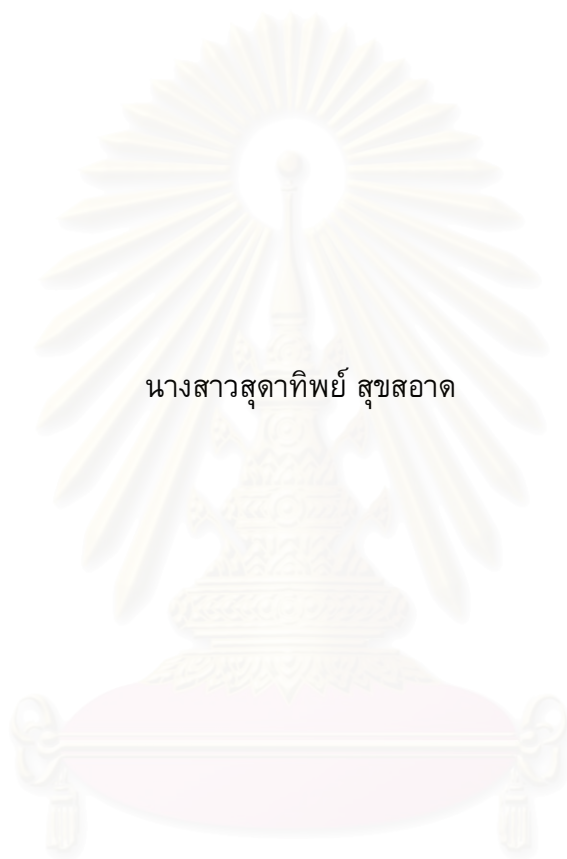


การค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้



นางสาวสุดาทิพย์ สุขสะอาด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

USE CASE RETRIEVAL USING USE CASE CLUSTERING
AND USER RELEVANCE FEEDBACK

Miss Sudathip Suksa-ard

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

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

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2007

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรง
ประเด็นจากผู้ใช้

โดย

นางสาวสุดาทิพย์ สุขสะอาด


สาขาวิชา

วิทยาศาสตร์คอมพิวเตอร์

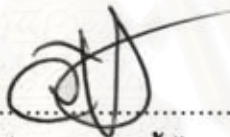
อาจารย์ที่ปรึกษา


ผู้ช่วยศาสตราจารย์ นครทิพย์ พร้อมพูล

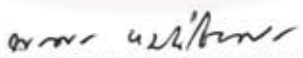
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็น
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

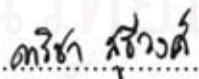

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร. บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ดร. วันชัย รั้วไพบูลย์)


..... อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ นครทิพย์ พร้อมพูล)


..... กรรมการ
(รองศาสตราจารย์ ดร. พรศิริ หมั่นไชยศรี)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. ดาริชา สุวีวงศ์)

สุดาทิพย์ สุขสะอาด : การค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรง
ประเด็นจากผู้ใช้. (USE CASE RETRIEVAL USING USE CASE CLUSTERING AND
USER RELEVANCE FEEDBACK) อ. ที่ปรึกษา : ผศ. นครทิพย์ พร้อมพูล, 179 หน้า.

การนำซอฟต์แวร์ที่มีคุณภาพและตรงกับวัตถุประสงค์ของการใช้งานกลับมาใช้ใหม่ เป็น
วิธีการหนึ่งที่ยิมนำมาใช้เพื่อเพิ่มอัตราการผลิต คุณภาพ และช่วยลดเวลากับค่าใช้จ่ายที่เกิดขึ้นใน
กระบวนการพัฒนาซอฟต์แวร์ โดยเฉพาะขั้นตอนการระบุความต้องการซอฟต์แวร์นั้นเป็นขั้นตอน
สำคัญ เพื่อให้ได้ข้อกำหนดความต้องการซอฟต์แวร์ที่มีคุณภาพและตรงกับความต้องการของผู้ที่
เกี่ยวข้อง คำอธิบายยูสเคสเป็นเครื่องมือหนึ่งที่ยิมนำมาใช้บันทึกข้อกำหนดความต้องการซอฟต์แวร์ การ
นำยูสเคสจากระบบงานเดิมที่มีการบันทึกไว้ก่อนหน้ากลับมาใช้ใหม่ เพื่อพัฒนาระบบที่มีลักษณะ
คล้ายคลึงกันนั้น สามารถประยุกต์ใช้หลักการจัดเก็บและค้นคืนสารสนเทศ

งานวิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อ นำเสนอวิธีการค้นคืนเอกสารคำอธิบายยูสเคส โดยใน
ขั้นตอนการจัดเก็บ ใช้การจัดกลุ่มเอกสารคำอธิบายยูสเคสตามโดเมนทางธุรกิจ ส่วนขั้นตอนการค้น
คืน จะคำนวณค่าความคล้ายโดยพิจารณาจากโครงสร้างของเอกสารคำอธิบายยูสเคสและค่าที่
ปรากฏในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคสระหว่างข้อความกับตัวแทนของเอกสาร
ทุกกลุ่ม จากนั้นคำนวณค่าความคล้ายระหว่างข้อความกับเอกสารคำอธิบายยูสเคสจากกลุ่มที่ถูก
เลือกซึ่งมีค่าความคล้ายสูงกว่าค่าขีดแบ่งเริ่มต้นที่กำหนด ข้อความเดิมและผลป้อนกลับที่ตรง
ประเด็นจากผู้ใช้ จะนำมาสร้างข้อความใหม่โดยใช้แบบจำลองปริภูมิเวกเตอร์และแบบจำลองความ
น่าจะเป็น จากนั้นพัฒนาเครื่องมือเพื่อทดสอบแนวคิดที่น่าเสนอและออกแบบการทดลองเพื่อประเมิน
ประสิทธิผลของระบบ

ผลการทดลองที่ได้จากงานวิทยานิพนธ์นี้แสดงให้เห็นว่า การค้นคืนยูสเคสโดยการจัดกลุ่ม
ยูสเคส และการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ให้ผลค่าความแม่นยำมากกว่าแบบไม่ให้ผล
ป้อนกลับที่ตรงประเด็นจากผู้ใช้ ทั้งแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น โดย
ในแบบจำลองปริภูมิเวกเตอร์ ให้ผลค่าความแม่นยำมากกว่าแบบจำลองความน่าจะเป็น 19.79
เปอร์เซ็นต์ เนื่องจากสามารถเปลี่ยนแปลงได้ทั้งค่าและน้ำหนักของค่าในข้อความ ในขณะที่
แบบจำลองความน่าจะเป็นเปลี่ยนแปลงเฉพาะน้ำหนักของค่าในข้อความเท่านั้น

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต.....สุดาทิพย์ สุขสะอาด.....
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา.....นครทิพย์ พร้อมพูล.....
ปีการศึกษา 2550

4770504221 : MAJOR COMPUTER SCIENCE

KEY WORD : USE CASE RETRIEVAL / USE CASE CLUSTERING / USER RELEVANCE FEEDBACK / USE CASE DESCRIPTION / QUERY REFORMULATION / TERM REWEIGHTING / VECTOR SPACE MODEL / PROBABILISTIC MODEL

SUDATHIP SUKSA-ARD : USE CASE RETRIEVAL USING USE CASE CLUSTERING AND USER RELEVANCE FEEDBACK. THESIS ADVISOR : ASST.PROF. NAKORNTHIP PROMPOON, 179 pp.

Reusing the quality software is one of the well known approach for increasing productivity and quality, and reducing costs and time of the software development process. Especially, specifying software requirements engineering is the significant process aims to earn the quality software requirements specification meets the stakeholders' requirements. Use case description is a popular tool used for keeping the software requirements specification. Reusing use cases which are recorded from the previous systems to develop the similar system can apply the principles of information and storage retrieval.

The objective of this thesis is to present an approach for use case description retrieval. In the storage phase, use case descriptions are clustered by their business domains. In the retrieval step, the similarity value is computed by considering the structure of use case description and terms appeared in each element of use case description between query and all centroids, represent each use case cluster. Then, similarity value between query and use case descriptions from any selected groups which has similarity score higher than predefined threshold is computed. In addition, the previous query and user relevance feedback is used to produce a new query using vector space model and probabilistic model. A supporting tool is developed to validate our proposed method, and the experiment is designed to evaluate the effectiveness of our approach.

The results of our experiment indicate that the precision value of our approach, use case retrieval using use case clustering and user relevance feedback, is greater than use case retrieval without user relevance feedback based on both the vector space model and probabilistic model. The precision value of the vector space model is 19.79 % more than that of the probabilistic model. Because the first model can adjust both terms and weight in the query whereas the second model can adjust only weight of term in that query.

Department : Computer Engineering

Field of Study : Computer Science

Academic Year : 2007

Student's Signature...*Sudathip Suksa-ard*

Advisor's Signature...*Nakornthip Prompoon*

กิตติกรรมประกาศ

ขอขอบพระคุณอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ นครทิพย์ พร้อมพูล ผู้เสียสละเวลาช่วยเหลือและให้คำปรึกษา คำแนะนำที่มีประโยชน์ต่องานวิจัย และความรู้ทางวิชาการอื่นๆ รวมถึงคำสั่งสอนด้านคุณธรรมและจริยธรรม ทำให้งานวิทยานิพนธ์ฉบับนี้สำเร็จ ลุล่วงไปได้ด้วยดี

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ทุกท่าน รองศาสตราจารย์ ดร. วันชัย รั้วไพฑูริย์ รองศาสตราจารย์ ดร.พรศิริ หมั่นไชยศรี และผู้ช่วยศาสตราจารย์ ดร.ดาริชา สุธีวงศ์ ที่กรุณาสละเวลาในการให้คำแนะนำเกี่ยวกับงานวิจัย และตรวจสอบความถูกต้อง สมบูรณ์ของวิทยานิพนธ์ฉบับนี้

ขอขอบพระคุณคณาจารย์ทุกท่านในภาควิชาวิศวกรรมคอมพิวเตอร์ คณะ วิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ความรู้ คำแนะนำในการเรียน และการทำวิจัย

ขอขอบคุณสมาชิกในห้องปฏิบัติการวิศวกรรมซอฟต์แวร์ รวมถึงเพื่อนๆ ชาว จุฬาย ทุกคนที่มอบกำลังใจ ความช่วยเหลือ และคำแนะนำต่างๆ อย่างสม่ำเสมอ

สุดท้ายนี้ ขอกราบขอบพระคุณบิดา มารดา และสมาชิกในครอบครัวทุกท่าน ผู้สนับสนุน ให้กำลังใจ แรงใจ ในการทำงานที่ต้องใช้ความมานะและอดทน แก่ผู้วิจัยตลอดมาจน สำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ต
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของงานวิจัย.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	3
1.3 ขอบเขตของงานวิจัย.....	3
1.4 ประโยชน์ของงานวิจัย.....	4
1.5 ขั้นตอนและวิธีการวิจัย.....	4
1.6 โครงสร้างของเนื้อหางานวิจัย.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 แผนภาพยูสเคส.....	5
2.1.2 ระบบการจัดเก็บและค้นคืนสารสนเทศ.....	8
2.1.3 การจัดกลุ่มเอกสาร.....	13
2.1.4 การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	16
2.2 งานวิจัยที่เกี่ยวข้อง.....	21
2.2.1 Use Case Retrieval Using Terms and Use Case Structure Similarity Computation.....	21
2.2.2 Improving Relevance Feedback in the Vector Space Model.....	24
2.2.3 A New Algorithm to Get the Initial Centroids.....	25
2.2.4 Advancing Candidate Link Generation for Requirements Tracing: The Study of Methods.....	25
บทที่ 3 วิธีการวิจัย.....	26
3.1 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	27
3.2 รวบรวมยูสเคสและวิเคราะห์แนวทางในการจัดกลุ่มยูสเคส.....	27
3.3 วิเคราะห์การค้นคืนยูสเคส.....	28

