนใจคือแนวคิดในการสกัดสิ่งที่สนใจหรือสรุปจากบทวิจารณ์ เพื่อรู้ว่าบทวิจารณ์นั้นกล่าวถึงอะไรบ้าง ซึ่งมีหลายงานวิจัยที่เสนอในแนวนี้ [3, 4, 25] โดยมีสองขั้นตอนคือการระบุลักษณะเฉพาะของสิ่งของว่ามีอะไรบ้าง แล้วหาคำแสดงความคิดเห็นต่อลักษณะเฉพาะของสิ่งของเหล่านั้น ซึ่งในสองขั้นตอนนี้ก็มีงานวิจัยที่เสนอวิธีไว้หลายวิธี [4, 6, 8]

2.1.4 แบบจำลองคุณภาพซอฟต์แวร์ (Software quality model)

จากไอเอสโอ 9000 (ISO 9000) [26] คุณภาพหมายถึง ระดับของกลุ่มคุณลักษณะที่เป็นไปตามที่คาดหวังไว้ ในรายละเอียดของคุณภาพซอฟต์แวร์จะขึ้นอยู่กับแต่ละแบบจำลองที่แต่ละแบบจำลองจะมีคุณลักษณะที่พิจารณาแตกต่างกันออกไป [27] ตัวอย่างแบบจำลองเช่น แบบจำลองคุณภาพของแมคคอลล (McCall's Quality Model) แบบจำลองคุณภาพของเบม (Boehm's Quality Model) แบบจำลองคุณภาพของโดรมีย์ (Dromey's Quality Model) แบบจำลองคุณภาพของ เอฟยูอาร์พีเอส (FURPS Quality Model) และแบบจำลองของไอเอสโอ 9126 (ISO 9126 Model) ในแบบจำลองทุกแบบจำลองนี้จะกำหนดคุณลักษณะของซอฟต์แวร์สำหรับพิจารณาในระหว่างพัฒนาหรือการทดสอบจริง เพื่อให้ซอฟต์แวร์ที่พัฒนาขึ้นมีคุณภาพ แต่มีเพียงแบบจำลองเดียวที่มีส่วนที่เน้นและระบุคุณลักษณะของคุณภาพในมุมมองของผู้ใช้งานโดยเฉพาะคือ แบบจำลองของไอเอสโอ 9126 ดังนั้นในงานวิจัยนี้จึงเลือกใช้แบบจำลองของไอเอสโอ 9126 ในการวัดคุณภาพซอฟต์แวร์

ไอเอสโอ 9126 [12] เป็นแบบจำลองที่เป็นมาตรฐานสำหรับการวัดคุณภาพของซอฟต์แวร์ที่ถูกกำหนดขึ้นจากองค์การมาตรฐานสากล (International Organization for Standardization) ในแบบจำลองนี้จะประกอบด้วยสองส่วน ส่วนแรกประกอบด้วยคุณภาพภายใน (internal quality) และคุณภาพภายนอก (external quality) ในส่วนนี้ได้กำหนดคุณลักษณะ (characteristic) หกคุณลักษณะและในแต่ละอันจะแยกออกเป็นคุณลักษณะย่อย (subcharacteristic) คุณภาพภายในและคุณภาพภายนอกได้ถูกกำหนดให้มีคุณลักษณะและคุณลักษณะย่อยเหมือนกัน แต่จะแตกต่างกันที่คุณภาพภายในจะถูกใช้วัดหรือประเมินผลในระหว่างการพัฒนา คุณภาพภายนอกจะถูกใช้วัดหรือประเมินผลขณะที่กำลังทดสอบในสถานการณ์จำลองด้วยข้อมูลทดสอบ อีกส่วนของแบบจำลองนี้คือคุณภาพในการใช้งาน (quality in use) ส่วนนี้จะวัดคุณภาพของซอฟต์แวร์ในมุมมองของผู้ใช้ ส่วนนี้ได้

กำหนดคุณลักษณะที่คุณลักษณะซึ่งมีความหมายต่างจากส่วนแรก ในการพัฒนาซอฟต์แวร์ ถ้าคุณภาพของซอฟต์แวร์ในระหว่างการพัฒนาดีก็ส่งผลให้คุณภาพของซอฟต์แวร์ในการทดสอบดีตาม และถ้าคุณภาพของซอฟต์แวร์ในการทดสอบดีจะส่งผลให้คุณภาพซอฟต์แวร์ขณะใช้งานจริงอาจดีตามด้วย โดยความสัมพันธ์ระหว่างสองส่วนนี้แสดงในรูปที่ 2.3



รูปที่ 2.3 คุณภาพของซอฟต์แวร์ในวงจรการพัฒนาซอฟต์แวร์ [12]

ในรายละเอียดส่วนแรกของแบบจำลอง คุณลักษณะและคุณลักษณะย่อยของคุณภาพภายใน และคุณภาพภายนอกมีชื่อเหมือนกัน ในงานวิจัยนี้จะให้ความสนใจกับคุณภาพภายนอกเนื่องจากงานวิจัยนี้พิจารณาคูณภาพซอฟต์แวร์ในขณะที่ซอฟต์แวร์ทำงาน คุณลักษณะของคุณภาพภายนอกมีรายละเอียดดังนี้

1) ฟังก์ชันการทำงาน (Functionality) หมายถึง ความสามารถของซอฟต์แวร์มีฟังก์ชัน (function) การทำงานที่ตรงตามความต้องการเมื่อซอฟต์แวร์ถูกใช้ในสถานการณ์ที่เฉพาะเจาะจง คุณลักษณะนี้ประเมินจากการทำงานของระบบที่อยู่ในซอฟต์แวร์ คุณลักษณะย่อยในคุณลักษณะนี้มีห้าคุณลักษณะย่อย คือ ความเหมาะสม (Suitability) ความแม่นยำ (Accuracy) ความสามารถในการทำงานร่วมกัน (Interoperability) ความปลอดภัย (Security) และความสอดคล้องกับมาตรฐานของฟังก์ชันการทำงาน (Functionality compliance)

2) ความเชื่อถือได้ (Reliability) หมายถึง ความสามารถของซอฟต์แวร์ที่รักษาการทำงานให้มีประสิทธิภาพอย่างสม่ำเสมอเมื่ออยู่ในสถานการณ์ที่เฉพาะเจาะจง ซอฟต์แวร์จะถูกวัดโดยพิจารณาลักษณะของพฤติกรรมการทำงานของระบบที่บอกความเชื่อถือได้ของระบบระหว่างการทำงาน คุณลักษณะย่อยในคุณลักษณะนี้มีสี่คุณลักษณะย่อย คือ วุฒิภาวะ (Maturity) ความทนทานต่อความ

ผิดพลาด (Fault tolerance) ความสามารถในการฟื้นคืน (Recoverability) และความสอดคล้องกับมาตรฐานของความเชื่อถือได้ (Reliability compliance)

3) การใช้งาน (Usability) หมายถึง ความสามารถของซอฟต์แวร์ที่ถูกเข้าใจและเรียนรู้วิธีการใช้งานได้ง่ายโดยผู้ใช้ อีกทั้งดึงดูดผู้ใช้เมื่ออยู่ในสถานการณ์ที่เฉพาะเจาะจง ซอฟต์แวร์ถูกวัดโดยการประเมินจากผู้พยายามใช้ซอฟต์แวร์ คุณลักษณะย่อยในคุณลักษณะนี้มีห้าคุณลักษณะย่อย คือ ความสามารถในการเข้าใจ (Understandability) ความสามารถในการเรียนรู้ (Learnability) ความสามารถในการทำงาน (Operability) ความน่าดึงดูด (Attractiveness) และ ความสอดคล้องกับมาตรฐานของการใช้งาน (Usability compliance)

4) ประสิทธิภาพ (Efficiency) หมายถึง ความสามารถของซอฟต์แวร์ที่มีประสิทธิภาพที่เหมาะสมกับทรัพยากรที่ถูกใช้ไปเมื่ออยู่ในสถานการณ์ที่เฉพาะเจาะจง ซอฟต์แวร์ถูกวัดโดยระยะเวลาที่เสียไปและการใช้ทรัพยากรให้มีประโยชน์ระหว่างทดสอบหรือการทำงาน คุณลักษณะย่อยในคุณลักษณะนี้มีสามคุณลักษณะย่อย คือ พฤติกรรมด้านเวลา (Time behavior) ความสามารถในการใช้ทรัพยากร (Resource utilization) และความสอดคล้องกับมาตรฐานของประสิทธิภาพ (Efficiency compliance)

5) ความสามารถในการดูแลรักษา (Maintainability) หมายถึง ความสามารถของซอฟต์แวร์ที่จะถูกปรับปรุงแก้ไข เพิ่มเติมหรือดัดแปลงซอฟต์แวร์ให้เปลี่ยนตามที่กำหนด ซอฟต์แวร์จะถูกวัดโดยพิจารณาพฤติกรรมผู้ที่แก้ไข ผู้ใช้งาน หรือระบบซอฟต์แวร์เมื่อซอฟต์แวร์ถูกแก้ไขระหว่างการทดสอบหรือการรักษาสภาพ คุณลักษณะย่อยในคุณลักษณะนี้มีห้าคุณลักษณะย่อย คือ ความสามารถด้านการวิเคราะห์ (Analyzability) ความสามารถในการปรับเปลี่ยน (Changeability) เสถียรภาพ (Stability) ความสามารถในการถูกทดสอบ (Testability) และความสอดคล้องกับมาตรฐานของความสามารถในการดูแลรักษา (Maintainability compliance)

6) ความสามารถในการเคลื่อนย้าย (Portability) หมายถึง ความสามารถของซอฟต์แวร์ที่สามารถย้ายจากสภาพแวดล้อมหนึ่งไปยังอีกสภาพแวดล้อมอื่นได้ ซอฟต์แวร์จะถูกวัดโดยพิจารณาพฤติกรรมของการทำงานของระบบตลอดการย้าย คุณลักษณะย่อยที่อยู่ในคุณลักษณะนี้มีห้าคุณลักษณะย่อย คือ ความสามารถในการปรับตัว (Adaptability) ความสามารถในการติดตั้ง (Installability) ความสามารถ

ในการทำงานร่วม (Co-existence) ความสามารถในการทดแทน (Replaceability) และความสามารถคล้อยกับมาตรฐานของความสามารถในการเคลื่อนย้าย (Portability compliance)

ในรายละเอียดส่วนที่สองของแบบจำลอง คุณภาพในการใช้งาน (Quality in use) [28] หมายถึง ความสามารถของซอฟต์แวร์ที่ทำให้ผู้ใช้ใช้งานได้ตามจุดมุ่งหมายอย่างมีประสิทธิภาพ (Effectiveness) ผลผลิตภาพ (Productivity) ความปลอดภัย (Safety) และ ความพึงพอใจ (Satisfaction) ลักษณะในคุณภาพในการใช้งานไม่มีการแบ่งเป็นคุณลักษณะย่อย โดยลักษณะจะมีทั้งหมดสี่ลักษณะ คือ

1) ประสิทธิภาพ (Effectiveness) หมายถึง ความสามารถของซอฟต์แวร์ที่ทำให้ผู้ใช้สามารถใช้งานได้ตามจุดมุ่งหมายด้วยความแม่นยำและถูกต้องในบริบทของการใช้ที่เฉพาะเจาะจง ซอฟต์แวร์จะถูกวัดที่ความแม่นยำและความถูกต้องของการทำงานของระบบ ในคุณลักษณะนี้มีมาตรวัดหลายมาตรวัด ตัวอย่างเช่น ประสิทธิภาพของการทำงาน ความสมบูรณ์ในการทำงาน ความถี่ของการเกิดข้อผิดพลาด เป็นต้น

2) ผลผลิตภาพ (Productivity) หมายถึง ความสามารถของซอฟต์แวร์ที่ทำให้ผู้ใช้ใช้ทรัพยากรอย่างมีประสิทธิภาพเพื่อให้ทำงานได้ตามจุดมุ่งหมายในบริบทของการใช้ที่เฉพาะเจาะจง ซอฟต์แวร์จะถูกวัดโดยวัดที่การใช้ทรัพยากร ในคุณลักษณะนี้มีมาตรวัดหลายมาตรวัด ตัวอย่างเช่น ระยะเวลาการทำงาน ประสิทธิภาพของการทำงาน อัตราส่วนทรัพยากรที่ใช้กับผลลัพธ์ที่ได้ เป็นต้น

3) ความปลอดภัย (Safety) หมายถึง ความสามารถของซอฟต์แวร์เกี่ยวกับระดับความเสี่ยงที่ยอมรับที่มีผลร้ายต่อคน ธุรกิจ ซอฟต์แวร์ ทรัพย์สินหรือสภาพแวดล้อมในบริบทของการใช้ที่เฉพาะเจาะจง ซอฟต์แวร์จะถูกวัดโดยการประเมินที่ระดับความเสี่ยงที่มีผลร้ายต่อคน ธุรกิจ ซอฟต์แวร์ ทรัพย์สินหรือสภาพแวดล้อม รวมถึงระดับความปลอดภัยในการใช้งาน ในคุณลักษณะนี้มีมาตรวัดหลายมาตรวัด ตัวอย่างเช่น ความปลอดภัยและสุขภาพของผู้ใช้ ความเสียหายทางด้านเศรษฐกิจ ความเสียหายของซอฟต์แวร์ เป็นต้น

4) ความพึงพอใจ (Satisfaction) หมายถึง ความสามารถของซอฟต์แวร์ที่สร้างความพึงพอใจให้ผู้ใช้ในการใช้งาน ซึ่งความพึงพอใจนี้รวมถึงทัศนคติของผู้ใช้ต่อการใช้ซอฟต์แวร์ด้วย ในคุณลักษณะนี้มี

มาตรวัดหลายมาตรวัด ตัวอย่างเช่น อัตราส่วนความพึงพอใจ แบบสอบถามความพึงพอใจให้ผู้ใช้กรอก เป็นต้น

2.2 งานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะนำแนวคิดในงานวิจัยที่ใกล้เคียงกับแนวคิดที่จะนำเสนอมาอธิบายปัจจัยสำคัญที่สำคัญและน่าสนใจ และสุดท้ายจะแสดงความแตกต่างของแนวคิดในงานวิจัยและแนวคิดที่นำเสนอ

2.2.1 ระบบแนะนำซึ่งมีความเชี่ยวชาญโดยการแนะนำนั้นพิจารณาจากบทวิจารณ์สินค้าของลูกค้า (Informed Recommender: Basing Recommendations on Consumer Product Reviews) [6]

งานวิจัยนี้เสนอกรอบงาน (framework) ของระบบแนะนำ (recommender system) ที่ใช้บทวิจารณ์ที่อยู่ในรูปแบบข้อความ กรอบงานที่นำเสนอมีข้อมูลนำเข้าคือบทวิจารณ์ กรอบงานนี้จะสร้างรายการสินค้าแนะนำจากคำขอของผู้ใช้ระบบ โดยกรอบงานนี้มีกระบวนการทำงานหลักอยู่สองกระบวนการ กระบวนการทำงานส่วนแรกเกี่ยวข้องการทำเหมืองข้อความ คือ บทวิจารณ์จะถูกแบ่งเป็นข้อความแล้วถูกจับคู่เข้ากับออนโทโลยี โดยออนโทโลยีที่กำหนดขึ้นแบ่งเป็นสองส่วนหลัก คุณภาพของความคิดเห็น (opinion quality) และคุณภาพของสินค้า (product quality) คุณภาพของความคิดเห็นคือข้อมูลเกี่ยวกับความสามารถหรือประสบการณ์ของผู้วิจารณ์ รายละเอียดในออนโทโลยีกำหนดขึ้นเองโดยผู้เชี่ยวชาญ การจับคู่ออนโทโลยีจะทำโดยอัตโนมัติหลังจากการเรียนรู้ของระบบ และแต่ละประโยคจะถูกจำแนกไปว่าเป็นประเภท “good” “bad” หรือ “quality” ด้วยวิธีการจำแนกด้วยกฎ (rule-based classification) กระบวนการทำงานในส่วนที่สองเกี่ยวกับการสร้างรายการสินค้าแนะนำ สินค้าที่ถูกแนะนำจะเกิดจากการประมวลผลจากข้อความคำร้องขอของผู้ใช้ที่สอดคล้องกับข้อมูลประสบการณ์ของผู้เขียนบทวิจารณ์และข้อมูลในบทวิจารณ์ ข้อมูลสองข้อมูลนี้เป็นข้อมูลสกัดจากกระบวนการแรก คือข้อมูลประโยคที่จับคู่กับออนโทโลยีและที่ถูกจำแนกแล้ว บทวิจารณ์ที่ใช้ในการทดลองในงานวิจัยนี้คือบทวิจารณ์เกี่ยวข้องกับกล้องถ่ายรูป

แนวคิดในงานวิจัยนี้มีความใกล้เคียงกับแนวคิดที่ได้นำเสนอในส่วนของการทำเหมืองข้อความ อย่างไรก็ตาม แนวคิดในงานวิจัยนี้มีจุดอ่อนในการสร้างออนโทโลยีที่มีการจับคู่ตรงที่ต้องกำหนดคำศัพท์ที่เกี่ยวข้องเองทั้งหมด เมื่อสินค้ามีจำนวนคุณสมบัติมากหรือรายการสินค้าที่ต้องนำมา

ประมวลผลมีความหลากหลาย ทำให้การหาค่าที่เกี่ยวข้องในการจับคู่ออนโทโลยีเป็นเรื่องยากมากขึ้น แนวคิดที่จะนำเสนอคือการกำหนดออนโทโลยีที่มีความชัดเจนตามมาตรฐานของคุณภาพซอฟต์แวร์ และมีการลงขยายออนโทโลยีหรือเพิ่มคำศัพท์โดยใช้เครื่องมือเวิร์ดเน็ต [21] ในการเข้ามาช่วยหาคำศัพท์ที่เกี่ยวข้องด้วยการหาค่าพ้องความหมายหรือคำความหมายตรงข้าม

2.2.2 การทำเหมืองเชิงขั้วโดยมีออนโทโลยีสนับสนุน (Ontology-Supported Polarity Mining) [10]

ในงานวิจัยนี้เสนอแนวคิดการทำเหมืองเชิงขั้วโดยมีออนโทโลยีสนับสนุน (Ontology-Supported Polarity Mining Approach) และจากการทดลอง ประสิทธิภาพของการทำเหมืองเชิงขั้วสามารถเพิ่มขึ้นด้วยการใช้ออนโทโลยี ในการทดลอง บทวิจารณ์จะถูกทดลองในสองแบบ แบบที่ถูกแบ่งเป็นส่วนข้อความตามออนโทโลยีและแบบที่ไม่ถูกแบ่ง ข้อมูลที่ใช้เป็นบทวิจารณ์ของภาพยนตร์ และการกำหนดออนโทโลยีเป็นการกำหนดโดยข้อมูลเกี่ยวกับภาพยนตร์โดยผู้เชี่ยวชาญสองคน หลังจากนั้นข้อความในบทวิจารณ์จะถูกแบ่งออกเป็นส่วนๆ ตามออนโทโลยีที่กำหนดขึ้นโดยผู้เชี่ยวชาญเป็นผู้ดำเนินการ เพื่อที่จะพิสูจน์ประสิทธิภาพของแนวคิดที่มีการแบ่งข้อความตามออนโทโลยีดีกว่าไม่มีการแบ่งข้อความทั้งในการทำเหมืองเชิงขั้วทั้งแบบมีผู้สอน (supervised) และแบบไม่มีผู้สอน (unsupervised) ในการประเมินผล การทำเหมืองเชิงขั้วแบบมีผู้สอนจะใช้วิธีการของแบบจำลองภาษาเอ็นแกรม (n-gram language modeling) สำหรับการทดสอบ โดยวิธีนี้จะใช้การเรียนรู้จากข้อมูลสอนแล้วนำมาพยากรณ์บทวิจารณ์ในข้อมูลทดสอบ ส่วนการทำเหมืองเชิงขั้วแบบไม่มีผู้สอนจะใช้เครื่องมือเจอนเนอร์อล อินควเออร์ [24] สำหรับการทดสอบ ในทั้งสองวิธีนี้ บทวิจารณ์จะถูกวิเคราะห์ว่าเป็นบทวิจารณ์ที่แสดงความคิดเห็นเชิงบวกหรือเชิงลบตามผลพยากรณ์ของส่วนข้อความส่วนมากที่ถูกแบ่งโดยออนโทโลยีที่กำหนด และเปรียบเทียบกับผลพยากรณ์ของข้อความทั้งหมดที่ไม่มีการแบ่งผลของการทำเหมืองเชิงขั้วแสดงให้เห็นว่าการทำเหมืองเชิงขั้วที่แบ่งข้อความออกเป็นส่วนๆตามออนโทโลยีมีความแม่นยำขึ้นในทั้งแบบมีผู้สอนและไม่มีผู้สอน

งานวิจัยนี้แสดงให้เห็นแนวคิดการนำออนโทโลยีเข้ามาช่วยเพิ่มความแม่นยำในการทำเหมืองเชิงขั้วสำหรับการหาความคิดเห็นที่เป็นไปทางเชิงบวกหรือเชิงลบ แนวคิดที่จะนำเสนอจึงนำแนวคิดออนโทโลยีเข้ามาใช้ แต่แนวคิดในงานวิจัยนี้การสร้างออนโทโลยีและการจับคู่ออนโทโลียังเป็นการทำ

ด้วยผู้เชี่ยวชาญ แนวคิดที่จะนำเสนอจึงถูกพัฒนาขึ้นเพื่อให้เป็นวิธีอัตโนมัติซึ่งสามารถทำได้ด้วยการหาคำศัพท์ที่เกี่ยวข้องกับออนโทโลยีที่ถูกกำหนดขึ้นเพื่อให้การจับคู่โดยไม่ใช่ผู้เชี่ยวชาญ การลองขยายออนโทโลยีหรือการเพิ่มคำศัพท์ที่พ้องความหมายเข้าไปในออนโทโลยีนั้นจะใช้เครื่องมือเวิร์ดเน็ตในแนวคิดที่จะนำเสนอ

2.2.3 การทำเหมืองความคิดเห็นที่ใช้ออนโทโลยีสำหรับบทวิจารณ์เกี่ยวกับภาพยนตร์ (Ontology Based Opinion Mining for Movie Reviews) [7]

ความคิดหลักในงานวิจัยนี้คือการพัฒนาการทำเหมืองความคิดเห็นเกี่ยวกับลักษณะเฉพาะ (feature) โดยใช้ออนโทโลยี เฟรมเวิร์คที่นำเสนอในงานวิจัยนี้ประกอบด้วยสี่ส่วนคือ ส่วนเตรียมข้อมูล (data preprocessing) ส่วนระบุลักษณะเฉพาะ (feature identification) ส่วนระบุเชิงขั้ว (polarity identification) และส่วนการวิเคราะห์ความรู้สึก (sentiment analysis)

ส่วนเตรียมข้อมูลมีการแยกคำศัพท์ (words segmentation) และการระบุหน้าที่ของคำ (part-of-speech tagging) ออนโทโลยีจะถูกนำมาช่วยเพิ่มความแม่นยำในการแยกคำศัพท์ในส่วนเตรียมข้อมูล และนอกจากนี้ยังช่วยเพิ่มความแม่นยำในการสกัดลักษณะเฉพาะ (feature extraction) ในส่วนระบุลักษณะเฉพาะ เนื่องจากงานวิจัยนี้เป็นการทำเหมืองเชิงขั้วของบทวิจารณ์ของภาพยนตร์ งานวิจัยนี้สนใจลักษณะเฉพาะของภาพยนตร์ที่เป็นคำนาม (noun) เท่านั้น การสร้างออนโทโลยีจะสร้างจากการกำหนดคำศัพท์ขึ้นมาเป็นคำศัพท์เริ่มต้นที่เป็นคำนาม แล้วอาศัยประโยคที่มีคำเชื่อมประโยค (conjunction word) และประโยคนั้นมีคำศัพท์ที่รู้ว่าเป็นลักษณะเฉพาะ คำนามใหม่ที่อยู่ประโยคนั้นจะถูกนำมาเพิ่มเป็นคำศัพท์ที่รู้ว่าเป็นลักษณะเฉพาะ ในการระบุเชิงขั้ว งานวิจัยนี้นำเครื่องมือเซนต์เวิร์ดเน็ต [22, 23] มาประยุกต์ใช้ โดยเสนอวิธีรวมค่าเชิงขั้วสามค่าของแต่ละคำศัพท์คนหนึ่งที่มียุทธศาสตร์หลายความหมาย และคำนวณจากลักษณะเฉพาะแต่ละลักษณะตามออนโทโลยีและเชิงขั้วของแต่ละส่วนที่วิเคราะห์มาได้ จะแสดงผลความคิดเห็นโดยรวมว่ามีความคิดเห็นทางบวกหรือทางลบ

งานวิจัยนี้สนับสนุนการทำเหมืองเชิงขั้วที่ใช้ออนโทโลยีเข้าช่วยในการแยกข้อความไปตามลักษณะเฉพาะที่ข้อความนั้นเกี่ยวข้อง ซึ่งสนับสนุนแนวคิดที่จะนำเสนอที่ใช้ออนโทโลยีเพื่อแยกข้อความไปตามลักษณะเฉพาะ แต่แนวคิดที่จะนำเสนอจะแตกต่างจากแนวคิดในงานวิจัยนี้คือการสร้างออนโทโลยี เนื่องจากงานวิจัยนี้พิจารณาบทวิจารณ์เกี่ยวกับภาพยนตร์และลักษณะเฉพาะที่เป็น

คำนาม แต่ในขณะที่คำศัพท์ในออนโทโลยีที่เกี่ยวข้องกับคุณภาพซอฟต์แวร์นั้นมีทั้งคำคุณศัพท์ (adjective) คำวิเศษ (adverb) คำนาม (noun) และกริยา (verb) ยกตัวอย่างเช่น “fast” “easy” “easily” “error” “install” แนวคิดในงานวิจัยนี้ไม่สามารถนำมาใช้ได้ ดังนั้นแนวคิดที่นำเสนอจะใช้ผู้เชี่ยวชาญและลองใช้เครื่องมือเวิร์ดเน็ตต์ช่วยในการสร้างออนโทโลยีแทน และส่วนที่แตกต่างกันอีกประการหนึ่งคือการใช้เซนต์เวิร์ดเน็ตต์ที่มีประสิทธิภาพนั้นจำเป็นต้องรู้ความหมายของคำนั้น เพราะค่าเชิงขั้วที่เครื่องมือให้มานั้นให้ตามความหมายของคำศัพท์ซึ่งคำศัพท์คำหนึ่งสามารถตีความได้หลากหลายความหมาย เหตุนี้ทำให้เป็นเรื่องยากที่จะใช้เครื่องมือนี้ช่วยในการตีความให้มีประสิทธิภาพ แนวคิดที่จะนำเสนอเลยนำวิธีการจำแนกด้วยกฎมาใช้แทน วิธีนี้มีความเหมาะสมกับข้อมูลที่นำมาทดสอบเนื่องจากโครงสร้างของบทวิจารณ์มีส่วนประกอบที่บอกว่าเป็นความคิดเห็นเชิงบวกและความคิดเห็นเชิงลบอยู่ ซึ่งสามารถนำข้อมูลส่วนนี้มาเป็นข้อมูลสอนได้

2.2.4 การทำเหมืองความคิดเห็นที่ใช้ลักษณะเฉพาะโดยกระบวนการที่แนะนำโดยออนโทโลยี (Ontology-Guided Approach to Feature-Based Opinion Mining) [5]

งานวิจัยนี้นำเสนอวิธีการทำเหมืองความคิดเห็นโดยนำความรู้เกี่ยวกับการประมวลผลภาษาธรรมชาติ (natural language processing) ความรู้เกี่ยวกับการวิเคราะห์ความรู้สึก (sentimental analysis) และเว็บเชิงความหมาย (semantic web) เข้ามาประยุกต์ใช้ วิธีการในงานวิจัยนี้แบ่งออกเป็นสี่ขั้นตอน ขั้นตอนแรกคือการใช้การประมวลผลภาษาธรรมชาติ ขั้นตอนนี้จะมีการแยกคำออกจากกันและระบุหน้าที่ของคำ ขั้นตอนที่สองคือการระบุลักษณะเฉพาะด้วยออนโทโลยี (ontology-based feature identification) ขั้นตอนนี้จะสร้างออนโทโลยีจากโครงสร้างข้อมูลเกี่ยวกับภาพยนตร์บนเว็บไซต์โดยใช้ความรู้เกี่ยวกับเว็บเชิงความหมาย จากออนโทโลยีที่รู้ว่าคำศัพท์ใดเป็นลักษณะเฉพาะของภาพยนตร์ เช่น “Actor” “Soundtrack” “Producer” เป็นต้น ในงานวิจัยนี้เชื่อว่า แต่ละลักษณะเฉพาะนั้นส่งผลต่อความคิดเห็นโดยรวมไม่เท่ากัน ตำแหน่งของคำในบทวิจารณ์และจำนวนคำหรือความถี่ที่เขียนเกี่ยวกับลักษณะเฉพาะนั้น ตำแหน่งตอนต้นและตำแหน่งตอนท้ายจะบ่งบอกความรู้สึกโดยรวมมากกว่าตำแหน่งตรงกลาง อีกทั้งยิ่งความถี่สูงก็จะมีผลต่อความรู้สึกโดยรวมมากกว่าความถี่น้อย จึงมีการคำนวณคะแนนของลักษณะเฉพาะที่ส่งผลต่อความคิดเห็นโดยรวม ต่อมาขั้นตอนที่สามคือขั้นตอนระบุเชิงขั้ว (polarity identification) ขั้นตอนนี้จะระบุว่าลักษณะเฉพาะนั้นถูกกล่าวถึงในมุมมองเชิงบวกหรือเชิงลบด้วยการใช้เครื่องมือเซนต์เวิร์ดเน็ตต์ [22, 23] คำที่ถูกนำมา

พิจารณาคือกริยา (verb) หรือคำคุณศัพท์ (adjective) ที่อยู่ติดกับลักษณะเฉพาะนั้นและใช้คำเชิงชี้ว่าที่ จากคำที่ความหมายร่วมกันส่วนมากที่ได้จากเครื่องมือ ขั้นตอนสุดท้ายคือการวิเคราะห์ความคิดเห็น ขั้นตอนนี้จะสรุปว่าลักษณะเฉพาะและภาพรวมของความคิดเห็นโดยรวมทั้งหมดในบทวิจารณ์เป็นไปในทางบวกหรือลบ

จากขั้นตอนในงานวิจัยนี้จะคล้ายกับขั้นตอนแนวคิดที่จะนำเสนอ คือ ขั้นแรกเป็นการเตรียม ข้อมูลให้พร้อมสำหรับขั้นตอนต่อไป ขั้นที่สองเป็นการกระบวนการเขียนกับลักษณะเฉพาะหรือออนโทโลยีเพื่อให้รู้ว่าในบทวิจารณ์นั้นเขียนอธิบายเกี่ยวกับอะไร ขั้นที่สามคือการวิเคราะห์หว่าเพื่อหาว่า ข้อความหรือลักษณะเฉพาะนั้นเป็นไปในทางเชิงบวกหรือลบ และขั้นสุดท้ายคือสรุปความคิดเห็น ทั้งหมด แต่ในรายละเอียดของแต่ละขั้นตอนแตกต่างกันอย่างสิ้นเชิง แนวคิดที่จะนำเสนอได้พัฒนาใน ส่วนการสร้างออนโทโลยีด้วยผู้เชี่ยวชาญและลองขยายคำที่มีความหมายใกล้เคียงโดยใช้เวิร์ดเน็ต โดย ไม่ได้อาศัยจากข้อมูลที่ถูกกำหนดไว้แต่แรกเพียงอย่างเดียว นอกจากนี้เรายังเลือกแนวคิดการจำแนก ด้วยกฎที่ต้องใช้ข้อมูลสอนแทนการใช้แทนเครื่องมือเช่นดีเวิร์ดเน็ต ที่สามารถให้คำเชิงชี้ของคำนั้นได้ทันที เนื่องจากคำในแต่ละบทวิจารณ์แต่ละประเภทอาจจะแสดงความคิดเห็นเชิงบวกหรือลบต่างกัน เช่น “long life time” และ “wait long time” จะเห็นว่าคำว่า “long” นั้นให้ความคิดเห็นที่ แตกต่างกันในส่วนสรุปมีความแตกต่างกันที่แนวคิดที่จะนำเสนอมีการสรุปนอกจากบทวิจารณ์บทเดียว จะมีการแสดงสรุปผลทุกบทวิจารณ์รวมกันเพื่อเป็นข้อมูลเพิ่มเติมที่เปรียบเสมือนระดับดาวหรือระดับ ความพึงพอใจโดยรวม