3.4	วิเคราะห์การคั่นคั้นยูสเคสแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	28
3.5	ภาพรวมการทำงานของวิธีที่นำเสนอ	29
3.5.1	ขั้นตอนการจัดเก็บยูสเคส	29
3.5.2	ขั้นตอนการคั่นคั้นยูสเคส.....	31
3.5.3	ขั้นตอนการประเมินประสิทธิผลของระบบการคั่นคั้นยูสเคส	36
บทที่ 4	การพัฒนาเครื่องมือ.....	37
4.1	สภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือ	37
4.1.1	ฮาร์ดแวร์	37
4.1.2	ซอฟต์แวร์	37
4.2	การออกแบบสถาปัตยกรรมของเครื่องมือ	38
4.3	โครงสร้างของเครื่องมือ	38
4.3.1	ส่วนการจัดเก็บเอกสารคำอธิบายยูสเคส	40
4.3.2	ส่วนการคั่นคั้นเอกสารคำอธิบายยูสเคส.....	40
4.4	แบบจำลองข้อมูล	44
บทที่ 5	การทดลอง.....	50
5.1	วัตถุประสงค์ของการทดลอง	50
5.2	วิธีการทดลอง	50
5.2.1	ยูสเคส	50
5.2.2	หน่วยตัวอย่าง.....	50
5.2.3	วิธีการคั่นคั้นยูสเคส.....	51
5.2.4	โจทย์ปัญหาที่ใช้ในการทดลอง.....	52
5.2.5	ข้อคำถาม.....	53
5.3	ขั้นตอนการทดลอง.....	54
5.4	การกำหนดค่าขีดแบ่งเริ่มต้นความคล้าย	55
5.5	ผลการทดลอง.....	56
5.6	การวิเคราะห์ผลการทดลองโดยทางสถิติ	66
5.7	สรุปผลการทดลอง	71
5.8	ข้ออภิปราย.....	73
บทที่ 6	สรุปผลงานวิจัย	75
6.1	สรุปผลงานวิจัย.....	75

	หน้า
6.2 งานวิจัยในอนาคต.....	78
6.3 บทความวิชาการที่ตีพิมพ์.....	78
รายการอ้างอิง	80
ภาคผนวก	82
ภาคผนวก ก ยูสเคสของระบบที่เลือกใช้.....	83
ภาคผนวก ข ข้อคำถามจากหน่วยตัวอย่าง.....	115
ภาคผนวก ค ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค ที่ได้จากการทดลอง..	136
ภาคผนวก ง สรุปสูตรที่ใช้ในงานวิจัย.....	157
ภาคผนวก จ บทความวิชาการที่ตีพิมพ์	160
จ.1 บทความวิชาการเรื่อง "การค้นคืนยูสเคสด้วยการจำแนกประเภทยูสเคสและ การค้นคืนย้อนกลับจากผู้ใช้"	161
จ.2 บทความวิชาการเรื่อง "การจัดเก็บและค้นคืนยูสเคสที่มีการจัดกลุ่มยูสเคส ร่วมกับการค้นคืนย้อนกลับจากผู้ใช้"	171
ประวัติผู้เขียนวิทยานิพนธ์.....	179

สารบัญตาราง

ตาราง	หน้า
ตารางที่ 2.1	23
ตารางที่ 4.1	46
ตารางที่ 4.2	47
ตารางที่ 4.3	47
ตารางที่ 4.4	47
ตารางที่ 4.5	47
ตารางที่ 4.6	48
ตารางที่ 4.7	48
ตารางที่ 4.8	48
ตารางที่ 4.9	48
ตารางที่ 4.10	49
ตารางที่ 4.11	49
ตารางที่ 4.12	49
ตารางที่ 4.13	49
ตารางที่ 4.14	49
ตารางที่ 5.1	53
ตารางที่ 5.2	57
ตารางที่ 5.3	58
ตารางที่ 5.4	60
ตารางที่ 5.5	62

ตาราง	หน้า
ตารางที่ 5.6 สรุปผลค่าความแม่นยำเฉลี่ยจากข้อคำถามทั้งหมด 250 ข้อคำถามเรียงลำดับตามค่าเรียกคืนทั้ง 11 ค่า จาก 0 ถึง 1 ในแต่ละวิธีการค้นคืน.....	63
ตารางที่ 5.7 แสดงค่าสถิติทดสอบการแจกแจงปกติของค่าความแม่นยำโดยเฉลี่ยในแต่ละวิธีการค้นคืน.....	67
ตารางที่ 5.8 แสดงค่าสถิติทดสอบแบบจับคู่ระหว่างค่าเฉลี่ยสองประชากรของการทดสอบสมมุติฐานที่ 1 ถึง 8.....	71
ตารางที่ ก.1 ระบบงานที่เลือกใช้ 40 ระบบงาน.....	83
ตารางที่ ก.2 ระบบงาน ICT Blocking and Filtering.....	85
ตารางที่ ก.3 ระบบงาน Departmental Information System for Curriculum and Course Offering Management.....	86
ตารางที่ ก.4 ระบบงาน Resource Management.....	87
ตารางที่ ก.5 ระบบงาน Automatic Restaurant lookup.....	88
ตารางที่ ก.6 ระบบงาน Library Management.....	89
ตารางที่ ก.7 ระบบงาน Mobile Phone Selling on the Internet.....	89
ตารางที่ ก.8 ระบบงาน Payroll.....	90
ตารางที่ ก.9 ระบบงาน Personal Stock Manager on PDA.....	91
ตารางที่ ก.10 ระบบงาน Basic Product Trading.....	92
ตารางที่ ก.11 ระบบงาน Mobile Phone Service Paying.....	93
ตารางที่ ก.12 ระบบงาน Basic MLM Computerize.....	93
ตารางที่ ก.13 ระบบงาน Content Management and Platform Upgrade for the Official Website of Nation Channel.....	94
ตารางที่ ก.14 ระบบงาน Purchase, Producing and Export.....	95
ตารางที่ ก.15 ระบบงาน Management Information System for Insurance Company.....	96
ตารางที่ ก.16 ระบบงาน E-Learning.....	96
ตารางที่ ก.17 ระบบงาน Management Information System for an Academic.....	97
ตารางที่ ก.18 ระบบงาน SME Production Planning.....	99
ตารางที่ ก.19 ระบบงาน E-Document.....	100
ตารางที่ ก.20 ระบบงาน Purchase and Sale Real Estate Management Online.....	101
ตารางที่ ก.21 ระบบงาน Simulation Office.....	101
ตารางที่ ก.22 ระบบงาน Thai Spell Checking and Suggesting.....	103
ตารางที่ ก.23 ระบบงาน MRP-Material Resource Planning.....	103

ตาราง	หน้า
ตารางที่ ก.24 ระบบงาน Hided Information Security	104
ตารางที่ ก.25 ระบบงาน Next Generation Requirements Management	105
ตารางที่ ก.26 ระบบงาน News Operation News Reporting.....	105
ตารางที่ ก.27 ระบบงาน Customer Follow-up and Appointment.....	106
ตารางที่ ก.28 ระบบงาน Support and Management Information System of Online Car Insurance	106
ตารางที่ ก.29 ระบบงาน Central Help Desk	107
ตารางที่ ก.30 ระบบงาน Cruise Control and Monitoring: Case Study	108
ตารางที่ ก.31 ระบบงาน Data converter between bank	108
ตารางที่ ก.32 ระบบงาน ATM.....	109
ตารางที่ ก.33 ระบบงาน Public Relation Online	109
ตารางที่ ก.34 ระบบงาน News Operation Staff Training	110
ตารางที่ ก.35 ระบบงาน Intelligent Dictionary	110
ตารางที่ ก.36 ระบบงาน Job Assignment Management.....	111
ตารางที่ ก.37 ระบบงาน The Primary Information System to Plan Education and Check Graduation for Undergraduate Students of Engineering Faculty, Chulalongkorn University.....	111
ตารางที่ ก.38 ระบบงาน Departmental Information System for Research Management	112
ตารางที่ ก.39 ระบบงาน Real Time News Feeding.....	112
ตารางที่ ก.40 ระบบงาน Staff Procurement.....	113
ตารางที่ ก.41 ระบบงาน Software Subcontract Management.....	114
ตารางที่ ข.1 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ ปัญหาเรื่อง “Educational System”	115
ตารางที่ ข.2 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ ปัญหาเรื่อง “Material Requirements Planning/Manufacturing Planning” ..	120
ตารางที่ ข.3 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ ปัญหาเรื่อง “Member/User’s Information Management”	124
ตารางที่ ข.4 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ ปัญหาเรื่อง “Product Trading Operation”	129
ตารางที่ ข.5 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ ปัญหาเรื่อง “Report Generation”	132

ตาราง	หน้า
ตารางที่ ค.1 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4	137
ตารางที่ ค.2 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4	139
ตารางที่ ค.3 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4	141
ตารางที่ ค.4 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4	143
ตารางที่ ค.5 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4	145
ตารางที่ ค.6 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8	147
ตารางที่ ค.7 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8	149
ตารางที่ ค.8 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8	151
ตารางที่ ค.9 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8	153
ตารางที่ ค.10 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8	155
ตารางที่ ง.1 สรุปลสูตรและคำอธิบายสูตรทั้งหมดที่ใช้ในงานวิจัย.....	157

สารบัญญภาพ

ภาพประกอบ	หน้า
รูปที่ 2.1 องค์ประกอบสำคัญของแผนภาพยูสเคส	5
รูปที่ 2.2 ตัวอย่างแผนภาพยูสเคสในระบบการสั่งซื้อสินค้า.....	6
รูปที่ 2.3 รูปแบบของเอกสารคำอธิบายยูสเคส	7
รูปที่ 2.4 กระบวนการในระบบการจัดเก็บและคั่นคืนสารสนเทศ	9
รูปที่ 2.5 แผนภาพกิจกรรมการทำธุรชนี้อัตโนมติ	11
รูปที่ 2.6 ค่าเรียกคืนและค่าความแม่นยำจากข้อคำถามหนึ่ง ๆ	13
รูปที่ 2.7 ตัวอย่างชุดข้อมูลที่มีการจัดกลุ่มเอกสาร.....	15
รูปที่ 2.8 ผลที่เกิดจากการดำเนินการโดยการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	17
รูปที่ 2.9 แผนภาพแสดงชุดของเอกสารที่ใช้ในแบบจำลองปริภูมิเวกเตอร์.....	18
รูปที่ 2.10 แสดงค่าน้ำหนักที่เป็นไปได้ในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส.....	22
รูปที่ 3.1 ขั้นตอนการดำเนินการวิจัย	26
รูปที่ 3.2 แผนภาพกิจกรรมการจัดเก็บเอกสารยูสเคสเพื่อทำธุรชนีและน้ำหนักคำอัติโนมติ...29	29
รูปที่ 3.3 แผนภาพกิจกรรมของขั้นตอนโดยรวมของวิธีที่นำเสนอ	30
รูปที่ 3.4 แผนภาพกิจกรรมของการหาตัวแทนกลุ่มเอกสารยูสเคส.....	31
รูปที่ 3.5 แผนภาพกิจกรรมของการคั่นคืนยูสเคส.....	31
รูปที่ 3.6 แผนภาพกิจกรรมของการคั่นคืนแบบให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	35
รูปที่ 3.7 แผนภาพกิจกรรมการประเมินประสิทธิผลของระบบ	36
รูปที่ 4.1 สถาปัตยกรรมของเครื่องมือสำหรับการคั่นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและ การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	38
รูปที่ 4.2 แผนภาพส่วนประกอบโครงสร้างของเครื่องมือ	39
รูปที่ 4.3 หน้าจอแสดงรายการหลักของเครื่องมือ	39
รูปที่ 4.4 หน้าจอแสดงส่วนสำหรับการจัดเก็บเอกสารคำอธิบายยูสเคส	41
รูปที่ 4.5 หน้าจอแสดงส่วนสำหรับการคั่นคืนเอกสารคำอธิบายยูสเคส	42
รูปที่ 4.6 หน้าจอแสดงผลลัพธ์เอกสารคำอธิบายยูสเคสที่คั่นคืนแบบไม่มีผลป้อนกลับที่ตรง ประเด็นจากผู้ใช้.....	42
รูปที่ 4.7 หน้าจอแสดงรายละเอียดของเอกสารคำอธิบายยูสเคสที่ผู้ใช้เลือกเข้าไปดู.....	43
รูปที่ 4.8 หน้าจอแสดงผลลัพธ์เอกสารคำอธิบายยูสเคสที่คั่นคืนหลังให้ผลป้อนกลับที่ตรง ประเด็นจากผู้ใช้.....	44
รูปที่ 4.9 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูลของระบบการคั่นคืนยูสเคสโดยการจัด กลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้.....	45
รูปที่ 5.1 แผนภาพกิจกรรมขั้นตอนการทดลอง	54

ภาพประกอบ	หน้า
รูปที่ 5.2 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 1 และ 2 ตามลำดับ.....	64
รูปที่ 5.3 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 3 และ 4 ตามลำดับ.....	64
รูปที่ 5.4 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 5 และ 6 ตามลำดับ.....	65
รูปที่ 5.5 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 7 และ 8 ตามลำดับ.....	65
รูปที่ 5.6 กราฟค่าเรียกคืน และค่าความแม่นยำของวิธีการค้นคืน 4 วิธี เฉพาะการค้นคืนแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส.....	66

บทที่ 1

บทนำ

ในบทนี้ จะกล่าวถึงแนวคิดหลักของงานวิจัย อันประกอบไปด้วย ที่มาและความสำคัญ วัตถุประสงค์ ขอบเขต ประโยชน์ ขั้นตอนและโครงสร้างของเนื้อหาของงานวิจัย ซึ่งมีเนื้อหา ดังต่อไปนี้

1.1 ที่มาและความสำคัญของงานวิจัย

ในกระบวนการพัฒนาซอฟต์แวร์หนึ่งๆ นั้นมีขั้นตอนต่างๆ มากมาย แต่ละขั้นตอนจะ เสียค่าใช้จ่ายแตกต่างกันไป การลดค่าใช้จ่ายเท่าที่จะสามารถทำได้จึงเป็นเป้าหมายหนึ่งสำหรับ ทุกองค์กร และยังเป็นความท้าทายในกระบวนการพัฒนาซอฟต์แวร์ด้วย วิธีการหนึ่งเพื่อลด ค่าใช้จ่ายในกระบวนการดังกล่าว ที่นิยมใช้ และมีผลดีก็คือ การนำซอฟต์แวร์กลับมาใช้ใหม่ (Software Reuse) [1] โดยในกระบวนการพัฒนาซอฟต์แวร์ ขั้นตอนการกำหนดความต้องการ เป็นขั้นตอนหนึ่งที่มีความสำคัญมากเพราะเป็นขั้นตอนที่ใช้เวลา และค่าใช้จ่ายสูง เพื่อให้ได้มา ซึ่งข้อกำหนดความต้องการซอฟต์แวร์ (Software Requirements Specification) ที่รวบรวมไว้ นั้น มีความครบถ้วนและตรงกับความต้องการของผู้เกี่ยวข้องทั้งหลาย

ยูสเคส (Use Case) เป็นเครื่องมือที่ผู้บันทึกข้อกำหนดความต้องการ หรือวิศวกร ซอฟต์แวร์ นำมาใช้ในการบันทึกข้อกำหนดความต้องการของระบบอย่างแพร่หลายในวงการ พัฒนาซอฟต์แวร์ [2] ดังนั้นการนำยูสเคสกลับมาใช้ใหม่ จึงมีความสำคัญอย่างยิ่ง กล่าวคือ วิศวกรซอฟต์แวร์สามารถใช้ข้อมูลเดิมของระบบที่มีความคล้ายกัน และทำการเพิ่มเติมข้อมูล อื่นๆ ตามความเหมาะสม ซึ่งไม่ต้องเริ่มต้นใหม่โดยปราศจากข้อมูล และจะเป็นประโยชน์มาก หากยูสเคสที่มีการบันทึกไว้ก่อนหน้านี้ มีความสมบูรณ์โดยได้ระบุถึงฟังก์ชันต่างๆ อย่าง ครบถ้วน ผู้ที่นำยูสเคสกลับมาใช้อีกครั้งเพียงแคปรับแต่งให้เหมาะสมกับความต้องการใหม่ ดังนั้นการนำยูสเคสกลับมาใช้ใหม่จึงนับเป็นประโยชน์อย่างยิ่ง สามารถช่วยลดเวลา และ ค่าใช้จ่ายในกระบวนการพัฒนาซอฟต์แวร์

ในการที่จะนำยูสเคสต่างๆ ซึ่งเก็บไว้ในไลบรารี (Library) มาใช้ใหม่ได้นั้น ต้องอาศัย กระบวนการในระบบการจัดเก็บและค้นคืนสารสนเทศ (Information Storage and Retrieval System) เข้ามาสนับสนุนด้วย ซึ่งในระบบการจัดเก็บและค้นคืนสารสนเทศ มีประเด็นพื้นฐาน 4 ประการที่สำคัญอันได้แก่ การกำหนดลักษณะที่เหมาะสมสำหรับใช้เป็นตัวแทนของเอกสาร (Document Representation) ภาษาที่ใช้ในการคิวรี (Query Language) การใช้ฟังก์ชันในการ ทดสอบ หรือเปรียบเทียบค่าความคล้ายกัน (Similarity Function) ระหว่างข้อความกับลักษณะ ที่ใช้เป็นตัวแทนเอกสารดังกล่าว ตลอดจนการจัดอันดับผลของเอกสารที่ค้นคืนได้จากการ ทดสอบความคล้ายกันนั้น โดยใช้ฟังก์ชันในการจัดอันดับ (Ranking Function) ที่เหมาะสม

การจัดเก็บและการค้นคืนยูสเคส ก็สามารถประยุกต์ใช้หลักการเดียวกันกับการจัดเก็บและการค้นคืนเอกสารทั่วไป ซึ่งจำเป็นต้องมีการกำหนดรูปแบบที่เหมาะสมที่จะเป็นตัวแทนให้กับยูสเคส การกำหนดรูปแบบข้อความ รวมทั้งฟังก์ชันในการหาความคล้ายกันระหว่างข้อความจากผู้ใช้และตัวแทนของยูสเคส และฟังก์ชันในการจัดอันดับผลลัพธ์ที่ได้จากการค้นคืน เพื่อให้ได้มาซึ่งยูสเคสที่ตรงกับความต้องการของผู้พัฒนาซอฟต์แวร์มากที่สุด

อย่างไรก็ตามเนื่องด้วยงานในระบบการค้นคืนสารสนเทศทุกวันนี้ ต้องสามารถรองรับการจัดการเกี่ยวกับข้อมูลที่มีจำนวนเพิ่มมากขึ้น นั้นหมายความว่า นอกจากจะต้องจัดเก็บเอกสารเพิ่มขึ้นแล้ว จำนวนเทอมที่ใช้เป็นดรรชนีชี้โยงไปยังเอกสารก็มากขึ้นด้วย ดังนั้นจะเป็นประโยชน์อย่างยิ่ง หากมีการเพิ่มเติมในส่วนการจัดองค์ประกอบให้กับการจัดเก็บเอกสารเพื่อให้ง่ายต่อการจัดการ และการเข้าถึงตัวเอกสาร [3] โดยวิธีการหนึ่งที่ทำได้คือ การจัดกลุ่มเอกสาร (Document Clustering) ซึ่งการจัดกลุ่มให้กับเอกสารนี้ นับเป็นวิธีการที่ช่วยเพิ่มศักยภาพให้กับระบบการค้นคืนสารสนเทศ กล่าวคือ ข้อคำถามที่เข้าสู่ระบบจะถูกนำมาเปรียบเทียบกับดรรชนีของตัวแทนกลุ่มเอกสารแต่ละกลุ่มเสียก่อน แทนที่จะเปรียบเทียบกับดรรชนีของแต่ละเอกสาร ซึ่งนอกจากจะทำให้เข้าถึงกลุ่มเอกสารที่มีความใกล้เคียง หรือคล้ายกับเอกสารที่ผู้ใช้ต้องการได้เร็วขึ้นแล้ว ยังช่วยให้ระบบสามารถค้นหาเอกสารที่ตรงกับความต้องการของผู้ใช้เพิ่มมากขึ้นด้วย

แม้ว่าจะมีการจัดองค์ประกอบต่าง ๆ เพื่อให้ง่ายต่อผู้ใช้ในการเข้าถึงข้อมูลที่ต้องการเป็นอย่างดีแล้วก็ตาม แต่เนื่องด้วยความต้องการที่แตกต่างกันไปของผู้ใช้แต่ละคน ยังเป็นปัญหาสำคัญ ซึ่งวิธีการแก้ไขปัญหาก็เพื่อช่วยให้ระบบค้นคืนข้อมูล สามารถดึงเอกสารที่ตรงกับความต้องการ หรือเป็นที่พอใจของผู้ใช้ วิธีการหนึ่งที่ได้ผลดี และช่วยปรับปรุงประสิทธิภาพให้กับระบบได้ นั่นคือ การให้ผลป้อนกลับที่ตรงประเด็น (Relevance Feedback) [4] เพราะวิธีการใช้ผลป้อนกลับที่ตรงประเด็นนั้นจะใช้ข้อมูลที่ได้กลับมาจากผู้ใช้งาน มาทำการหาข้อคำถามใหม่ที่เหมาะสมกว่าข้อคำถามเดิม ซึ่งจะช่วยให้ระบบรู้ว่าเอกสารใดบ้างที่มีความเกี่ยวข้อง ตรงประเด็นกับที่ผู้ใช้ต้องการ เพื่อที่ระบบจะได้นำข้อมูลเหล่านั้นไปใช้ในการค้นคืนครั้งถัดไป ส่งผลให้ได้มาซึ่งเอกสารที่ตรงตามความต้องการ และสามารถทำการจัดอันดับได้อย่างเหมาะสมมากยิ่งขึ้น

ส่วนเรื่องการนำยูสเคสกลับมาใช้ใหม่นั้น มีหลายงานวิจัยที่นำเสนอแนวคิดในประเด็นดังกล่าวแตกต่างกันไป เช่น ในการนำยูสเคสกลับมาใช้ใหม่ จากงานวิจัยของ Motoshi Saeki [5] ได้เสนอแบบรูป (Pattern) ต่าง ๆ ที่สามารถนำกลับมาใช้ใหม่ เพื่อให้ได้มาซึ่งข้อกำหนดความต้องการ โดยที่ข้อกำหนดความต้องการนั้น จะสามารถนำมาแปลงเป็นคำอธิบายยูสเคส (Use Case Description) ได้ แต่ข้อจำกัดของงานวิจัยนี้คือ จะไม่สามารถนำข้อมูลภายในคำอธิบายยูสเคสที่มีอยู่เดิมกลับมาใช้ใหม่ได้ หรือในงานวิจัยของ Han G. Woo and William N. Robinson [6] ได้เสนอวิธีการนำแผนภาพยูสเคสกลับมาใช้ใหม่ โดยการพัฒนาเครื่องมือที่ชื่อว่า "ScenAsst" เพื่อแปลงยูสเคสให้เป็นกราฟ แล้วจัดกลุ่มเก็บไว้ในไลบรารี หลังจากนั้นทำการค้น

คืนด้วยข้อความที่ถูกแปลงให้อยู่ในรูปแบบของกราฟเช่นเดียวกัน แล้วเปรียบเทียบค่าความคล้ายกัน (Similarity) กับยูสเคสที่เก็บไว้ ซึ่งเทคนิคในงานวิจัยนี้ต้องใช้เวลาในการประมวลผลค่อนข้างมากเนื่องจากมีความซับซ้อนของโครงสร้างสูง แต่ภาพรวมจากผลของงานวิจัยทั้งสองต่างก็ถือว่าให้ผลลัพธ์ที่ดีในการนำยูสเคสกลับมาใช้ใหม่

นอกจากนี้ งานวิจัยของ อัครเดช อุดมชัยพร และคณะ [7] ใช้การจัดเก็บเอกสารที่เป็นคำอธิบายยูสเคส และใช้การค้นคืนด้วยการหาค่าความคล้ายกันของคำในแต่ละส่วนประกอบ (Elements) ของเอกสารคำอธิบายยูสเคสกับข้อความที่มีลักษณะเป็นคำอธิบายยูสเคสอย่างง่าย (Simple Use Case Description) โดยที่ผู้ใช้สามารถระบุค่าน้ำหนักในแต่ละส่วนประกอบได้เพื่อเป็นการกำหนดความสำคัญให้กับแต่ละส่วนประกอบแตกต่างกันไป แล้วจึงหาค่าผลรวมความคล้ายกันที่ได้จากแต่ละส่วนประกอบนั้น จากนั้นระบบจะเลือกยูสเคสที่มีค่าคะแนนความคล้ายกันตามค่าขีดแบ่ง (Threshold Value) ที่กำหนดไว้ ก่อนแสดงเอกสารยูสเคสที่มีความคล้ายกันกับข้อความ โดยการจัดอันดับของเอกสารตามคะแนนความคล้ายกันจากมากไปหาน้อย จุดเด่นของงานวิจัยนี้คือ การนำข้อมูลและรายละเอียดต่าง ๆ ที่อยู่ภายในคำอธิบายยูสเคสมาใช้ใหม่ได้โดยที่ผลการค้นคืนคำอธิบายยูสเคสจะมีความแม่นยำสูงขึ้น เนื่องจากมีการพิจารณาด้วยคำและโครงสร้างของยูสเคสเป็นหลัก แต่งานวิจัยนี้ ไม่ได้พิจารณาถึงการจัดกลุ่มยูสเคส ก่อนที่จะนำข้อความมาเปรียบเทียบเพื่อหาค่าความคล้ายกัน และไม่มีขั้นตอนสำหรับให้ผู้ใช้พิจารณาเลือกยูสเคสที่ตรงกับความต้องการเพื่อนำไปใช้ในขั้นตอนการค้นคืนโดยมีผลป้อนกลับ