2.2.5 การทำเหมืองและสรุปบทวิจารณ์ของลูกค้า (Mining and Summarizing Customer Reviews) [8]

งานวิจัยนี้เสนอวิธีการสรุปความคิดเห็นด้วยวิธีการค้นหาลักษณะเฉพาะของสินค้าในบทวิจารณ์แล้วจึงหาว่าความคิดเห็นต่อลักษณะเฉพาะนั้นเป็นไปในทางบวกหรือลบ ผลที่ได้คือความคิดเห็นต่อลักษณะเฉพาะของสินค้าว่าเป็นบวกเป็นจำนวนเท่าไรและเป็นลบเป็นจำนวนเท่าไร ซึ่งจะมีผลแสดงในทุกลักษณะเฉพาะของสินค้า งานวิจัยนี้เป็นงานวิจัยต่อจาก [3] โดยงานวิจัยก่อนหน้านี้จะแสดงวิธีการค้นหาลักษณะเฉพาะของสินค้าออกมาจากบทวิจารณ์ และในงานวิจัยนี้จะนำวิธีนั้นมาใช้

และเพิ่มในส่วนการวิเคราะห์ว่าลักษณะเฉพาะของสินค้านั้นถูกกล่าวถึงไปในทางบวกหรือทางลบ
วิธีการที่งานวิจัยนี้เสนอมีการทำงานแบ่งออกได้เป็นเจ็ดส่วนดังนี้

1) ส่วนระบุหน้าที่ของคำ (part-of-speech tagging) ส่วนนี้จะระบุหน้าที่ของคำทุกคำในบท
วิจารณ์เป็นการเตรียมข้อมูลเพื่อส่วนอื่นต่อไป

2) ส่วนที่ระบุลักษณะเฉพาะที่เจอบ่อย ส่วนนี้จะระบุลักษณะเฉพาะที่เป็นคำนาม (noun) เท่านั้น
โดยมีการตรวจสอบลักษณะเฉพาะที่เป็นกลุ่มคำ (phrase) และการตรวจสอบความซ้ำซ้อนของคำด้วย

3) ส่วนค้นหาคำที่แสดงความคิดเห็น งานวิจัยนี้สนใจเชื่อว่าคำคุณศัพท์เท่านั้นคือคำที่บ่งบอกว่า
ลักษณะเฉพาะนั้นไปในทางบวกหรือลบ ส่วนนี้จึงสกัดคำคุณศัพท์ทุกคำจากประโยคที่มีการกล่าวถึงคำ
ที่เป็นลักษณะเฉพาะที่ถูกระบุจากส่วนก่อนหน้า

4) ส่วนที่ระบุว่าคำที่แสดงความคิดเห็นนั้นมีความหมายไปในทางบวกหรือลบ ในส่วนนี้จะใช้
เครื่องมือเวิร์ดเน็ตเข้าช่วยในประมวลผลโดยสร้างเครือข่ายของคำจากคำเริ่มต้นที่รู้ความหมายไป
ทางบวกหรือลบ แล้วจะใช้คำพ้องความหมาย (synonym) ในการเพิ่มคำที่มีความหมายไปในทาง
เดียวกัน และใช้คำตรงกันข้าม (antonym) ในการเพิ่มคำที่มีความหมายตรงกันข้าม ในงานวิจัยนี้
กำหนดคำเริ่มต้น 30 คำแล้วขยายเพิ่มคำขึ้น ผลลัพธ์จากส่วนนี้จะได้รายการคำคุณศัพท์ที่สามารถบอก
ได้ว่ามีความหมายไปทางบวกหรือลบ

5) ส่วนที่ระบุลักษณะเฉพาะที่ไม่พบบ่อย เนื่องจากลักษณะเฉพาะบางลักษณะไม่ได้ถูกพูดถึงบ่อย
การค้นหาลักษณะเฉพาะเหล่านั้นจึงต้องใช้คำที่แสดงความคิดเห็นเข้าช่วย คำที่แสดงความคิดเห็นอยู่
กับคำนามคำใด คำนามคำนั้นน่าจะเป็นลักษณะเฉพาะด้วย

6) ส่วนที่ระบุประโยคว่าประโยคนั้นมีความหมายไปในทางบวกหรือทางลบ ซึ่งจะประมวลผลจาก
จำนวนคำที่แสดงความคิดเห็นไปทางบวกและจำนวนคำที่แสดงความคิดเห็นไปทางลบพร้อมทั้งมีการ
พิจารณาคำนิเสธ ถ้าคำนิเสธใกล้คำใดคำนั้นจะมีความหมายตรงกันข้าม ถ้าจำนวนใดว่ามากกว่า
ประโยคจะมีความหมายไปในทางนั้น แต่ถ้าเท่ากันจะคิดตามลักษณะเฉพาะที่อยู่ในประโยคนั้น จำนวน
ลักษณะเฉพาะถูกบ่งบอกไปในทางบวกมากกว่าจำนวนลักษณะเฉพาะถูกบ่งบอกไปในทางลบ ประโยค
นั้นก็จะให้ความหมายไปในทางบวก และถ้าน้อยกว่าก็ให้ความหมายไปในทางลบ แต่ถ้าเท่ากันประโยค
นั้นจะถูกพิจารณาให้มีความหมายไปในทางเดียวกันกับประโยคก่อนหน้า

7) ส่วนที่สร้างข้อสรุป ส่วนนี้จะแสดงผลเรียงจากลักษณะเฉพาะที่มีความถี่สูงไปยังความถี่ต่ำ โดยแต่ละลักษณะเฉพาะที่แสดงจะพร้อมทั้งจำนวนประโยคที่ถูกพิจารณาว่าให้ความหมายไปในทางบวก และจำนวนประโยคที่ถูกพิจารณาว่าให้ความหมายไปในทางลบ นอกจากนี้ยังมีประโยคที่เกี่ยวกับลักษณะเฉพาะนั้นแสดงด้วย

งานวิจัยนี้จะเน้นลักษณะเฉพาะที่เป็นคำนามและคำที่แสดงความคิดเห็นเป็นคำคุณศัพท์เพียงเท่านั้น ในการที่จะหาคุณภาพของซอฟต์แวร์นั้น คำที่บ่งบอกคุณลักษณะของคุณภาพซอฟต์แวร์อาจจะไม่ได้อยู่ในรูปของคำนามเพียงอย่างเดียว เช่น “It is easy to use” เป็นต้น การสร้างออนโทโลยีในแนวที่จะนำเสนอจึงกำหนดจากมาตรฐานของซอฟต์แวร์แล้วลองขยายเพิ่มคำศัพท์ที่เกี่ยวข้องด้วยเวิร์ดเน็ต [21] อีกส่วนหนึ่งที่แนวคิดที่นำเสนอแตกต่างคือส่วนที่วิเคราะห์ประโยคว่ามีความหมายไปในทางบวกหรือทางลบ แนวคิดที่นำเสนอจะใช้คำศัพท์ที่เกี่ยวข้องกับความรู้สึกที่เสนอในงานวิจัยนี้ ประยุกต์กับการเรียนรู้จากข้อมูลสอนแล้วสร้างกฎขึ้นมา เนื่องจากบทวิจารณ์ที่นำมาใช้เป็นข้อมูลทดสอบนั้นมีโครงสร้างที่บอกว่าส่วนใดเป็นความคิดเห็นเชิงบวกและส่วนใดเป็นความคิดเห็นเชิงลบ ซึ่งการเรียนรู้จากข้อมูลสอนในบทวิจารณ์เฉพาะด้านใดด้านหนึ่งน่าจะให้ความแม่นยำมากกว่า

2.2.6 สรุปและวิเคราะห์แนวคิดในงานวิจัยต่างๆ

สรุปและวิเคราะห์แนวคิดในงานวิจัยต่างๆพร้อมทั้งแนวคิดที่จะนำเสนอ โดยแสดงถึงปัจจัยสำคัญในส่วนสำคัญของแนวคิดสามส่วน ส่วนแรกคือการสร้างออนโทโลยีหรือการระบุลักษณะเฉพาะว่าสิ่งที่บทวิจารณ์เขียนถึงนั้นมีสิ่งที่เกี่ยวข้องอะไรบ้าง ส่วนที่สองคือการหาว่าประโยคหรือส่วนของข้อความหรือทั้งบทวิจารณ์นั้นแสดงความคิดเห็นไปในทางบวกหรือลบ และส่วนสุดท้ายคือส่วนของประเภทบทวิจารณ์ที่แต่ละแนวคิดนำมาพิจารณา สรุปได้ดังตารางที่ 2.1

ตารางที่ 2.1 ปัจจัยสำคัญของแนวคิดในงานวิจัยต่างๆ

แนวคิด	การหาลักษณะเฉพาะและกำหนดออนโทโลยี	วิธีการทำเหมืองเชิงข้อความ	ประเภทข้อมูล
งานวิจัย 2.2.1 [6] (Debbie et al., 2007)	กำหนดออนโทโลยีขึ้นโดยผู้เชี่ยวชาญ	ใช้วิธีการจำแนกด้วยกฎ (rule-based classification) ซึ่งสร้างด้วยเครื่องมือที่เอ็มเอสเค (TMSK) และริกเท็กซ์ (RIKTEXT)	กล้องถ่ายภาพ
งานวิจัย 2.2.2 [10] (Lina & Pimwadee, 2007)	กำหนดและแยกส่วนของข้อความกับออนโทโลยีโดยผู้เชี่ยวชาญ	มีแสดงวิธีสองวิธีคือ 1) ใช้แบบจำลองภาษาเอ็นแกรม (N-gram Language modeling) โดยกำหนดกลุ่มคำตามแบบจำลองภาษาไบแกรม (bi-gram language model) แล้วนำไปเรียนรู้ แล้วจะได้เครื่องมือทดสอบ 2) ใช้เครื่องมือเจนเนอร์อลอินควเอร์รี่สำหรับหาว่าส่วนของข้อความนั้นแสดงความคิดเห็นในทางบวกหรือลบ	ภาพยนตร์
งานวิจัย 2.2.3 [7] (Lili & Chunping, 2009)	นำคำศัพท์ที่รู้เป็นคำศัพท์เริ่มต้นแล้ว ขยายเพิ่มคำศัพท์ด้วยการนำคำนามที่อยู่ในประโยค ที่มีคำเชื่อมประโยคกับคำนามที่อยู่ในกลุ่มคำศัพท์เริ่มต้น	ใช้เครื่องมือเซนติเวิร์ดเน็ตในการหาคำคุณศัพท์ที่มีความหมายไปทางบวกหรือทางลบ	ภาพยนตร์

ตารางที่ 2.1 ปัจจัยสำคัญของแนวคิดในงานวิจัยต่างๆ (ต่อ)

แนวคิด	การหาลักษณะเฉพาะและกำหนดออนโทโลยี	วิธีการทำเหมืองเชิงข้อความ	ประเภทข้อมูล
งานวิจัย 2.2.4 [5] (Isidro et al., 2011)	สร้างออนโทโลยีหรือกำหนดลักษณะเฉพาะจากเว็บเชิงความหมายที่มีลักษณะเฉพาะบอกอยู่ในโครงสร้างที่สามารถนำมาใช้ได้ของเว็บไอเอ็มดีบี (IMDB)	ใช้เครื่องมือเซนต์เวิร์ดเน็ตพร้อมวิธีคำนวณการให้คะแนนของแต่ละลักษณะเฉพาะไม่เท่ากัน เพราะลักษณะเฉพาะที่แสดงความรู้สึกของผู้เขียนบทวิจารณ์มากหรือน้อยนั้นมาจากปัจจัยสองปัจจัยคือ ตำแหน่งในบทวิจารณ์ของลักษณะเฉพาะนั้นและจำนวนครั้งที่ถูกกล่าวถึง	ภาพยนตร์
งานวิจัย 2.2.5 [8] (Minqing & Bing, 2004)	นำความถี่ของคำนำมาใช้เพื่อหาลักษณะเฉพาะ และหาคำที่ถูกลบโดยคำคุณศัพท์ที่เกี่ยวข้องกับการแสดงความคิดเห็น	กำหนดคำคุณศัพท์ที่รู้ความหมายแล้วใช้เวิร์ดเน็ตช่วยในการขยายด้วยคำพ้องความหมายและคำความหมายตรงข้าม เพื่อจะได้รายการของคำคุณศัพท์ที่แสดงความคิดเห็น	อุปกรณ์ และเครื่องใช้ไฟฟ้า
แนวคิดที่จะนำเสนอ	สร้างออนโทโลยีโดยพิจารณาจากแบบจำลองคุณภาพซอฟต์แวร์ที่มีมาตรฐาน และลองขยายเพิ่มเติมคำที่พ้องความหมายและคำความหมายตรงกันข้ามด้วยเครื่องมือเวิร์ดเน็ต	ใช้วิธีประยุกต์คำศัพท์ [8] กับการจำแนกด้วยกฎ (rule-based classification) ซึ่งสร้างด้วยเครื่องมือทีเอ็มเอสเค (TMSK) และริคเท็ค (RIKTEXT)	ซอฟต์แวร์

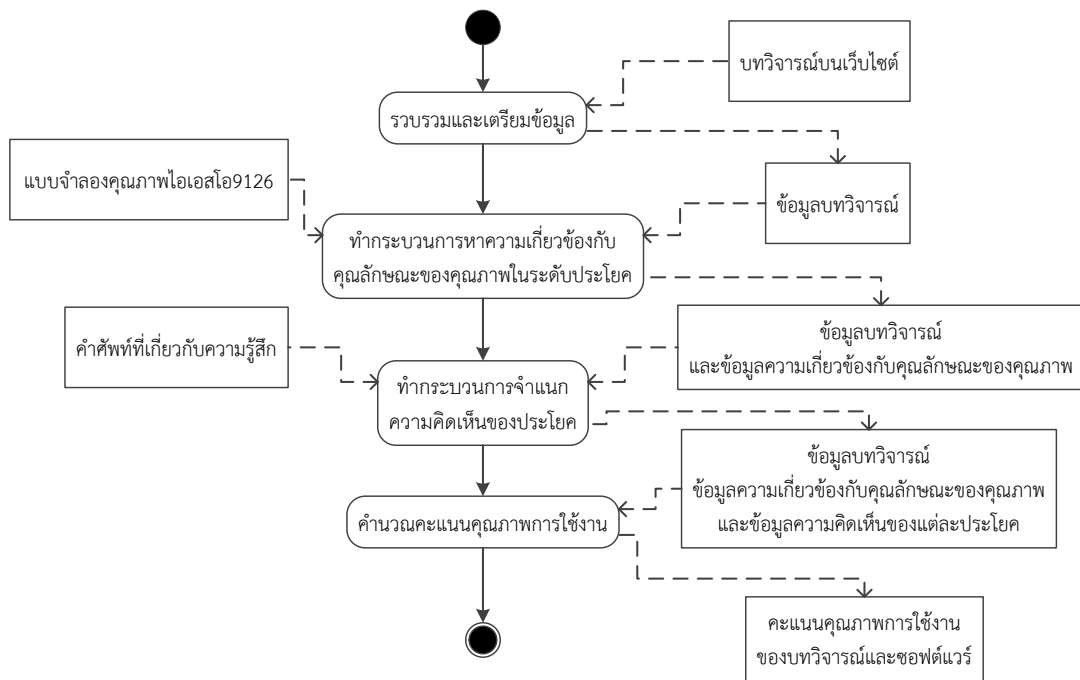
บทที่ 3

แนวคิดที่นำเสนอ

ในบทนี้จะกล่าวถึง กระบวนการทำเหมืองบทวิจารณ์ด้านคุณภาพซอฟต์แวร์ โดยเริ่มจาก การรวบรวมและการจัดเตรียมข้อมูล กระบวนการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของ คุณภาพการใช้งาน กระบวนการจำแนกความคิดเห็นของประโยค และการคำนวณคุณภาพการใช้งาน ตามลำดับ

3.1 ภาพรวมกระบวนการทำเหมืองบทวิจารณ์ด้านคุณภาพซอฟต์แวร์

กระบวนการทำเหมืองคุณภาพซอฟต์แวร์ที่ได้นำเสนอมีขั้นตอนการทำงานดังรูปที่ 3.1 กระบวนการจะมีข้อมูลนำเข้าสามส่วน ส่วนแรกคือข้อมูลทั้งหมดของบทวิจารณ์ ส่วนนี้จะเป็นส่วนที่ เก็บข้อมูลจากเว็บไซต์อีคอมเมิร์ซ เป็นข้อมูลตั้งต้นสำหรับการประมวลผลสำหรับประเมินคุณภาพ ซอฟต์แวร์ เว็บไซต์ที่อยู่ในขอบเขตคือ www.cnet.com ส่วนที่สองคือแบบจำลองคุณภาพซอฟต์แวร์ ไอเอสโอ 9126 ซึ่งนำมาจากเอกสารที่ออกโดยองค์การมาตรฐานสากล ส่วนนี้นำมาใช้สำหรับสร้างออน โทโลยีสำหรับจำแนกคุณภาพออกเป็นคุณภาพด้านต่างๆ ในขั้นต้นของกระบวนการ ส่วนที่สามคือ รายการคำศัพท์ที่แสดงความคิดเห็นเชิงบวกหรือเชิงลบ (sentiment words) ส่วนนี้นำมาสร้าง เครื่องมือจำแนกประโยคว่าประโยคเป็นประโยคเชิงบวกหรือลบ ขั้นตอนหลักของกระบวนการทำ เหมืองคุณภาพซอฟต์แวร์ที่ได้นำเสนอมีสี่กระบวนการ



รูปที่ 3.1 แผนภาพกิจกรรมของภาพรวมของแนวคิดที่นำเสนอ

1) รวบรวมและจัดเตรียมข้อมูล ในขั้นตอนนี้ทำหน้าที่สองอย่าง อย่างแรกคือเก็บรวบรวมข้อมูลที่เกี่ยวข้องกับซอฟต์แวร์รวมถึงข้อมูลบทวิจารณ์ของซอฟต์แวร์นั้นจากเว็บไซต์ ในกรณีของข้อมูลที่เรานำมาทดสอบนี้ ข้อมูลถูกรวบรวมมาจาก www.cnet.com และหน้าที่อย่างที่สองคือการเตรียมข้อมูลที่ได้รวบรวมมานั้นให้พร้อมหรือเหมาะสมเพื่อที่จะนำไปใช้ในส่วนต่อไป

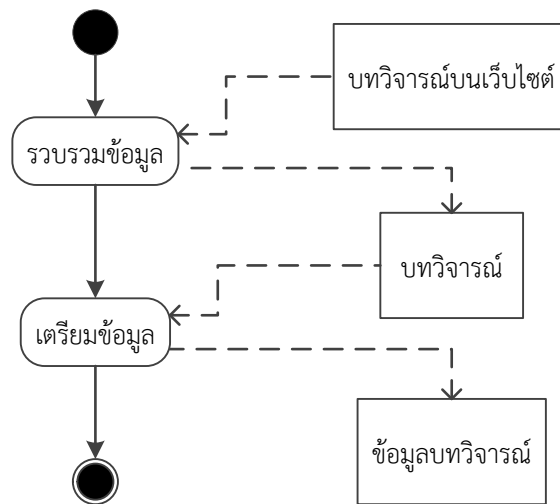
2) ทำกระบวนการหาความเกี่ยวข้องกับคุณลักษณะของคุณภาพ ขั้นตอนนี้ทำการหาความเกี่ยวข้องของข้อมูลระดับประโยคในบทวิจารณ์ว่าประโยคใดมีความเกี่ยวข้องกับคุณลักษณะด้านใดของคุณภาพ โดยอ้างอิงคุณภาพการใช้งาน (quality in use) จากแบบจำลองไอเอสไอ 9126

3) ทำกระบวนการจำแนกความคิดเห็นของประโยค ในขั้นตอนนี้ทำหน้าที่หาว่าข้อมูลระดับประโยคในบทวิจารณ์ว่าประโยคนั้นมีความคิดเห็นไปเชิงบวกหรือลบ ด้วยแนวคิดที่นำเสนอและคำศัพท์เกี่ยวกับความรู้สึก (sentiment words)

4) คำนวณคะแนนคุณภาพในการใช้งาน ในขั้นตอนนี้ทำการคำนวณคะแนนคุณภาพของบทวิจารณ์จากประโยคทุกประโยคที่ถูกระบุว่าเกี่ยวข้องกับคุณลักษณะด้านใดของคุณภาพและประโยคนั้นมีความคิดเห็นเชิงบวกหรือเชิงลบ และคำนวณคะแนนคุณภาพการใช้งานของซอฟต์แวร์โดยรวมจากคะแนนคุณภาพการใช้งานของบทวิจารณ์

3.2 การรวบรวมและจัดเตรียมข้อมูล

ส่วนนี้ประกอบด้วยสองขั้นตอน ขั้นตอนแรกคือการเก็บข้อมูลเกี่ยวกับซอฟต์แวร์และบทวิจารณ์จากเว็บไซต์ และขั้นตอนที่สองคือการเตรียมข้อมูลจากการวิเคราะห์แต่ละประโยคในบทวิจารณ์ โดยข้อมูลของแต่ละประโยคที่ถูกวิเคราะห์จะถูกกำหนดขึ้นจะมีรายละเอียดในลำดับถัดไป ขั้นตอนของส่วนนี้แสดงดังรูปที่ 3.2



รูปที่ 3.2 แผนภาพกิจกรรมของการรวบรวมและจัดเตรียมข้อมูล

1) การเก็บข้อมูลเกี่ยวกับซอฟต์แวร์และบทวิจารณ์จากเว็บไซต์ ข้อมูลเกี่ยวกับซอฟต์แวร์และบทวิจารณ์จะเป็นข้อมูลดิบ (raw data) ที่ต้องจัดเก็บในขั้นแรกของกระบวนการ ในส่วนข้อมูลเกี่ยวกับซอฟต์แวร์ แต่ละซอฟต์แวร์จะถูกเก็บข้อมูลสามอย่างคือ ชื่อซอฟต์แวร์ ประเภทของซอฟต์แวร์และระดับความพึงพอใจโดยรวมของซอฟต์แวร์ ในส่วนบทวิจารณ์ ข้อมูลจากแต่ละบทวิจารณ์จะถูกเก็บทำรายการคือ ระดับความพึงพอใจต่อซอฟต์แวร์ สรุปย่อ (one-line summary) ความคิดเห็นเชิงบวก (pros) ความคิดเห็นเชิงลบ (cons) และสรุป (summary)

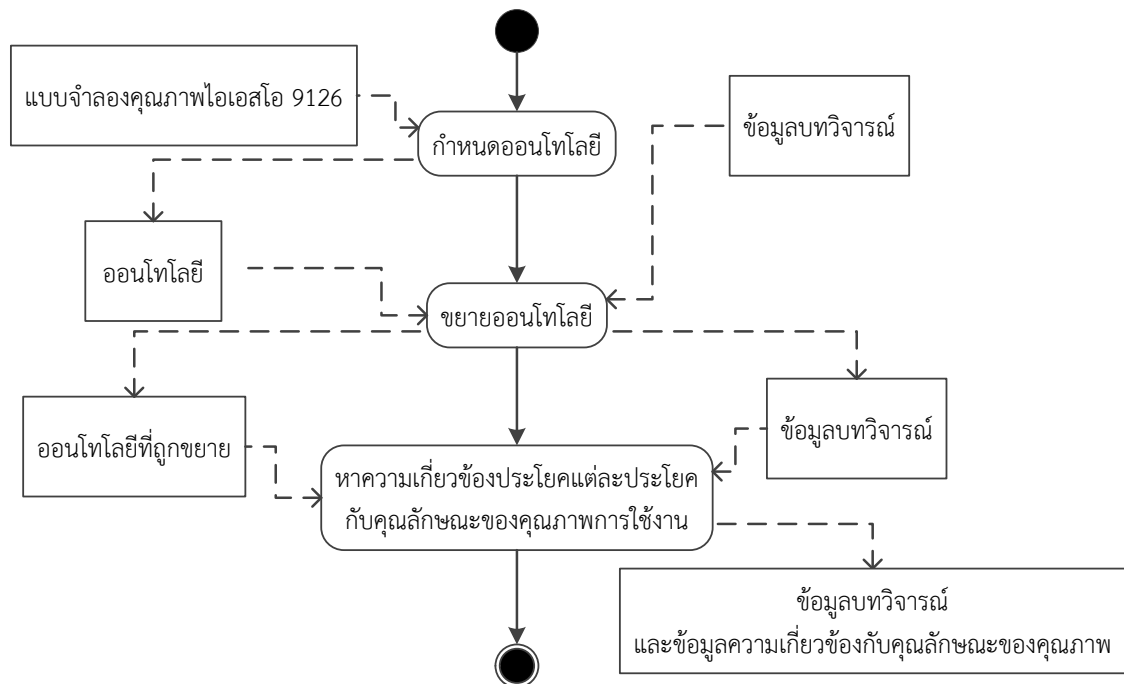
2) การเตรียมข้อมูลจากการวิเคราะห์ประโยคในบทวิจารณ์ ประโยคในบทวิจารณ์จะถูกแยกโดย “.” “?” “!” และแต่ละประโยคจะถูกแยกเป็นรายการคำศัพท์ รายการคำศัพท์นี้จะถูกตัดคำศัพท์ประเภทสตอปลิส (stop list) เครื่องมือที่ใช้ในการทำส่วนนี้คือ ลูซีน (Lucene) ตัวอย่างประโยคหลังผ่านการเตรียมข้อมูลแสดงในตารางที่ 3.1 จากประโยคตัวอย่างมีคำศัพท์ประเภทสตอปลิสคือ “it” “is” “to” “and”

ตารางที่ 3.1 ตัวอย่างรายการคำศัพท์หลังการตัดคำ

ประโยค	คำศัพท์ (ข้อมูลของประโยค)
It is easy to use.	“easy” “use”
Slow and very hard to understand	“slow” “very” “hard” “understand”

3.3 กระบวนการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน

จุดประสงค์ของกระบวนการนี้คือการหาว่าแต่ละประโยคในบทวิจารณ์นั้นเกี่ยวข้องกับคุณลักษณะด้านใดของคุณภาพ ในกระบวนการนี้ประกอบด้วยสามขั้นตอนดังรูปที่ 3.3

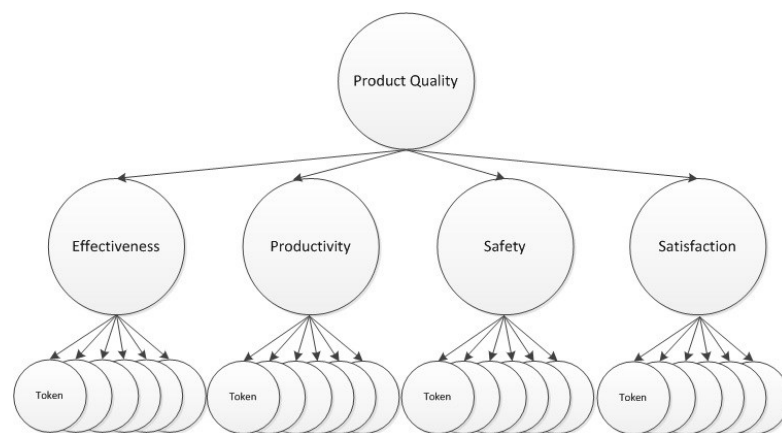


รูปที่ 3.3 แผนภาพกิจกรรมของกระบวนการหาความเกี่ยวข้องกับคุณลักษณะของคุณภาพในระดับประโยค

1) การกำหนดออนโทโลยี (ontology definition) คุณภาพในการใช้งาน (quality in use) ในแบบจำลองคุณภาพซอฟต์แวร์ไอเอสโอ 9126 ถูกพิจารณาเพื่อสร้างออนโทโลยีในตอนแรก เพราะคุณภาพในการใช้งานเป็นคุณภาพที่แสดงในมุมมองของผู้ใช้งาน ซึ่งสอดคล้องกับบทวิจารณ์เป็นความคิดเห็นในมุมมองของผู้ใช้งานเช่นกัน การกำหนดออนโทโลยีตอนแรกจะพิจารณาจากรายละเอียด

และมาตรวัดของคุณภาพในการใช้งานเพื่อมาสร้างรายการคำศัพท์ คุณภาพที่นำมาใช้เพิ่มคือ คุณภาพภายนอก (external quality) ในแบบจำลองคุณภาพซอฟต์แวร์ไอเอสโอ 9126 ซึ่งมีความเกี่ยวข้องกับคุณภาพในการใช้งาน [12, 28, 29] ทั้งโดยตรงและทางอ้อม แต่ในการพิจารณาบทวิจารณ์กับคุณลักษณะนั้นจะอ้างอิงจากมาตรวัดที่ใช้ในการวัดคุณลักษณะนั้นๆ คุณลักษณะบางคุณลักษณะในคุณภาพภายนอกที่มีมาตรวัดลักษณะใกล้เคียงกับคุณลักษณะของคุณภาพการใช้งาน จะถูกพิจารณาในรายละเอียดเข้าไปเพิ่มคำศัพท์ในออนโทโลยีด้วย

2) การขยายออนโทโลยี (ontology expansion) ขั้นตอนนี้ใช้เครื่องมือเวอร์ดเน็ต 3.0 เป็นอรรถาภิธาน (thesaurus) และข้อมูลประโยชน์จากการวิเคราะห์ที่ในขั้นตอนเตรียมข้อมูล ออนโทโลยีที่กำหนดในขั้นตอนย่อยแรกจะถูกขยายหรือเพิ่มคำศัพท์จากคำพ้องความหมาย (synonym) และคำตรงกันข้าม (antonym) เพื่อให้ได้คำที่เกี่ยวข้องมากที่สุดและครอบคลุมถึงคำที่มีความหมายใกล้เคียงกัน คำศัพท์ที่อยู่ในออนโทโลยีที่มีแก่นคำ (stemmed word) เหมือนกันกับแก่นคำของคำศัพท์ที่อยู่ในบทวิจารณ์จะถูกนำมาขยายเพิ่มคำศัพท์ในออนโทโลยีต่อ ออนโทโลยีจะแสดงดังรูปที่ 3.4



รูปที่ 3.4 ออนโทโลยี

ให้ $W_i = \{w_{i0}, w_{i1}, w_{i2}, \dots, w_{in}\}$ เป็นเซตของคำศัพท์ของคุณลักษณะ i ของคุณภาพในการใช้งาน (ประสิทธิผล (effectiveness), ผลิตผล (productivity), ความปลอดภัย (safety) และความพึงพอใจ (satisfaction)) ในออนโทโลยีตอนแรก และ $R = \{r_0, r_1, r_2, \dots, r_n\}$ เป็นเซตของแก่นคำของคำศัพท์ในบทวิจารณ์และ $F_w = \{f \mid f \text{ is token that is synonym or antonym of } w\}$ $S_i = \{s_{i0}, s_{i1}, s_{i2}, \dots, s_{im}\}$ เป็นเซตของคำศัพท์ของคุณลักษณะ i ของคุณภาพการใช้งานในออนโทโลยีที่ถูกขยายแล้ว การขยายออนโทโลยีแสดงในอัลกอริทึม 1.