ดังนั้นในงานวิทยานิพนธ์นี้ จะใช้แนวคิดในการนำยูสเคสกลับมาใช้ใหม่จากงานวิจัยของ อัครเดช อุดมชัยพร มาเพิ่มเติมในส่วนการจัดกลุ่มให้กับยูสเคส และการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ โดยการเปลี่ยนแปลงเทอมและน้ำหนักของคำในข้อความบนแบบจำลองปริภูมิเวกเตอร์ และการเปลี่ยนแปลงเฉพาะค่าน้ำหนักของคำในข้อความบนแบบจำลองความน่าจะเป็น เพื่อปรับปรุงประสิทธิภาพของระบบดังกล่าวให้ดีขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

- 1) เพื่อเสนอวิธีการการค้นคืนยูสเคสที่พิจารณาโครงสร้างของยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้
- 2) เพื่อพัฒนาเครื่องมือในการค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

1.3 ขอบเขตของงานวิจัย

- 1) พัฒนาเครื่องมือที่ใช้ในการค้นคืนยูสเคส ซึ่งมีวิธีการจัดกลุ่มยูสเคส และให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้
- 2) ระบบค้นคืนยูสเคสนี้สามารถใช้ได้กับเอกสารยูสเคสที่เป็นภาษาอังกฤษเท่านั้น
- 3) ระบบค้นคืนยูสเคสนี้สามารถค้นคืนแบบวิธีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ได้

- 4) ระบบคั่นคินยูสเคสนี้หาตัวแทนกลุ่มเอกสารยูสเคสแต่ละกลุ่มได้
- 5) ในการคั่นคินของระบบนี้ จะไม่มีการพิจารณาเรื่องความกำกวมของคำ
- 6) งานวิจัยนี้ใช้ค่าของความแม่นยำ ค่าเรียกคิน และค่าเฉลี่ยฮาร์โมนิก ในการวัดประสิทธิผลของการคั่นคินยูสเคส ของระบบ
- 7) ใช้ชุดข้อมูลยูสเคสอย่างน้อย 40 ระบบงาน

1.4 ประโยชน์ของงานวิจัย

ทำให้ได้เครื่องมือที่ใช้ในการคั่นคินยูสเคส ซึ่งมีวิธีการจัดกลุ่มเอกสาร และมีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน เพื่อให้ผู้ใช้งานสามารถนำยูสเคสที่ตรงกับความต้องการ กลับมาใช้ใหม่ได้

1.5 ขั้นตอนและวิธีการวิจัย

- 1) ศึกษาวิธีการจัดเตรียมเอกสาร และรายละเอียดต่างๆ ที่เกี่ยวกับยูสเคส
- 2) ศึกษากระบวนการจัดเก็บ และคั่นคินสารสนเทศที่เหมาะสมกับการคั่นคินยูสเคส
- 3) ศึกษาทฤษฎีที่เกี่ยวข้องกับการจัดกลุ่มเอกสารที่เหมาะสมกับเอกสารยูสเคส
- 4) ศึกษาทฤษฎีที่เกี่ยวข้องกับการดำเนินการกับข้อความในส่วนของการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน
- 5) วิเคราะห์และออกแบบโครงสร้างของข้อมูล และโครงสร้างแฟ้มข้อมูลที่ใช้ตามแนววิธีนี้
- 6) วิเคราะห์และออกแบบขั้นตอนในการจัดเก็บ การคั่นคิน และประเมินผลการคั่นคิน
- 7) วิเคราะห์และออกแบบขั้นตอนในการคั่นคินแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน และการประเมินผลการคั่นคินแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน
- 8) พัฒนาโปรแกรมที่ได้ออกแบบไว้
- 9) ทดสอบการทำงานของโปรแกรมที่พัฒนา
- 10) เตรียมการทดลอง และทำการทดลอง
- 11) ประเมินผลลัพธ์ที่ได้จากการทดลอง
- 12) จัดทำเอกสารสรุปงานวิจัย และข้อเสนอแนะ

1.6 โครงสร้างของเนื้อหางานวิจัย

เนื้อหาโครงสร้างของวิทยานิพนธ์ฉบับนี้ถูกแบ่งออกเป็น 6 บทดังนี้ คือ บทที่ 1 เป็นบทนำ บทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยต่างๆ ที่เกี่ยวข้อง บทที่ 3 กล่าวถึงขั้นตอนวิธีการวิจัยที่ใช้ในการคั่นคินยูสเคส การจัดกลุ่มของยูสเคส และวิธีคั่นคินแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน รวมถึงการประเมินประสิทธิผลของแนวคิดในงานวิจัย บทที่ 4 จะกล่าวถึงการออกแบบและพัฒนาเครื่องมือที่สนับสนุนตามขั้นตอนวิธีที่นำเสนอ บทที่ 5 กล่าวถึงการทดลองและผลการทดลองของขั้นตอนวิธีที่นำเสนอ รวมถึงการวิเคราะห์ผลการทดลอง และบทที่ 6 ซึ่งเป็นบทสุดท้ายจะเป็นบทสรุปของงานวิจัย รวมทั้งงานวิจัยในอนาคต และบทความวิชาการที่ตีพิมพ์

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

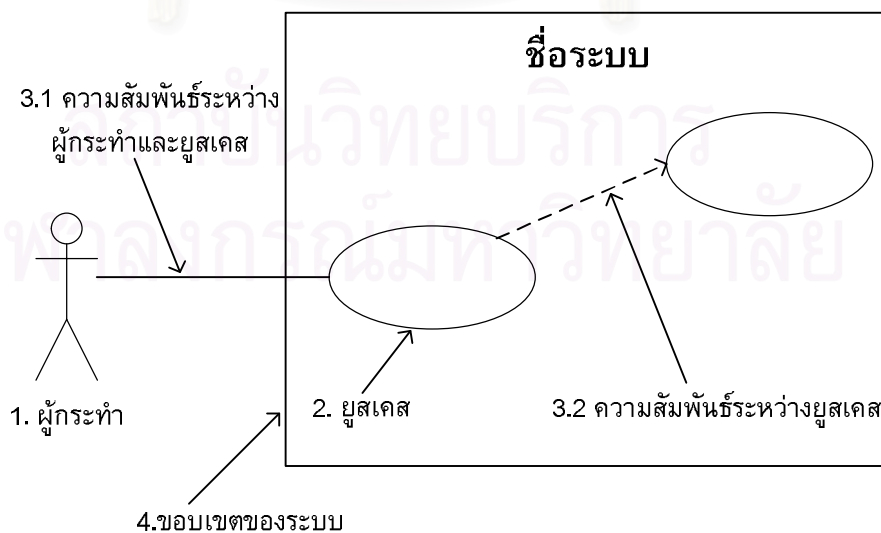
ในบทนี้ จะกล่าวถึงทฤษฎีที่สำคัญ ซึ่งได้นำมาประยุกต์ สันนิษฐาน และใช้อ้างอิงในการทำงานวิทยานิพนธ์ รวมถึงข้อดีและข้อจำกัดของงานวิจัยต่างๆ ที่เกี่ยวข้อง โดยมีเนื้อหา ดังต่อไปนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

งานวิทยานิพนธ์นี้เกี่ยวข้องกับการเพิ่มประสิทธิภาพในการค้นคืนยูสเคสโดยการจัดกลุ่ม ยูสเคสในส่วนการจัดเก็บยูสเคส ร่วมกับการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ในแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น โดยมีแนวคิด และทฤษฎีที่เกี่ยวข้อง ดังต่อไปนี้

2.1.1 แผนภาพยูสเคส (Use Case Diagram)

ในงานด้านวิศวกรรมซอฟต์แวร์และวิศวกรรมระบบนั้น ยูสเคส (Use Case) คือ เทคนิคหรือวิธีการหนึ่งที่ใช้เพื่อตีความความต้องการเชิงหน้าที่ต่างๆ ที่จำเป็นต้องมีในระบบธุรกิจและระบบเทคโนโลยีสารสนเทศที่สนับสนุนระบบธุรกิจนั้นๆ ในแบบจำลองยูสเคส (Use Case Model) จึงมีผู้กระทำ (Actor) และยูสเคสหลายๆ ยูสเคส (Use Cases) มีผู้กระทำเป็นตัวแทนของบุคคลหรือระบบภายนอก และยูสเคสต่างๆ เป็นฟังก์ชันภายในระบบนั้น หรือกล่าวได้ว่ายูสเคสเป็นชุดของแผนงานที่อธิบายการทำงานร่วมกันระหว่างผู้ใช้ (User) กับระบบ (System) โดยมีแผนภาพยูสเคส ในการนำเสนอความสัมพันธ์ระหว่างผู้กระทำกับยูสเคสต่างๆ โดย 4 องค์ประกอบหลักของแผนภาพยูสเคส [8] แสดงดังรูปที่ 2.1

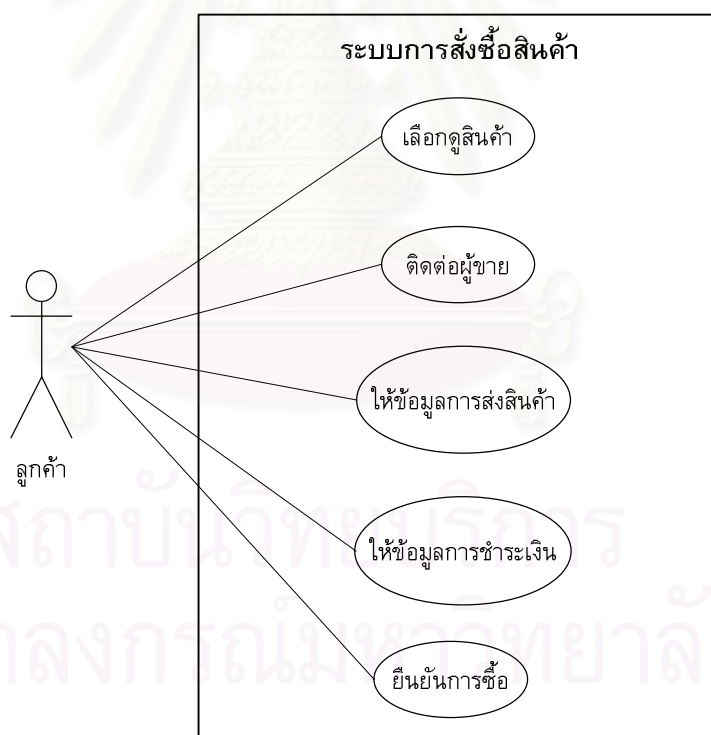


รูปที่ 2.1 องค์ประกอบสำคัญของแผนภาพยูสเคส

4 องค์ประกอบสำคัญของแผนภาพยูสเคส ได้แก่

- 1) ผู้กระทำ (Actor) เป็นตัวแทนของผู้ใช้ หรือระบบภายนอกที่มีการกระทำ หรือมีผลกระทบต่อระบบงานภายในแผนภาพ
- 2) ยูสเคส (Use Case) ใช้อธิบายการกระทำต่างๆ (Actions) ที่ถูกกระทำโดยผู้กระทำ
- 3) ความสัมพันธ์ (Relationship) สามารถแบ่งได้ 2 ประเภทใหญ่ๆ ได้แก่ ความสัมพันธ์ระหว่างผู้กระทำกับยูสเคส และความสัมพันธ์ระหว่างยูสเคสด้วยกัน
- 4) ขอบเขตของระบบ (System Boundary) ส่วนที่เป็นขอบเขตของระบบคือ กรอบสี่เหลี่ยมที่มียูสเคสอยู่ภายในซึ่งแสดงความสามารถ หรือหน้าที่ (Functionality) ที่ระบบสามารถทำได้

แผนภาพยูสเคสจึงเป็นมุมมองภายนอก (External View) ของระบบซึ่งอธิบายการกระทำที่ผู้ใช้จะปฏิบัติเพื่อให้บรรลุภารกิจนั้นๆ ตัวอย่างแผนภาพยูสเคสในระบบการสั่งซื้อ (Ordering System) แสดงดังรูปที่ 2.2 สามารถอธิบายรายละเอียดของระบบคือ ลูกค้าสามารถกระทำการดังต่อไปนี้ ได้แก่ การเลือกดูสินค้า การติดต่อผู้ขาย การให้ข้อมูลการส่งสินค้า การให้ข้อมูลการชำระเงิน และการยืนยันการซื้อ



รูปที่ 2.2 ตัวอย่างแผนภาพยูสเคสในระบบการสั่งซื้อสินค้า

สำหรับรายละเอียดภายในแผนภาพยูสเคสนั้นจะอธิบายด้วย คำอธิบายยูสเคส [9] ซึ่งมีรูปแบบ (Template) ดังแสดงในรูปที่ 2.3

Use Case Name:	
Objective:	
Actor:	
Relationship:	Association:
	Include:
	Extend:
	Generalization:
Precondition:	
Postcondition:	
Normal Flow of Events:	1. 2. 3. ...
Subflow:	
Alternative or Exceptional Flow of Events:	1. 2. 3. ...