Algorithm 1 Ontology Expansion

 Require: Set of tokens in initial ontology W , Set of tokens in reviews R

 Return: Sets of tokens in ontology S

begin:

 $S \leftarrow W$
 $count \leftarrow 0$
 $i \leftarrow 1$

 while $i \leq 4$ do // for 4 characteristics

 while $count < NumberOfToken(S_i)$ do

 if $Stem(S_{i,count}) \in R$ then

 $S_i \leftarrow append F_{S_{i,count}} to S_i$

end if

 $count \leftarrow count + 1$

end while

 $i \leftarrow i + 1$

end while

 return S

 end

3) หาความเกี่ยวข้องประโยคแต่ละประโยคกับคุณลักษณะของคุณภาพในการทำงานด้วยออนโทโลยี ถ้าแก่นคำของคำศัพท์ที่อยู่ในประโยคนั้นเหมือนกับแก่นคำของคำศัพท์ที่อยู่ในออนโทโลยี ประโยคนั้นก็จะถูกระบุว่ามีความเกี่ยวข้องกับคุณลักษณะที่มีคำศัพท์นั้นอยู่ ประโยคหนึ่งประโยคมีความเกี่ยวข้องได้มากกว่าหนึ่งคุณลักษณะ การหาแก่นคำของคำศัพท์จะใช้พอร์ตเตอร์อัลกอริทึม (Porter's stemming algorithm) [17]

3.4 กระบวนการจำแนกความคิดเห็นของประโยค

กระบวนการจำแนกความคิดเห็นของประโยคมีสองแบบ คือกระบวนการจำแนกความคิดเห็นของประโยคแบบพื้นฐานและกระบวนการจำแนกความคิดเห็นของประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยมีรายละเอียดดังนี้

3.4.1 กระบวนการจำแนกความคิดเห็นของประโยคแบบพื้นฐาน

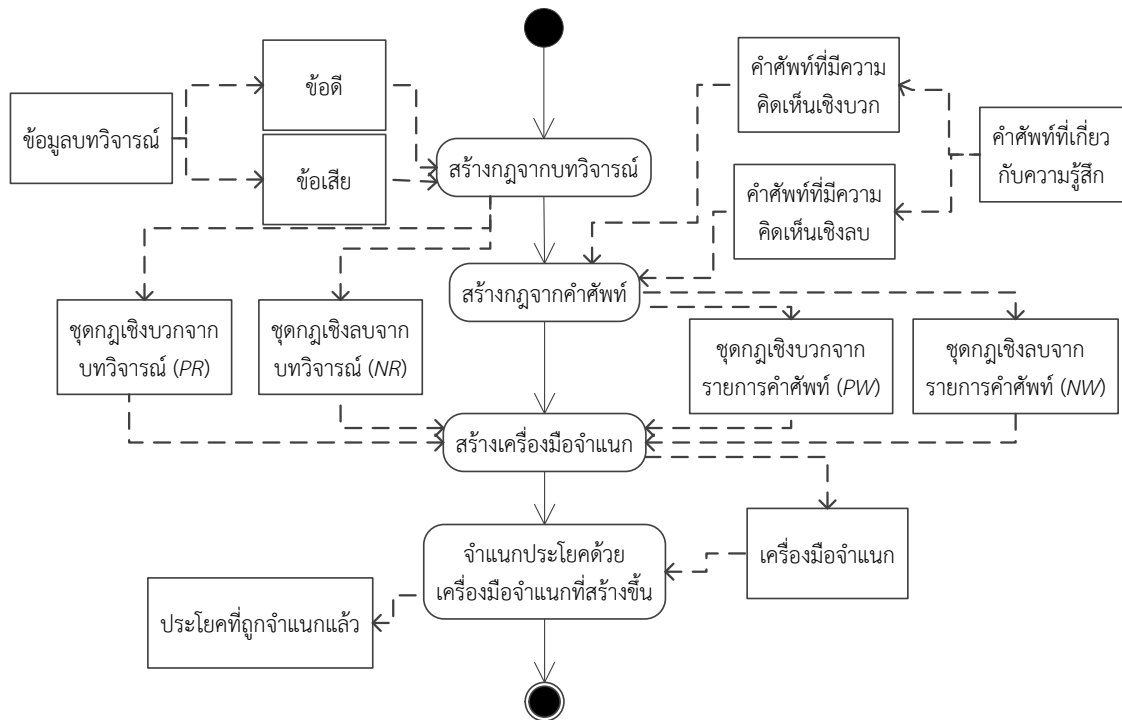
กระบวนการจำแนกความคิดเห็นของประโยคแบบพื้นฐานประกอบด้วยเครื่องมือจำแนกสองอย่าง เครื่องมือจำแนกอย่างแรกสร้างจากชุดกฎที่เป็นเชิงบวก แต่ละกฎสร้างจากคำศัพท์แต่ละคำศัพท์เกี่ยวกับความรู้สึกเชิงบวก ประโยคนั้นตรงตามกฎก็ต่อเมื่อประโยคประกอบด้วยคำศัพท์เกี่ยวกับความรู้สึกเชิงบวก เครื่องมือจำแนกอีกอย่างหนึ่งสร้างจากชุดกฎที่เป็นเชิงลบและแต่ละกฎสร้างจากคำศัพท์แต่ละคำศัพท์เกี่ยวกับความรู้สึกเชิงลบ ถ้าประโยคตรงตามกฎใดกฎหนึ่งในชุดกฎของประเภทนั้น ประโยคจะถูกจำแนกไปยังประเภทนั้น ผลของการจำแนกประโยคด้วยสองเครื่องมือจำแนกแสดงในตารางที่ 3.2 ในกรณีที่ประโยคประกอบด้วยคำนิเสธ เช่น “not” “don’t” “never” เป็นต้น ผลของการจำแนกประโยคนั้นจะตรงกันข้ามกัน ในตาราง ค่าประโยคที่มีความคิดเห็นเชิงบวกเท่ากับ “1” ค่าประโยคที่มีความคิดเห็นเป็นกลาง “0” ค่าประโยคที่มีความคิดเห็นเชิงลบ “-1” ยกตัวอย่างเช่น ประโยค “It does not have a protection.” เครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกมีกฎที่มีคำศัพท์ตรงกับคำศัพท์ในประโยคคือ “protection” แต่เครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบไม่มีกฎที่มีคำศัพท์ที่ตรงกับคำศัพท์ในประโยค ผลจากการจำแนกความคิดเห็นของประโยคเบื้องต้นคือเป็นประโยคที่มีความคิดเห็นเชิงบวก แต่ในประโยคมีคำนิเสธคือ “not” ทำให้ผลสุดท้ายจากจำแนกความคิดเห็นของประโยคคือเป็นประโยคที่มีความคิดเห็นเชิงลบ

ตารางที่ 3.2 ผลของการจำแนกประโยคด้วยสองเครื่องมือจำแนก

เครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวก	เครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบ	ค่าความคิดเห็นของประโยค
มีความคิดเห็นเชิงบวก	มีความคิดเห็นเชิงลบ	0
มีความคิดเห็นเชิงบวก	ไม่มีความคิดเห็นเชิงลบ	1
ไม่มีความคิดเห็นเชิงบวก	มีความคิดเห็นเชิงลบ	-1
ไม่มีความคิดเห็นเชิงบวก	ไม่มีความคิดเห็นเชิงลบ	0

3.4.2 กระบวนการจำแนกความคิดเห็นของประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน

เพื่อระบุว่าประโยคนั้นมีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลาง ส่วนนี้ประกอบด้วยสี่ขั้นตอน ซึ่งอธิบายตามรูปที่ 3.5



รูปที่ 3.5 แผนภาพกิจกรรมของกระบวนการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน

1) สร้างกฎจากบทวิจารณ์ โดยสร้างจากประโยคในข้อดีข้อเสียของบทวิจารณ์ เครื่องมือที่คามาเนอร์ซอฟต์แวร์คิด (Text-Miner Software Kit, TMSK) และเครื่องมือรูลอินดักชันคิตฟอร์เท็กซ์ (Rule Induction Kit for Text, RIKTEXT) ถูกนำมาใช้ในขั้นตอนนี้ ทีเอ็มเอสเค (TMSK) สร้างพจนานุกรม (dictionary) จากประโยคในข้อดี (pros) และข้อเสีย (cons) และปรับเปลี่ยนประโยคเหล่านั้นให้อยู่ในรูปเวกเตอร์ (vector) ตามพจนานุกรม พจนานุกรม เครื่องมือริคเท็กซ์ (RIKTEXT) จะใช้พจนานุกรมและเวกเตอร์ที่เป็นตัวแทนของประโยคที่มีความคิดเห็นเชิงบวกและที่ไม่มีความคิดเห็นเชิงบวกสำหรับเรียนรู้เพื่อสร้างกฎขึ้นมาสำหรับการจำแนก ประโยคในข้อดีถูกสมมติให้เป็นประโยคที่มีความคิดเห็นเชิงบวก ในทางตรงกันข้าม ประโยคในข้อเสียถูกสมมติให้เป็นประโยคที่มีความคิดเห็นเชิงลบ ดังนั้นเวกเตอร์ที่เป็นตัวแทนของประโยคที่มีความคิดเห็นเชิงบวกจะมาจากประโยคที่อยู่ในข้อดี ในทางตรงกันข้ามเวกเตอร์ที่เป็นตัวแทนของประโยคที่ไม่มีความคิดเห็นเชิงบวกจะมาจากประโยคที่อยู่ในความคิดเห็นข้อเสีย ดังนั้นผลที่ได้จากเครื่องมือนี้คือชุดกฎที่ไว้สำหรับจำแนกประเภทประโยคที่มีความคิดเห็นเชิงบวกกับที่ไม่มีความคิดเห็นเชิงบวกหลายชุดและกฎหนึ่งชุดที่มีความผิดพลาดน้อยจะถูกลเลือก ชุดกฎที่ไว้สำหรับจำแนกประเภทประโยคที่มีความคิดเห็นเชิงลบกับที่ไม่มีความคิดเห็นเชิงลบก็สร้างด้วยวิธีเดียวกันแตกต่างกันที่เวกเตอร์ที่มีความคิดเห็นเชิงลบจะมาจากประโยคที่อยู่ในข้อเสีย ส่วนเวกเตอร์ที่

เป็นตัวแทนของประโยคที่ไม่มีความคิดเห็นเชิงลบจะมาจากประโยคที่อยู่ในข้อดี กฎที่สร้างอยู่ในรูป $x \rightarrow y$ ประโยคที่ตรงตามกฎ x จะถูกจำแนกไปยังประเภท y ในกรณีนี้ การตรงตามกฎหมายถึงประโยคที่ประกอบด้วยคำศัพท์ในกฎ รายการคำศัพท์ที่ได้มานั้นมีความเฉพาะกับบทวิจารณ์นั้นๆ ผลลัพธ์ที่ได้จากขั้นตอนนี้คือ ชุดกฎเชิงบวกที่มาจากบทวิจารณ์ (PR) และชุดกฎเชิงลบที่มาจากบทวิจารณ์ (NR)

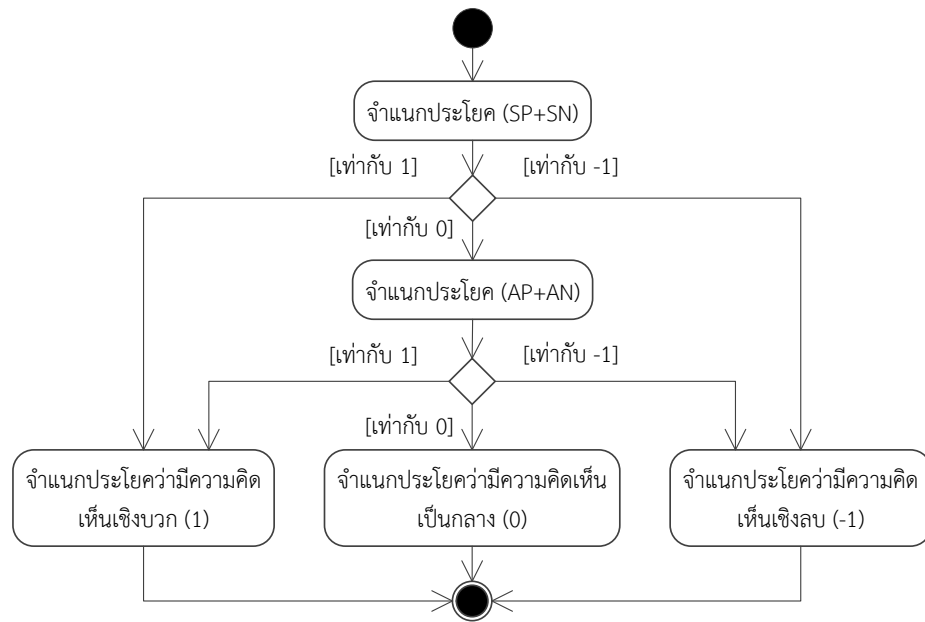
2) สร้างกฎจากรายการคำศัพท์ในเชิงบวกและในเชิงลบ [8] มีคำเชิงบวก 2,006 คำ มีคำเชิงลบ 4,783 คำ กฎหนึ่งกฎถูกสร้างขึ้นตามคำศัพท์ 1 คำ ดังนั้น ชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงบวก (PW) ประกอบด้วย 2,006 กฎ และชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงลบ (NW) ประกอบด้วย 4,783 กฎ (ตัวอย่างคำศัพท์ในภาคผนวก ค)

3) สร้างเครื่องมือจำแนกจากชุดกฎ ในแนวคิดที่นำเสนอสี่เครื่องจำแนกถูกสร้างขึ้นจากสี่ชุดกฎ เครื่องมือแรกคือเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกแบบเน้น (SP) ประกอบด้วยกฎที่ปรากฏอยู่ในชุดกฎเชิงบวกที่มาจากบทวิจารณ์ (PR) และชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงบวก (PW) เครื่องมือที่สองคือเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบแบบเน้น (SN) ประกอบด้วยกฎที่ปรากฏอยู่ในชุดกฎเชิงลบที่มาจากบทวิจารณ์ (NR) และชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงลบ (NW) เครื่องมือที่สามคือเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกที่ถูกปรับปรุง (AP) ประกอบด้วยกฎที่ปรากฏอยู่ในชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงบวก (PW) แต่ไม่ปรากฏอยู่ในชุดกฎเชิงลบที่มาจากบทวิจารณ์ (NR) เครื่องมือที่สี่คือเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบที่ถูกปรับปรุง (AN) ประกอบด้วยกฎที่ปรากฏอยู่ในชุดกฎที่มาจากคำศัพท์ที่มีความรู้สึกเชิงลบ (NW) แต่ไม่ปรากฏอยู่ในชุดกฎเชิงบวกที่มาจากบทวิจารณ์ (PR) กฎในเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกแบบเน้นเป็นกฎที่มีคำศัพท์ที่มีความรู้สึกเชิงบวกอย่างแน่นอน และกฎในเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบแบบเน้นเป็นกฎที่มีคำศัพท์ที่มีความรู้สึกเชิงลบอย่างแน่นอน คำศัพท์ที่มาจากรายการคำศัพท์ที่มีความรู้สึกเชิงบวกและเชิงลบเป็นคำที่ใช้แสดงความรู้สึกโดยทั่วไป เพื่อที่จะปรับปรุงให้เข้ากับประเภทของบทวิจารณ์ ถ้ากฎที่สร้างขึ้นจากรายการคำศัพท์ที่มีความรู้สึกเชิงบวกและเชิงลบขัดแย้งกับกฎที่สร้างจากข้อดีข้อเสีย กฎนั้นจะถูกตัดออก ยกตัวอย่างเช่นกฎที่สร้างจากรายการคำศัพท์ที่มีความรู้สึกเชิงบวกแต่เหมือนกับกฎที่สร้างจากข้อเสีย กฎนั้นจะถูกตัดออก เป็นต้น ตัวอย่างการสร้างเครื่องมือแสดงในตารางที่ 3.3 แต่ละกฎประกอบด้วยคำศัพท์หนึ่งคำ

ตารางที่ 3.3 ตัวอย่างการสร้างเครื่องมือจากสี่ชุดกฎ

ขั้นตอนที่	ชุดกฎและเครื่องมือ	คำศัพท์ (กฎแต่ละกฎ)
หนึ่ง	ชุดกฎ PR	{cheap} {easy}
	ชุดกฎ NR	{refund} {crash}
สอง	ชุดกฎ PW	{refund} {easy} {efficient}
	ชุดกฎ NW	{cheap} {crash} {awful}
สาม	เครื่องมือ SP	{easy}
	เครื่องมือ SN	{crash}
	เครื่องมือ AP	{easy} {efficient}
	เครื่องมือ AN	{crash} {awful}

4) จำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง ขั้นตอนนี้ใช้เครื่องมือจำแนกสี่เครื่องมือโดยแบ่งเป็นคู่สองคู่ คู่แรกเป็นเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกแบบเน้นและเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบแบบเน้น และคู่ที่สองเป็นเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงบวกที่ถูกปรับปรุงและเครื่องมือจำแนกประโยคที่มีความคิดเห็นเชิงลบที่ถูกปรับปรุง แต่ละประโยคจะถูกจำแนกด้วยเครื่องมือจำแนกคู่แรกและคู่สองตามลำดับ รูปที่ 3.6 แสดงขั้นตอนการจำแนกประโยค สามารถแสดงผลได้ตามตารางที่ 3.4 โดย “-” คือไม่ได้ผ่านการจำแนกประโยคด้วยเครื่องมือ หลังจากได้ผลการจำแนกด้วยเครื่องมือ ถ้าประโยคเป็นประโยคปฏิเสธที่ประกอบด้วย “not” “don’t” “never” ผลของการจำแนกประโยคจะกลับตรงกันข้ามคือบวกเป็นลบและลบเป็นบวก หรือ 1 เป็น -1 และ -1 เป็น 1



รูปที่ 3.6 ขั้นตอนการหาค่าความคิดเห็นของประโยค

ตารางที่ 3.4 ผลการจำแนกประโยคโดยวิธีการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน

ผลเครื่องมือจำแนก ประโยคเชิงบวก แบบเน้น (SP)	ผลเครื่องมือ จำแนกประโยคเชิง ลบแบบเน้น (SN)	ผลเครื่องมือ จำแนกประโยค เชิงบวกที่ถูก ปรับปรุง (AP)	ผลเครื่องมือ จำแนกประโยค เชิงลบที่ถูก ปรับปรุง (AN)	ผลของการ จำแนก ประโยค
เป็นบวก	ไม่เป็นลบ	-	-	บวก (1)
ไม่เป็นบวก	เป็นลบ	-	-	ลบ (-1)
เป็นบวก	เป็นลบ	เป็นบวก	เป็นลบ	กลาง (0)
เป็นบวก	เป็นลบ	เป็นบวก	ไม่เป็นลบ	บวก (1)
เป็นบวก	เป็นลบ	ไม่เป็นบวก	เป็นลบ	ลบ (-1)
เป็นบวก	เป็นลบ	ไม่เป็นบวก	ไม่เป็นลบ	กลาง (0)
ไม่เป็นบวก	ไม่เป็นลบ	เป็นบวก	เป็นลบ	กลาง (0)
ไม่เป็นบวก	ไม่เป็นลบ	เป็นบวก	ไม่เป็นลบ	บวก (1)
ไม่เป็นบวก	ไม่เป็นลบ	ไม่เป็นบวก	เป็นลบ	ลบ (-1)
ไม่เป็นบวก	ไม่เป็นลบ	ไม่เป็นบวก	ไม่เป็นลบ	กลาง (0)

กระบวนการจำแนกความคิดเห็นของประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน มีประสิทธิผลอยู่กับประโยคในข้อดีข้อเสีย เพราะประโยคที่อยู่ในข้อดีถูกสมมติว่าเป็นเชิงบวกและประโยคที่อยู่ในข้อเสียถูกสมมติว่าเป็นเชิงลบ ประโยคเหล่านี้ถูกนำไปสร้างกฎ ในความเป็นจริงประโยคในข้อดีก็มีทั้งประโยคในเชิงบวก เชิงลบและเป็นกลาง ประโยคในข้อเสียก็เช่นกัน ทำให้กฎที่สร้างขึ้นอาจจะไม่ถูกต้อง ถ้าทำให้ประโยคที่ถูกสมมติมีอัตราส่วนความถูกต้องที่สูงขึ้น กฎที่สร้างขึ้นก็มีความถูกต้องมากขึ้น ในการทดลองเพื่อให้เห็นความแตกต่างจึงประกอบด้วยสามรูปแบบในการสร้างกฎ

3.4.2.1 รูปแบบการสร้างกฎโดยใช้ทุกประโยค (Rule Construction using All Sentences, ALL) คือ การสร้างกฎที่ใช้ประโยคทุกประโยคในข้อดีและประโยคทุกประโยคในข้อเสีย เป็นชุดข้อมูลสอน (training set)

3.4.2.2 รูปแบบการสร้างกฎโดยใช้ประโยคที่เหลือจากการคัดออกอัตโนมัติ (Rule Construction using Automatic Sentence Filtering, ASF) เพื่อเพิ่มอัตราความถูกต้องของข้อมูลที่น่ามาสร้างกฎ เราจึงสร้างเงื่อนไขขึ้นเพื่อคัดประโยคออกทำให้อัตราความถูกต้องของประโยคที่เหลือที่ตรงสมมติฐานมากขึ้น ประโยคที่ถูกคัดออกจะสอดคล้องกับเงื่อนไขสามประการ

1) เงื่อนไขแรกคือประโยคที่เป็นประโยคเงื่อนไข ประโยคที่เป็นประโยคเงื่อนไขจะประกอบด้วยคำเชื่อม เช่น “if” และ “unless” เป็นต้น สองคำนี้เป็นคำที่ปรากฏบ่อยในประโยคเงื่อนไขในบทวิจารณ์ ประโยคเงื่อนไขส่วนใหญ่จะมีความคิดเห็นที่เป็นกลาง [30]

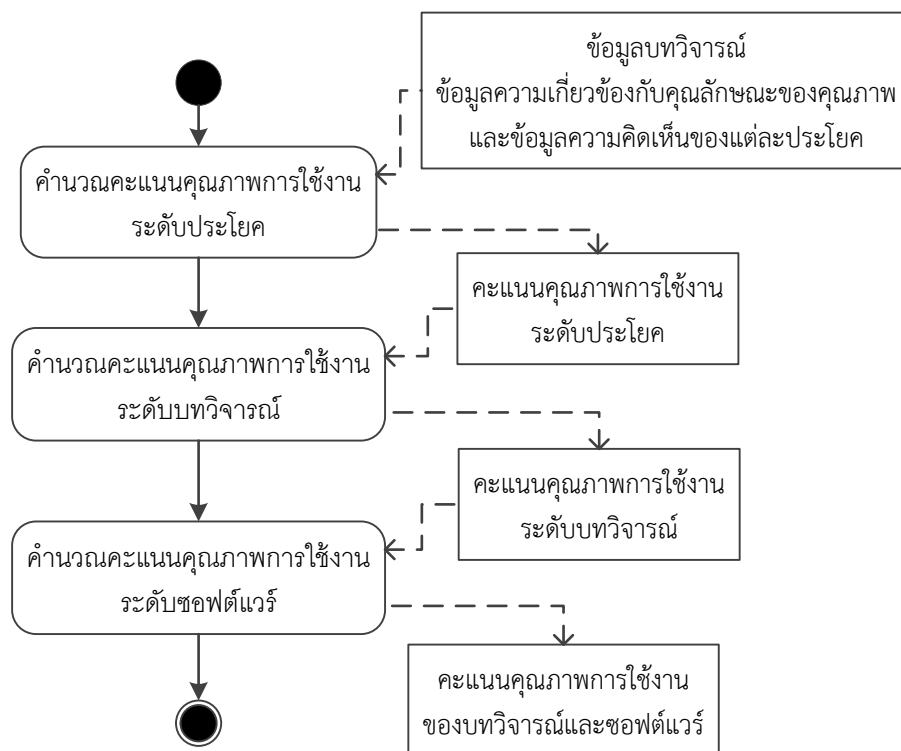
2) เงื่อนไขที่สองคือประโยคที่อยู่ในรูปอดีต ประโยคที่เป็นประโยคอยู่ในรูปอดีตจะมีคำกริยาเป็นตัวบ่งบอก ดังนั้นเครื่องมือที่ใช้หาว่าประโยคนั้นมีกริยาในรูปอดีตคือเครื่องมือระบุหน้าที่ของคำที่ใช้ความรู้เกี่ยวกับประมวลผลภาษาธรรมชาติ (natural language processing) ตัวอย่างเครื่องมือคือสแตนฟอร์ดล็อกลิเนียร์พาร์ทออฟสปีชแท็กเกอร์ (Stanford log-linear part-of-speech tagger) [19] เป็นต้น หลังจากผ่านเครื่องมือแล้วระบุคำกริยาของประโยคว่าเป็น “VBD” คือ “Verb, past tense” ประโยคนั้นจะเป็นประโยคที่อยู่ในรูปอดีต

3) เงื่อนไขที่สามคือประโยคที่เป็นประโยคปฏิเสธ ประโยคปฏิเสธคือประโยคที่ประกอบด้วยคำปฏิเสธ เช่น “not” “never” “no” เป็นต้น [31]

3.4.2.3 รูปแบบการสร้างกฎโดยใช้ประโยคที่เหลือจากการคัดออกด้วยมือ (Rule Construction using Manual Sentence Filtering, MSF) ในความเป็นจริง ประโยคในข้อดีไม่ควรจะมีประโยคที่มีความคิดเห็นเชิงลบ และประโยคในข้อเสียก็ไม่ควรมีประโยคที่มีความคิดเห็นเชิงบวก ประโยคที่ขัดแย้งดังกล่าวจึงถูกคัดออก หลังจากนั้นก็ใช้สามเงื่อนไขในการคัดเลือกประโยคเหมือนที่ใช้ใน 3.4.2.2 ดังนั้นสุดท้ายประโยคในข้อดีประกอบด้วยประโยคที่มีความเห็นเชิงบวกและเป็นกลาง และประโยคในข้อเสียประกอบด้วยประโยคที่มีความคิดเห็นเชิงลบและเป็นกลาง

3.5 การคำนวณคะแนนคุณภาพในการใช้งาน

จุดประสงค์ของกระบวนการนี้คือการหาว่าแต่ละบทวิจารณ์นั้นมีความคิดเห็นเกี่ยวกับคุณลักษณะของคุณภาพด้านต่างๆในเชิงบวกหรือเชิงลบแล้วนำมาหาคะแนนคุณภาพการใช้งานของซอฟต์แวร์นั้น การคำนวณคะแนนคุณภาพในการใช้งานประกอบด้วยสามขั้นตอนดังรูปที่ 3.7



รูปที่ 3.7 แผนภาพกิจกรรมของการคำนวณคะแนนคุณภาพในการใช้งานผู้ใช้งาน

3.5.1 การคำนวณคะแนนคุณภาพในการใช้งานในระดับประโยค

คะแนนคุณภาพในการใช้งานของประโยคคำนวณจากค่าความคิดเห็นของประโยคนั้นกับความเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านต่างๆ คะแนนคุณภาพในการใช้งานของประโยคประกอบด้วย คะแนนคุณภาพด้านประสิทธิผล (sentence-leveled effectiveness score, sq_0) คะแนนคุณภาพด้านผลิตผล (sentence-leveled productivity score, sq_1) คะแนนคุณภาพด้านความปลอดภัย (sentence-leveled safety score, sq_2) คะแนนคุณภาพด้านความพึงพอใจ (sentence-leveled satisfaction score, sq_3) โดยแต่ละคะแนนมีค่าคือ -1, 0 หรือ 1 และการคำนวณคะแนนคุณภาพการใช้งานระดับประโยคแสดงในสมการ (3.1) โดย p คือค่าความคิดเห็นของประโยคที่มีค่า -1, 0 หรือ 1 q_i เป็น true เมื่อประโยคนั้นมีความเกี่ยวข้องกับคุณลักษณะของคุณภาพด้าน i และ q_i เป็น false เมื่อประโยคนั้นไม่มีความเกี่ยวข้องกับคุณลักษณะคุณภาพด้าน i

$$sq_i = \begin{cases} p & \text{if } q_i = \text{true} \\ 0 & \text{if } q_i = \text{false} \end{cases} \quad (3.1)$$

3.5.2 การคำนวณคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์

คะแนนคุณภาพในการใช้งานของบทวิจารณ์ถูกคำนวณจากคะแนนคุณภาพการใช้งานของประโยคทุกประโยคในบทวิจารณ์ คะแนนคุณภาพการใช้งานของบทวิจารณ์ประกอบด้วย คะแนนคุณภาพด้านประสิทธิผล (review-leveled effectiveness score, rq_0) คะแนนคุณภาพด้านผลิตผล (review-leveled productivity score, rq_1) คะแนนคุณภาพด้านความปลอดภัย (review-leveled safety score, rq_2) คะแนนคุณภาพด้านความพึงพอใจ (review-leveled satisfaction score, rq_3) โดยแต่ละคะแนนมีค่าคือ -1, 0 หรือ 1 และการคำนวณคะแนนคุณภาพการใช้งานระดับบทวิจารณ์เป็นการรวมคะแนนประโยคที่เกี่ยวข้องกับคุณภาพด้านนั้น ถ้าคะแนนรวมระดับประโยคมีค่ามากกว่า 0 คะแนนคุณภาพด้านนั้นของบทวิจารณ์เป็น 1 แต่ถ้าคะแนนรวมระดับประโยคมีค่าน้อยกว่า 0 คะแนนคุณภาพด้านนั้นของบทวิจารณ์เป็น -1 การคำนวณนี้แสดงในสมการ (3.2) โดย sq_{ji} คือคะแนนคุณภาพการใช้งานของคุณลักษณะของคุณภาพด้าน i ของประโยค j และ N คือจำนวนประโยคที่อยู่ในบทวิจารณ์

$$rq_i = \begin{cases} 1 & \text{if } \sum_{j=1}^N sq_{ji} > 0 \\ 0 & \text{if } \sum_{j=1}^N sq_{ji} = 0 \\ -1 & \text{if } \sum_{j=1}^N sq_{ji} < 0 \end{cases} \quad (3.2)$$

3.5.3 การคำนวณคะแนนคุณภาพการใช้งานในระดับซอฟต์แวร์

คะแนนคุณภาพการใช้งานของซอฟต์แวร์จะถูกคำนวณขึ้นเพื่อเป็นตัวแทนคุณภาพของซอฟต์แวร์ที่มาจากบทวิจารณ์ คะแนนคุณภาพการใช้งานของซอฟต์แวร์จะประกอบด้วยคะแนนสี่ด้าน คือ คะแนนประสิทธิผล (effectiveness score) คะแนนผลิตผล (productivity score) คะแนนความปลอดภัย (safety score) และคะแนนความพึงพอใจ (satisfaction score) โดยแต่ละคะแนนมีค่าตั้งแต่ -1 ถึง 1 คะแนนแต่ละอย่างจะถูกคำนวณจากการรวมคะแนนประเภทนั้นที่อยู่ในทุกบทวิจารณ์แล้วหารด้วยจำนวนบทวิจารณ์ การคำนวณไม่ได้คำนึงถึงเวลาที่บทวิจารณ์นั้นเขียนขึ้น เนื่องจากถ้าในสภาพแวดล้อมเดิมคุณภาพของซอฟต์แวร์จะเปลี่ยนแปลงถ้าซอฟต์แวร์นั้นได้รับการแก้ไขปรับปรุง ซึ่งส่วนมากการปรับปรุงของซอฟต์แวร์ที่สังเกตได้ชัดหรือคุณภาพเปลี่ยนแปลงไปจะเป็นการเปลี่ยนแปลงรุ่น (version) ซึ่งเปรียบเสมือนเป็นสินค้าใหม่ การเขียนบทวิจารณ์ซอฟต์แวร์รุ่นใหม่เป็นข้อมูลของซอฟต์แวร์รุ่นใหม่ การคำนวณจะแสดงอยู่ในสมการ (3.3) (3.4) (3.5) และ (3.6)

$$effectiveness_score = \frac{\sum_{i=1}^{N_0} rq_{i0}}{N_0} \quad (3.3)$$

$$productivity_score = \frac{\sum_{i=1}^{N_1} rq_{i1}}{N_1} \quad (3.4)$$

$$safety_score = \frac{\sum_{i=1}^{N_2} rq_{i2}}{N_2} \quad (3.5)$$

$$satisfaction_score = \frac{\sum_{i=1}^{N_3} rq_{i3}}{N_3} \quad (3.6)$$

N_0 คือ จำนวนของบทวิจารณ์ที่เกี่ยวข้องกับคุณภาพด้านประสิทธิผล

N_1 คือ จำนวนของบทวิจารณ์ที่เกี่ยวข้องกับคุณภาพด้านผลิตผล

N_2 คือ จำนวนของบทวิจารณ์ที่เกี่ยวข้องกับคุณภาพด้านความปลอดภัย

N_3 คือ จำนวนของบทวิจารณ์ที่เกี่ยวข้องกับคุณภาพด้านความพึงพอใจ

rq_{i0} คือ คะแนนประสิทธิผลของบทวิจารณ์ i

rq_{i1} คือ คะแนนผลิตผลของบทวิจารณ์ i

rq_{i2} คือ คะแนนความปลอดภัยของบทวิจารณ์ i

rq_{i3} คือ คะแนนความพึงพอใจของบทวิจารณ์ i

บทที่ 4

การทดลองและการวิเคราะห์ผล

ในบทนี้ประกอบด้วย การทดลองเพื่อวัดประสิทธิผลของแนวคิดที่ได้นำเสนอในบทที่ 3 โดยจะเริ่มอธิบายจากการวางแผนการทดลอง การจัดเตรียมข้อมูลไว้สำหรับการทดลอง การวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน การวัดประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือเชิงลบ การวัดประสิทธิผลส่วนการคำนวณคะแนนคุณภาพการใช้งาน และการวัดประสิทธิผลของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง

4.1 การวางแผนการทดลอง

4.1.1 จุดประสงค์การทดลอง

จุดประสงค์ของการทดลองคือการวัดประสิทธิผลของวิธีการที่นำเสนอสี่ด้าน คือ

- 1) ด้านของการหาความเกี่ยวข้องของประโยคกับลักษณะของคุณภาพ
- 2) ด้านของการจำแนกประโยคว่ามีความคิดเห็นเชิงบวกหรือลบ
- 3) ด้านของการคำนวณคะแนนตามลักษณะคุณภาพ
- 4) ด้านของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือลบ

4.1.2 ออกแบบการทดลอง

การทดลองถูกแบ่งออกเป็นสี่ส่วนตามการวัดประสิทธิผลของวิธีการที่นำเสนอในจุดประสงค์ โดยมีรายละเอียดคือ

4.1.2.1 การวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน

จากแนวคิดที่นำเสนอในบทที่ 3 มีการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งานด้วยการใช้ออนโทโลยี ออนโทโลยีที่ถูกสร้างขึ้นและถูกขยายตามแนวคิดในหัวข้อ

3.3 ในส่วนนี้มีจุดประสงค์คือ การวัดประสิทธิผลของการใช้ออนโทโลยีแบบไม่ขยายเปรียบเทียบกับ การใช้ออนโทโลยีแบบขยาย ในการทดลองใช้ข้อมูลของประโยคที่ถูกระบุว่าเกี่ยวข้องกับคุณลักษณะของ คุณภาพการใช้งานด้านใดโดยผู้เชี่ยวชาญ

4.1.2.2 การวัดประสิทธิผลส่วนการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวก หรือเชิงลบหรือเป็นกลาง

จากแนวคิดที่นำเสนอในบทที่ 3 มีการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวก หรือเชิงลบหรือเป็นกลางด้วยเครื่องมือจำแนก แนวคิดในการสร้างเครื่องมือจำแนกมีสองแนวคิดคือ แนวคิดในหัวข้อ 3.4.1 และแนวคิดในหัวข้อ 3.4.2 ในส่วนนี้มีสองจุดประสงค์คือ

1) การเปรียบเทียบประสิทธิผลการจำแนกประโยคด้วยแนวคิดในหัวข้อ 3.4.1 และแนวคิดในหัวข้อ 3.4.2

2) การพิจารณาประสิทธิผลของการจำแนกโดยใช้แนวคิดในหัวข้อ 3.4.2 เมื่ออัตราความ ถูกต้องของข้อมูลที่ถูกสมมติสูงขึ้น (ประโยคในข้อดีสมมติให้เป็นประโยคที่มีความคิดเห็นเชิงบวกและ ประโยคในข้อเสียสมมติให้เป็นประโยคที่มีความคิดเห็นเชิงลบ)

เมื่อนำข้อมูลประโยคเหล่านั้นขึ้นมาสรางกฎ ประสิทธิภาพของการจำแนกด้วยเครื่องมือที่ ประกอบด้วยกฎที่สร้างขึ้นก็จะสูงขึ้น โดยใช้ประโยคในบทวิจารณ์เป็นข้อมูลทดสอบ ความคิดเห็น ของประโยคเหล่านั้นจะถูกพิจารณาโดยผู้เชี่ยวชาญว่าประโยคนั้นมีความคิดเห็นเชิงบวก เชิงลบหรือ เป็นกลาง

4.1.2.3 การวัดประสิทธิผลส่วนการคำนวณคะแนนคุณภาพการใช้งาน

จากแนวคิดที่นำเสนอในบทที่ 3 มีการคำนวณคะแนนคุณภาพการใช้งานของบทวิจารณ์ แนวคิดในการคำนวณคะแนนคุณภาพการใช้งานแสดงในหัวข้อ 3.5 ในส่วนนี้มีจุดประสงค์คือ การวัด ประสิทธิภาพของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบในคุณภาพสี่ด้านด้วยวิธีที่ นำเสนอเปรียบเทียบกับวิธีจำแนกประโยคแบบพื้นฐาน โดยใช้บทวิจารณ์เป็นข้อมูลทดสอบ ความ คิดเห็นของบทวิจารณ์เหล่านั้นจะถูกพิจารณาโดยผู้เชี่ยวชาญว่ามีความคิดเห็นเชิงบวก เชิงลบ หรือเป็น กลางในคุณลักษณะของคุณภาพด้านประสิทธิผล (effectiveness) ผลิตผล (productivity) ความ ปลอดภัย (safety) และความพึงพอใจ (satisfaction)

4.1.2.4 การวัดประสิทธิผลของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือ ความคิดเห็นเชิงลบ

จากแนวคิดที่นำเสนอในบทที่ 3 ผลลัพธ์ของแนวคิดที่เป็นคะแนนคุณภาพการใช้งานของบทวิจารณ์ควรสอดคล้องกับความคิดเห็นในบทวิจารณ์ ทั้งนี้ความคิดเห็นของบทวิจารณ์นั้นจะสอดคล้องกับจำนวนดาวที่ผู้เขียนวิจารณ์ให้กับบทวิจารณ์นั้น ดังนั้นความคิดเห็นที่คิดจากคะแนนคุณภาพการใช้งานของบทวิจารณ์นั้นควรจะสอดคล้องกับความคิดเห็นที่พิจารณาจากจำนวนดาว โดยจำนวนดาวมากกว่าสาม บทวิจารณ์นั้นมีความคิดเห็นเชิงบวก จำนวนดาวเท่ากับสามแสดงว่าความคิดเห็นของบทวิจารณ์นั้นเป็นกลาง และจำนวนดาวน้อยกว่าสามแสดงว่าความคิดเห็นของบทวิจารณ์นั้นเป็นลบ ความสอดคล้องนี้เปรียบเสมือนการคาดการณ์ความคิดเห็นของบทวิจารณ์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งาน ในส่วนนี้มีจุดประสงค์คือเปรียบเทียบประสิทธิผลของการคาดการณ์ความคิดเห็นจากจำนวนดาวในบทวิจารณ์จากคะแนนคุณภาพการใช้งานกับการคาดการณ์ความคิดเห็นจากจำนวนดาวในบทวิจารณ์จากความคิดเห็นแต่ละประโยค

4.1.3 มาตรฐาน

ในการทดลอง มีการวัดประสิทธิผลของการจำแนกสองประเภทและการจำแนกสามประเภท จำนวนของสิ่งที่ถูกจำแนกสองประเภทและสามประเภทโดยระบบเปรียบเทียบกับผู้เชี่ยวชาญ แสดงในตาราง 4.1 และ 4.2 ค่า A ถึง M คือจำนวนเต็มใดๆ และจำนวนสิ่งของที่ต้องการจำแนกทั้งหมดคือ N สมมติให้สิ่งที่ถูกจำแนกมีสามประเภทคือประเภท (ก) (ข) และ (ค) โดยมาตรฐานที่ถูกนำมาใช้มีดังนี้

ตารางที่ 4.1 จำนวนของสิ่งที่ถูกจำแนกได้สองประเภทโดยระบบเปรียบเทียบกับผู้เชี่ยวชาญ

	ระบบจำแนกว่า ประเภท (ก)	ระบบจำแนกว่าไม่ใช่ ประเภท (ก)
ผู้เชี่ยวชาญจำแนกว่า ประเภท (ก)	A	B
ผู้เชี่ยวชาญจำแนกว่า ไม่ใช่ประเภท (ก)	C	D

ตารางที่ 4.2 จำนวนของสิ่งที่ถูกจำแนกได้สามประเภทโดยระบบเปรียบเทียบกับผู้เชี่ยวชาญ

	ระบบจำแนกว่า ประเภท (ก)	ระบบจำแนกว่า ประเภท (ข)	ระบบจำแนกว่า ประเภท (ค)
ผู้เชี่ยวชาญจำแนกว่า ประเภท (ก)	E	F	G
ผู้เชี่ยวชาญจำแนกว่า ประเภท (ข)	H	I	J
ผู้เชี่ยวชาญจำแนกว่า ประเภท (ค)	K	L	M

4.1.3.1 ความแม่นยำ (precision)

ความแม่นยำ (precision) คือ อัตราส่วนของจำนวนสิ่งที่จำแนกเป็นประเภท (ก) โดยระบบตรงกับผู้เชี่ยวชาญ (A) ต่อจำนวนสิ่งที่จำแนกได้โดยระบบทั้งหมด (A+C) จากตารางที่ 4.1 จะได้สมการ (4.1)

$$precision = \frac{A}{A+C} \quad (4.1)$$

4.1.3.2 ความครบถ้วน (recall)

ความครบถ้วน (recall) คือ อัตราส่วนของจำนวนสิ่งที่ยืนยันว่าเป็นประเภท (ก) โดยระบบตรงกับผู้เชี่ยวชาญ (A) ต่อจำนวนสิ่งที่ยืนยันได้โดยผู้เชี่ยวชาญ (A+B) จากตารางที่ 4.1 จะได้สมการ (4.2)

$$recall = \frac{A}{A+B} \quad (4.2)$$

4.1.3.3 เอฟ-สกอร์ (F-score)

เอฟ-สกอร์ (F-score) คือ ค่าเฉลี่ยฮาร์โมนิกระหว่างความแม่นยำและความครบถ้วนตามสมการ (4.3)

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (4.3)$$

4.1.3.4 ความถูกต้องของการจำแนกของสองสิ่ง (accuracy_2)

ความถูกต้องของการจำแนกของสองสิ่ง (accuracy_2) อัตราส่วนของจำนวนสิ่งที่ยืนยันว่าเป็นประเภท (ก) โดยระบบตรงกับผู้เชี่ยวชาญ (A) รวมกับจำนวนสิ่งที่ยืนยันเป็นไม่ใช่ประเภท (ก) โดยระบบตรงกับผู้เชี่ยวชาญ (D) ต่อจำนวนทั้งหมด (N) จากตารางที่ 4.1 จะได้สมการ (4.4)

$$accuracy_2 = \frac{A+D}{N} \quad (4.4)$$

4.1.3.5 ความถูกต้องของการจำแนกของสามสิ่ง (accuracy_3)

ความถูกต้องของการจำแนกของสามสิ่ง (accuracy_3) อัตราส่วนของจำนวนสิ่งที่ยืนยันเป็นประเภท (ก) โดยระบบตรงกับผู้เชี่ยวชาญ (E) รวมกับจำนวนสิ่งที่ยืนยันเป็นประเภท (ข) โดยระบบตรงกับผู้เชี่ยวชาญ (I) รวมกับจำนวนสิ่งที่ยืนยันเป็นประเภท (ค) โดยระบบตรงกับผู้เชี่ยวชาญ (M) ต่อจำนวนทั้งหมด (N) จากตารางที่ 4.2 จะได้สมการ (4.5)

$$accuracy_3 = \frac{E+I+M}{N} \quad (4.5)$$

4.2 การเตรียมการทดลองและข้อมูลทดลอง

ข้อมูลที่ใช้ในการทดลองนี้มีรายละเอียดดังตารางที่ 4.3

ตารางที่ 4.3 ข้อมูลบทวิจารณ์

เว็บไซต์ของข้อมูล	www.cnet.com
วันที่เก็บข้อมูล	วันที่ 18 พฤษภาคม พ.ศ. 2555
จำนวนประเภทซอฟต์แวร์	3 ประเภทซอฟต์แวร์
จำนวนซอฟต์แวร์	30 ซอฟต์แวร์ (10 ซอฟต์แวร์ต่อ 1 ประเภทซอฟต์แวร์)
จำนวนบทวิจารณ์	3,000 บทวิจารณ์ (100 บทวิจารณ์ล่าสุดต่อ 1 ซอฟต์แวร์)
จำนวนบทวิจารณ์ตามความคิดเห็น	ความคิดเห็นเชิงบวกมีจำนวน 2,005 บทวิจารณ์ (มากกว่า 3 ดาว) ความคิดเห็นเป็นกลางมีจำนวน 228 บทวิจารณ์ (เท่ากับ 3 ดาว) ความคิดเห็นเชิงลบมีจำนวน 767 บทวิจารณ์ (น้อยกว่า 3 ดาว)

ในบทวิจารณ์จำนวน 3,000 บทวิจารณ์ ประโยคในบทวิจารณ์จะถูกแบ่งโดย “.” “?” “!” ทำให้ข้อมูลทั้งหมดประกอบด้วยประโยคทั้งหมด 18,679 ประโยค โดยประโยคแต่ละประโยคที่เก็บจะถูกระบุมาจากส่วนสรุปสั้น (one-line summary) ข้อดี (pros) ข้อเสีย (cons) และสรุป (summary) ประโยคทุกประโยคจะถูกพิจารณาโดยผู้เชี่ยวชาญว่าเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดและความคิดเห็นของประโยคมีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลาง สรุปข้อมูลที่ใช้ในการทดลองมีดังนี้

1) ข้อมูลของบทวิจารณ์ ประกอบด้วย ชื่อซอฟต์แวร์ที่เขียนวิจารณ์ จำนวนดาวที่ผู้วิจารณ์ให้ ประโยคในส่วนสรุปสั้น ประโยคในข้อดี ประโยคในข้อเสีย ประโยคในสรุป ความคิดเห็นของคุณลักษณะของคุณภาพด้านประสิทธิภาพ ประสิทธิภาพ ผลผลิต ความปลอดภัยและความพึงพอใจที่ถูกพิจารณาโดยผู้เชี่ยวชาญว่ามีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลาง

2) ข้อมูลของประโยค ประกอบด้วย คำศัพท์ในประโยค ประเภทประโยค (ประโยคในส่วนสรุปสั้น ในข้อดี ในข้อเสียหรือในสรุป) ความเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดที่พิจารณาโดยผู้เชี่ยวชาญ ความคิดเห็น และความคิดเห็นของประโยคที่พิจารณาโดยผู้เชี่ยวชาญว่ามีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลาง

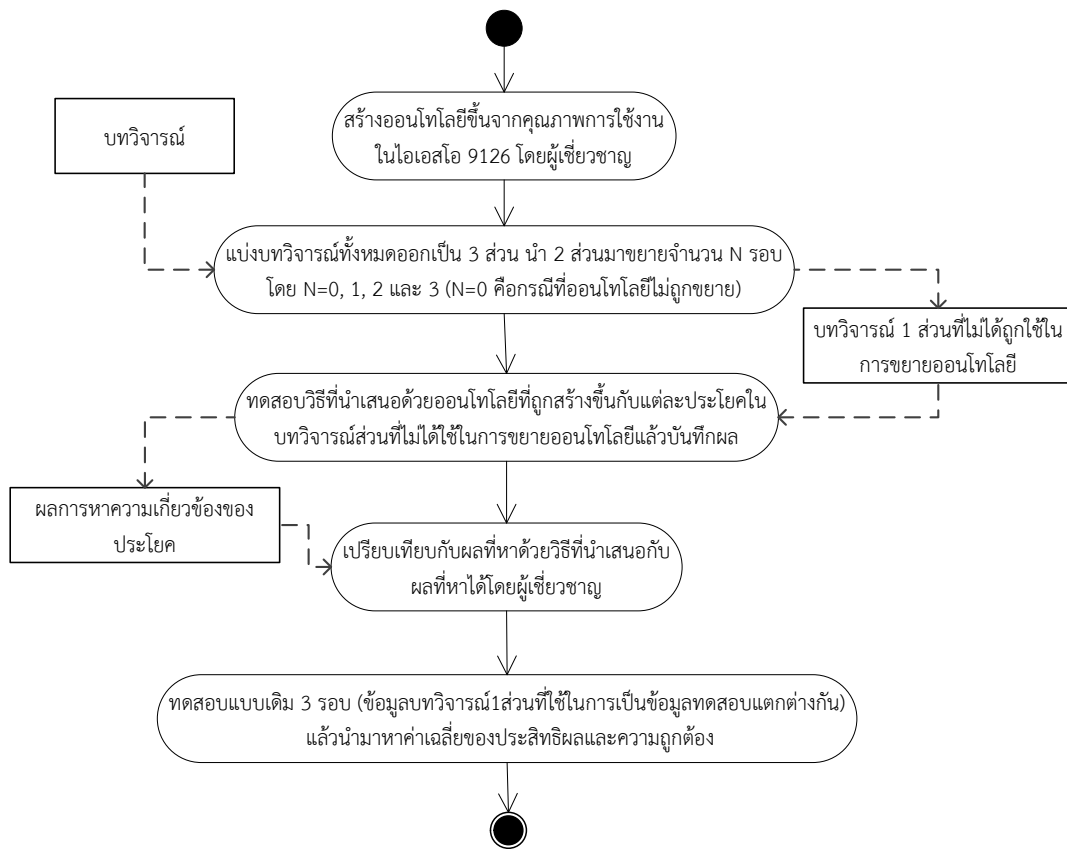
4.3 การวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน

ในส่วนนี้มีจุดประสงค์คือการเปรียบเทียบประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้วยอนโทโลยีที่ไม่ถูกขยายและด้วยอนโทโลยีที่ถูกขยายด้วยรรณภพฐานซึ่งใช้เครื่องมือเวิร์ดเน็ต [21] จำนวน 1 – 3 ครั้ง วิธีสร้างอนโทโลยีที่นำมาเปรียบเทียบประสิทธิผลมีดังนี้

- 1) การวิเคราะห์ประโยคด้วยอนโทโลยีระดับ 0 คือ การวิเคราะห์ด้วยอนโทโลยีที่กำหนดขึ้นแล้วไม่ได้ถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 (O_0)
- 2) การวิเคราะห์ประโยคด้วยอนโทโลยีระดับ 1 คือ การวิเคราะห์ด้วยอนโทโลยีที่กำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนหนึ่งรอบ (O_1)
- 3) การวิเคราะห์ประโยคด้วยอนโทโลยีระดับ 2 คือการวิเคราะห์ด้วยอนโทโลยีที่กำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนสองรอบ (O_2)
- 4) การวิเคราะห์ประโยคด้วยอนโทโลยีระดับ 3 คือการวิเคราะห์ด้วยอนโทโลยีที่กำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนสามรอบ (O_3)

4.3.1 ขั้นตอนการทดลอง

ในการทดลองนี้ใช้ข้อมูลทดสอบทั้งหมดประกอบด้วยประโยคทุกประโยคในทุกบทวิจารณ์ภาพรวมของการทดลองในส่วนของการวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งานแสดงในรูปที่ 4.1



รูปที่ 4.1 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิผลของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งาน

การทดลองมีรายละเอียดแต่ละขั้นตอนดังนี้

1) อนโทโลยีจะถูกกำหนดโดยผู้เชี่ยวชาญ โดยผู้เชี่ยวชาญจะพิจารณารายละเอียดและมาตรฐานวัดของแต่ละคุณลักษณะในคุณภาพการใช้งานของไอเอสโอ 9126 พร้อมทั้งพิจารณาไปถึงรายละเอียดและมาตรฐานวัดของแต่ละคุณลักษณะในคุณภาพภายนอกของไอเอสโอ 9126 ที่มีความสัมพันธ์กับคุณลักษณะในคุณภาพการใช้งาน

2) บทวิจารณ์ทั้งหมดถูกแบ่งออกเป็นสามส่วน (ส่วนที่หนึ่ง ส่วนที่สองและส่วนที่สาม) แล้วนำข้อมูลบทวิจารณ์สองส่วน (ส่วนที่หนึ่งและส่วนที่สอง) มาใช้ในการขยายอนโทโลยี ในการทดลองนี้มีข้อมูลบทวิจารณ์ทั้งหมด 3,000 บทวิจารณ์ จะใช้ข้อมูลบทวิจารณ์จำนวน 2,000 บทวิจารณ์นำขยายอนโทโลยีด้วยอัลกอริทึมที่ 1 ในหัวข้อ 3.3 จำนวน N รอบ โดย $N = 0, 1, 2$ และ 3 ($N=0$ คือกรณีที่ไม่ได้ขยายอนโทโลยี (O_0), $N=1$ (O_1)) โดยสาเหตุที่การทดลองใช้ข้อมูลเพียงสองในสามส่วน เพราะ

เมื่อนำไปใช้จริง ไม่สามารถนำข้อมูลทั้งหมดมาใช้ในการขยายได้ เครื่องมือที่ใช้ในการหาค่าที่จะนำมาขยายคือเวิร์ดเน็ต (Wordnet) (ภาคผนวก ก แสดงออนโทโลยีที่ใช้ในการทดลอง)

3) ทดสอบกับประโยคทั้งหมดในบทวิจารณ์ส่วนที่สาม โดยแต่ละประโยคถูกหาความเกี่ยวข้องกับคุณลักษณะของคุณภาพโดยวิธีการที่นำเสนอด้วยออนโทโลยีที่สร้างขึ้นทั้งสื่อออนโทโลยีและบันทึกผล

4) ผลการหาความเกี่ยวข้องของแต่ละประโยคกับคุณลักษณะของคุณภาพโดยวิธีการที่นำเสนอถูกนำไปเปรียบเทียบกับผลของการหาความเกี่ยวข้องของแต่ละประโยคกับคุณลักษณะของคุณภาพโดยของผู้เชี่ยวชาญหาความถูกต้อง

5) เปลี่ยนข้อมูลบทวิจารณ์ที่นำมาขยายออนโทโลยีเป็นส่วนที่สองและส่วนสาม และทดสอบกับประโยคในบทวิจารณ์ส่วนที่หนึ่ง ทำตามขั้นตอนที่ 2 ถึง 4

6) เปลี่ยนข้อมูลบทวิจารณ์ที่นำมาขยายออนโทโลยีเป็นส่วนที่หนึ่งและส่วนที่สาม และทดสอบกับประโยคในบทวิจารณ์ส่วนที่สอง ทำตามขั้นตอนที่ 2 ถึง 4

7) นำผลที่ได้มาหาค่าเฉลี่ยของประสิทธิผลและความถูกต้องของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพในวิธีการที่นำเสนอด้วยออนโทโลยีที่ไม่ได้ขยายหรือที่ขยายหนึ่งรอบ สองรอบและสามรอบ พร้อมเปรียบเทียบผล

4.3.2 ผลการทดลองและการวิเคราะห์ผล

การเปรียบเทียบความถูกต้องของผลการทดลองและประสิทธิผลในการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพการใช้งานแสดงตามตารางที่ 4.4 และ 4.5

ตารางที่ 4.4 ความถูกต้องของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้านต่างๆ

ออนโทโลยี	ความถูกต้อง				
	ประสิทธิผล	ผลิตผล	ความปลอดภัย	ความพึงพอใจ	เฉลี่ย
O ₀	90.74%	88.27%	91.44%	70.43%	85.22%
O ₁	82.77%	86.88%	88.98%	65.64%	81.07%
O ₂	56.54%	62.27%	77.91%	50.47%	61.80%
O ₃	31.10%	41.89%	55.57%	42.43%	42.75%

ตารางที่ 4.5 ประสิทธิภาพของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้านต่างๆ

ออนโทโลยี	ประสิทธิผล			ผลิตผล			ความปลอดภัย			ความพึงพอใจ			เฉลี่ย		
	Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1
O ₀	0.67	0.44	0.53	0.76	0.55	0.64	0.85	0.39	0.53	0.64	0.32	0.43	0.73	0.42	0.53
O ₁	0.35	0.53	0.42	0.64	0.69	0.66	0.58	0.44	0.50	0.50	0.53	0.52	0.52	0.55	0.53
O ₂	0.18	0.77	0.30	0.31	0.83	0.46	0.29	0.52	0.37	0.38	0.68	0.48	0.29	0.70	0.40
O ₃	0.14	0.93	0.24	0.24	0.94	0.38	0.18	0.73	0.29	0.36	0.86	0.51	0.23	0.87	0.36

จากตารางที่ 4.4 ค่าความถูกต้องของการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ไม่มีการขยาย (O₀) สูงกว่าการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่มีการขยาย (O₁, O₂, O₃) ในทุกคุณลักษณะของคุณภาพทั้งในด้านประสิทธิผล (effectiveness) ด้านผลิตผล (productivity) ด้านความปลอดภัย (safety) และด้านความพึงพอใจ (satisfaction) เนื่องจากคำศัพท์ที่เกี่ยวข้องกับคุณลักษณะคุณภavn นั้นเป็นคำลักษณะที่ค่อนข้างเฉพาะ การขยายโดยใช้เวิร์ดเน็ตไม่ได้ช่วยให้ได้คำศัพท์ที่เกี่ยวข้องกับคุณลักษณะคุณภavn ทั้งหมด แต่เมื่อวิเคราะห์ค่าความแม่นยำ ค่าความครบถ้วน และค่าเอฟสกออร์ของการหาความเกี่ยวข้องโดยใช้ออนโทโลยีของคุณลักษณะคุณภavn ด้านต่างๆ ในตารางที่ 4.5 สังเกตว่าออนโทโลยียิ่งถูกขยาย จะทำให้ค่าความแม่นยำลดลง ค่าความครบถ้วนสูงขึ้น เนื่องจากการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ถูกขยายมากทำให้ได้ผลลัพธ์ที่หาได้มีจำนวนมาก ส่งผลให้ค่าความแม่นยำลดลง แต่ค่าความครบถ้วนสูงขึ้น ส่วนเมื่อนำมาคิดค่าเอฟสกออร์ในคุณลักษณะของคุณภavn ด้านประสิทธิภavn และความพึงพอใจมีค่าเพิ่มขึ้นเมื่อใช้การหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ขยายหนึ่งรอบ (O₁) แทนการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ไม่ขยาย (O₀) เพราะค่าที่เกี่ยวข้องกับคุณลักษณะทั้งสองนี้มีลักษณะคำที่เชิงกว้าง เช่น “fast” “happy” เป็นต้น ซึ่งถ้านำคำเหล่านี้ไปขยายออนโทโลยีจะทำให้การหาความเกี่ยวข้องหรือการจำแนกดีขึ้น แต่ค่าเอฟสกออร์โดยเฉลี่ยของการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ขยายไม่ได้ให้ผลดีขึ้นเมื่อเปรียบเทียบการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ไม่ขยาย

4.3.3 สรุปผลการทดลอง

จากหัวข้อ 4.3.2 ความถูกต้อง ค่าความแม่นยำ ค่าความครบถ้วนและค่าเอฟสกออร์ของการหาความเกี่ยวข้องโดยใช้ออนโทโลยีที่ไม่มีการขยายในทุกคุณลักษณะของคุณภavn โดยเฉลี่ย สรุปได้ว่าการ

หาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพโดยใช้ออนโทโลยีที่ไม่มีการขยายมีความถูกต้องและมีประสิทธิผลมากที่สุด

4.3.4 ข้อมูลสนับสนุนผลการทดลอง

4.3.4.1 การขยายคำศัพท์ด้วยอรรถาภิธานโดยใช้เครื่องมือเวิร์ดเน็ต

การขยายคำศัพท์ที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้วยอรรถาภิธานโดยใช้เครื่องมือเวิร์ดเน็ตไม่ได้ให้คำศัพท์ที่เกี่ยวข้องกับคุณลักษณะของคุณภาพทั้งหมด เนื่องจากคำศัพท์นั้นมีหลายความหมายขึ้นอยู่กับบริบทที่ใช้ ทำให้คำศัพท์ที่พ้องความหมายหรือความหมายตรงข้ามนั้นมีความหมายต่างกัน ยกตัวอย่างเช่น คำว่า “fast” ที่อยู่ในคุณภาพด้านผลิตผล มีความหมายทั้งหมดสลับอย่างตามบริบทที่ใช้ โดยมีความหมายด้านคำนามหนึ่งอย่าง มีความหมายด้านกริยาสองอย่าง มีความหมายด้านคำคุณศัพท์สลับอย่างและมีความหมายด้านคำวิเศษณ์สองอย่าง คำที่เพิ่มขึ้นจากการขยายคำศัพท์ “fast” คือ “fasting” “slow” “tight” “debauched” “degenerate” “degraded” “dissipated” “dissolute” “libertine” “profligate” “riotous” “flying” “quick” “firm” “immobile” “loyal” และ “truehearted” คำศัพท์ที่เพิ่มขึ้นที่ไม่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านผลิตผลที่เห็นได้ชัดเจนได้แก่คำศัพท์ว่า “tight” “debauched” “degenerate” “degraded” “dissolute” “libertine” “profligate” “riotous” “flying” “firm” “loyal” และ “truehearted” แต่อย่างไรก็ตาม ความถูกต้องของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพนั้นไม่ได้มีปัจจัยจากความถูกต้องและความครบถ้วนของคำศัพท์ที่เกี่ยวข้องกับคุณลักษณะของคุณภาพเพียงอย่างเดียว แต่อีกปัจจัยหนึ่งคือบริบทของคำศัพท์นั้นก็สำคัญด้วยเช่นกันที่จะให้ความหมายของคำศัพท์ไปในทางที่เกี่ยวข้องกับคุณลักษณะของคุณภาพนั้นๆ

4.4 การจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง

ในส่วนนี้มีสองจุดประสงค์ จุดประสงค์แรกคือ การวัดประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือความคิดเห็นเชิงลบโดยแนวคิดที่นำเสนอ (แนวคิดที่เสนอในหัวข้อ 3.4.2) เปรียบเทียบวิธีการจำแนกประโยคแบบพื้นฐาน (วิธีในหัวข้อ 3.4.1) จุดประสงค์ที่สองคือ ประสิทธิภาพของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือความคิดเห็นเชิงลบโดยแนวคิดที่นำเสนอจะสูงขึ้น เมื่ออัตราความถูกต้องของข้อมูลที่ถูกสมมติสูงขึ้น (ประโยคในข้อดีสมมติให้

เป็นประโยชน์ที่มีความคิดเห็นเชิงบวกและประโยชน์ในข้อเสียสมมติให้เป็นประโยชน์ที่มีความคิดเห็นเชิงลบ)
วิธีการที่ใช้ในการทดลองนี้มีสี่วิธีดังนี้

ตารางที่ 4.6 วิธีในการทดลองส่วนการจำแนกประโยชน์

วิธีการ	รายละเอียด
1) วิธีจำแนกประโยชน์แบบพื้นฐาน (BL)	การใช้รายการคำศัพท์เกี่ยวกับความรู้สึกเท่านั้น รายละเอียดถูกอธิบายในหัวข้อ 3.4.1
2) วิธีจำแนกประโยชน์แบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยชน์ทุกประโยชน์ในข้อดีข้อเสีย (ALL)	การประยุกต์ใช้รายการคำศัพท์เกี่ยวกับความรู้สึกกับกฎที่สร้างขึ้นจากทุกประโยชน์ รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.1
3) วิธีจำแนกประโยชน์แบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยชน์ที่เหลือจากการคัดออกอัตโนมัติ (ASF)	การประยุกต์ใช้รายการคำศัพท์เกี่ยวกับความรู้สึกกับกฎที่สร้างขึ้นจากประโยชน์ที่เหลือจากการคัดออกอัตโนมัติ รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.2
4) วิธีจำแนกประโยชน์แบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยชน์ที่เหลือจากการคัดออกด้วยมือ (MSF)	การประยุกต์ใช้รายการคำศัพท์เกี่ยวกับความรู้สึกกับกฎที่สร้างขึ้นจากประโยชน์ที่เหลือจากการคัดออกด้วยมือ รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.3

4.4.1 ขั้นตอนการทดลอง

ในการทดลองในส่วนนี้จะมีการทดลองกับข้อมูลทดสอบสองชุด ชุดที่หนึ่งคือข้อมูลประโยชน์ที่เกี่ยวข้องกับลักษณะของคุณภาพและเป็นประโยชน์ที่อยู่ในข้อดีข้อเสีย ชุดที่สองคือข้อมูลประโยชน์ที่เกี่ยวข้องกับลักษณะของคุณภาพ

ในการทดลองกับข้อมูลทดสอบชุดแรก แสดงประสิทธิภาพของการจำแนกประโยชน์โดยแนวคิดที่นำเสนอเกี่ยวกับวิธีการจำแนกประโยชน์แบบพื้นฐาน เมื่อการจำแนกประโยชน์โดยแนวคิดที่นำเสนอใช้ประโยชน์ที่เป็นข้อมูลสอนเป็นข้อมูลคนละส่วนกับข้อมูลทดสอบ เริ่มจากแบ่งข้อมูลประโยชน์ที่เกี่ยวข้องกับ

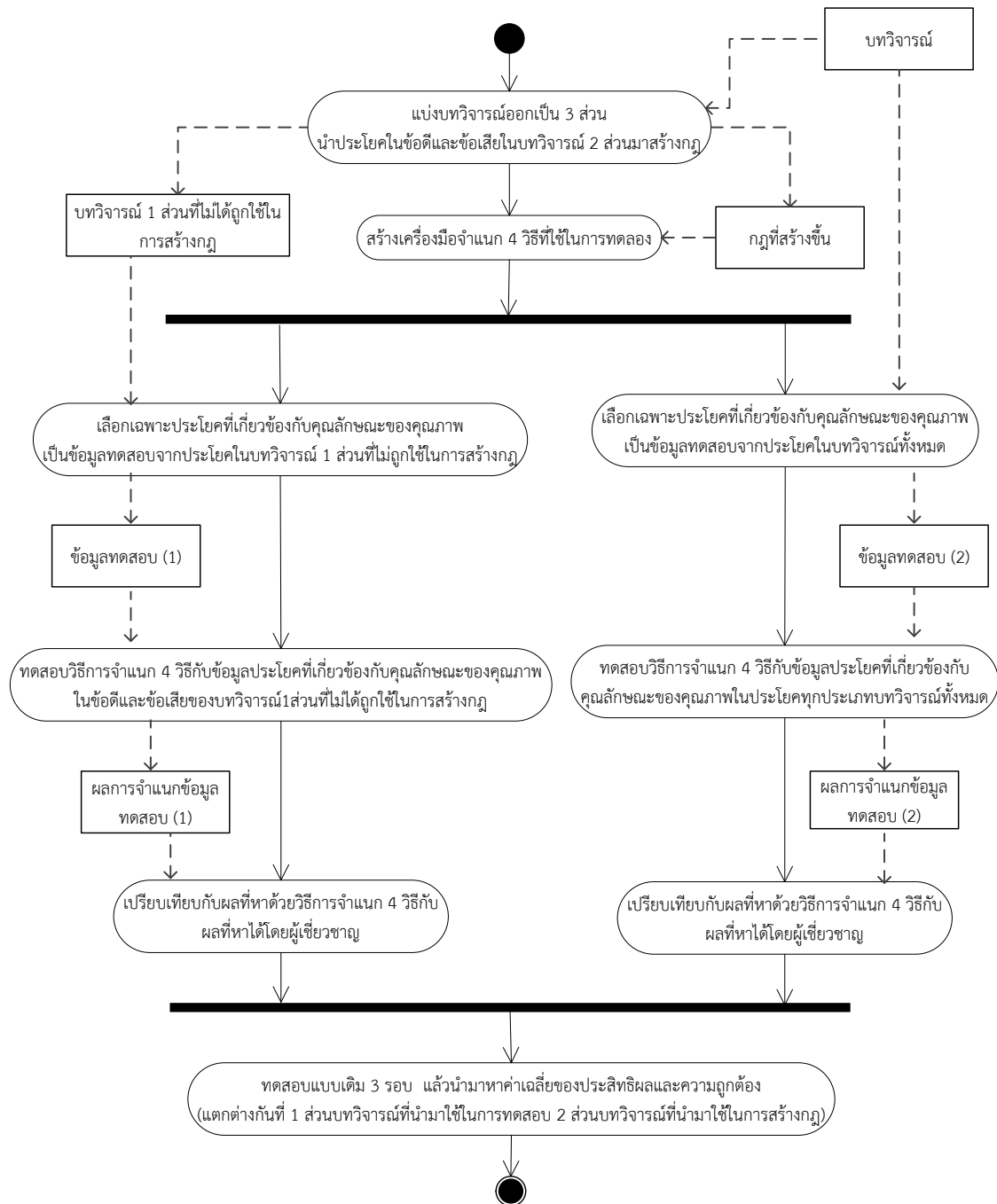
ลักษณะคุณภาพในข้อดีและข้อเสียของบทวิจารณ์ของออกเป็นสามส่วน จำนวนเท่าๆกัน โดยนำสองส่วนมาใช้ในการสร้างกฎ แล้วทดสอบกับข้อมูลประโยคอีกหนึ่งส่วน สรุปจำนวนประโยคทั้งหมดที่ใช้ในการทดสอบจำนวน 5,483 ประโยค

ในการทดลองกับข้อมูลทดสอบชุดที่สอง แสดงประสิทธิผลของการจำแนกประโยคโดยแนวคิดที่นำเสนอเกี่ยวกับวิธีการจำแนกประโยคแบบพื้นฐานในการใช้งานจริง เริ่มจากแบ่งข้อมูลประโยคที่เกี่ยวข้องเกี่ยวกับคุณลักษณะของคุณภาพทั้งหมดในข้อดีข้อเสียออกเป็นสามส่วน จำนวนเท่าๆกัน โดยนำประโยคสองส่วนที่อยู่ในข้อดีข้อเสียมาใช้ในการสร้างกฎ แล้วทดสอบกับข้อมูลประโยคทั้งหมดทุกประเภทในบทวิจารณ์ที่เกี่ยวข้องเกี่ยวกับคุณลักษณะของคุณภาพ สรุปจำนวนประโยคทั้งหมดที่ใช้ในการทดสอบจำนวน 11,217 ประโยค

ภาพรวมของการทดลองในส่วนของการวัดประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่ความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลางแสดงในรูปที่ 4.2

การทดลองมีรายละเอียดแต่ละขั้นตอนดังนี้

- 1) บทวิจารณ์ทั้งหมดถูกแบ่งออกเป็นสามส่วน (ส่วนที่หนึ่ง, ส่วนที่สองและส่วนที่สาม)
- 2) ข้อมูลประโยคในข้อดีและข้อเสียของบทวิจารณ์สองส่วน (ส่วนที่หนึ่งและส่วนที่สอง) ถูกนำมาใช้ในการสร้างกฎที่ใช้ในแนวคิดที่นำเสนอ ในหัวข้อ 3.4.2 ในการทดลองนี้มีข้อมูลบทวิจารณ์ทั้งหมด 3,000 บทวิจารณ์ จะใช้ข้อมูลประโยคในข้อดีและข้อเสียในบทวิจารณ์จำนวน 2,000 บทวิจารณ์นำมาสร้างกฎตามแนวคิดที่นำเสนอในหัวข้อ 3.4.2.1, 3.4.2.2 และ 3.4.2.3 โดยสาเหตุที่การทดลองใช้ข้อมูลเพียงสองในสามส่วน เพราะเมื่อนำไปใช้จริง ไม่สามารถนำข้อมูลทั้งหมดมาสร้างกฎได้ เครื่องมือที่ใช้ในการสร้างกฎคือ เท็กซ์มายเนอร์ซอฟต์แวร์คิต (Text-Miner Software Kit, TMSK) และ รูลอินดักชันคิตฟอ์เท็กซ์ (Rule Induction Kit for Text, RIKTEXT)



รูปที่ 4.2 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิผลของการจำแนกประโยค

3) เลือกเฉพาะประโยคที่เกี่ยวข้องกับลักษณะของคุณภาพเพื่อใช้ในการทดสอบ เพราะประโยคที่ต้องนำไปคิดคะแนนคุณภาพการใช้งานนั้นเฉพาะประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพเท่านั้น ดังนั้นผลการจำแนกประโยคที่ไม่เกี่ยวข้องกับคุณลักษณะของคุณภาพนั้นถูกหรือผิดจะไม่ส่งผลต่อการคิดคะแนนคุณภาพการใช้งาน

(1) ในข้อมูลทดสอบชุดที่หนึ่ง ประโยคในข้อดีและข้อเสียในข้อมูลบทวิจารณ์ในหนึ่งส่วนที่ไม่ใช้ในการสร้างกฎถูกนำมาเลือกเป็นข้อมูลทดสอบเฉพาะประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ

(2) ในข้อมูลทดสอบชุดที่สอง ประโยคทุกประโยคในข้อมูลทั้งหมดของบทวิจารณ์ถูกนำมาเลือกเป็นข้อมูลทดสอบเฉพาะประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ

4) ทดสอบข้อมูลชุดที่หนึ่งและชุดที่สอง ด้วยวิธีการจำแนกสี่วิธีคือ วิธีการจำแนกแบบพื้นฐาน (BL) และวิธีการจำแนกแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานอีกสามวิธี (ALL, ASF และ MSF)

5) เปรียบเทียบผลที่ได้จากวิธีการจำแนกทั้งสี่วิธีกับผลที่ถูกประเมินโดยผู้เชี่ยวชาญ โดยมีการวัดประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวก การวัดประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงลบและความถูกต้องของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง

6) เปลี่ยนข้อมูลบทวิจารณ์ที่นำมาสร้างกฎเป็นส่วนที่สองและส่วนสาม และเปลี่ยนข้อมูลทดสอบชุดที่หนึ่งเป็นข้อมูลบทวิจารณ์ส่วนที่หนึ่ง ชุดข้อมูลทดสอบชุดที่สองคงเดิม (จำนวนประโยคทั้งหมด 11,217 ประโยค) ทำตามขั้นตอนที่ 2 ถึง 5

7) เปลี่ยนข้อมูลบทวิจารณ์ที่นำมาสร้างกฎเป็นส่วนที่หนึ่งและส่วนที่สาม และเปลี่ยนข้อมูลทดสอบชุดที่หนึ่งเป็นข้อมูลบทวิจารณ์ส่วนที่สอง ชุดข้อมูลทดสอบชุดที่สองคงเดิม (จำนวนประโยคทั้งหมด 11,217 ประโยค) ทำตามขั้นตอนที่ 2 ถึง 5

8) นำผลที่ได้มาหาค่าเฉลี่ยของประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวก ประสิทธิผลของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงลบและความถูกต้องของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลางด้วยการจำแนกประโยคสี่วิธีข้างต้น

4.4.2 ผลการทดลองและการวิเคราะห์ผล

ประสิทธิผลและความถูกต้องของการจำแนกสี่วิธีดังแสดงในตารางที่ 4.7 และตารางที่ 4.8

ตารางที่ 4.7 ประสิทธิภาพของการจำแนกประโยคสี่วิธีเมื่อใช้ข้อมูลทดสอบชุดที่หนึ่ง

	การจำแนกประโยคเชิงบวก			การจำแนกประโยคเชิงลบ			ความถูกต้อง (accuracy_3)
	Precision	Recall	F1 Score	Precision	Recall	F1 Score	
BL	83.99%	68.65%	75.54%	57.28%	43.74%	49.60%	58.63%
ALL	81.74%	74.28%	77.83%	54.92%	52.94%	53.91%	62.88%
ASF	82.52%	74.60%	78.36%	55.88%	54.04%	54.94%	63.38%
MSF	82.18%	74.65%	78.23%	56.41%	54.73%	55.56%	63.52%

ตารางที่ 4.8 ประสิทธิภาพของการจำแนกประโยคสี่วิธีเมื่อใช้ข้อมูลทดสอบชุดที่สอง

	การจำแนกประโยคเชิงบวก			การจำแนกประโยคเชิงลบ			ความถูกต้อง (accuracy_3)
	Precision	Recall	F1 Score	Precision	Recall	F1 Score	
BL	84.92%	73.62%	78.87%	61.34%	45.80%	52.45%	62.17%
ALL	82.47%	78.48%	80.43%	58.32%	54.80%	56.51%	65.60%
ASF	82.77%	78.86%	80.76%	58.51%	55.26%	56.84%	65.86%
MSF	82.52%	79.10%	80.77%	59.84%	57.24%	58.51%	66.13%

จากผลการทดลองกับข้อมูลทดสอบชุดที่หนึ่งในตารางที่ 4.7 ค่าความครบถ้วน (recall) ของการจำแนกประโยคที่มีความคิดเห็นเชิงบวก (positive classification) และการจำแนกประโยคที่มีความคิดเห็นเชิงลบ (negative classification) โดยใช้วิธี BL น้อยกว่าวิธี ALL ประมาณ 6 % และ 9% เนื่องจาก วิธี BL มันไม่มีการให้ความสำคัญกับคำศัพท์ที่เกี่ยวข้องและเลือกคำศัพท์ที่ไม่เกี่ยวข้องออก แล้วประโยคที่มีคำศัพท์ทั้งบวกและลบจะถูกจำแนกว่าระบุไม่ได้ ทำให้วิธี BL นี้ประเมินประโยคที่เป็นเชิงบวกและเชิงลบได้จำนวนประยोन้อยกว่า ALL แต่เหตุนี้ทำให้ค่าความแม่นยำ (precision) ของ BL มีค่าสูงกว่าประมาณ 1-3% ซึ่งถ้าเรานำมาเปรียบเทียบประสิทธิภาพการจำแนกประโยคโดยใช้ค่าเอฟสกอร์ (F1 Score) วิธี ALL มีค่าสูงกว่า BL และเมื่อพิจารณาค่าความถูกต้อง (accuracy_3) ALL มีค่าสูงกว่า BL สอดคล้องกันกับค่าเอฟสกอร์

เมื่อพิจารณาระหว่างการจำแนกประโยคโดยวิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานทั้ง 3 รูปแบบคือ ALL, ASF และ MSF ค่าความแม่นยำ ค่าความครบถ้วน ค่าเอฟสกอร์ และความถูกต้องของ ASF สูงกว่า ALL ทุกค่า เนื่องจากกฎของ ASF ที่สร้างความถูกต้องของ

สมมติฐานมากกว่า ALL (สมมติฐานคือประโยคในข้อดีเป็นประโยคที่มีความคิดเห็นเชิงบวกและประโยคในข้อเสียเป็นประโยคที่มีความคิดเห็นเชิงลบ) ประสิทธิภาพของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกของ ASF และ MSF มีค่าใกล้เคียงกัน แต่ประสิทธิภาพของการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงลบของ MSF สูงกว่า ASF เนื่องจากกฎที่ถูกสร้างมาจากข้อมูลประโยคในข้อดีนั้นเป็นประโยคที่มีความคิดเห็นเชิงลบอยู่น้อย แต่ประโยคในข้อเสียนั้นมีประโยคที่มีความคิดเห็นดีอยู่จำนวนหนึ่ง การคัดประโยคเหล่านี้ออกทำให้การสร้างกฎที่มีค่าความคิดเห็นเชิงลบมีความถูกต้องมากขึ้น ส่งผลโดยรวมให้ความถูกต้องของการจำแนกประโยคของ MSF สูงกว่า ASF พิจารณาความถูกต้องของสมมติฐานของประโยคที่ใช้ในการสร้างกฎใน MSF ที่สูงกว่า ASF และ ALL ทำให้ MSF มีประสิทธิภาพและความถูกต้องของการจำแนกประโยคมากที่สุด

เมื่อนำมาทดลองในทุกประโยคของบทวิจารณ์จะได้ผลการทดลองกับข้อมูลทดสอบชุดที่สองแสดงในตารางที่ 4.8 สังเกตได้ว่าผลที่ได้สอดคล้องกับผลในการทดสอบในข้อมูลชุดที่แรก คือ ประสิทธิภาพและความถูกต้องของการจำแนกประโยคของ MSF สูงกว่า ASF และ ALL และการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกมีประสิทธิภาพและความถูกต้องดีกว่าการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงลบนั่น เนื่องจากประโยคที่มีความคิดเห็นเชิงบวกมีการใช้คำศัพท์ที่ไม่หลากหลายเท่ากับประโยคที่มีความคิดเห็นเชิงลบ คำศัพท์ที่ถูกใช้ในประโยคที่มีความคิดเห็นเชิงบวกเป็นคำที่ถูกใช้ในการแสดงความรู้สึกทั่วไปแต่คำศัพท์ที่ถูกใช้ในประโยคที่มีความคิดเห็นเชิงลบนั่นส่วนมากเป็นคำเฉพาะเกี่ยวกับสินค้านั้น เช่น “Use big Ram space” เป็นต้น

4.4.3 สรุปผลการทดลอง

จากผลการทดลอง เมื่อเปรียบเทียบค่าความแม่นยำ ค่าความครบถ้วน ค่าเอฟสกอว์และความถูกต้องของการจำแนกประโยคแบบพื้นฐาน (BL) น้อยกว่าการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ทั้งในการทดสอบกับข้อมูลประโยคในข้อดีและข้อเสีย (ข้อมูลชุดที่หนึ่ง) กับข้อมูลประโยคทั้งหมดในบทวิจารณ์ทั้งหมด (ข้อมูลชุดที่สอง) สรุปได้ว่าประสิทธิภาพของการจำแนกประโยคโดยวิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานให้ผลดีกว่าวิธีจำแนกประโยคแบบพื้นฐาน (BL) เมื่อพิจารณาวิธีการจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน ALL, ASF และ MSF ใน

การทดสอบทั้งสองข้อมูล ประสิทธิภาพและความถูกต้องของ MSF มีค่าสูงสุด เพราะความถูกต้องของสมมติฐานมีความถูกต้องมากขึ้น ทำให้ประสิทธิภาพและความถูกต้องของการจำแนกประโยคสูงขึ้น

4.4.4 ข้อมูลสนับสนุนผลการทดลอง

4.4.4.1 ลักษณะของข้อมูลเบื้องต้น

ในบทวิจารณ์ ประโยคที่เราจะนำมาช่วยพัฒนาวิธีการที่เราจะนำเสนอคือประโยคที่อยู่ในส่วนข้อดีข้อเสีย ประโยคในข้อดีมีจำนวน 4,938 ประโยคและประโยคในข้อเสียมีจำนวน 4,828 ประโยค จำนวนประโยคในข้อดีและข้อเสียถูกจำแนกเป็นประโยคที่มีความคิดเห็นเชิงบวก เชิงลบ หรือเป็นกลาง ดังตารางที่ 4.9

ตารางที่ 4.9 จำนวนประโยคที่มีความคิดเห็นเชิงบวก เชิงลบและเป็นกลางในข้อดีและข้อเสีย

ประโยค	จำนวนประโยคที่มีความคิดเห็น			จำนวนประโยคทั้งหมด
	เชิงบวก	เป็นกลาง	เชิงลบ	
ข้อดี	2,760	1,917	261	4,938
ข้อเสีย	669	3,049	1,110	4,828

โดยประโยคที่อยู่ในข้อดีถูกคาดว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกทั้งหมดและประโยคที่อยู่ในข้อเสียถูกคาดว่าเป็นประโยคที่มีความคิดเห็นเชิงลบทั้งหมด จากตารางข้างต้นมีความถูกต้องเพียง 28.26% เพื่อนำประโยคเหล่านี้ไปสร้างกฎที่มีความถูกต้องสูง จำเป็นต้องกรองประโยคให้ข้อมูลมีความถูกต้องที่มากขึ้น ประโยคสามประเภทที่จะถูกกรองออกคือ

1) ประโยคเงื่อนไข (conditional sentence) จาก [30] ประโยคส่วนใหญ่ที่เป็นประโยคเงื่อนไขจะประกอบด้วยคำว่า “if” และประโยคเหล่านั้นส่วนใหญ่จะเป็นประโยคที่มีความคิดเห็นเป็นกลาง ประโยคที่เป็นถ้าแล้วในข้อดีและข้อเสียมีจำนวน 346 ประโยค ซึ่งเป็นประโยคที่มีความคิดเห็นเป็นกลาง 226 ประโยค คิดเป็น 65.32%

2) ประโยคในรูปอดีต ประโยคที่อยู่รูปอดีตในข้อดีและข้อเสียมีจำนวน 2,290 ประโยค ซึ่งเป็นประโยคที่มีความคิดเห็นเป็นกลาง 1,259 ประโยค คิดเป็น 54.98%

3) ประโยคที่ประกอบด้วยคำนิเสธ ประโยคที่ประกอบด้วยคำนิเสธในข้อดีข้อเสียมีจำนวน 1,691 ประโยค ซึ่งเป็นประโยคที่มีความคิดเห็นเป็นกลาง 969 ประโยค คิดเป็น 57.30%

จากข้อมูลในตารางที่ 4.7 ประโยคในข้อดีไม่ควรมีประโยคที่มีความคิดเห็นเชิงลบและประโยคในข้อเสียไม่ควรมีประโยคที่มีความคิดเห็นเชิงบวก ในหัวข้อ 3.4.2.3 การตัดประโยคที่ผิดพลาดเหล่านี้ด้วยออกโดยผู้เชี่ยวชาญออกหลังจากกรองประโยคทั้ง 3 ประเภทออกนั้นเพื่อเพิ่มความถูกต้องของสมมติฐานที่ใช้ในแนวคิดที่น่าเสนอตามตารางที่ 4.10

ตารางที่ 4.10 ความถูกต้องของสมมติฐานในกรณีต่างๆ

กรณี	จำนวนประโยคที่ตรงสมมติฐาน	จำนวนประโยคทั้งหมด	ร้อยละของประโยคที่ตรงสมมติฐาน
กรณีไม่กรองประโยคออก	2,760	9,766	28.26%
กรณีกรองประโยคเงื่อนไข	2,713	9,420	28.80%
กรณีกรองประโยคในรูปอดีต	2,484	8,075	30.76%
กรณีกรองประโยคประกอบด้วยคำนิเสธ	2,422	7,476	32.40%
กรณีกรองประโยคทั้ง 3 ประเภทออก (ASF)	2,189	6,122	35.76%
กรณีกรองประโยคด้วยวิธี MSF	2,189	5,656	38.70%

4.4.4.2 ประโยคปฏิเสธ

ในการจำแนกที่น่าเสนอมีการปรับประโยคที่ประกอบด้วยคำนิเสธให้มีความคิดเห็นตรงกันข้ามจากเดิม การปรับประโยคที่ประกอบด้วยคำนิเสธให้มีความคิดเห็นตรงกันข้ามกับที่วิเคราะห์มาจะช่วยเพิ่มความถูกต้อง โดยการเปรียบเทียบความถูกต้องระหว่างประโยคที่วิเคราะห์ความคิดเห็นเชิงบวกหรือเชิงลบผ่านทั้งวิธีการจำแนกประโยคแบบพื้นฐาน (BL) และวิธีการจำแนกประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานแบบอัตโนมัติ (ASF) ด้วยการปรับประโยค (N1) และไม่มีการปรับความคิดเห็นในประโยค (N0) ในการทดลองนี้ ข้อมูลที่ใช้ในการทดสอบเป็นประโยคที่ประกอบด้วยคำนิเสธที่เกี่ยวข้องกับคุณลักษณะของคุณภาพในบทวิจารณ์ทั้งหมด 2,236 ประโยค โดยเป็นประโยคที่มีความ

คิดเห็นเชิงบวก 926 ประโยค ประโยคที่มีความคิดเห็นเชิงลบ 874 ประโยคและประโยคที่มีความคิดเห็นเป็นกลาง 436 ประโยค ผลการจำแนกแสดงดังตารางที่ 4.11

ตารางที่ 4.11 ความถูกต้องของการจำแนกประโยคที่ประกอบด้วยคำนิเสธ

วิธีการ	ความถูกต้อง		
	การจำแนกประโยคที่มีความคิดเห็นเชิงบวก	การจำแนกประโยคที่มีความคิดเห็นเชิงลบ	การจำแนกประโยค
BL+N0	46.06%	52.77%	26.79%
BL+N1	61.72%	56.44%	36.45%
ASF+N0	44.95%	51.14%	31.00%
ASF+N1	59.55%	53.02%	39.24%

จากตารางที่ 4.11 สังเกตได้ว่า ค่าสูงขึ้นในทุกค่าในค่าความถูกต้องของการจำแนกประโยคที่มีความคิดเห็นเชิงบวก (positive classification) ค่าความถูกต้องในการจำแนกประโยคที่มีความคิดเห็นเชิงลบ (negative classification) และค่าความถูกต้องในการจำแนกประโยคที่มีความคิดเห็นเชิงบวกเชิงลบและเป็นกลาง ทำให้การกลับประโยคที่ประกอบด้วยคำนิเสธถูกนำมาใช้ในแนวคิดที่นำเสนอ

4.5 การวัดประสิทธิผลส่วนการคำนวณคะแนนคุณภาพการใช้งาน

ในส่วนนี้มีจุดประสงค์คือ การวัดประสิทธิผลของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบในคุณภาพสี่ด้านด้วยวิธีที่นำเสนอเปรียบเทียบกับวิธีจำแนกประโยคแบบพื้นฐาน (BL) โดยใช้บทวิจารณ์เป็นข้อมูลทดสอบ ความคิดเห็นของบทวิจารณ์เหล่านั้นจะถูกพิจารณาโดยผู้เชี่ยวชาญว่ามีความคิดเห็นเชิงบวก เชิงลบ หรือเป็นกลางในคุณลักษณะของคุณภาพด้านประสิทธิผล (effectiveness) ผลิตผล (productivity) ความปลอดภัย (safety) และความพึงพอใจ (satisfaction) วิธีการจำแนกประโยคที่ใช้ในการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงใดในการทดลองนี้มีสามวิธีมาเปรียบเทียบกันคือ

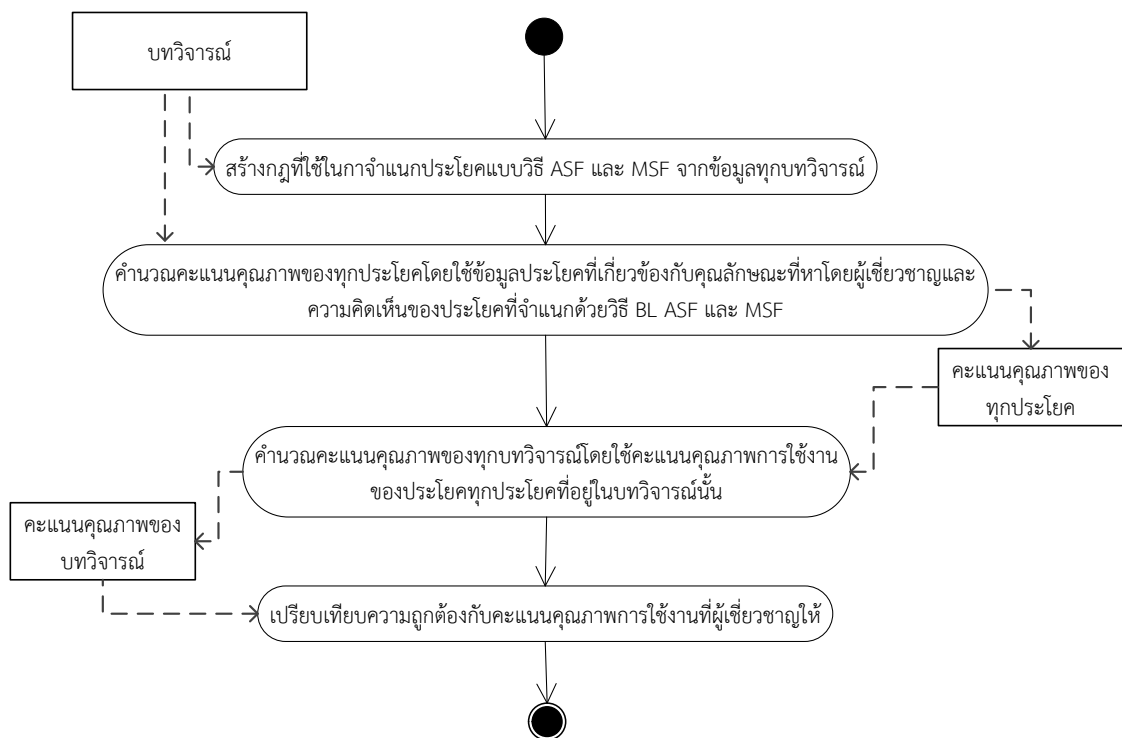
- 1) วิธีจำแนกประโยคแบบพื้นฐาน (BL) รายละเอียดถูกอธิบายในหัวข้อ 3.4.1

2) วิธีจำแนกประเภทแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกอัตโนมัติ (ASF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.2

3) วิธีจำแนกประเภทแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.3

4.5.1 ขั้นตอนการทดลอง

ในการทดลองนี้ใช้ข้อมูลทดสอบคือบทวิจารณ์ 3,000 บทวิจารณ์ ภาพรวมของการทดลองของการวัดประสิทธิผลส่วนการคำนวณคะแนนคุณภาพการใช้งานแสดงรูปที่ 4.3



รูปที่ 4.3 แผนภาพกิจกรรมของภาพรวมการวัดความถูกต้องของการคำนวณคะแนนคุณภาพการใช้งาน

การทดลองมีรายละเอียดแต่ละขั้นตอนดังนี้

1) ประโยคในข้อดีและข้อเสียในบทวิจารณ์ทั้งหมดถูกนำมาเป็นข้อมูลสอนของการสร้างกฎด้วยวิธี ASF และ MSF

2) คะแนนคุณภาพการใช้งานของทุกประโยคในบทวิจารณ์ทั้งหมดถูกคำนวณจากข้อมูลประโยคที่บอกว่าประโยคเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดที่หาโดยผู้เชี่ยวชาญ และข้อมูลความคิดเห็นของประโยคที่ถูกวิเคราะห์จำแนกด้วยวิธี BL ASF และ MSF รายละเอียดวิธีการคำนวณแสดงในขั้นตอนที่หนึ่งในหัวข้อ 3.5 สาเหตุที่การทดลองใช้ข้อมูลประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดที่หาจากผู้เชี่ยวชาญนั้น เพื่อวัดประสิทธิผลของวิธีการจำแนกประโยคที่นำเสนอว่ามีผลต่อการคำนวณคะแนนคุณภาพเท่าใด

3) คะแนนคุณภาพการใช้งานของแต่ละบทวิจารณ์จากคะแนนคุณภาพการใช้งานของแต่ละประโยคในบทวิจารณ์นั้น รายละเอียดวิธีการคำนวณแสดงในขั้นตอนที่สองในหัวข้อ 3.5

4) คะแนนคุณภาพการใช้งานของแต่ละบทวิจารณ์เปรียบเทียบความถูกต้องกับคะแนนคุณภาพการใช้งานที่ผู้เชี่ยวชาญให้ความถูกต้องของคะแนนคุณภาพการใช้งานที่คำนวณได้ถูกต้องร้อยละเท่าใด (คะแนนคุณภาพการใช้งานในแต่ละด้านมีค่า -1, 0 และ 1)

4.5.2 ผลการทดลองและการวิเคราะห์ผล

จากการทดลอง ความถูกต้องของการคำนวณคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์แสดงในตารางที่ 4.12

ตารางที่ 4.12 ความถูกต้องของการคำนวณคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์

วิธีการจำแนกประโยคที่ใช้ในการคำนวณคะแนน	ด้านประสิทธิผล	ด้านผลิตผล	ด้านความปลอดภัย	ด้านความพึงพอใจ	ค่าเฉลี่ย
BL	81.00%	79.77%	80.73%	79.77%	80.32%
ASF	82.57%	81.40%	82.77%	80.87%	81.90%
MSF	82.57%	81.83%	82.97%	80.83%	82.05%

จากตารางที่ 4.12 การคำนวณคะแนนคุณภาพการใช้งานด้วยการจำแนกประโยคที่มีความคิดเห็นเชิงใดด้วยวิธี ASF และ MSF จะให้ผลค่าความถูกต้อง (accuracy) ใกล้เคียงกันในทุกคุณลักษณะของคุณภาพ ค่าความถูกต้องของทั้งสองวิธีนั้นมากกว่าการจำแนกประโยคที่มีความคิดเห็นเชิงใดด้วยวิธี BL ในทุกคุณลักษณะของคุณภาพ ทั้งด้านประสิทธิผล (effectiveness) ผลิตผล