รูปที่ 2.3 รูปแบบของเอกสารคำอธิบายยูสเคส

ส่วนประกอบแต่ละส่วนของคำอธิบายยูสเคสได้แก่

- 1) ชื่อยูสเคส (Use Case Name) คือ ชื่อที่ใช้เรียกยูสเคส โดยแต่ละยูสเคสในระบบงานหนึ่งๆ ต้องมีชื่อที่ไม่ซ้ำกัน และสื่อถึงวัตถุประสงค์ได้อย่างเหมาะสม
- 2) วัตถุประสงค์ (Objective) คือ ส่วนที่อธิบายเป้าหมายทางธุรกิจหรือการทำงานหลักของยูสเคสนั้นๆ
- 3) ผู้กระทำ คือ คน สิ่งของ หรือระบบภายนอก ที่เกี่ยวข้องหรือมีบทบาทกับการทำงานของยูสเคส
- 4) ความสัมพันธ์ คือ ส่วนที่อธิบายความสัมพันธ์ระหว่างส่วนต่างๆ ภายในแผนภาพ ซึ่งความสัมพันธ์มีด้วยกัน 4 ประเภท ได้แก่ ความสัมพันธ์แบบแอสโซซิเอชัน (Association Relationship) ซึ่งใช้บอกความสัมพันธ์ระหว่างผู้กระทำกับยูสเคส ความสัมพันธ์แบบเจเนอรัลไลเซชัน (Generalization Relationship) ใช้บอกความสัมพันธ์ระหว่างผู้กระทำกับผู้กระทำ และความสัมพันธ์ระหว่างยูสเคสกับยูสเคส ซึ่งสามารถแบ่งออกเป็น ความสัมพันธ์แบบรวม (Include Relationship) และความสัมพันธ์แบบขยาย (Extend Relationship)
- 5) เงื่อนไขก่อนทำงาน (Precondition) คือ ส่วนที่กำหนดว่าทุกเงื่อนไขต่อไปนี้ต้องเป็นจริงก่อนที่ยูสเคสนี้จะถูกกระทำหรือเริ่มทำงาน
- 6) เงื่อนไขหลังทำงาน (Postcondition) คือ เงื่อนไขหรือสถานะของระบบที่ต้องเป็นจริงเสมอหลังจบการทำงานของยูสเคส

7) เหตุการณ์ปกติ (Normal Flow of Events) คือ การอธิบายถึงชุดเหตุการณ์หลักที่เกิดขึ้นในการปฏิบัติงานของยูสเคสหนึ่ง ๆ ของระบบตามลำดับ

8) สายงานย่อย (Subflow) คือ คำอธิบายลำดับเหตุการณ์ย่อยที่อาจเกิดขึ้นภายในเหตุการณ์ปกติ

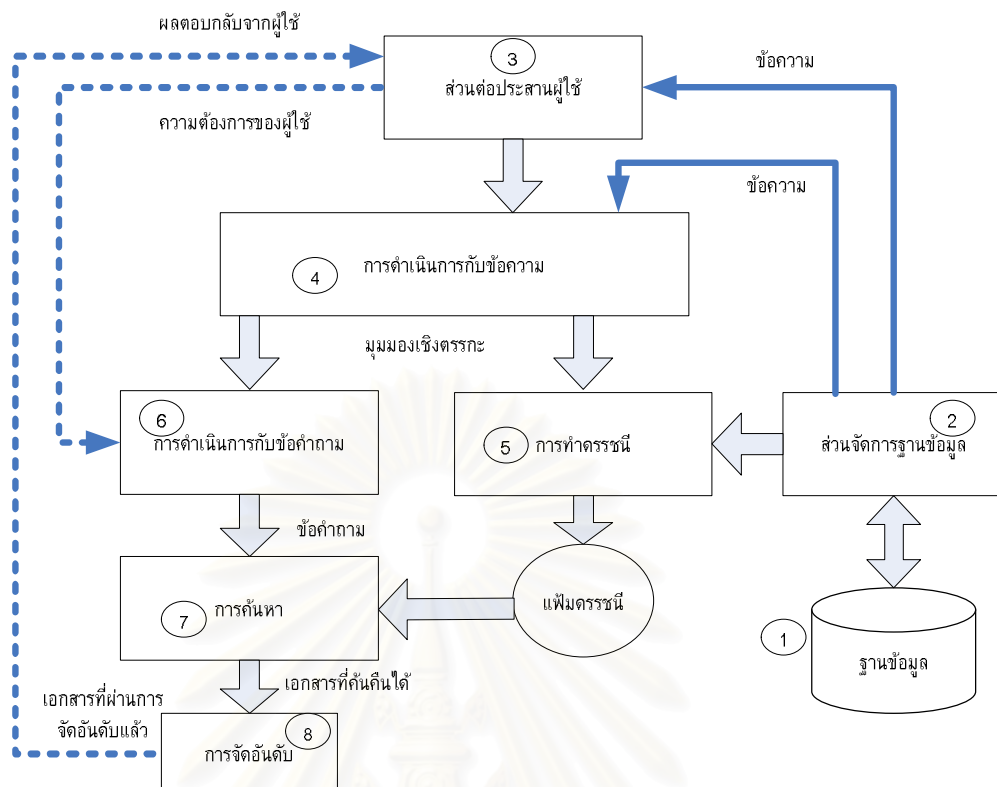
9) เหตุการณ์ทางเลือกหรือเหตุการณ์พิเศษ (Alternative or Exceptional Flow of Events) คือ คำอธิบายเหตุการณ์ทางเลือกหรือเหตุการณ์พิเศษ ซึ่งเป็นขั้นตอนการทำงานที่ต่างจากลำดับเหตุการณ์ปกติ โดยจะทำงานเมื่อเกิดกรณีที่เป็นไปตามเงื่อนไขบางอย่างของลำดับเหตุการณ์ปกติ

2.1.2 ระบบการจัดเก็บและค้นคืนสารสนเทศ (Information Storage and Retrieval System)

การจัดเก็บและค้นคืนสารสนเทศ [10] เป็นศาสตร์สาขาหนึ่งของวิทยาศาสตร์คอมพิวเตอร์ ซึ่งจัดการเกี่ยวกับการจัดเก็บและค้นคืนเอกสารแบบอัตโนมัติ โดยจะขึ้นอยู่กับเนื้อหาสาระและคำที่ปรากฏในเอกสารเหล่านั้น เนื่องจากในปัจจุบันสารสนเทศต่าง ๆ มีหลายประเภท และมีจำนวนมาก ทำให้ยากต่อการค้นหา การจัดเก็บและค้นคืนสารสนเทศ จึงเป็นงานสำคัญที่ถูกนำมาใช้ เพื่อให้สามารถค้นคืนเอกสารหรือสารสนเทศที่ตรงกับความต้องการของผู้ใช้มากที่สุด และช่วยให้ผู้ใช้สามารถค้นหาข้อมูลที่ตรงกับความต้องการได้สะดวกรวดเร็วขึ้น โดยผู้ใช้จะค้นหาข้อมูลด้วยการใส่ข้อความ (Query or Search Term) ซึ่งอธิบายถึงข้อมูลที่ผู้ใช้ต้องการค้นหา จากนั้นระบบจะทำการค้นคืนสารสนเทศจำนวนหนึ่งออกมา แต่ผลลัพธ์ที่ได้จะตรงกับความต้องการ หรือเป็นที่พอใจของผู้ใช้หรือไม่นั้น ขึ้นอยู่กับความสามารถในการออกแบบ และพัฒนาระบบการจัดเก็บและค้นคืนสารสนเทศด้วยเช่นกัน กระบวนการในระบบการจัดเก็บและค้นคืนสารสนเทศ แสดงดังรูปที่ 2.4

รายละเอียดของแต่ละส่วนประกอบในกระบวนการจัดเก็บและค้นคืนสารสนเทศ ได้แก่

- 1) ฐานข้อมูล (Database) คือ ส่วนจัดเก็บชุดข้อมูลหรือเอกสารทั้งหมดในระบบ
- 2) ส่วนจัดการฐานข้อมูล (Database Management Module) คือ ตัวจัดการงานต่าง ๆ ที่เกี่ยวข้องกับฐานข้อมูล
- 3) ส่วนต่อประสานผู้ใช้ (User Interface) คือ ส่วนที่เชื่อมต่อระหว่างผู้ใช้กับระบบการจัดเก็บและค้นคืนสารสนเทศ เพื่อให้ผู้ใช้ใส่ข้อความเข้าสู่ระบบ และเป็นส่วนแสดงรายการของเอกสารที่ค้นคืนได้จากระบบ นอกจากนี้ยังเป็นส่วนที่อำนวยความสะดวกให้แก่เจ้าหน้าที่ควบคุมระบบในการใส่ข้อมูล หรือเอกสารต่าง ๆ ไปเก็บไว้ในฐานข้อมูลอีกด้วย
- 4) การดำเนินการกับข้อความ (Text Operation) คือ ขั้นตอนในการจัดการกับข้อความ (Text) ซึ่งเป็นส่วนของการเตรียมเอกสารก่อนนำไปทำดรรชนี เช่น การตัดคำ (Word Segmentation) การกำจัดคำยกเว้น (Elimination of Stop Words) การลดรูปคำศัพท์ให้มีรากศัพท์เดียวกัน (Stemming) หรือการทำอรรถาภิธาน (Thesaurus) เป็นต้น



รูปที่ 2.4 กระบวนการในระบบการจัดเก็บและค้นคืนสารสนเทศ [10]

5) การทำดัชนี (Indexing) การทำดัชนีให้กับคำที่ปรากฏในเอกสารต่าง ๆ เริ่มต้นจากข้อความในเอกสารหนึ่ง ๆ ผ่านการดำเนินการกับข้อความเรียบร้อยแล้ว ได้คำต่าง ๆ (Keywords) ออกมา และจัดเก็บตามโครงสร้างที่ได้ออกแบบไว้ การทำดัชนีในปัจจุบันมีโครงสร้างหลายแบบ เช่น การใช้แฟ้มผกผัน (Inverted File) เป็นโครงสร้างข้อมูลที่อยู่อาศัยตารางเป็นกลไกในการเข้าถึงข้อมูล การใช้แฟ้มลายเซ็นต์ (Signature File) เป็นการค้นหาแบบแฮช (Hashing) หรือการใช้ต้นไม้แพ็ต (Pat Tree) ซึ่งมีโครงสร้างแบบต้นไม้เชิงเลขทวิภาค (Binary Digital Tree) เป็นต้น

6) การดำเนินการกับข้อความ (Query Operation) คือ ขั้นตอนในการแปลงความต้องการของผู้ใช้ให้เป็นข้อความ ก่อนป้อนเข้าสู่ระบบการค้นคืน โดยจะต้องผ่านขั้นตอนคล้ายกันกับการดำเนินการกับข้อความ เพื่อให้ได้คำสำคัญ หรือคำมาใช้ในขั้นตอนการค้นหาต่อไป นอกจากนี้ ในขั้นตอนนี้ยังสามารถเพิ่มเติมส่วนงานย่อยที่สำคัญคือ การเปลี่ยนแปลงรูปแบบข้อความ (Query Reformulation) ซึ่งได้มาจากผลป้อนกลับที่ตรงประเด็นจากผู้กลับสู่ระบบว่าเอกสารที่ระบบค้นคืนให้ นั้น เอกสารใดบ้างที่ตรงกับความต้องการของตน งานในส่วนนี้ จะกระทำหลังจากที่ระบบค้นคืนเอกสารให้กับผู้ใช้ในครั้งแรกเสร็จสิ้นแล้ว

7) การค้นหา (Searching) คือ ขั้นตอนการค้นหาเอกสาร ด้วยเซตของดัชนีที่ได้จากข้อความ

8) การจัดลำดับ (Ranking) คือ ขั้นตอนการจัดอันดับเอกสารที่ค้นคืนได้จากฐานข้อมูลตามค่าความคล้ายระหว่างข้อความกับเอกสาร ก่อนที่ระบบจะแสดงรายการเอกสารเหล่านี้บางส่วนต่อประสานผู้ใช้ต่อไป

จากที่กล่าวมาข้างต้นนั้นเป็นภาพรวมของแต่ละส่วนของระบบการจัดเก็บและค้นคืนสารสนเทศ โดยแต่ละส่วนนั้นเป็นองค์ประกอบหนึ่งของกระบวนการหลักๆ 3 กระบวนการดังต่อไปนี้

1. การจัดเก็บสารสนเทศ (Information Storage)