(productivity) ความปลอดภัย (safety) และความพึงพอใจ (satisfaction) ประมาณ 1-2% ทำให้ค่าเฉลี่ยของความถูกต้องของการคำนวณคะแนนคุณภาพการใช้งานด้วยวิธีการจำแนกประโยชน์กว่ามีความคิดเห็นเชิงใดด้วยวิธี MSF สูงกว่าวิธีการจำแนกประโยชน์กว่ามีความคิดเห็นเชิงใดด้วยวิธี BL 1.73% เนื่องจากวิธีการที่นำเสนอมีประสิทธิภาพการจำแนกประโยชน์กว่าเป็นประโยชน์ที่มีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลางได้ดีกว่าวิธีการจำแนกประโยชน์แบบพื้นฐาน (BL)

4.5.3 สรุปผลการทดลอง

จากผลการทดลอง การคำนวณคะแนนคุณภาพการใช้งานด้วยการจำแนกประโยชน์กว่ามีความคิดเห็นเชิงใดด้วยวิธี MSF ให้ความถูกต้อง 82.05% ซึ่งให้ผลความถูกต้องมากกว่าการคำนวณคะแนนคุณภาพการใช้งานด้วยการจำแนกประโยชน์กว่ามีความคิดเห็นเชิงใดด้วยวิธี BL จากค่าเฉลี่ยของความถูกต้องของคุณลักษณะของคุณภาพสี่ด้านที่สูงขึ้น 1.73%

4.6 การวัดประสิทธิผลของการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง

คะแนนคุณภาพการใช้งานของบทวิจารณ์ควรสอดคล้องกับความคิดเห็นในบทวิจารณ์ ทั้งนี้ความคิดเห็นของบทวิจารณ์นั้นจะสอดคล้องกับจำนวนดาวที่ผู้เขียนวิจารณ์ให้กับบทวิจารณ์นั้น ดังนั้นความคิดเห็นที่คิดจากคะแนนคุณภาพการใช้งานของบทวิจารณ์นั้นควรจะสอดคล้องกับความคิดเห็นที่พิจารณาจากจำนวนดาว โดยจำนวนดาวมากกว่าสาม บทวิจารณ์นั้นมีความคิดเห็นเชิงบวก จำนวนดาวเท่ากับสามแสดงว่าความคิดเห็นของบทวิจารณ์นั้นเป็นกลาง และจำนวนดาวน้อยกว่าสามแสดงว่าความคิดเห็นของบทวิจารณ์นั้นเป็นลบ ความสอดคล้องนี้เปรียบเสมือนการคาดการณ์ความคิดเห็นของบทวิจารณ์อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งาน ในส่วนนี้มีจุดประสงค์คือเปรียบเทียบประสิทธิผลของการคาดการณ์ความคิดเห็นจากจำนวนดาวในบทวิจารณ์จากคะแนนคุณภาพการใช้งานกับการคาดการณ์ความคิดเห็นจากจำนวนดาวในบทวิจารณ์จากความคิดเห็นแต่ละประโยค (ไม่ใช่คุณลักษณะของคุณภาพการใช้งานในการพิจารณา)

1) การคาดการณ์ความคิดเห็นของบทวิจารณ์ที่อ้างอิงจากจำนวนดาวโดยพิจารณาจากความคิดเห็นของแต่ละประโยค (P_0) คือการนับจำนวนประโยคที่มีความคิดเห็นเชิงบวกและลบในบทวิจารณ์

ถ้าจำนวนประโยชน์ส่วนมากมีความคิดเห็นเชิงใด แล้วความคิดเห็นของบทวิจารณ์จะถูกวิเคราะห์ที่ได้ไปว่า
มีความคิดเห็นไปเชิงนั้น สามารถเขียนในรูปสมการ (4.6)

$$rp = \begin{cases} 1 & \text{if } \sum sp_i > 0 \\ 0 & \text{if } \sum sp_i = 0 \\ -1 & \text{if } \sum sp_i < 0 \end{cases} \quad (4.6)$$

โดย rp คือค่าความคิดเห็นของบทวิจารณ์ที่คาดการณ์จากความคิดเห็นของแต่ละประโยชน์
 sp_i คือค่าความคิดเห็นของประโยชน์ที่ i (ความคิดเห็นเชิงบวกของประโยชน์มีค่า 1 ความคิดเห็นเป็น
กลางของประโยชน์มีค่า 0 และความคิดเห็นเชิงลบของประโยชน์มีค่า -1)

2) การคาดการณ์ความคิดเห็นของบทวิจารณ์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งาน
งานของบทวิจารณ์ (P_1) คือการคิดคะแนนจากแต่ละคุณลักษณะของคุณภาพซึ่งแต่ละคุณลักษณะของ
คุณภาพอาจจะมีน้ำหนักไม่เท่ากัน เมื่อคิดผลรวมของคะแนนจากทุกคุณลักษณะของคุณภาพแล้วมีค่า
เป็นบวก ลบหรือศูนย์ ความคิดเห็นของบทวิจารณ์ถูกคาดการณ์ว่ามีความคิดเห็นเชิงบวก ลบและเป็น
กลางตามลำดับ สามารถเขียนในรูปสมการ (4.7) และ (4.8)

$$r(w_0, w_1, w_2, w_3) = \sum_{i=0}^3 (w_i \times rq_i) \quad (4.7)$$

$$rpo(w_0, w_1, w_2, w_3) = \begin{cases} 1 & \text{if } r(w_0, w_1, w_2, w_3) > 0 \\ 0 & \text{if } r(w_0, w_1, w_2, w_3) = 0 \\ -1 & \text{if } r(w_0, w_1, w_2, w_3) < 0 \end{cases} \quad (4.8)$$

โดย r คือผลรวมคะแนนจากทุกคุณลักษณะของคุณภาพ w_i คือน้ำหนักความสำคัญของ
คุณลักษณะของคุณภาพ i ที่ส่งผลต่อความคิดเห็นของบทวิจารณ์ rq_i คือคะแนนคุณลักษณะของ
คุณภาพ i คะแนนคุณภาพการใช้งานของบทวิจารณ์ rpo คือค่าความคิดเห็นของบทวิจารณ์ที่คาดการณ์
จากคะแนนคุณภาพการใช้งานของบทวิจารณ์ คุณลักษณะของคุณภาพ $i = 0, 1, 2$ และ 3 หมายถึง
คุณลักษณะของคุณภาพด้านประสิทธิผล (effectiveness) ผลผลิต (productivity) ความปลอดภัย
(safety) และความพึงพอใจ (satisfaction)

3) การคาดการณ์ความคิดเห็นของบทวิจารณ์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งาน
งานของบทวิจารณ์และจำนวนประโยชน์ที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ (P_2) คือการคิดคะแนน

จากแต่ละคุณลักษณะของคุณภาพซึ่งแต่ละคุณลักษณะของคุณภาพอาจจะมีจำนวนประโยชน์ไม่เท่ากัน เมื่อคิดผลรวมของคะแนนจากทุกคุณลักษณะของคุณภาพแล้วมีค่าเป็นบวก ลบหรือศูนย์ ความคิดเห็นของบทวิจารณ์ถูกคาดการณ์ว่ามีความคิดเห็นเชิงบวก ลบและเป็นกลางตามลำดับ สามารถเขียนในรูปสมการ (4.9) และ (4.10)

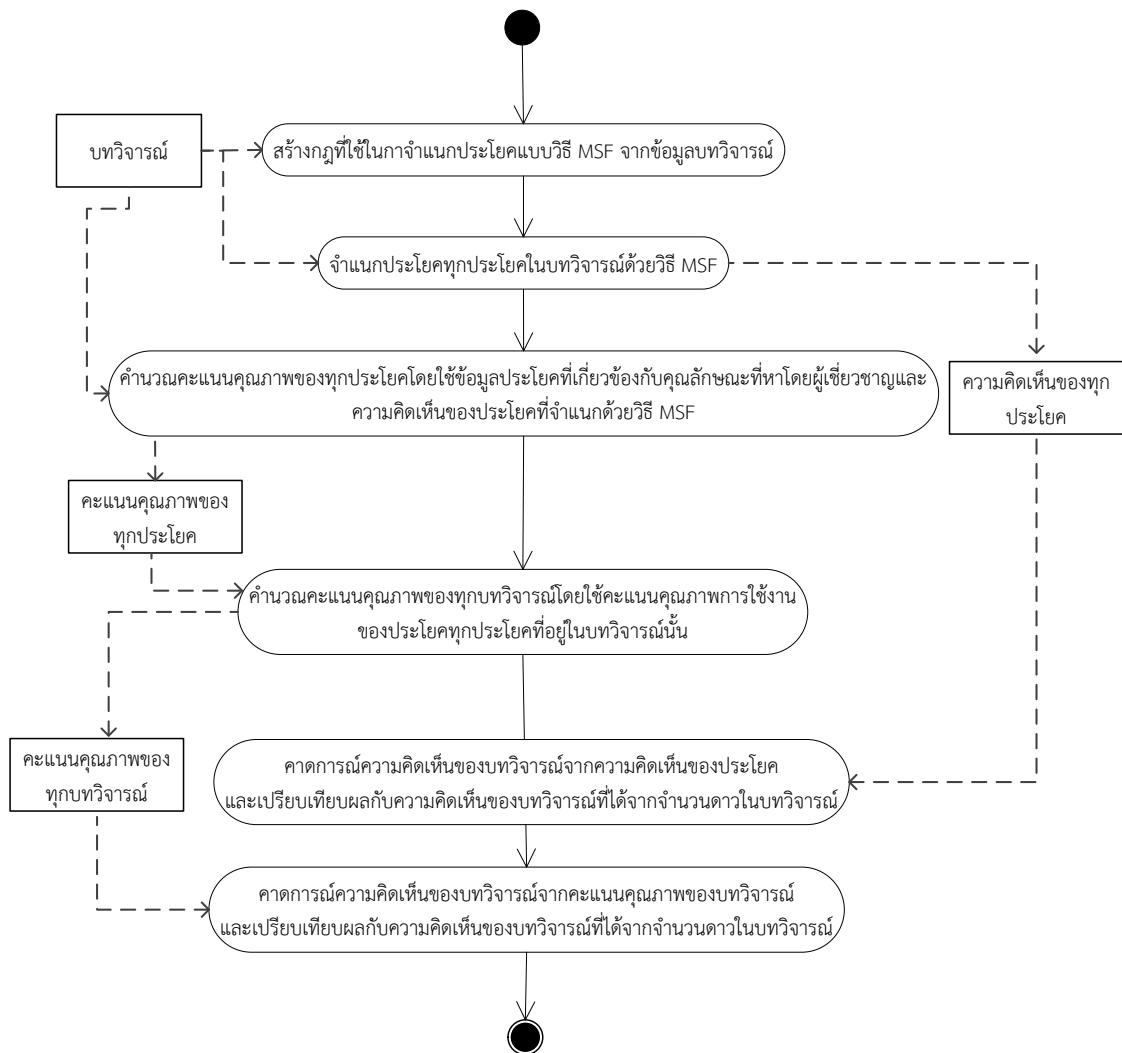
$$r = \sum_{i=0}^3 (n_i \times r q_i) \quad (4.9)$$

$$rpo = \begin{cases} 1 & \text{if } \sum r > 0 \\ 0 & \text{if } \sum r = 0 \\ -1 & \text{if } \sum r < 0 \end{cases} \quad (4.10)$$

โดย r คือผลรวมคะแนนจากทุกคุณลักษณะของคุณภาพ n_i คือจำนวนประโยชน์ที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ i ที่ส่งผลต่อความคิดเห็นของบทวิจารณ์ $r q_i$ คือคะแนนคุณลักษณะของคุณภาพ i คะแนนคุณภาพการใช้งานของบทวิจารณ์ rpo คือค่าความคิดเห็นของบทวิจารณ์ที่คาดการณ์จากคะแนนคุณภาพการใช้งานของบทวิจารณ์ คุณลักษณะของคุณภาพ $i = 0, 1, 2$ และ 3 หมายถึงคุณลักษณะของคุณภาพด้านประสิทธิภาพ, ผลผลิต, ความปลอดภัย และความพึงพอใจ

4.6.1 ขั้นตอนการทดลอง

ในการทดลอง เนื่องจากทดสอบความสัมพันธ์ของความคิดเห็นของบทวิจารณ์ที่คิดได้กับจำนวนดาว ซึ่งข้อมูลทดสอบส่วนใหญ่เป็นบทวิจารณ์ที่มีความคิดเห็นเชิงบวก การเลือกข้อมูลทดสอบจึงมีความจำเป็น โดยเลือกข้อมูลบทวิจารณ์ที่มีความคิดเห็นเชิงบวกที่พิจารณาจากจำนวนสี่และห้าดาวสองส่วน บทวิจารณ์ที่มีความคิดเห็นเป็นกลางที่พิจารณาจากจำนวนสามดาวหนึ่งส่วนและบทวิจารณ์ที่มีความคิดเห็นเชิงลบที่พิจารณาจากจำนวนสองและหนึ่งดาวหนึ่งส่วน สรุปในการทดลองจะใช้ข้อมูลทดสอบจำนวน 1,140 บทวิจารณ์ประกอบด้วย จำนวนบทวิจารณ์ที่มีความคิดเห็นเชิงบวกหรือจำนวนดาวมากกว่าสามดาวจำนวน 456 บทวิจารณ์ จำนวนบทวิจารณ์ที่มีความคิดเห็นเป็นกลางหรือจำนวนดาวเท่ากับสามดาวจำนวน 228 บทวิจารณ์และจำนวนบทวิจารณ์ที่มีความคิดเห็นเชิงลบหรือจำนวนดาวน้อยกว่าสามดาวจำนวน 456 บทวิจารณ์ ภาพรวมของการทดลองของการวัดประสิทธิภาพส่วนการคำนวณคะแนนคุณภาพการใช้งานแสดงรูปที่ 4.4



รูปที่ 4.4 แผนภาพกิจกรรมของภาพรวมการวัดประสิทธิภาพการวิเคราะห์ความคิดเห็นของบทวิจารณ์

การทดลองมีรายละเอียดแต่ละขั้นตอนดังนี้

- 1) ประโยคในข้อดีและข้อเสียในบทวิจารณ์ทั้งหมดถูกนำมาเป็นข้อมูลสอนของการสร้างกฎด้วยวิธี MSF
- 2) ประโยคทุกประโยคในบทวิจารณ์ถูกจำแนกว่ามีความคิดเห็นเชิงใดด้วยการจำแนกด้วยวิธี MSF
- 3) คะแนนคุณภาพการใช้งานของทุกประโยคในบทวิจารณ์ทั้งหมดถูกคำนวณจากข้อมูลประโยคที่บอกว่าประโยคเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดที่ทำโดยผู้เชี่ยวชาญ และข้อมูลความคิดเห็นของประโยคที่ถูกวิเคราะห์จำแนกด้วยวิธี MSF รายละเอียดวิธีการคำนวณแสดงในขั้นตอนที่หนึ่งในหัวข้อ 4.5 สาเหตุที่การทดลองใช้ข้อมูลประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านใดที่

หาจากผู้เชี่ยวชาญนั้น เพื่อวัดประสิทธิผลของวิธีการจำแนกประโยคที่นำเสนอว่ามีผลต่อการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงใด

4) คะแนนคุณภาพการใช้งานของแต่ละบทวิจารณ์จากคะแนนคุณภาพการใช้งานของแต่ละประโยคในบทวิจารณ์นั้น รายละเอียดวิธีการคำนวณแสดงในขั้นตอนที่สองในหัวข้อ 4.5

5) คาดการณ์ความคิดเห็นบทวิจารณ์หรือวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงใดจากความคิดเห็นของทุกประโยคในบทวิจารณ์นั้น โดยวิธี P_0 , P_1 และ P_2 แล้วเปรียบเทียบกับความถูกต้องกับความคิดเห็นของบทวิจารณ์ที่ได้จากจำนวนดาวที่ผู้วิจารณ์ให้ของบทวิจารณ์นั้น

4.6.2 ผลการทดลองและการวิเคราะห์ผล

จากการทดลองจะได้ค่าความถูกต้องของการวิเคราะห์หรือคาดการณ์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือความคิดเห็นเชิงลบหรือเป็นกลาง แสดงผลตามตารางที่ 4.13

ตารางที่ 4.13 ความถูกต้องของการวิเคราะห์ความคิดเห็นของบทวิจารณ์

วิธีการ	ความถูกต้อง
P_0	58.07%
P_1 กรณี $rpo(1,1,1,1)$	56.67%
P_1 กรณี $rpo(2,1,1,1)$	56.75%
P_1 กรณี $rpo(1,2,1,1)$	55.61%
P_1 กรณี $rpo(1,1,2,1)$	57.02%
P_1 กรณี $rpo(1,1,1,2)$	58.51%
P_2	58.60%

จากตารางที่ 4.13 ยกตัวอย่างกรณีศึกษาที่ให้น้ำหนักธรรมดาจะมีค่าน้ำหนักเป็น 1 และให้น้ำหนักแบบเน้นจะมีค่าน้ำหนักเป็น 2 การจำแนกบทวิจารณ์โดยใช้คุณลักษณะในกรณีที่ให้ค่าน้ำหนักความสำคัญทุกคุณลักษณะเท่ากัน P_1 ในกรณี $rpo(1,1,1,1)$ มีประสิทธิผลด้านความถูกต้องน้อยกว่าการจำแนกบทวิจารณ์โดยความคิดเห็นของประโยค P_0 เนื่องจากแต่ละคุณลักษณะของคุณภาพมีผลต่อความคิดเห็นของบทวิจารณ์นั้นไม่เท่ากัน และเมื่อลองให้น้ำหนักกับแต่ละคุณลักษณะของคุณภาพเพิ่มขึ้น กรณีที่ให้ค่าน้ำหนักความสำคัญในคุณภาพด้านความพึงพอใจ (satisfaction) P_1 กรณี

$rpo(1,1,1,2)$ มีประสิทธิผลด้านความถูกต้องมากกว่ากรณีที่ให้ค่าน้ำหนักความสำคัญในคุณภาพด้านอื่นๆ และมากกว่า P_0 เล็กน้อย เนื่องจากการจำแนกบทวิจารณ์โดยใช้คุณลักษณะของคุณภาพ กรณีที่ให้ค่าน้ำหนักความสำคัญในคุณภาพด้านประสิทธิผล ผลผลิต และความปลอดภัย มีค่าประสิทธิผลด้านความถูกต้องน้อยกว่ากรณีอื่นๆ เพราะบทวิจารณ์จะมีความคิดเห็นเชิงบวกหรือเชิงลบนั้นมาจากความชอบหรือพอใจของผู้วิจารณ์ต่อสินค้าหรือบริการนั้นเป็นส่วนสำคัญ ดังนั้นคุณภาพด้านความพึงพอใจ จึงมีผลต่อความคิดเห็นเชิงบวกหรือเชิงลบมากที่สุด แต่เมื่อคิดจำนวนประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพนั้นแทนน้ำหนักที่ให้ ใน P_2 จะมีประสิทธิผลด้านความถูกต้องสูงสุดแต่ใกล้เคียง P_1 กรณี $rpo(1,1,1,2)$ เนื่องจากวิธี P_2 คิดจากจำนวนประโยคซึ่งประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านความพึงพอใจมีจำนวนมากสุดเมื่อเทียบกับคุณลักษณะด้านอื่นๆ (ในจำนวนประโยค 18,679 ประโยค ประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านประสิทธิผลจำนวน 2,236 ประโยค ด้านผลผลิตจำนวน 3,529 ประโยค ด้านความปลอดภัย 2,351 ประโยคและด้านความพึงพอใจจำนวน 6,429 ประโยค)

4.6.3 สรุปผลการทดลอง

จากการทดลองสรุปได้ว่าการวิเคราะห์ความคิดเห็นของบทวิจารณ์ที่ได้จากแนวคิดที่นำเสนอโดยใช้คะแนนคุณภาพการใช้งานของบทวิจารณ์และจำนวนประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพให้ประสิทธิผลในด้านความถูกต้องมากกว่าการวิเคราะห์ความคิดเห็นของบทวิจารณ์ที่ได้จากแนวคิดที่นำเสนอโดยใช้เพียงความคิดเห็นของทุกประโยคและคุณลักษณะของคุณภาพด้านความพึงพอใจ (satisfaction) มีผลต่อจำนวนดาวที่ให้ในบทวิจารณ์มากที่สุด

4.6.4 ข้อมูลสนับสนุนผลการทดลอง

4.6.4.1 ความสัมพันธ์ระหว่างคุณภาพและความคิดเห็นที่แสดงในรูปแบบของจำนวนดาวในบทวิจารณ์

คุณลักษณะของคุณภาพแต่ละด้านของซอฟต์แวร์จะมีผลต่อจำนวนดาวหรือความคิดเห็นไม่เท่ากัน เพื่อหาว่าคุณภาพด้านใดมีผลต่อความคิดเห็นและจำนวนดาวของบทวิจารณ์มากที่สุด จึงมีการทดลองสองการทดลองคือ ความสัมพันธ์ระหว่างคะแนนคุณภาพและจำนวนดาวที่ให้โดยผู้วิจารณ์ในบท

วิจารณ์ และความสัมพันธ์ระหว่างคะแนนคุณภาพและความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลาง โดยรวมของบทวิจารณ์ ใช้ข้อมูลทดสอบทั้งหมด 3,000 บทวิจารณ์ในการทดลอง

ความสัมพันธ์ระหว่างคะแนนคุณภาพการใช้งานและจำนวนดาวที่ให้โดยผู้วิจารณ์ในบทวิจารณ์ ถูกแสดงในรูปสมการ 4.11

$$s_i = \beta_0 + \beta_1 q_{i1} + \beta_2 q_{i2} + \beta_3 q_{i3} + \beta_4 q_{i4} + e_i \quad (4.11)$$

s_i คือจำนวนดาวของบทวิจารณ์ที่ผู้วิจารณ์ให้ ($s_i \in \{1, 2, 3, 4, 5\}$) q_{ij} เป็นคะแนนคุณลักษณะของคุณภาพ j ของบทวิจารณ์ i (effectiveness, productivity, safety และ satisfaction) โดยที่ q_{ij} มีค่า -1, 0 และ 1

เมื่อคิดคำนวณบทวิจารณ์จำนวน 3,000 บทวิจารณ์ด้วยวิธีการประมาณค่าพารามิเตอร์ของสมการความถดถอยเพื่อหาค่าสัมประสิทธิ์ของสมการเมื่อค่าเฉลี่ยของความผิดพลาดน้อยที่สุด ผลลัพธ์ที่ได้แสดงในตารางที่ 4.14

ตารางที่ 4.14 สัมประสิทธิ์ของสมการและค่าเฉลี่ยของความผิดพลาดของสมการ 4.11

Base	Effectiveness	Productivity	Safety	Satisfaction	ค่าเฉลี่ยของ
(β_0)	(β_1)	(β_2)	(β_3)	(β_4)	$ e_i $
2.8888	0.3228	0.3160	0.4872	1.1933	0.68056

จะสังเกตได้ว่าคุณภาพด้านความพึงพอใจ จะมีผลต่อจำนวนดาวที่ให้โดยผู้วิจารณ์มากที่สุด รองลงมาคือ ความปลอดภัย ประสิทธิภาพ และผลิตผลตามลำดับ

ความสัมพันธ์ระหว่างคะแนนคุณภาพการใช้งานและความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลางโดยรวมของบทวิจารณ์ถูกแสดงในรูปสมการ 4.12

$$p_i = \beta_0 + \beta_1 q_{i1} + \beta_2 q_{i2} + \beta_3 q_{i3} + \beta_4 q_{i4} + e_i \quad (4.12)$$

p_i คือความคิดเห็นเชิงบวกหรือเชิงลบหรือเป็นกลางที่พิจารณาจากจำนวนดาวของบทวิจารณ์ที่ผู้วิจารณ์ให้ ($p_i \in \{1, 2, 3, 4, 5\}$) q_{ij} เป็นคะแนนคุณลักษณะของคุณภาพ j ของบทวิจารณ์ i (effectiveness, productivity, safety และ satisfaction) โดยที่ q_{ij} มีค่า -1, 0 และ 1

เมื่อคิดคำนวณบทวิจารณ์จำนวน 3,000 บทวิจารณ์ด้วยวิธีการประมาณค่าพารามิเตอร์ของสมการความถดถอยเพื่อหาค่าสัมประสิทธิ์ของสมการเมื่อค่าเฉลี่ยของความผิดพลาดน้อยที่สุด ผลลัพธ์ที่ได้แสดงในตารางที่ 4.15

ตารางที่ 4.15 สัมประสิทธิ์ของสมการและค่าเฉลี่ยของความผิดพลาดของสมการ 4.12

Base (β_0)	Effectiveness (β_1)	Productivity (β_2)	Safety (β_3)	Satisfaction (β_4)	ค่าเฉลี่ยของ $ e_i $
-0.004	0.1563	0.1391	0.2536	0.6395	0.3010

จะสังเกตเห็นว่าคุณภาพด้านความพึงพอใจ จะมีผลต่อจำนวนดาวที่ให้โดยผู้วิจารณ์มากที่สุด รองลงมาคือ ความปลอดภัย ประสิทธิภาพ และผลิตผลตามลำดับ ซึ่งสอดคล้องกับตารางข้างต้น

4.6.4.2 ความสัมพันธ์ระหว่างคุณภาพและความคิดเห็นในบทวิจารณ์ตามซอฟต์แวร์แต่ละประเภท

เมื่อคิดความสัมพันธ์ระหว่างคะแนนคุณภาพการใช้งานและจำนวนดาวที่ให้โดยผู้วิจารณ์ในบทวิจารณ์ตามวิธีการ 4.6.4.1 ในซอฟต์แวร์แต่ละประเภท สัมประสิทธิ์ของสมการความสัมพันธ์ของคุณลักษณะของคุณภาพแต่ละด้านแสดงในตารางที่ 4.16

ตารางที่ 4.16 ความสัมพันธ์ระหว่างคุณภาพด้านแต่ละด้านกับจำนวนดาวในซอฟต์แวร์แต่ละประเภท

Software	Base	Effective-ness	Productivity	Safety	Satisfac-tion	ค่าเฉลี่ยความผิดพลาด
Antivirus	2.969	0.254	0.233	0.516	1.145	0.588
Video Player	2.861	0.334	0.246	0.653	1.224	0.746
Maintenance & Optimization	2.798	0.365	0.509	0.330	1.188	0.670

ความพึงพอใจเป็นคุณลักษณะของคุณภาพที่มีผลต่อจำนวนดาวมากที่สุดในทุกสามประเภทซอฟต์แวร์ แต่คุณภาพด้านอื่นๆจะมีความสำคัญแตกต่างกันไป โดยโปรแกรมป้องกันไวรัส (Antivirus) และโปรแกรมเล่นวิดีโอ (Video Player) จะให้ความสำคัญกับความปลอดภัยรองลงมา เนื่องจากการ

เล่นไฟล์ (file) ทุกประเภทต้องไม่ทำให้โปรแกรมล้มละลาย (crash) ส่วนโปรแกรมการบำรุงรักษาและการหาค่าเหมาะสม (Maintenance & Optimization) จะให้ความสำคัญกับผลิตรองจากความพึงพอใจ

4.7 สรุปผลทั้งหมดของทุกการทดลอง

จากการทดลองทั้งหมด 4 ส่วนการทดลองสามารถสร้างตารางสรุปได้ดังนี้

ตารางที่ 4.17 สรุปการทดลองในหัวข้อ 4.3

จุดประสงค์	วัดประสิทธิผลการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้วยออนโทโลยี
ข้อมูลที่ใช้ทดสอบ	ประโยคทุกประโยคในทุกบทวิจารณ์ รวมทั้งหมดจำนวน 18,679 ประโยค
วิธีที่นำมาเปรียบเทียบในการทดลอง	<ol style="list-style-type: none"> 1) การวิเคราะห์ประโยคด้วยออนโทโลยีระดับ 0 คือ การวิเคราะห์ด้วยออนโทโลยีที่ถูกกำหนดขึ้นแล้วไม่ได้ถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 (O_0) 2) การวิเคราะห์ประโยคด้วยออนโทโลยีระดับ 1 คือ การวิเคราะห์ด้วยออนโทโลยีที่ถูกกำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนหนึ่งรอบ (O_1) 3) การวิเคราะห์ประโยคด้วยออนโทโลยีระดับ 2 คือการวิเคราะห์ด้วยออนโทโลยีที่ถูกกำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนสองรอบ (O_2) 4) การวิเคราะห์ประโยคด้วยออนโทโลยีระดับ 3 คือการวิเคราะห์ด้วยออนโทโลยีที่ถูกกำหนดขึ้นแล้วถูกขยายด้วยอัลกอริธึม 1 ใน 3.3 จำนวนสามรอบ (O_3)
สรุปผลการทดลอง	การหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพโดยใช้ออนโทโลยีที่ไม่มีการขยาย (O_0) มีความถูกต้องและมีประสิทธิผลมากที่สุด

ตารางที่ 4.18 สรุปการทดลองในหัวข้อ 4.4

จุดประสงค์	วัดประสิทธิผลการจำแนกประโยชน์ว่าเป็นประโยชน์ที่มีความคิดเห็นเชิงบวกหรือความคิดเห็นเชิงลบ
ข้อมูลที่ใช้ทดสอบ	ข้อมูลทดสอบสองชุด 1) ประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพทั้งสิ้นในด้านในข้อดีข้อเสียของทุกบทวิจารณ์ รวมทั้งหมดจำนวน 5,483 ประโยค (ทดสอบครั้งละหนึ่งในสามของจำนวนทั้งหมด) 2) ประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพทั้งสิ้นด้านในทุกบทวิจารณ์ รวมทั้งหมดจำนวน 11,217 ประโยค (ทดสอบสามครั้ง แต่แต่ละครั้งใช้ทั้งหมดประโยค 11,217 ประโยค)
วิธีที่นำมาเปรียบเทียบในการทดลอง	1) วิธีจำแนกประโยคแบบพื้นฐาน (BL) รายละเอียดดูกอธิบายในหัวข้อ 3.4.1 2) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคทุกประโยคในข้อดีข้อเสีย (ALL) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.1 3) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกอัตโนมัติ (ASF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.2 4) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.3
สรุปผลการทดลอง	1) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานมีประสิทธิภาพดีกว่าวิธีการจำแนกประโยคแบบพื้นฐาน 2) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF) มีประสิทธิภาพสูงสุด แต่ถ้าจะสร้างระบบอัตโนมัติ ต้องใช้วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกอัตโนมัติ (ASF)

ตารางที่ 4.19 สรุปการทดลองในหัวข้อ 4.5

จุดประสงค์	วัดประสิทธิผลการคำนวณคะแนนคุณภาพการใช้งาน
ข้อมูลที่ใช้ทดสอบ	บทวิจารณ์ทั้งหมด 3,000 บทวิจารณ์ โดยประโยคในบทวิจารณ์ทุกประโยคถูกหาความเกี่ยวข้องกับคุณลักษณะของคุณภาพการใช้งานโดยผู้เชี่ยวชาญ
วิธีที่นำมาเปรียบเทียบในการทดลอง	<p>1) วิธีจำแนกประโยคแบบพื้นฐาน (BL) รายละเอียดถูกอธิบายในหัวข้อ 3.4.1</p> <p>2) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกอัตโนมัติ (ASF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.2</p> <p>3) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF) รายละเอียดของวิธีการถูกอธิบายในหัวข้อ 3.4.2 และรายละเอียดของวิธีการสร้างกฎถูกอธิบายในหัวข้อ 3.4.2.3</p>
สรุปผลการทดลอง	วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF) มีประสิทธิผลด้านความถูกต้องสูงกว่าวิธีจำแนกประโยคแบบพื้นฐาน (BL) ดังนั้นการคำนวณคะแนนคุณภาพการใช้งานจึงเลือกใช้วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่เหลือจากการคัดออกด้วยมือ (MSF)