การจัดเก็บสารสนเทศ เป็นกระบวนการในการนำเนื้อหาหรือข้อความในเอกสารที่ต้องการจัดเก็บเข้าสู่ฐานข้อมูล ผ่านการดำเนินการกับข้อความ การทำดรรชนี การกำหนดค่าน้ำหนัก และจัดเก็บตามโครงสร้างที่ผู้พัฒนาได้ออกแบบไว้ สำหรับการจัดเก็บสารสนเทศในสมัยใหม่จึงใช้การทำงานที่เป็นอัตโนมัติ ได้แก่ การทำดรรชนีอัตโนมัติ และการหาค่าน้ำหนักให้กับค่าแบบอัตโนมัติ ซึ่งเป็นการทำงานที่สำคัญ และมีรายละเอียดของการทำงานดังต่อไปนี้

1) การทำดรรชนีอัตโนมัติ (Automatic Indexing)

เมื่อเอกสารถูกส่งเข้าสู่ระบบผ่านทางส่วนต่อประสานผู้ใช้ เนื้อหาหรือข้อความในเอกสารนั้นจะถูกนำไปประมวลผลเพื่อทำดรรชนีอัตโนมัติ โดยมีขั้นตอนดังนี้

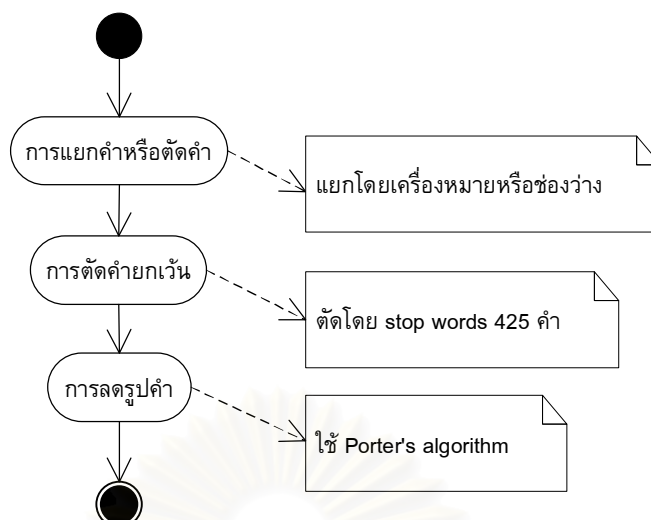
(1) การแยกคำหรือการตัดคำ (Word Segmentation) พิจารณาจากเครื่องหมายวรรคตอน (Punctuation Marks) หรือช่องว่างระหว่างคำ เป็นต้น

(2) การตัดคำบางคำที่ไม่มีประโยชน์ในการค้นคืนเอกสารทิ้งไป (Elimination of Stop words) เช่น คำกำกับนาม (Articles) คำบุพบท (Preposition) และคำสันธาน (Conjunction) เป็นต้น โดยคำที่นิยมใช้ในขั้นตอนนี้มีทั้งหมด 425 คำ [11] เนื่องจากเป็นคำที่ไม่สามารถสื่อความหมายไปยังเอกสารที่ผู้ใช้ต้องการได้ ขั้นตอนนี้จึงช่วยลดขนาดของแฟ้มดรรชนีได้

(3) การลดรูปคำ (Stemming) เป็นการตัดคำนำหน้า (Prefix) และคำต่อท้าย (Suffix) ออกจากคำสำคัญหรือคำที่เหลือเพียงรากคำศัพท์ (Root Word) โดยอัลกอริทึมที่นิยมใช้คืออัลกอริทึมของพอตเตอร์ (Porter's algorithm) [12]

(4) การกำจัดคำที่มีจำนวนคำหรือความถี่ในเอกสารสูง (Elimination High Frequency Term) เนื่องจากคำที่มีความถี่ของคำๆ นั้นมากเกินไป อาจไม่มีความสำคัญกับเอกสารนั้นๆ ก็เป็นไปได้

คำที่ไม่ถูกตัดทิ้งไปจึงถูกนำมาเป็นเซตของดรรชนี (Set of Indexes) ที่จะอ้างอิงไปยังเอกสารที่เกี่ยวข้องได้ โดยขั้นตอนทั้งหมดที่กล่าวมา แสดงดังรูปที่ 2.5



รูปที่ 2.5 แผนภาพกิจกรรมการทำดรรชนีอัตโนมัติ

2) การหาค่าน้ำหนักคำอัตโนมัติ (Automatic Term Weighting)

ในเซตของดรรชนีที่ประกอบด้วยคำหลายๆ คำ ต้องมีการหาค่าน้ำหนักของคำแต่ละคำ เพื่อเป็นการบอกถึงความสำคัญของคำๆ นั้น ในการเป็นตัวแทนของเอกสาร ในการหาค่าน้ำหนักหรือความสำคัญจากหนังสือ [10] กล่าวถึงวิธีการคำนวณไว้ 4 วิธี ได้แก่ ค่าน้ำหนักทั่วไป (General Consideration) ค่าน้ำหนักความถี่ของเอกสารแบบผกผัน (The Inverse Document Frequency Weight: IDF) ค่าอัตราส่วนของสัญญาณรบกวน (The Signal-Noise Ratio) และการให้ค่ากับคำสำคัญที่ทำให้เอกสารมีความแตกต่างกัน (The Term Discrimination Value)

สำหรับงานวิทยานิพนธ์นี้เลือกการคำนวณน้ำหนักด้วยความถี่ของเอกสารแบบผกผัน ซึ่งมีสมมติฐาน คือ ความสำคัญของคำ (Term Importance) เป็นสัดส่วนตามกับความถี่ของคำหรือคำที่ k ที่เกิดขึ้นใน เอกสารที่ i ($Freq_{ik}$) และเป็นสัดส่วนผกผันกับจำนวนเอกสารทั้งหมดที่มีคำหรือคำที่ k ปรากฏอยู่ ($TotFreq_k$) ดังนั้นค่าน้ำหนักหรือค่าความสำคัญ หาได้จากสูตรในสมการที่ 1

$$Weight_{ik} = \frac{Freq_{ik}}{TotFreq_k} \quad (1)$$

กำหนดให้ $Weight_{ik}$ คือ ค่าน้ำหนักของคำหรือคำ k ในเอกสาร i

$Freq_{ik}$ คือ ค่าความถี่หรือจำนวนครั้งที่พบคำหรือคำ k ในเอกสาร i

$TotFreq_k$ คือความถี่หรือจำนวนครั้งที่พบคำ k ปรากฏในชุดข้อมูล (Collection)

2. การค้นคืนสารสนเทศ (Information Retrieval)

การค้นคืนสารสนเทศเป็นขั้นตอนในการค้นหาสารสนเทศให้มีความใกล้เคียงหรือเกี่ยวข้องกับความต้องการของผู้ทำการค้นคืนสารสนเทศมากที่สุดออกมา เริ่มจากการที่ผู้ค้นคืนใส่ความต้องการหรือข้อความผ่านส่วนต่อประสานผู้ใช้ จากนั้นจะทำการแปลงข้อความเป็นเซตตรรกะของข้อความ และถูกนำไปเปรียบเทียบความคล้ายกับเซตของตรรกะที่เป็นตัวแทนของเอกสาร สูตรที่ใช้ในการเปรียบเทียบความคล้ายกันระหว่างความต้องการกับเอกสารมีหลายสูตร โดยสูตรการคำนวณที่ใช้ในวิทยานิพนธ์นี้คือ ค่าสัมประสิทธิ์โคซายน์ (Cosine Coefficient) ดังสมการที่ 2

$$\text{Similarity}(DOC_i, Query_j) = \frac{\sum_{k=1}^t (Term_{ik} \cdot QTerm_{jk})}{\sqrt{\sum_{k=1}^t (Term_{ik})^2 \cdot \sum_{k=1}^t (QTerm_{jk})^2}} \quad (2)$$

กำหนดให้ $\text{Similarity}(DOC_i, Query_j)$ คือ ความคล้ายระหว่างเอกสารที่ i กับข้อความที่ j

$Term_{ik}$ คือ ค่าน้ำหนักของคำ k ในเอกสารที่ i

$QTerm_{jk}$ คือ ค่าน้ำหนักของคำ k ในข้อความที่ j

3. การประเมินประสิทธิผลของระบบการค้นคืนสารสนเทศ (Evaluation of Information Retrieval Effectiveness)

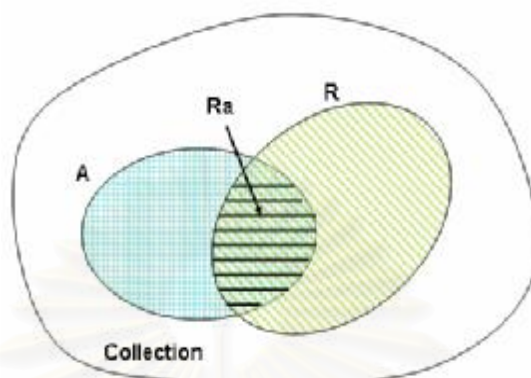
การวัดหรือประเมินประสิทธิผลของระบบโดยทั่วไปนั้น ส่วนใหญ่วัดกันที่ระยะเวลาในการประมวลผลและการใช้เนื้อที่หน่วยความจำน้อยกว่า ถือว่าเป็นระบบที่มีประสิทธิผลดีกว่า แต่หากระบบนั้นถูกออกแบบสำหรับการค้นคืนเอกสาร การวัดประสิทธิผลของการค้นคืนเอกสารได้ตรงกับความต้องการหรือเป็นเอกสารที่ตรงประเด็นกับความต้องการของผู้ค้นคืนเอกสารนับว่ามีบทบาทมากกว่า โดยการประเมินผลการค้นคืนเอกสารของระบบนั้น นิยมใช้วิธีการวัดด้วยค่าเรียกคืน (Recall) และค่าความแม่นยำ (Precision) [11] กล่าวคือ ในการวัดค่าเรียกคืน จะคิดสัดส่วนของเอกสารที่ตรงประเด็นที่ค้นคืนมาได้ต่อเอกสารที่ตรงประเด็นทั้งหมดที่เก็บไว้ในฐานข้อมูล ส่วนค่าความแม่นยำ คิดสัดส่วนของเอกสารที่ตรงประเด็นที่ค้นคืนมาได้ต่อเอกสารที่ค้นคืนออกมาทั้งหมด แสดงดังสมการที่ 3 และ 4 ตามลำดับ หรือแสดงด้วยแผนภาพในรูปที่

2.6

$$\text{ค่าเรียกคืน (Recall)} = \frac{|Ra|}{|R|} \quad (3)$$

$$\text{ค่าความแม่นยำ (Precision)} = \frac{|Ra|}{|A|} \quad (4)$$

กำหนดให้	Ra	คือ จำนวนเอกสารที่ตรงประเด็นที่ค้นคืนออกมาได้
	R	คือ จำนวนเอกสารที่ตรงประเด็นทั้งหมดในฐานข้อมูล
	A	คือ จำนวนเอกสารทั้งหมดที่ค้นคืนออกมา



รูปที่ 2.6 ค่าเรียกคืนและค่าความแม่นยำจากข้อความหนึ่ง ๆ

นอกจากการประเมินผลด้วยมาตรวัดทั้ง 2 ดังกล่าวแล้ว มีมาตรวัดอีกวิธีหนึ่งที่ใช้ในงานวิทยานิพนธ์นี้ ได้แก่ ค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) โดยค่าเฉลี่ยฮาร์โมนิก คือ มาตรวัดรวมค่าเรียกคืน และค่าความแม่นยำ โดยมีค่าอยู่ในช่วง 0 ถึง 1 ค่าเฉลี่ยฮาร์โมนิกจะมีค่าเป็น 0 เมื่อเอกสารที่ค้นคืนมาได้นั้น ไม่มีเอกสารใดเกี่ยวข้อง หรือตรงประเด็นเลย และจะมีค่าเป็น 1 เมื่อทุกเอกสารในที่ค้นคืนมาได้เป็นเอกสารที่ตรงกับความต้องการทั้งหมด สามารถคำนวณได้ดังสูตรในสมการที่ 5

$$F(j) = \frac{2}{\frac{1}{r(j)} + \frac{1}{p(j)}} \quad (5)$$

กำหนดให้	F(j)	คือ ค่าเฉลี่ยฮาร์โมนิกของ r(j) และ p(j)
	r(j)	คือ ค่าเรียกคืนของเอกสารที่ j
	p(j)	คือ ค่าความแม่นยำของเอกสารที่ j

2.1.3 การจัดกลุ่มเอกสาร (Document Clustering)

การจัดกลุ่มเอกสาร [4] เป็นการดำเนินการจัดกลุ่มให้กับเอกสารที่มีความคล้าย หรือเกี่ยวข้องกัน โดยทั่วไปใช้การเปรียบเทียบความคล้ายกันของเอกสารที่ละคู่ ซึ่งวิธีการนี้จะทำให้ได้แฟ้มเอกสารที่ถูกจัดเป็นกลุ่ม ๆ (Clustered Document File) เอกสารที่อยู่ในกลุ่มเดียวกันก็มักจะมีเนื้อหา หรือใจความสำคัญในเอกสารนั้น ๆ คล้ายคลึงกัน นอกจากนี้เอกสารหนึ่ง ๆ อาจปรากฏอยู่ในกลุ่มได้มากกว่าหนึ่งกลุ่มอีกด้วย

โดยในส่วนของวิธีการ หรือแนวทางในการจัดกลุ่มให้กับเอกสารนั้นมีหลายวิธีด้วยกัน ขึ้นอยู่กับการออกแบบของผู้พัฒนาระบบ เช่น ใช้วิธีการเปรียบเทียบความคล้ายของเอกสารทีละคู่ และใช้เกณฑ์ในการจัดกลุ่มตามค่าขีดแบ่ง (Threshold Value) ที่กำหนดไว้ เป็นต้น ส่วนฟังก์ชันที่นิยมใช้ในการเปรียบเทียบความคล้ายกันของเอกสาร ได้แก่ การเปรียบเทียบความคล้ายด้วยค่าสัมประสิทธิ์โคซายน์ เนื่องจากมีช่วงของค่าความคล้ายอยู่ระหว่าง 0 ถึง 1 เหมาะแก่การกำหนดค่าขีดแบ่ง ในการเปรียบเทียบความคล้ายของเอกสารแต่ละคู่ซึ่งอยู่ในรูปแบบเวกเตอร์ด้วยค่าสัมประสิทธิ์โคซายน์ มีสูตรการคำนวณดังสมการที่ 6

$$\text{Cos}(DOC_i, DOC_j) = \frac{\sum_{k=1}^t (\text{Term}_{ik} \cdot \text{Term}_{jk})}{\sqrt{\sum_{k=1}^t (\text{Term}_{ik})^2 \cdot \sum_{k=1}^t (\text{Term}_{jk})^2}} \quad (6)$$

กำหนดให้ $\text{Cos}(DOC_i, DOC_j)$ คือ ค่าความคล้ายที่วัดด้วยสัมประสิทธิ์โคซายน์ระหว่างเอกสารที่ i และเอกสารที่ j

Term_{ik} คือ ค่าน้ำหนักหรือค่าความสำคัญของคำ k ในเอกสารที่ i

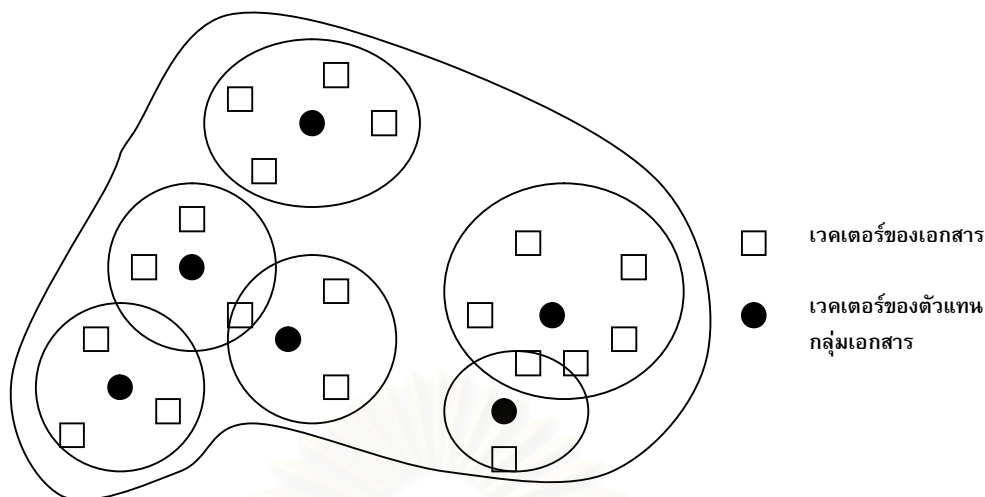
Term_{jk} คือ ค่าน้ำหนักหรือค่าความสำคัญของคำ k ในเอกสารที่ j

หลังจากนั้นเอกสารทั้งหมดในชุดข้อมูลจะถูกจัดเป็นกลุ่มๆ ดังรูปที่ 2.7 ซึ่งแสดงตัวอย่างชุดข้อมูลที่มีการจัดเอกสารเป็นกลุ่มแล้ว จากตัวอย่างมีเอกสารถูกแบ่งกลุ่มได้ 6 กลุ่ม เอกสารแต่ละเอกสารแทนด้วยรูปสี่เหลี่ยม กระจายอยู่ในวงกลมหรือวงรีซึ่งเป็นขอบเขตของแต่ละกลุ่ม และมีวงกลมที่ขนาดเล็กแทนเอกสารที่เป็นตัวแทนของแต่ละกลุ่ม (Centroid) ซึ่งมักจะอยู่ตรงกลางภายในวงกลมกลุ่มของตนเอง จะเห็นว่ากลุ่มของเอกสารสามารถซ้อนทับ (Overlap) กันได้ และมีเอกสารบางเอกสารที่เป็นสมาชิก หรือปรากฏอยู่ในกลุ่มมากกว่าหนึ่งกลุ่ม อันเนื่องมาจากการกำหนดช่วงของค่าขีดแบ่งว่าเอกสารจะถูกจัดให้อยู่ในกลุ่มเดียวกันจะมีค่าความคล้ายกันอยู่ในช่วงใด เอกสารทั่วไปและเอกสารตัวแทนกลุ่มนั้นแสดงอยู่ในรูปแบบของเวกเตอร์ ซึ่งมีรูปแบบดังนี้

รูปแบบของเอกสารแบบเวกเตอร์

$$DOC_i = (\text{Term}_{i1}, \text{Term}_{i2}, \dots, \text{Term}_{in})$$

กำหนดให้ Term_{ik} คือ ค่าน้ำหนักหรือค่าความสำคัญของคำ k ในเอกสารที่ i



รูปที่ 2.7 ตัวอย่างชุดข้อมูลที่มีการจัดกลุ่มเอกสาร

เวกเตอร์ของเอกสารตัวแทนกลุ่ม (Centroid Vector) ซึ่งเป็นเวกเตอร์โดยเฉลี่ยของเวกเตอร์เอกสารในกลุ่ม ได้แก่

$$CENTROID_p = (CTerm_{p1}, CTerm_{p2}, \dots, CTerm_{pt})$$

กำหนดให้ $CENTROID_p$ คือ เวกเตอร์ของตัวแทนเอกสารในกลุ่ม p

สูตรที่ใช้ในการคำนวณเพื่อหาค่าน้ำหนักของคำในเอกสารที่เป็นตัวแทนของแต่ละกลุ่ม ดังสมการที่ 7

$$CTerm_{pk} = \frac{1}{m} \sum_{i=1}^m Term_{ik} \quad (7)$$

กำหนดให้ $CTerm_{pk}$ คือ ค่าน้ำหนักของคำ k เฉลี่ยในเอกสารของกลุ่ม p

m คือ จำนวนเอกสารทั้งหมดในกลุ่ม p

$Term_{ik}$ คือ ค่าน้ำหนักของคำ k ในเอกสารที่ i

ด้วยเหตุนี้ ในการค้นหาเอกสารด้วยข้อความหนึ่งๆ จะมีขั้นตอนการทำงาน 2 ขั้นตอนด้วยกัน ได้แก่

1) การเปรียบเทียบความคล้ายกันระหว่างข้อความกับเวกเตอร์เอกสารตัวแทนกลุ่มทุกกลุ่ม ซึ่งจะพบเอกสารตัวแทนกลุ่มที่มีค่าความคล้ายสูงพอค่าขีดแบ่งเริ่มต้นกับข้อความ

2) เมื่อทราบว่าเวกเตอร์เอกสารตัวแทนตัวใดมีค่าความคล้ายกับข้อความแล้ว จะทราบกลุ่มของเอกสาร เพื่อที่จะนำข้อความดังกล่าวมาเปรียบเทียบกับเอกสารแต่ละเอกสารภายในกลุ่มนั้นๆ

ส่วนการจัดกลุ่มให้กับเอกสารใหม่ที่เข้ามาในระบบก็สามารถจัดกลุ่มได้โดยง่ายซึ่งมีวิธีการเช่นเดียวกับการใช้ข้อความในการค้นคืนเอกสาร นั่นคือ การนำเอกสารที่เข้ามาใหม่เปรียบเทียบกับเอกสารตัวแทนกลุ่มแต่ละกลุ่มจนครบทุกกลุ่ม เมื่อพบว่าเอกสารใหม่นั้นมีค่าความคล้ายสูงมากพอกับเอกสารกลุ่มใด จึงทำการเพิ่มเอกสารนั้นในกลุ่มดังกล่าวได้

ดังนั้นจึงนับได้ว่า การจัดกลุ่มเอกสาร และหาตัวแทนเอกสารของแต่ละกลุ่ม เป็นประโยชน์อย่างยิ่งในการช่วยลดระยะเวลาในการค้นหาเอกสารหนึ่งๆ เนื่องจากไม่ต้องนำข้อความมาเปรียบเทียบกับเอกสารทุกๆ เอกสาร

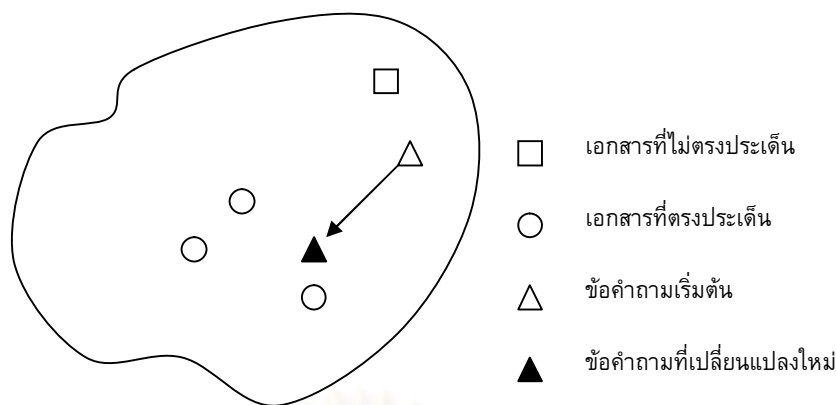
2.1.4 ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน (User Relevance Feedback)

การให้ผลป้อนกลับที่ตรงประเด็น (Relevance Feedback) เป็นขั้นตอนย่อยในขั้นตอนการดำเนินการกับข้อความ (Query Operation) โดยการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน [13] เป็นวิธีที่นิยมมาก ในการเปลี่ยนแปลงรูปแบบข้อความใหม่ (Query Reformulation) ในการให้ผลป้อนกลับนั้น ผู้ใช้จะเห็นรายการเอกสารที่ค้นคืนมาจากข้อความที่ป้อนเข้าสู่ระบบในครั้งแรก หลังจากที่ผู้ใช้ตรวจสอบเอกสารเหล่านั้นแล้ว จึงเลือกเอกสารที่เห็นว่าตรงประเด็นกับความต้องการของตน จากนั้นระบบจะนำข้อมูลที่ได้กลับมาจากผู้ใช้งาน ไปดำเนินการสร้างข้อความใหม่ เพื่อป้อนกลับสู่ระบบซ้ำอีกครั้ง ซึ่งในการค้นหาซ้ำนี้ ข้อความใหม่ที่ได้อาจนำไปใช้เปรียบเทียบกับเอกสารแทนข้อความเริ่มต้น จุดมุ่งหมายสำคัญของการสร้างข้อความขึ้นมาใหม่นี้คือ เพื่อแสดงให้เห็นว่า ข้อความใหม่มีค่าความคล้าย (Similarity) กับเอกสารที่ผู้ใช้ระบุว่าตรงประเด็น (Relevant) มากกว่าเอกสารที่ได้จากข้อความเริ่มต้น ในขณะที่เดียวกัน ข้อความใหม่ก็ถูกคาดหวังว่าจะมีความคล้ายกับเอกสารที่ไม่ถูกผู้ใช้ระบุว่าตรงประเด็น (Non-Relevant) น้อยกว่าเอกสารที่ได้จากข้อความเริ่มต้นเช่นกัน