ตารางที่ 4.20 สรุปการทดลองในหัวข้อ 4.6

จุดประสงค์	วัดประสิทธิผลการวิเคราะห์บทรวิจารย์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบ
ข้อมูลที่ใช้ทดสอบ	บทรวิจารย์จำนวน 1,140 บทรวิจารย์ โดยประโยคในบทรวิจารย์ทุกประโยคถูกหาความเกี่ยวข้องกับคุณลักษณะของคุณภาพการใช้งานโดยผู้เชี่ยวชาญ และหาแต่ละประโยคว่ามีความคิดเห็นเชิงใดโดยวิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทรวิจารย์ผู้ใช้งาน โดยกฎถูกสร้างจากประโยคที่หลีกเลี่ยงการคัดออกด้วยมือ (MSF)
วิธีที่นำมาเปรียบเทียบในการทดลอง	<ol style="list-style-type: none"> 1) การคาดการณ์ความคิดเห็นของบทรวิจารย์ที่อ้างอิงจากจำนวนดาวโดยพิจารณาจากความคิดเห็นของแต่ละประโยค (P_0) 2) การคาดการณ์ความคิดเห็นของบทรวิจารย์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งานของบทรวิจารย์ (P_1) 3) การคาดการณ์ความคิดเห็นของบทรวิจารย์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งานของบทรวิจารย์และจำนวนประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ (P_2)
สรุปผลการทดลอง	<ol style="list-style-type: none"> 1) การคาดการณ์ความคิดเห็นของบทรวิจารย์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งานของบทรวิจารย์และจำนวนประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพ (P_2) มีผลความถูกต้องมากที่สุด และการคาดการณ์ความคิดเห็นของบทรวิจารย์ที่อ้างอิงจากจำนวนดาวโดยใช้คะแนนคุณภาพการใช้งานและการให้ความสำคัญกับคุณภาพด้านความพึงพอใจให้ผลความถูกต้องรองลงมา (จำนวนประโยคที่เกี่ยวข้องกับคุณลักษณะของคุณภาพด้านความพึงพอใจมากกว่าคุณภาพด้านอีกสามด้าน) 2) คุณภาพด้านความพึงพอใจ (satisfaction) มีผลต่อจำนวนดาวที่ให้ในบทรวิจารย์มากที่สุด

บทที่ 5

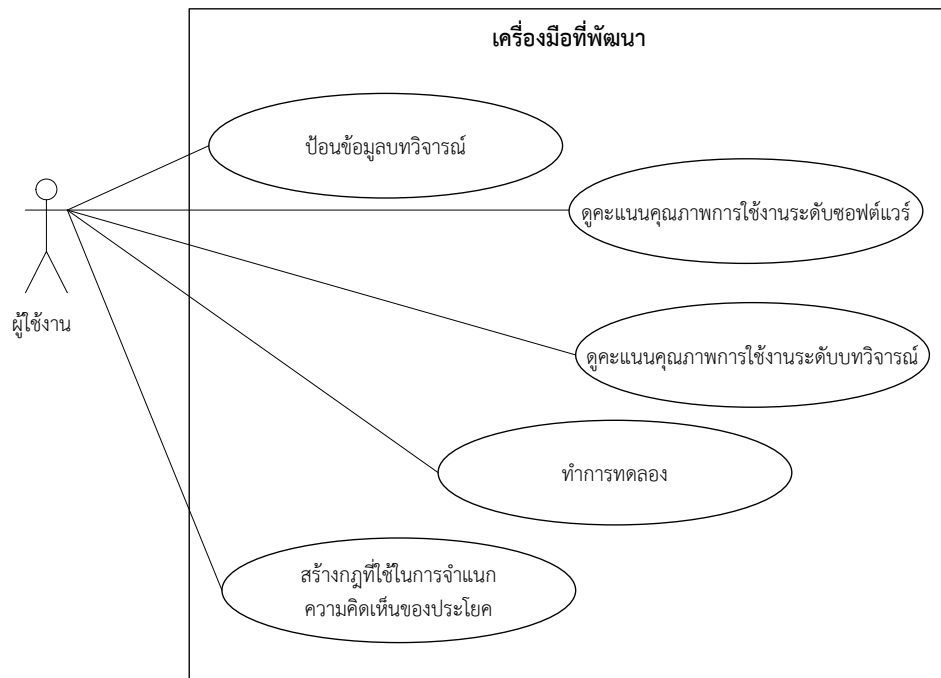
การพัฒนาเครื่องมือ

ในบทนี้อธิบายวิธีการพัฒนาเครื่องมือซึ่งเริ่มอธิบายจากความต้องการเชิงหน้าที่ การออกแบบการทำงานและส่วนต่อประสานผู้ใช้งานตามลำดับ

5.1 ความต้องการเชิงหน้าที่

เครื่องมือถูกพัฒนาขึ้นตามความต้องการเชิงหน้าที่ ดังรูปที่ 5.1 โดยมีรายละเอียดคือ

- 1) ป้อนข้อมูลบทวิจารณ์ ระบบสามารถรองรับการเก็บข้อมูลบทวิจารณ์จากเว็บไซต์ได้โดยให้ผู้ใช้งานกรอกข้อมูลของบทวิจารณ์เอง หรือกรอกที่อยู่เว็บไซต์ของบทวิจารณ์แล้วเครื่องมือสามารถไปดึงข้อมูลของบทวิจารณ์ในหน้าเว็บไซต์นั้น
- 2) ดูคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์ ระบบสามารถคำนวณคะแนนคุณภาพการใช้งานของซอฟต์แวร์และสามารถนำมาแสดงต่อผู้ใช้งานเมื่อผู้ใช้งานเรียกดูคะแนน
- 3) ดูคะแนนคุณภาพคุณภาพการใช้งานในระดับบทวิจารณ์ ระบบสามารถคำนวณคะแนนคุณภาพการใช้งานของแต่ละบทวิจารณ์และสามารถนำมาแสดงต่อผู้ใช้งานเมื่อผู้ใช้งานเรียกดูคะแนน
- 4) ทำการทดลอง ระบบสามารถรองรับการเตรียมข้อมูลเบื้องต้นสำหรับการทดลองที่ทำโดยผู้เชี่ยวชาญ และรองรับการทดลองที่ถูกอธิบายในบทที่ 4
- 5) สร้างกฎที่ใช้ในการจำแนกความคิดเห็นของประโยค เนื่องจากการสร้างที่ใช้ในการจำแนกต้องใช้เวลานาน เครื่องมือจึงถูกพัฒนาให้ผู้ใช้งานเลือกสร้างกฎได้ก่อน



รูปที่ 5.1 แผนภาพยูสเคสของเครื่องมือ

5.2 การออกแบบเครื่องมือ

5.2.1 การออกแบบสถาปัตยกรรมของเครื่องมือ

เครื่องมือถูกออกแบบให้เป็นทั้งวินโดว์แอปพลิเคชันและเว็บแอปพลิเคชัน เว็บแอปพลิเคชันใช้สำหรับการทดลองถูกพัฒนาขึ้นเพื่อให้ผู้เชี่ยวชาญอ่านและจัดทำข้อมูลซึ่งไม่มีความซับซ้อน ส่วนวินโดว์แอปพลิเคชันจะอธิบายรายละเอียดในหัวข้อ 5.2.5 ในการออกแบบเครื่องมือถูกแบ่งออกเป็นห้าชั้น ได้แก่ ชั้นส่วนนำเสนอ ชั้นส่วนควบคุม ชั้นส่วนแบบจำลองข้อมูล ชั้นส่วนดึงข้อมูลและ ชั้นส่วนหน่วยจัดเก็บข้อมูล สถาปัตยกรรมของเครื่องมือถูกแสดงในรูปที่ 5.2



รูปที่ 5.2 สถาปัตยกรรมของเครื่องมือ

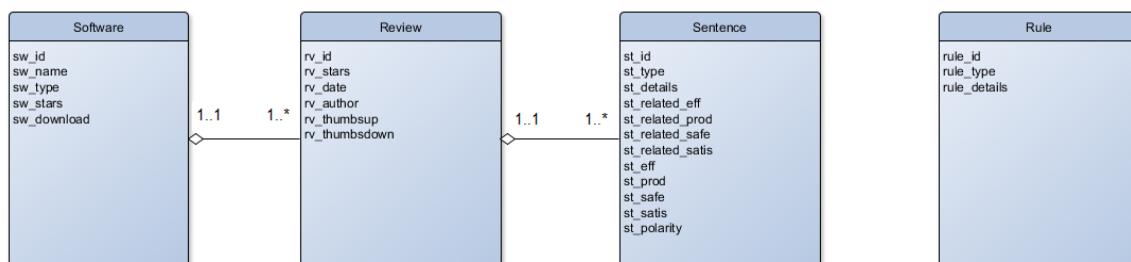
5.2.2 ความต้องการของระบบ

โปรแกรมที่ต้องติดตั้งก่อนใช้เครื่องมือมีดังนี้

- 1) Java Virtual Machine (JVM) เนื่องจากเครื่องมือเครื่องมือที่ถูกออกแบบและพัฒนาในภาษาจาวา (Java) และใช้โปรแกรมจัดการฐานข้อมูลชื่อ
- 2) MySQL เป็นระบบจัดการฐานข้อมูลเชิงสัมพันธ์ (relational database management system) ที่ใช้ในการพัฒนาเครื่องมือนี้
- 3) PHP server เป็นเว็บเซิร์ฟเวอร์ที่เครื่องมือนี้ใช้เพื่อทำเว็บไซต์ให้ผู้เชี่ยวชาญทำการทดลอง

5.2.3 แบบจำลองข้อมูล

การออกแบบแบบจำลองข้อมูลของเครื่องมือที่สร้างขึ้นเป็นการออกแบบโครงสร้างของข้อมูลและความสัมพันธ์ระหว่างข้อมูลภายในระบบ โดยแบบจำลองข้อมูลทั้งหมดสามารถออกแบบเป็นแผนภาพแสดงความสัมพันธ์ระหว่างเอนทิตี (entities relationship diagram: ER diagram) ดังรูปที่ 5.3 และตารางสคีมา (Schema) มีรายละเอียดแสดงในตารางที่ 5.1, 5.2, 5.3 และ 5.4



รูปที่ 5.3 แผนภาพแสดงความสัมพันธ์ระหว่างเอนทิตี

ตารางที่ 5.1 โครงสร้างตารางข้อมูล Software

ตาราง Software : ตารางจัดเก็บข้อมูลเกี่ยวกับซอฟต์แวร์			
คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
sw_id	รหัสซอฟต์แวร์	int	PK
sw_name	ชื่อของซอฟต์แวร์	text	
sw_type	ประเภทของซอฟต์แวร์	text	
sw_stars	จำนวนดาวเฉลี่ยที่ผู้วิจารณ์ให้กับซอฟต์แวร์	text	
sw_download	จำนวนครั้งที่ซอฟต์แวร์ถูกดาวน์โหลด	text	

ตารางที่ 5.2 โครงสร้างตารางข้อมูล Review

ตาราง Review : ตารางจัดเก็บข้อมูลเกี่ยวกับซอฟต์แวร์			
คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
rv_id	รหัสของบทวิจารณ์	int	PK
sw_id	รหัสของซอฟต์แวร์	int	FK
rv_stars	จำนวนดาวที่ผู้วิจารณ์ให้กับซอฟต์แวร์	text	
rv_date	วันที่ที่บทวิจารณ์ถูกเขียน	text	
rv_author	ชื่อของผู้วิจารณ์	text	
rv_thumbsup	จำนวนคนที่เห็นด้วยกับบทวิจารณ์นี้	int	
rv_thumbsdown	จำนวนคนที่ไม่เห็นด้วยกับบทวิจารณ์นี้	int	

ตารางที่ 5.3 โครงสร้างตารางข้อมูล Sentence

ตาราง Review : ตารางจัดเก็บข้อมูลเกี่ยวกับซอฟต์แวร์			
คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
st_id	รหัสของประโยค	int	PK
rv_id	รหัสของบทวิจารณ์	int	FK
st_type	ประเภทของโยค (สรุปล้น, ข้อดี, ข้อเสียและสรุป)	text	
st_details	รายละเอียดของประโยค	text	
st_related_eff	เกี่ยวข้องกับคุณภาพด้านประสิทธิภาพหรือไม่	boolean	
st_related_prod	เกี่ยวข้องกับคุณภาพด้านผลิตภัณฑ์หรือไม่	boolean	
st_related_safe	เกี่ยวข้องกับคุณภาพด้านความปลอดภัยหรือไม่	boolean	
st_related_satis	เกี่ยวข้องกับคุณภาพด้านความพึงพอใจหรือไม่	boolean	
st_eff	ความคิดเห็นในคุณภาพด้านประสิทธิภาพ	int	
st_prod	ความคิดเห็นในคุณภาพด้านผลิตภัณฑ์	int	
st_safe	ความคิดเห็นในคุณภาพด้านความปลอดภัย	int	
st_satis	ความคิดเห็นในคุณภาพด้านความพึงพอใจ	int	
st_polarity	ความคิดเห็นโดยรวมของประโยค	int	

ตารางที่ 5.4 โครงสร้างตารางข้อมูล Rule

ตาราง Review : ตารางจัดเก็บข้อมูลเกี่ยวกับซอฟต์แวร์			
คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
rule_id	รหัสของกฎ	int	PK
rule_type	ประเภทของกฎ	text	
rule_details	รายละเอียดของกฎ	text	

5.2.4 เครื่องมือสนับสนุนที่ใช้

เครื่องมือหรือคำสั่งที่เครื่องมือนี้ใช้มีดังนี้

- 1) คำสั่งภาษาจาวาชื่อ mysql-connector-java เป็นคำสั่งสำหรับเชื่อมต่อให้โปรแกรมที่เขียนบนภาษาจาวาสามารถติดต่อกับฐานข้อมูล mysql ได้ ในเครื่องมือนี้คำสั่งนี้จึงมีความจำเป็นที่ต้องถูกใช้งาน คำสั่งนี้หาได้จากเว็บไซต์ <http://www.mysql.com/products/connector/> [32]

2) คลังคำสั้งภาษาจาวาชื่อ jsoup เป็นคลังคำสั้งสำหรับจัดการกับไฟล์รูปแบบ HTML ในเครื่องมือนี้คลังคำสั้งนี้ถูกใช้สำหรับการสกัดข้อมูลจากหน้าเว็บไซต์ที่ถูกดาวโหลดเก็บไว้ในรูปแบบไฟล์ HTML คลังคำสั้งนี้หาได้จากเว็บไซต์ <http://jsoup.org/> [33]

3) ซอฟต์แวร์ลูซีน (Lucene) เป็นซอฟต์แวร์ที่ถูกพัฒนาขึ้นเพื่อใช้ในการค้นคืนสารสนเทศ ในเครื่องมือนี้ คำสั้งบางคำสั้งถูกเลือกใช้เพื่อคัดกรองและตัดคำประเภทสต่อปลิส (stop list) ออกซอฟต์แวร์ลูซีนหาได้จากเว็บไซต์ <http://lucene.apache.org/core/> [18]

4) เวิร์ดเน็ต คือฐานข้อมูลคำศัพท์ภาษาอังกฤษขนาดใหญ่ เครื่องมือนี้ใช้ฐานข้อมูลเวิร์ดเน็ตในการหาคำพ้องความหมาย และคำที่มีความหมายตรงกันข้าม เวิร์ดเน็ตนี้หาได้จากเว็บไซต์ <http://wordnet.princeton.edu/> [34]

5) เอพีไอชื่อ jaws-bin เป็นเอพีไอสำหรับโปรแกรมภาษาจาวาเรียกใช้ในการค้นหาคำศัพท์ในเวิร์ดเน็ต เอพีไอนี้หาได้จากเว็บไซต์ <http://lyle.smu.edu/~tspell/jaws/index.html> [35]

6) สแตนด์ฟอร์ดล็อกลิเนียร์พาร์ทออฟสปีชแท็กเกอร์ (Stanford log-linear part-of-speech tagger) เป็นโปรแกรมสำหรับหาหน้าที่ของแต่ละคำในประโยค โปรแกรมนี้ถูกพัฒนาด้วยภาษาจาวาดังนั้นเครื่องมือนี้สามารถนำเรียกใช้คำสั้งจากซอฟต์แวร์นี้ได้โดยตรง ซอฟต์แวร์นี้สามารถหาได้จาก <http://nlp.stanford.edu/software/tagger.shtml> [19]

7) คลังคำสั้งภาษาจาวาชื่อ javacsv เป็นคลังคำสั้งสำหรับการจัดการไฟล์รูปแบบ .csv ในเครื่องมือนี้สร้างไฟล์รูปแบบ .csv ในการแสดงผลในการทดลอง คลังคำสั้งนี้สามารถหาได้จาก <http://sourceforge.net/projects/javacsv/> [36]

8) คลังคำสั้งภาษาจาวาทางการคำนวณเชิงวิทยาศาสตร์ เครื่องมือนี้ได้ใช้คลังคำสั้งของ Michael Thomas Flanagan มาใช้ในการทดลองหาสมการการถดถอย (regression analysis) คลังคำสั้งนี้หาได้จาก www.ee.ucl.ac.uk/~mflanaga [37]

5.2.5 รายละเอียดในการออกแบบ

แนวคิดในการออกแบบจะเริ่มอธิบายจากชั้นส่วนที่ล่างสุด ได้แก่ ข้อมูลเกี่ยวกับซอฟต์แวร์ ข้อมูลเกี่ยวกับบทวิจารณ์ ข้อมูลเกี่ยวกับประโยคและกฎ ซึ่งเป็นข้อมูลที่ถูกเก็บในฐานข้อมูลเชิงสัมพันธ์

ขั้นส่วนต่อมาเป็นขั้นส่วนควบคุมทำหน้าที่เก็บข้อมูลและเรียกข้อมูลจากฐานข้อมูล โดยมีการเรียกใช้คำสั่งคำสั่งจาวาชื่อ `mysql-connector-java` สำหรับเชื่อมต่อระหว่างโปรแกรมจาวาและฐานข้อมูล `mysql` ข้อมูลซอฟต์แวร์ บทความ ประโยคและกฎจะถูกเก็บลงไปในฐานข้อมูล และเมื่อถูกเรียกขึ้นมาจะอยู่ในรูปของตารางที่ยังไม่พร้อมใช้งาน

ในส่วนถัดมาคือขั้นส่วนแบบจำลองข้อมูลทำหน้าที่ทำข้อมูลที่เรียกจากฐานข้อมูลให้อยู่ในรูปแบบพร้อมใช้งาน ข้อมูลซอฟต์แวร์ บทความ ประโยคและกฎจะถูกจัดเก็บอยู่รูปแบบคลาส (class) สำหรับนำไปใช้งานต่อไป

ขั้นส่วนดึงข้อมูลทำหน้าที่ดาวน์โหลดหน้าเว็บไซต์มาเก็บไว้แล้วทำการสกัดข้อมูลจากหน้าเว็บไซต์นั้น โดยใช้คำสั่งคำสั่งภาษาจาวาชื่อ `jsoup` ข้อมูลที่สกัดได้จะถูกส่งไปขั้นส่วนแบบจำลองข้อมูลเพื่อจัดการให้อยู่ในรูปแบบคลาส

ขั้นส่วนควบคุมเป็นส่วนที่มีความซับซ้อน มีหน้าที่ประมวลผลความคิดเห็นและคะแนนตามแนวคิดที่ได้นำเสนอ อีกทั้งควบคุมขั้นตอนการใช้เครื่องมือ โดยในส่วนนี้มีการใช้เครื่องมือสนับสนุนได้แก่ การใช้ซอฟต์แวร์ลูชัน เวิร์ดเน็ต เอพีไอชื่อ `jaws-bin` สแตนด์ฟอร์ตลือกิลีเนียร์พาร์ทออฟสปีชแท็คเกอร์ คำสั่งคำสั่งภาษาจาวาชื่อ `javacsv` คำสั่งคำสั่งภาษาจาวาทางการคำนวณเชิงวิทยาศาสตร์

ขั้นส่วนนำเสนอเป็นส่วนที่จัดหน้าแสดงผลให้กับผู้ใช้งาน โดยมีความครอบคลุมความต้องการเชิงหน้าที่ที่ได้กล่าวไว้ในหัวข้อ 5.1 โดยรายละเอียดและรูปของเครื่องมือจะอธิบายในหัวข้อต่อไป

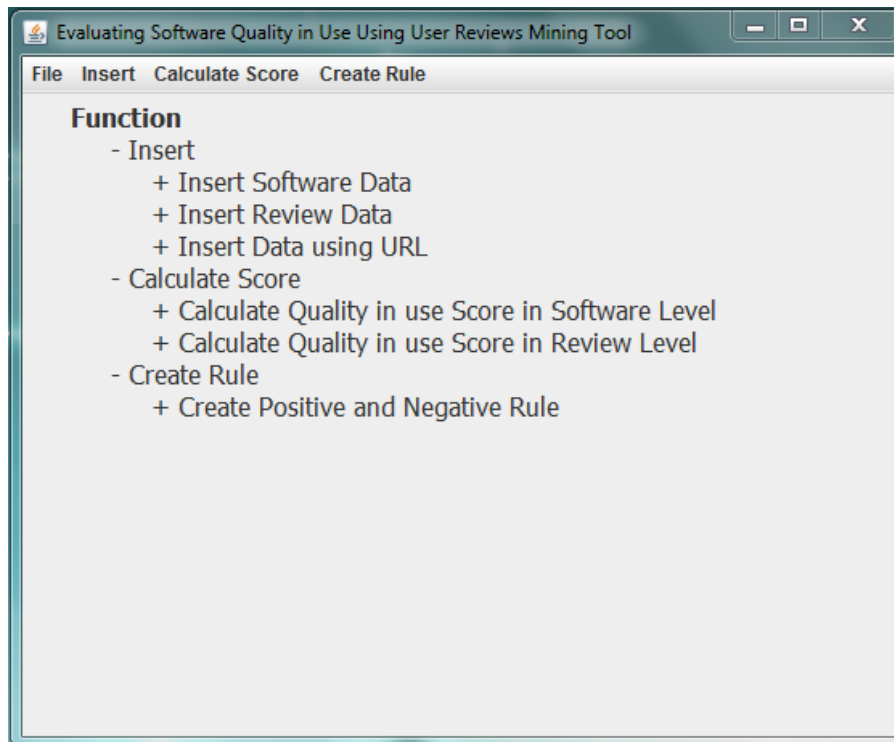
5.3 การทำงานและส่วนต่อประสานผู้ใช้งาน

เครื่องมือที่พัฒนาขึ้นมีหน้าหลักดังรูปที่ 5.4 ในหัวข้อนี้จะอธิบายรายละเอียดการทำงานตามความต้องการเชิงหน้าที่ดังต่อไปนี้

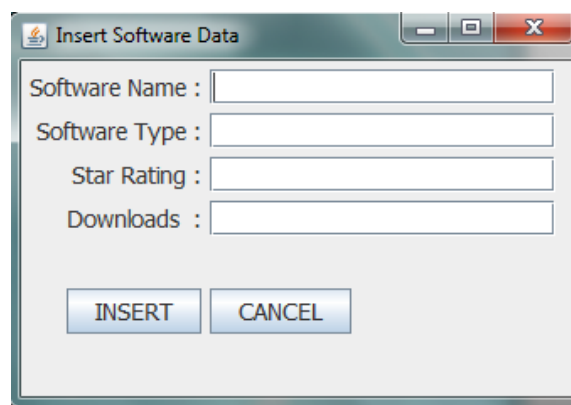
5.3.1 ป้อนข้อมูลบทความ

การป้อนข้อมูลบทความนั้นมีสองวิธีคือ กรอกข้อมูลเองและกรอกที่อยู่ของเว็บไซต์ กรอกข้อมูลเองเป็นนั่นเป็นการกรอกข้อมูลซอฟต์แวร์และกรอกข้อมูลบทความ ซึ่งมีรายละเอียดของหน้าจอของเครื่องมือในรูปที่ 5.5 และ 5.6 ตามลำดับ วิธีอีกวิธีหนึ่งคือกรอกที่อยู่ของเว็บไซต์และจำนวนบท

วิจารณ์ที่ต้องการ โดยเครื่องมือจะทำการดึงข้อมูลมาจัดเก็บให้อัตโนมัติ ซึ่งมีรายละเอียดของหน้าจอของเครื่องมือในรูปที่ 5.7



รูปที่ 5.4 หน้าจอหลักของเครื่องมือ



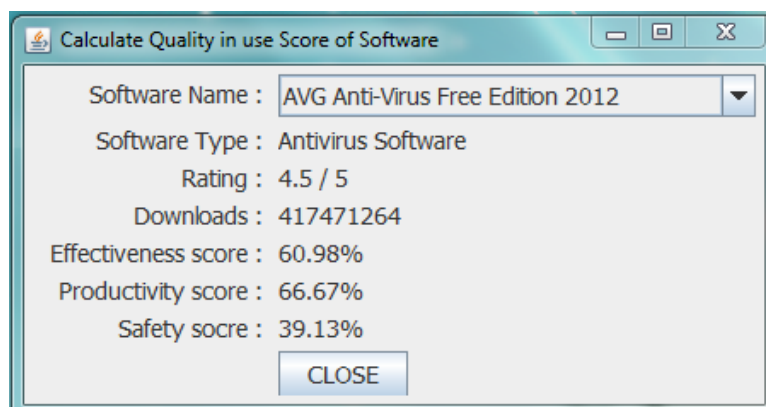
รูปที่ 5.5 หน้าจอส่วนการกรอกข้อมูลซอฟต์แวร์

รูปที่ 5.6 หน้าจอส่วนการกรอกข้อมูลบทวิจารณ์

รูปที่ 5.7 หน้าจอส่วนการกรอกข้อมูลโดยใช้ที่อยู่ของเว็บไซต์

5.3.2 ตูคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์และบทวิจารณ์

เครื่องมือคำนวณคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์จากบทวิจารณ์ทั้งหมดของซอฟต์แวร์นั้นโดยใช้แนวคิดที่นำเสนอในบทที่ 3 คือการหาความเกี่ยวข้องกับคุณภาพโดยใช้ออนโทโลยีที่ไม่ได้ขยายและการหาความคิดเห็นของประโยคที่มีความคิดเห็นเชิงบวกโดยวิธีกระบวนการจำแนกความคิดเห็นของประโยคจากการใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานในหัวข้อที่ 3.4.2 หลังจากนั้นนำข้อมูลที่ได้อมาคิดคะแนนคุณภาพการใช้งานระดับประโยค ระดับบทวิจารณ์แล้วระดับซอฟต์แวร์ตามลำดับในหัวข้อที่ 3.5 แล้วจึงนำมาแสดงผลดังรูปที่ 5.8



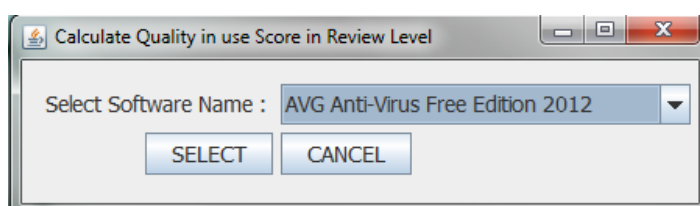
รูปที่ 5.8 หน้าจอส่วนการดูคะแนนคุณภาพการใช้งานระดับซอฟต์แวร์

5.3.3 ดูคะแนนคุณภาพการใช้งานระดับบทวิจารณ์

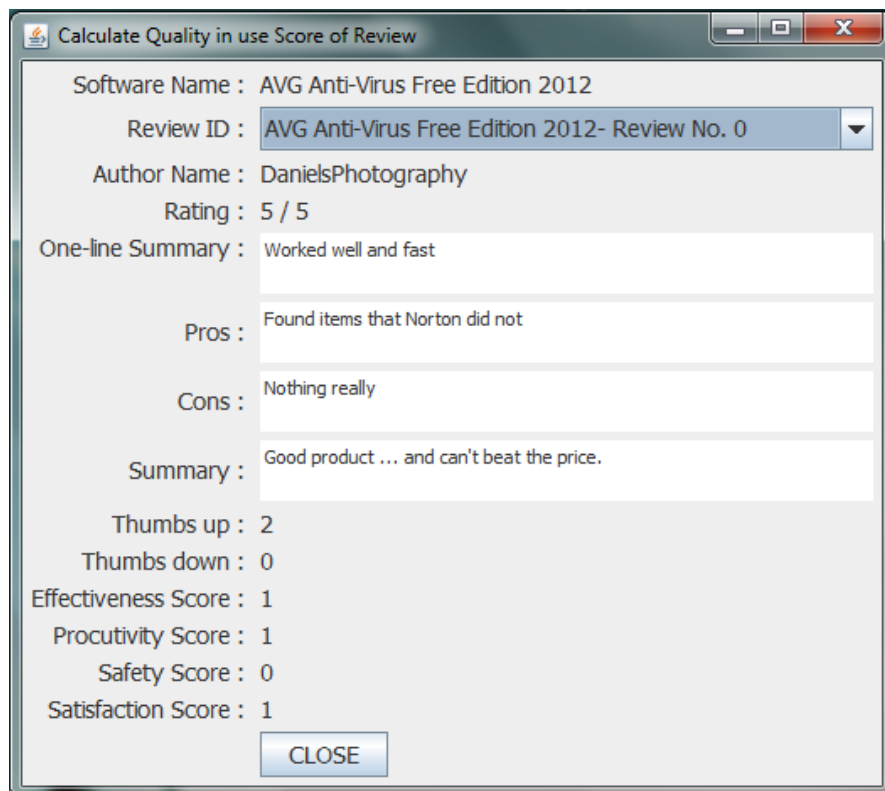
เครื่องมือจัดเตรียมให้สามารถดูคะแนนคุณภาพการใช้งานระดับบทวิจารณ์ โดยต้องเลือกซอฟต์แวร์ก่อนแล้วจึงเลือกบทวิจารณ์ที่ต้องการ แสดงดังรูปที่ 5.9 และ 5.10

5.3.4 สร้างกฎ

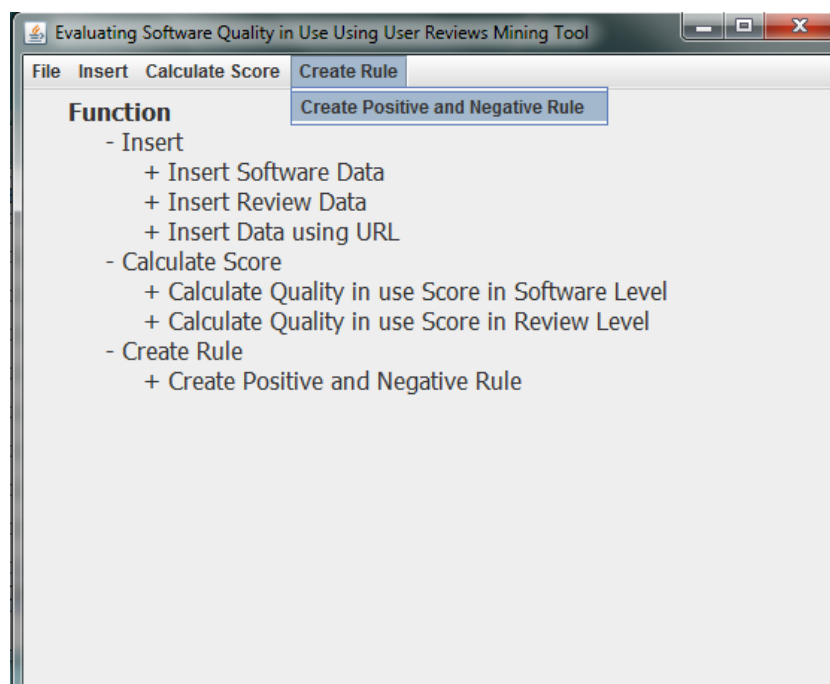
เนื่องจากการสร้างกฎจากข้อมูลบทวิจารณ์จำนวนมากต้องใช้เวลาาน เครื่องมือจึงมีให้เลือกสร้างกฎก่อนดังรูปที่ 5.11 โดยวิธีสร้างกฎนี้ใช้วิธีในหัวข้อ 3.4.2.2



รูปที่ 5.9 หน้าจอส่วนการเลือกซอฟต์แวร์เพื่อดูคะแนนคุณภาพการใช้งานของบทวิจารณ์



รูปที่ 5.10 หน้าจอส่วนการดูคะแนนคุณภาพการใช้งานระดับบทวิจารณ์



รูปที่ 5.11 หน้าจอส่วนการสร้างกฎ

5.3.5 ทำการทดลอง

ในส่วนทำการทดลองมีสองส่วนการเตรียมข้อมูลโดยผู้เชี่ยวชาญและส่วนการทดลอง ส่วนการเตรียมข้อมูลโดยผู้เชี่ยวชาญ เครื่องมือได้พัฒนาขึ้นให้ผู้เชี่ยวชาญตัดสินใจว่าประโยคนั้นเกี่ยวข้องกับคุณลักษณะของคุณภาพด้านหรือไม่ แต่ละคุณลักษณะของคุณภาพด้านต่างๆมีตัวเลือกสำหรับแต่ละประโยค คือ ไม่เกี่ยวข้องกับคุณลักษณะของคุณภาพนั้น (X) เกี่ยวข้องกับคุณลักษณะของคุณภาพนั้น และมีความคิดเห็นไปในเชิงบวก (+) เป็นกลาง (0) หรือเชิงลบ (-) ส่วนนี้จะแสดงดังรูปที่ 5.12

สำหรับส่วนการทดลอง การทดลองมีรายละเอียดในบทที่ 4 ซึ่งเครื่องมือถูกพัฒนาให้ประมวลผลออกมาในรูปแบบ .txt สำหรับผลสรุปและ .csv สำหรับตารางที่แสดงรายละเอียดของผลการทดลอง

Review ID

Software ID	1	Review ID	1
Review SMS	10855809	Rating	5
Header	Worked well and fast		
Pros	Found items that Norton did not		
Cons	Nothing really		
Summary	Good product ... and can't beat the price.		
Review Date	2012-05-17	Review Author	DanielsPhotography
Thumbs up	2	Thumbs down	0

No.	Sentence details	Effectiveness	Productivity	Safety	Satisfaction
1.	Worked well and fast	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -
2.	Found items that Norton did not	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -
3.	Nothing really	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -
4.	Good product	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -
5.	and can't beat the price	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -	<input checked="" type="radio"/> X <input type="radio"/> + <input type="radio"/> 0 <input type="radio"/> -

รูปที่ 5.12 หน้าจอส่วนการเตรียมข้อมูลทดลองโดยผู้เชี่ยวชาญ

บทที่ 6

สรุปผลการวิจัย

6.1 สรุปผลการวิจัย

บทวิจารณ์เป็นข้อมูลที่มีประโยชน์และที่มีผลต่อการตัดสินใจเลือกซื้อหรือใช้สินค้า ในงานวิจัยนี้สังเกตเห็นถึงความสำคัญของข้อมูลในบทวิจารณ์ จึงมีแนวคิดที่สกัดข้อมูลจากบทวิจารณ์ โดยข้อมูลนี้เปรียบเสมือนข้อมูลที่ได้จากการประเมินสินค้าเหล่านั้นจากมุมมองของผู้เขียนบทวิจารณ์ ในแนวคิดประกอบด้วยสี่ขั้นตอนย่อย ขั้นตอนแรกคือการรวบรวมและจัดเตรียมข้อมูล ขั้นตอนนี้เป็นการเก็บรวบรวมข้อมูลบทวิจารณ์และจัดเตรียมข้อมูลให้พร้อมโดยแยกประโยคในบทวิจารณ์ ขั้นตอนที่สองคือการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้านต่างๆ ในขั้นตอนนี้ในงานวิจัยนี้ได้เสนอการใช้ออนโทโลยีเพื่อช่วยในการหาความเกี่ยวข้องโดยการกำหนดคำศัพท์เริ่มต้นที่หามาจากแบบจำลองคุณภาพไอเอสโอ 9126 และขยายหรือเพิ่มเติมคำศัพท์ด้วยอรรถาภิธานโดยใช้เครื่องมือเวิร์ดเน็ต หลังจากหาประโยคเกี่ยวข้องกับคุณลักษณะด้านใดแล้ว ต่อมาขั้นตอนที่สามคือการหาว่าประโยคนั้นมีความคิดเห็นที่เป็นบวก เป็นลบหรือเป็นกลาง ในขั้นตอนนี้มีการประยุกต์ใช้คำศัพท์เกี่ยวกับความรู้สึกกับการสร้างกฎจากประโยคในบทวิจารณ์เพื่อให้เพิ่มความแม่นยำในการวิเคราะห์ความคิดเห็นของประโยค และขั้นตอนสุดท้ายคือการนำประโยคที่ผ่านการวิเคราะห์เหล่านั้นมาคำนวณเป็นคะแนนคุณภาพการใช้งานในระดับบทวิจารณ์และระดับซอฟต์แวร์ที่เข้าใจได้ง่าย

แนวคิดที่นำเสนอในงานวิจัยนี้ถูกทดสอบกับบทวิจารณ์จำนวน 3,000 บทวิจารณ์ จากเว็บไซต์ www.cnet.com ประโยคในบทวิจารณ์จะถูกแบ่งโดย “.” ทำให้ข้อมูลทั้งหมดประกอบด้วยประโยคทั้งหมด 18,679 ประโยค โดยประโยคแต่ละประโยคที่เก็บจะถูกระบุว่ามีมาจากส่วนสรุปสั้น (one-line summary) ข้อดี (pros) ข้อเสีย (cons) และสรุป (summary) ประโยคทุกประโยคจะถูกพิจารณาโดยผู้เชี่ยวชาญว่าเกี่ยวข้องของคุณลักษณะของคุณภาพด้านใดและความคิดเห็นของประโยคว่ามีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลาง มาตรฐานที่ใช้วัดประสิทธิผลในงานวิจัยนี้มีความแม่นยำ (precision) ค่าความครบถ้วน (recall) และค่าความถูกต้อง (accuracy)

ในการวัดประสิทธิผลของแนวคิดที่ได้นำเสนอในงานวิจัยนี้ได้แบ่งออกเป็นสี่ส่วน

- 1) การวัดประสิทธิผลการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้วยออนโทโลยีโดยเปรียบเทียบระหว่างออนโทโลยีที่ไม่ถูกขยายและที่ถูกขยายด้วยอรรถาภิธาน ผลที่ได้คือออนโทโลยีที่ถูกขยายจะมีคำศัพท์ที่ครอบคลุมมากขึ้นแต่ก็ได้คำศัพท์ที่ไม่เกี่ยวข้องมาด้วยเช่นกัน ทำให้

ประสิทธิผลและความถูกต้องของการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้วย
 ออนโทโลยีที่ถูกขยายนั้นมีค่าต่ำกว่าการหาความเกี่ยวข้องของประโยคกับคุณลักษณะของคุณภาพด้วย
 ออนโทโลยีที่ไม่ถูกขยาย

2) การวัดประสิทธิผลการจำแนกประโยคว่าเป็นประโยคที่มีความคิดเห็นเชิงบวกหรือความคิดเห็น
 เชิงลบโดยเปรียบเทียบระหว่างการใช้คำศัพท์เกี่ยวกับความรู้สึกเพียงอย่างเดียวกับการใช้กฎที่สร้างจาก
 บทวิจารณ์ประยุกต์เข้ากับคำศัพท์เกี่ยวกับความรู้สึก ผลที่ได้คือการใช้กฎประยุกต์เข้ากับคำศัพท์นั้นทำ
 ให้การจำแนกนั้นมีความเหมาะสมกับประโยคประเภทนั้นมากขึ้น ทำให้การจำแนกประโยคด้วยการ
 ประยุกต์กฎที่สร้างจากบทวิจารณ์กับคำศัพท์ที่เกี่ยวกับความรู้สึกมีประสิทธิผลและความถูกต้องที่สูง
 กว่า การจำแนกประโยคด้วยการใช้คำศัพท์ที่เกี่ยวกับความรู้สึกเพียงอย่างเดียว

3) การวัดประสิทธิผลการคำนวณคะแนนคุณภาพการใช้งาน โดยเปรียบเทียบระหว่างวิธีจำแนก
 ประโยคแบบวิธีการจำแนกประโยคด้วยการประยุกต์กฎที่สร้างจากบทวิจารณ์กับคำศัพท์ที่เกี่ยวกับ
 ความรู้สึก (วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานในหัวข้อ 3.4.2) กับวิธีการ
 จำแนกประโยคด้วยคำศัพท์ที่เกี่ยวกับความรู้สึกเพียงอย่างเดียว (วิธีการจำแนกประโยคแบบพื้นฐานใน
 หัวข้อ 3.4.1) วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานมีประสิทธิผลด้านความถูกต้อง
 สูงกว่าวิธีจำแนกประโยคแบบพื้นฐาน ดังนั้นการคำนวณคะแนนคุณภาพการใช้งานจึงเลือกใช้วิธีจำแนก
 ประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งาน

4) การวัดประสิทธิผลการวิเคราะห์บทวิจารณ์ว่ามีความคิดเห็นเชิงบวกหรือเชิงลบ โดย
 เปรียบเทียบการนำคะแนนคุณภาพการใช้งานของบทวิจารณ์กับความคิดเห็นของประโยคแต่ละ
 ประโยคเพียงอย่างเดียว การเปรียบเทียบนี้เพื่อทดสอบความสอดคล้องของคะแนนคุณภาพการใช้งาน
 ผลที่ได้คือถ้าให้น้ำหนักของคุณภาพด้านความพึงพอใจในการใช้คะแนนคุณภาพจะมีความสอดคล้อง
 หรือวิเคราะห์ได้ตรงกว่าการใช้ความคิดเห็นของประโยคเพียงอย่างเดียว

สรุปแนวคิดที่นำเสนอ การหาความเกี่ยวข้องของประโยคโดยใช้ออนโทโลยีที่ไม่ถูกขยายและ
 การจำแนกประโยคว่ามีความคิดเห็นเชิงบวก เชิงลบหรือเป็นกลางใช้การประยุกต์กฎที่สร้างจากบท
 วิจารณ์กับคำศัพท์ที่เกี่ยวกับความรู้สึก (วิธีจำแนกประโยคแบบใช้กฎที่สร้างจากบทวิจารณ์ผู้ใช้งานใน
 หัวข้อ 3.4.2) ให้ความถูกต้องสูงและถูกนำไปใช้ในการคำนวณคะแนนคุณภาพการใช้งาน อีกทั้งใน
 งานวิจัยนี้ยังได้พัฒนาเครื่องมือตามแนวคิดที่นำเสนอ

6.2 ข้อจำกัดของงานวิจัย

ในงานวิจัยนี้ มีข้อจำกัดในบางอย่างดังนี้

- 1) บทความที่นำเข้ามาเป็นข้อมูลสำหรับสร้างกฎต้องมีโครงสร้างที่ประกอบด้วยส่วนที่เป็นข้อดีและข้อเสีย เนื่องจากการสร้างกฎอาศัยข้อมูลจากส่วนเหล่านี้มาเป็นข้อมูลส่วน
- 2) การสร้างกฎต้องใช้เวลาเพิ่มขึ้นเมื่อมีข้อมูลจำนวนมากขึ้นทำให้การคำนวณคะแนนคุณภาพการใช้งานของซอฟต์แวร์ที่มีบทความจำนวนมากใช้เวลานาน

6.3 งานวิจัยในอนาคต

งานวิจัยนี้ยังสามารถพัฒนาได้อีกหลายจุด ได้แก่

- 1) แบบจำลองคุณภาพซอฟต์แวร์ที่ถูกนำมาใช้สามารถปรับเปลี่ยนให้มีความเหมาะสมและทันสมัยมากขึ้นได้ เช่น แบบจำลองคุณภาพซอฟต์แวร์ไอเอสโอ 25010 เป็นต้น
- 2) ในความสภาพแวดล้อมจริง ความคิดเห็นสามารถแบ่งได้หลายระดับ เช่น มีความคิดเห็นที่เป็นบวกมาก เป็นบวก เป็นกลาง เป็นลบ เป็นลบมาก เป็นต้น คะแนนคุณภาพการใช้งานที่คำนวณได้ของประโยชน์อาจจะมีค่าไม่ใช่แค่ -1, 0 และ 1 ซึ่งต่อไปจะถูกใช้ในการคำนวณคะแนนการใช้งานของบทความ การวิเคราะห์ความคิดเห็นของประโยคได้หลากหลายระดับสามารถนำมาประยุกต์กับงานวิจัยนี้ในอนาคตเพื่อความแม่นยำของการคำนวณคะแนนคุณภาพการใช้งานของบทความ
- 3) สำหรับการวิเคราะห์ความคิดเห็นของบทความ ตำแหน่งของประโยคส่งผลกระทบต่อความคิดเห็นของบทความไม่เท่ากัน เช่น ตำแหน่งประโยคช่วงต้นและช่วงท้ายของบทความส่วนมากจะมีผลต่อความคิดเห็นโดยรวมมาก เป็นต้น การวิเคราะห์ตำแหน่งประโยคให้มีความเหมาะสมกับน้ำหนักของแต่ละประโยคน่าจะสนับสนุนการคำนวณคะแนนคุณภาพการใช้งาน
- 4) ในการคำนวณคะแนนคุณภาพการใช้งานของซอฟต์แวร์จากบทความ บทความบางบทความถูกเขียนขึ้นเป็นสแปม (spam) หรือให้ข้อมูลไม่ตรงกับความจริงของซอฟต์แวร์นั้น บทความเหล่านี้จึงควรถูกคัดออกเพื่อทำให้การคำนวณคะแนนคุณภาพการใช้งานใกล้เคียงความจริงมากขึ้น วิธีการกำจัดบทความเหล่านั้น เช่น ถ้าบทความนั้นมีคนไม่เห็นด้วยเยอะ บทความนั้นก็อาจจะเป็นสแปมได้ เป็นต้น
- 5) ในการวิเคราะห์ความคิดเห็นของประโยคของแนวคิดที่นำเสนอในข้อความคำศัพท์ที่อยู่ในประโยค โดยคำศัพท์คำเดียวกันจะให้ความหมายเชิงความคิดเห็นในบริบทหรือสภาพแวดล้อมที่แตกต่างกัน สามารถนำจุดนี้มาพัฒนาวิธีประยุกต์การสร้างกฎจากบทความกับคำศัพท์ที่เกี่ยวข้องกับความรู้สึกได้

6.4 ผลงานตีพิมพ์จากวิทยานิพนธ์

1) Warit Leoparote, Athasit Surarerks, Nakornthip Prompoon, “Software Quality in Use Characteristic Mining from Customer Reviews” presented at Second International Conference on Digital Information and Communication Technology and it's Applications (DICTAP, 2012) on 16-18 May 2012, Bangkok, Thailand

2) Warit Leoparote, Athasit Surarerks, Nakornthip Prompoon, “Evaluating Software Quality in Use Using Reviews Mining” presented at The 10th International Joint Conference on Computer Science and Software Engineering (JCSSE, 2013) on 29-31 May 2013, Khonkhaen, Thailand

รายการอ้างอิง

- [1] P. Resnick and H. R. Varian. Recommender systems. Commun. ACM 40 (1997): 56-58.
- [2] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions 17 (2005): 734-749.
- [3] M. Hu and B. Liu. Mining Opinion Features in Customer Reviews The 19th National Conference on Artificial Intelligence (AAAI'04) 2004: 755-760.
- [4] V. Ramkumar, S. Rajasekar, and S. Swamynathan. Scoring products from reviews through application of fuzzy techniques. Expert Systems with Applications 37 (2010): 6862-6867.
- [5] I. Peñalver-Martínez, R. Valencia-García, and F. García-Sánchez. Ontology-Guided Approach to Feature-Based Opinion Mining. Natural Language Processing and Information Systems. vol. 6716, R. Muñoz, *et al.*, Eds., ed: Springer Berlin / Heidelberg, 2011: 193-200.
- [6] S. Aciar, D. Zhang, S. Simoff, and J. Debenham. Informed Recommender: Basing Recommendations on Consumer Product Reviews. Intelligent Systems, IEEE 22 (2007): 39-47.
- [7] L. Zhao and C. Li. Ontology Based Opinion Mining for Movie Reviews. Knowledge Science, Engineering and Management 5914 (2009): 204-214.
- [8] M. Hu and B. Liu. Mining and summarizing customer reviews. Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining 2004: 168-177.
- [9] P. D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics 2002: 417-424.
- [10] L. Zhou and P. Chaovalit. Ontology-supported polarity mining. Journal of the American Society for Information Science and Technology 59 (2008): 98-110.

- [11] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. Proceedings of the ACL-02 conference on Empirical methods in natural language processing 2002: 79-86.
- [12] ISO/IEC 9126-1: Software Engineering - Product Quality - Part 1: Quality model. Geneva: International Organization for Standardization, 2001.
- [13] K. Dave, S. Lawrence, and D. M. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. Proceedings of the 12th international conference on World Wide Web 2003: 519-528.
- [14] B. Pang and L. Lee. Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr. 2 (2008): 1-135.
- [15] cnet. Product reviews and prices, software downloads, and tech news - CNET [Online]. 2010. Available from: <http://www.cnet.com> [2010, May 18]
- [16] S. M. Weiss, N. Indurkha, T. Zhang, and F. J. Damerau. Text Mining: Predictive Methods for Analyzing Unstructured Information. Springer, 2004.
- [17] M. F. Porter. An algorithm for suffix stripping. Program: electronic library and information systems 14 (1980): 130-137.
- [18] Apache. Apache Lucene - Overview [Online]. 2010. Available from: <http://lucene.apache.org/java/docs/index.html> [2010, October 25]
- [19] StanfordNLPGroup. Stanford Log-linear Part-Of-Speech Tagger [Online]. 2010. Available from: <http://nlp.stanford.edu/software/tagger.shtml> [2010, October 25]
- [20] G. Salton and M. J. McGill. Introduction to Modern Information Retrieval. McGraw-Hill Publishing Company, 1983.
- [21] G. A. Miller. WordNet: a lexical database for English. Commun. ACM 38 (1995): 39-41.
- [22] A. Esuli and F. Sebastiani. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. The International Conference on Language Resources and Evaluation (LREC) 2006: 417-422.

- [23] S. Baccianella, A. Esuli, and F. Sebastiani. SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. The International Conference on Language Resources and Evaluation (LREC) 2008: 2200-2204.
- [24] P. J. Stone, D. C. Dunphy, and M. S. Smith. Text Analysis for the Social Sciences. Lawrence Erlbaum Associates, 1997.
- [25] R. Ghani, K. Probst, Y. Liu, M. Crema, and A. Fano. Text mining for product attribute extraction. SIGKDD Explor. NewsL 8 (2006): 41-48.
- [26] ISO 9001:2000, Quality Management Systems-Requirements. International Organization for Standardization, 2000.
- [27] D. Samadhiya, W. Su-Hua, and C. Dengjie. Quality models: Role and value in software engineering. Software Technology and Engineering (ICSTE), 2010 2nd International Conference 2010:320-324.
- [28] ISO/IEC 9126-4: Software Engineering - Product Quality - Part 4: Quality in use metrics. Switzerland: International Organization for Standardization, 2004.
- [29] M. A. Moraga, M. F. Bertoa, M. C. Morcillo, C. Calero, and A. Vallecillo. Evaluating Quality-in-Use Using Bayesian Networks. Proc. of OAOOSE 2008: 31-40.
- [30] R. Narayanan, B. Liu, and A. Choudhary. Sentiment analysis of conditional sentences. Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing 2009: 180-189.
- [31] X. Ding, B. Liu, and P. S. Yu. A holistic lexicon-based approach to opinion mining. Proceedings of the 2008 International Conference on Web Search and Data Mining 2008: 231-240.
- [32] Oracle. MySQL Connectors [Online]. 2010. Available from: <http://www.mysql.com/products/connector> [2010, May 18]
- [33] J. Hedley. jsoup: Java HTML Parser [Online]. 2009. Available from: <http://jsoup.org> [2010, May 18]

- [34] PrincetonUniversity. Downloading WordNet and associated packages and tools [Online]. 2010. Available from:
<http://wordnet.princeton.edu/wordnet/download/> [2010, May 18]
- [35] B. Spell. Java API for WordNet Searching (JAWS) [Online]. 2010. Available from:
<http://lyle.smu.edu/~tspell/jaws/index.html> [2010, May 18]
- [36] shriop. Java CSV Library [Online]. 2010. Available from:
<http://sourceforge.net/projects/javacsv> [2010, May 20]
- [37] M. T. Flanagan. Michael Thomas Flanagan's Java Scientific Library [Online]. 2010. Available from: <http://www.ee.ucl.ac.uk/~mflanaga/java/> [2010, May 22]

ภาคผนวก

ภาคผนวก ก

คำศัพท์ที่ใช้ในการสร้างและการขยายออนโทโลยี

ตารางที่ ก.1 คำศัพท์ของออนโทโลยีที่ไม่ได้ถูกขยาย

	คำศัพท์
ประสิทธิผล (effectiveness)	effectiveness, accuracy, completeness, correctly, completion, error, inaccurate, work
ผลิตผล (productivity)	resources, time, effort, material, financial, cost, fast, efficient, free, expensive, easy, spend
ความปลอดภัย (safety)	safety, harm, health, damage, virus, malware, protect, security, recovery, crash
ความพึงพอใจ (satisfaction)	satisfy, attitude, happy, good, better, like, love, install, feature

ตารางที่ ก.2 คำศัพท์ของออนไลน์ที่ถูกขยายหนึ่งรอบ

	คำศัพท์
ประสิทธิผล (effectiveness)	effectiveness, accuracy, completeness, correctly, completion, error, inaccurate, work, effectivity, effectualness, effectuality, ineffectiveness, potency, strength, truth, inaccuracy, incompleteness, right, aright, incorrectly, mistake, fault, erroneousness, misplay, wrongdoing, accurate, employment, study, workplace, oeuvre, idle, act, function, operate, go, run, process, exercise, make, bring, play, wreak, cultivate, crop, influence, shape, form, mold, mould, forge, knead, exploit, solve, lick, ferment, sour, turn
ผลิตผล (productivity)	resources, time, effort, material, financial, cost, fast, efficient, free, expensive, easy, spend, resource, resourcefulness, imagination, clip, meter, metre, sentence, clock, attempt, endeavor, endeavor, try, exertion, travail, sweat, feat, exploit, campaign, cause, crusade, drive, movement, stuff, fabric, cloth, textile, immaterial, corporeal, substantial, real, fiscal, nonfinancial, price, toll, fasting, slow, tight, debauched, degenerate, degraded, dissipated, dissolute, libertine, profligate, riotous, flying, quick, firm, immobile, loyal, truehearted, inefficient, effective, liberate, release, unloose, unloosen, loose, confinerid, disembarass, dislodge, exempt, relieve, discharge, disengage, obstruct, absolve, justify, relinquish, resign, unblock, unfreeze, unfree, bound, complimentary, costless, gratis, gratuitous, detached, spare, barren, destitute, devoid, innocent, liberal, cheap, difficult, uneasy, easily, slowly, tardily, soft, easygoing, leisurely, gentle, comfortable, prosperous, well-fixed, well-heeled, well-off, well-situated, well-to-do, light, promiscuous, sluttish, wanton, pass, expend, drop

ตารางที่ ก.2 คำศัพท์ของออนโทโลยีที่ถูกขยายหนึ่งรอบ (ต่อ)

	คำศัพท์
ความปลอดภัย (safety)	safety, harm, health, damage, virus, malware, protect, security, recovery, crashdanger, refuge, guard, condom, rubber, safe, prophylactic, injury, hurt, trauma, impairment, scathe, wellness, price, terms, wrong, insecurity, protection, certificate, surety, convalescence, recuperation, retrieval, clang, clangor, clangour, clanging, clank, clash, wreck, collapse, , mash, ram, gate-crash, dash, doss
ความพึงพอใจ (satisfaction)	satisfy, attitude, happy, good, better, like, love, install, feature, fulfill, fulfil, gratify, dissatisfy, meet, fill, position, posture, unhappy, felicitous, glad, well-chosen, goodness, evil, bad, commodity, well, thoroughly, soundly, full, estimable, honorable, respectable, beneficial, just, upright, adept, expert, practiced, proficient, skillful, skilful, dear, near, dependable, safe, secure, right, ripe, effective, serious, sound, salutary, honest, undecomposed, unspoiled, unspoilt, bettor, wagerer, punter, break, improve, amend, ameliorate, meliorate, worsen, worse, best, easily, considerably, substantially, intimately, advantageously, comfortably, ilk, wish, care, dislike, similar, unlike, same, alike, comparable, corresponding, hate, passion, beloved, dearest, honey, lovemaking, enjoy, know, screw, fuck, jazz, eff, hump, bed, bang, bonk, install, establish, characteristic, lineament, have, sport, boast

หมายเหตุ ออนโทโลยีที่ถูกขยายสองและสามรอบ มีคำศัพท์ปริมาณมากขึ้นมากและคำศัพท์เหล่านั้นมีความเกี่ยวข้องกับคุณลักษณะของคุณภาพของซอฟต์แวร์น้อย

ภาคผนวก ข

การใช้เครื่องมือเครื่องมือเทคมาเยเนอร์ซอฟต์แวร์คิด (Text-Miner Software Kit) และเครื่องมือรูลอินดักชันคิตฟอร์เท็กซ์ (Rule Induction Kit for Text)

เพื่อให้ได้ปฏิบัติตามแนวคิด จึงมีการใช้เครื่องมือเครื่องมือเทคมาเยเนอร์ซอฟต์แวร์คิด (Text-Miner Software Kit, TMSK) และเครื่องมือรูลอินดักชันคิตฟอร์เท็กซ์ (Rule Induction Kit for Text, RIKTEXT) เครื่องมือถูกพัฒนาให้ทำงานบน command line เครื่องมือในงานวิจัยนี้พัฒนาด้วยภาษาจาวา จึงมีการเรียกใช้ Class Runtime เพื่อเรียกใช้คำสั่งในเครื่องมือต่อ โดยขั้นตอนและรายละเอียดของคำสั่ง command line ของการใช้เครื่องมือถูกอธิบายได้ดังนี้

1) จัดไฟล์ที่นำเข้าจะประกอบด้วยเนื้อหาและหัวข้อที่เนื้อหานั้นเกี่ยวข้อง (หัวข้อสามารถมีได้มากกว่า 1) แต่ละเอกสารจะแสดงดังรูปด้านล่าง โดยไฟล์ที่นำเข้ามีหนึ่งไฟล์ที่รวมเอกสารทั้งหมด

```
<DOC>
  <BODY>
    เนื้อหา 1
  </BODY>
  <TOPICS>
    <TOPIC> หัวข้อ 1 </TOPIC>
    <TOPIC> หัวข้อ 2 </TOPIC>
  </TOPICS>
</DOC>
```

2) สร้างพจนานุกรม (dictionary) เพื่อใช้ในการแปลงเนื้อหาในเอกสารให้เป็นเวกเตอร์ (vector) โดยการสร้างพจนานุกรมจะตัดสต็อปเวิร์ด (stopwords) ออกด้วย ในขั้นตอนนี้ใช้เครื่องมือ TMSK ซึ่งคำสั่งที่ใช้คือ

```
java mkdict "จำนวนคำในพจนานุกรม" "ชื่อของไฟล์พจนานุกรม"
```

จำนวนคำในพจนานุกรมที่ใช้คือ 4,000 คำเนื่องจากถ้าใช้คำจำนวนเยอะจะใช้เวลาในการสร้างกฎนาน ชื่อของไฟล์พจนานุกรมต้องเหมือนกันกับในไฟล์ tmsk.properties

3) สร้างเวกเตอร์ (vector) เป็นตัวแทนของแต่ละเอกสาร จากพจนานุกรมที่สร้างขึ้นในขั้นตอนนี้แล้ว ในขั้นตอนนี้ใช้เครื่องมือ TMSK ซึ่งคำสั่งที่ใช้คือ

```
java vectorize "ชื่อหัวข้อ" "ชื่อไฟล์ของเวกเตอร์"
```

ผลลัพธ์ที่ได้จากขั้นตอนนี้เป็นไฟล์หนึ่งไฟล์ที่บอกว่าเอกสารแต่ละเอกสารประกอบด้วยคำศัพท์ในพจนานุกรมค่าใดบ้างและเอกสารที่มีชื่อหัวข้อตรงจะมีค่าในคอลัมน์แรกเป็น 1 และถ้ามีชื่อหัวข้อไม่ตรงจะมีค่าเป็น 0 แสดงดังรูป

```

0 101 401 501 2701 3101 4201 4802 5702 6501 7201 9201 11701 35501 50801 68701 132601
0 601 41601
1 2401 4001 7401 15701
0 3101 23301
0 109101 117701 152601
1 1901
0 11601 13601 24901 100501
0 801 1601 11002 53001 83301 99201 110501 140401
0 9301 15001
0 201 301 3101 11601 21601 36801 51701
1 101 4001 41801
1 1301 16201
1 24001 157301
0 9301 10101 26801 31801
0 101 1001 6701 10501 11601 22401 23101 28901 34101 43101 100601
0 24301 56401 69201 149401
1 601 2901 37801 53301
0 901 39901

```

รูปที่ ข.1 ตัวอย่างเวกเตอร์ที่สร้างจากเครื่องมือ TMSK

4) สร้างกฎจากเครื่องมือ RIKTEXT จากเวกเตอร์ที่สร้างจากขั้นตอนที่แล้ว เครื่องมือนี้สามารถสร้างกฎแล้วทดสอบวัดความถูกต้องได้ แต่เนื่องจากต้องการสร้างกฎอย่างเดียวและมีข้อจำกัดด้านเวลาในงานวิจัยนี้จึงใช้คำสั่ง

```
riktext -q "ชื่อของไฟล์พจนานุกรม" earn "ชื่อของไฟล์เวกเตอร์" > "ชื่อของไฟล์กฎ"
```

ผลลัพธ์ที่ได้จากขั้นตอนนี้เป็นไฟล์ 1 ไฟล์ที่มีกฎที่สามารถนำไปใช้จำแนกประเภทหัวข้อเอกสารข้อความในไฟล์ตัวอย่างมีดังนี้

```

earn
~earn

annoying --> earn

cons --> earn

expensive --> earn

[TRUE] --> ~earn

```

ภาคผนวก ค

ตัวอย่างของรายการคำศัพท์ในเชิงบวกและเชิงลบ

ตัวอย่างรายการคำศัพท์ในเชิงบวก 2,006 คำ ได้แก่ a+, abound, abounds, abundance, abundant, accessible, accessible, acclaim, acclaimed, acclamation, accolade, accolades, accommodative, accommodative, accomplish, accomplished, accomplishment, accomplishments, accurate, accurately, achievable, achievement, achievements, achievable, acumen, adaptable, adaptive, adequate, adjustable, admirable, admirably, admiration, admire, admirer, admiring, admiringly, adorable, adore, adored, adorer, adoring, adoringly, adroit, adroitly, adulate, adulation, adulatory, advanced, advantage, advantageous เป็นต้น

ตัวอย่างรายการคำศัพท์ในเชิงลบ 4,783 คำ ได้แก่ abnormal, abolish, abominable, abominably, abominate, abomination, abort, aborted, aborts, abrade, abrasive, abrupt, abruptly, abscond, absence, absent-minded, absentee, absurd, absurdity, absurdly, absurdness, abuse, abused, abuses, abusive, abysmal, abysmally, abyss, accidental, accost, accursed, accusation, accusations, accuse, accuses, accusing, accusingly, acerbate, acerbic, acerbically, ache, ached, aches, achey, aching, acrid, acridly, acridness, acrimonious, acrimoniously, acrimony, adamant, adamantly, addict, , addicted, addicting, addicts เป็นต้น

ประวัติผู้เขียนวิทยานิพนธ์

นายวฤธ เลี้ยวไพโรจน์ เกิดเมื่อวันที่ 23 มกราคม พ.ศ. 2532 ที่โรงพยาบาลจุฬาลงกรณ์ กรุงเทพมหานคร มีภูมิลำเนาอยู่จังหวัดสระบุรี สำเร็จการศึกษาระดับประถมศึกษาจากโรงเรียนอนุบาลสระบุรี จังหวัดสระบุรี สำเร็จการศึกษาระดับมัธยมศึกษาตอนต้นจากโรงเรียนสระบุรีวิทยาคม จังหวัดสระบุรี สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายจากโรงเรียนเตรียมอุดมศึกษา จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2553 (เกียรตินิยมอันดับ 2) และเข้าศึกษาในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2554