ในกระบวนการเปลี่ยนแปลงรูปแบบข้อความใหม่นี้อาศัยพื้นฐานการดำเนินการได้แก่

- 1) การพิจารณาเพื่อเพิ่มคำที่ปรากฏในเอกสารที่ผู้ใช้ระบุว่าตรงประเด็นลงในเวกเตอร์ของข้อความเริ่มต้น หรือเพิ่มค่าน้ำหนักของคำเหล่านั้น
- 2) การพิจารณาเพื่อลบคำที่ปรากฏในเอกสารที่ไม่ถูกผู้ใช้ระบุว่าตรงประเด็นออกจากข้อความเริ่มต้น หรือลดค่าน้ำหนักคำนั้นๆ

โดยสรุปคือ การเปลี่ยนแปลงรูปแบบข้อความใหม่เพื่อการค้นคืนแบบให้ผลป้อนกลับที่ตรงประเด็น ใช้การเปลี่ยนแปลงคำในข้อความกับการเปลี่ยนแปลงค่าน้ำหนักของคำในข้อความนั่นเอง ซึ่งการเปลี่ยนแปลงคำในข้อความ อาจช่วยให้เพิ่มเอกสารที่ตรงประเด็นมากขึ้น ส่วนการเปลี่ยนแปลงค่าน้ำหนักคำในข้อความ อาจช่วยให้มีการจัดอันดับเอกสารที่ดีขึ้น ดังรูปที่ 2.8 แสดงผลที่เกิดจากการดำเนินการแบบมีผลป้อนกลับที่ตรงประเด็น



รูปที่ 2.8 ผลที่เกิดจากการค้นคืนแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน

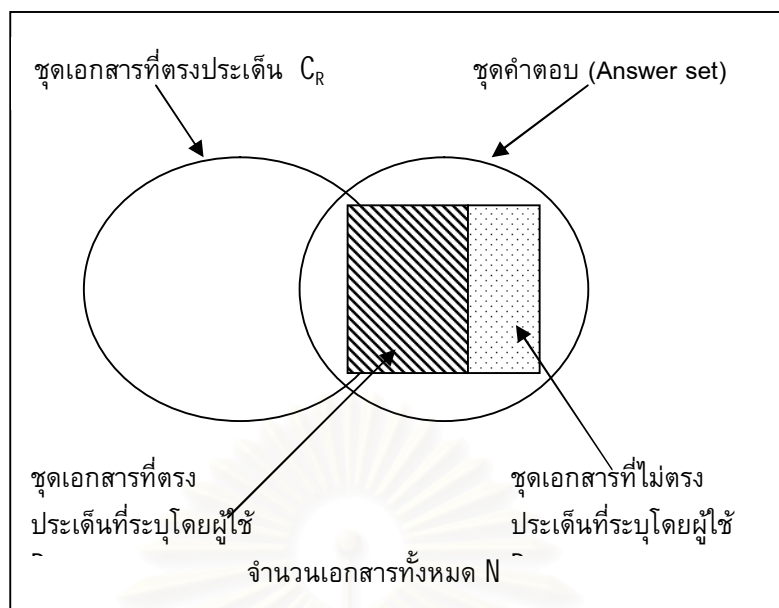
ในงานวิทยานิพนธ์นี้จะทำการทดลองโดยใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน 2 วิธีการด้วยกันได้แก่

- 1) การเปลี่ยนแปลงข้อความ หรือการขยายคำในข้อความ ร่วมกับการเปลี่ยนแปลงค่าน้ำหนักคำในข้อความ ด้วยแบบจำลองปริภูมิเวกเตอร์ (Query Expansion and Term Reweighting for the Vector Space Model)
- 2) การเปลี่ยนแปลงค่าน้ำหนักของคำในข้อความใหม่ ด้วยแบบจำลองความน่าจะเป็น (Term Reweighting for the Probabilistic Model)

(1) การเปลี่ยนแปลงข้อความ หรือการขยายคำในข้อความ และการเปลี่ยนแปลงน้ำหนักคำในข้อความ ด้วยแบบจำลองปริภูมิเวกเตอร์

การเปลี่ยนแปลงข้อความ และการเปลี่ยนแปลงน้ำหนักคำ สำหรับแบบจำลองปริภูมิเวกเตอร์ มีข้อสมมติฐานว่า เอกสารที่ผู้ใช้ระบุว่ามีความเกี่ยวข้องหรือตรงประเด็นกับข้อความ (Relevant Documents) จะมีเวกเตอร์น้ำหนักของคำหนึ่งๆ คล้ายหรือใกล้เคียงกัน และเอกสารที่ไม่ถูกระบุว่ามีความเกี่ยวข้องหรือตรงประเด็นนั้น (Non-Relevant Documents) จะมีเวกเตอร์น้ำหนักของคำหนึ่งๆ แตกต่างจากเอกสารที่ถูกระบุว่าตรงประเด็น ข้อความใหม่ที่เปลี่ยนแปลงได้นั้นจึงมีความคล้าย หรืออยู่ใกล้ชิดกับเอกสารที่ถูกระบุว่าตรงประเด็นนั่นเอง ชุดของเอกสารที่ใช้ในแบบจำลองปริภูมิเวกเตอร์ มีนิยามความหมายดังรูปที่ 2.9

จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.9 แผนภาพแสดงชุดของเอกสารที่ใช้ในแบบจำลองปริภูมิเวกเตอร์

สูตรการคำนวณเพื่อเปลี่ยนแปลงรูปแบบข้อความใหม่ เป็นดังสมการที่ 8

$$Q' = \alpha Q + \beta \frac{1}{R'} \sum_{i \in D_{R'}} DOC_i - \gamma \frac{1}{N'} \sum_{i \in D_{N'}} DOC_i \quad (8)$$

กำหนดให้ Q' คือ ข้อความที่เปลี่ยนแปลงรูปแบบใหม่
 Q คือ ข้อความเริ่มต้น
 R' คือ จำนวนเอกสารที่ตรงประเด็น
 N' คือ จำนวนเอกสารที่ไม่ตรงประเด็น
 DOC_i คือ เวกเตอร์ของเอกสารที่ i
 $D_{R'}$ คือ ชุดของเอกสารที่ตรงประเด็นในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด
 $D_{N'}$ คือ ชุดของเอกสารที่ไม่ตรงประเด็นในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด
 a, b, g คือ ค่าคงที่สำหรับการปรับค่ามีค่าอยู่ระหว่าง 0 ถึง 1 โดยงานวิจัยนี้

กำหนดให้ $a = 1$, $b = 0.5$ และ $g = 0$

การกำหนด $a = 1$ เพื่อกำหนดค่าน้ำหนักกับค่าในข้อความเดิมให้คงเดิมไว้ $b = 0.5$ เพื่อกำหนดว่าจะนำค่าน้ำหนักในเอกสารที่ผู้ใช้ระบุว่าตรงกับความต้องการ มาใช้ในการเปลี่ยนแปลงเป็นข้อความใหม่ ในอัตราส่วนที่พอดี ไม่มากหรือน้อยเกินไปคือ 0.5 และ $g = 0$ เพราะจะไม่พิจารณาค่าน้ำหนักในเอกสารที่ผู้ใช้ไม่ได้ระบุว่าตรงกับความต้องการมาเปลี่ยนแปลงในข้อความใหม่ ซึ่งจะมีผลทำให้ข้อความใหม่ตรงกับข้อความเดิมน้อยลง

(2) การเปลี่ยนแปลงค่าน้ำหนักคำ ด้วยแบบจำลองความน่าจะเป็น

การเปลี่ยนแปลงค่าน้ำหนักของคำในแบบจำลองความน่าจะเป็น [10] ซึ่งเป็นแบบจำลองที่มีแนวคิดในการค้นคืนเอกสารโดยพิจารณาความน่าจะเป็นที่เอกสารชิ้นหนึ่งมีความเกี่ยวข้องหรือตรงประเด็นกับข้อความ โดยใช้กฎการจัดอันดับตามความน่าจะเป็น (Probabilistic Ranking Principle) ในการคำนวณความคล้ายกันของเอกสาร d_j กับข้อความ q ไปพร้อมๆ กับการเปลี่ยนแปลงค่าน้ำหนักคำในข้อความ ดังนี้

$$\text{sim}(d_j, q) = \sum_{i=1}^t w_{i,q} \times w_{i,j} \times \log \frac{P(k_i | R)}{1 - P(k_i | R)} + \log \frac{1 - P(k_i | R)}{P(k_i | R)} \quad (9)$$

กำหนดให้ $P(k_i | R)$ คือ ความน่าจะเป็นที่จะพบคำ k_i ในเซตของเอกสารที่ตรงประเด็น (R)
 $P(k_i | \bar{R})$ คือ ความน่าจะเป็นที่จะพบคำ k_i ในเซตของเอกสารที่ไม่ตรงประเด็น (\bar{R})

$w_{i,j}$ คือ ค่าน้ำหนักของคำ k_i ในเอกสาร d_j ซึ่ง $w_{i,j} \in \{0,1\}$

$w_{i,q}$ คือ ค่าน้ำหนักของคำ k_i ในข้อความ q ซึ่ง $w_{i,q} \in \{0,1\}$

โดยสูตรในสมการที่ 9 นั้น ยังไม่สามารถใช้ได้ทันทีเนื่องจากยังไม่ทราบค่า $P(k_i | R)$ และ $P(k_i | \bar{R})$ ทำให้ไม่สามารถค้นคืนเอกสารใดๆ ออกมาได้ จึงกำหนดให้การค้นหาเอกสารในครั้งแรก (Initial Search) มีสมมุติฐานดังนี้

1) $P(k_i | R)$ เป็นค่าคงที่สำหรับทุกคำ k_i โดยปกติมีค่าตั้งแต่ 0 ถึง 0.5

2) $P(k_i | \bar{R})$ กำหนดให้การกระจายความน่าจะเป็นของคำในเอกสารที่ไม่ตรงประเด็นมีค่าเท่ากับการกระจายความน่าจะเป็นของคำในเอกสารทั้งหมด

ผลของสมมุติฐานทั้งสองที่ได้คือ

$$P(k_i | R) = 0.5$$

$$P(k_i | \bar{R}) = \frac{n_i}{N}$$

กำหนดให้ n_i คือ จำนวนเอกสารในชุดข้อมูลทั้งหมด ที่มีคำ k_i ปรากฏอยู่

N คือ เอกสารทั้งหมดในชุดข้อมูล

แทนค่าที่ได้ในสมการที่ 9 จะได้สูตรใหม่คือ สมการที่ 10 ดังนี้

$$sim_{initial}(d_j, q) = \sum_i^t w_{i,q} \times w_{i,j} \times \log \frac{N - n_i}{n_i} \quad (10)$$

ในการค้นหาครั้งถัดมาจะมีค่าความถี่สะสมของเอกสารที่ตรงประเด็นและไม่ตรงประเด็นของการค้นคืนก่อนหน้าที่เกิดขึ้น ซึ่งจะถูกนำมาใช้ประเมินค่าความน่าจะเป็นของ $P(k_i | R)$ และ $P(k_j | \bar{R})$ จะได้เป็น

$$P(k_i | R) = \frac{|D_{r,i}|}{|D_r|} \quad \text{และ} \quad P(k_i | \bar{R}) = \frac{n_i - |D_{r,i}|}{N - |D_r|}$$

กำหนดให้ D_r คือ ชุดของเอกสารที่ตรงประเด็นที่ค้นคืนได้ซึ่งระบุโดยผู้ค้นคืน
 $D_{r,i}$ คือ ชุดย่อยของ D_r ที่ประกอบด้วยเอกสารที่มีคำ k_i ปรากฏอยู่
 จากสมการที่ 9 เมื่อใช้ค่าประมาณนี้แล้ว จะได้สูตรเป็นดังสมการที่ 11 ดังนี้

$$sim(d_j, q) = \sum_{i=1}^t w_{i,q} \times w_{i,j} \times \log \frac{|D_{r,i}|}{|D_r| - |D_{r,i}|} \times \frac{N - |D_r| - (n_i - |D_{r,i}|)}{n_i - |D_{r,i}|} \quad (11)$$

หากเกิดกรณีที่พบว่าจำนวนเอกสารที่ตรงประเด็นซึ่งระบุโดยผู้ค้นคืนนั้น มีจำนวนน้อยมาก เช่นในกรณีที่ผู้ใช้ระบุเอกสารเพียง 1 ชิ้นเท่านั้นเป็นเอกสารที่ตรงประเด็น จะได้ว่า $D_r = 1$ และเอกสารชิ้นนั้นอาจจะมีคำ k_i ปรากฏอยู่ จะได้ว่า $D_{r,i} = 0$ เป็นผลให้ ไม่สามารถค้นคืนเอกสารใดๆ ออกมา ปัญหานี้สามารถหลีกเลี่ยงด้วยการปรับค่าตัวประกอบ (Factor) ของค่าความน่าจะเป็นของ $P(k_i | R)$ และ $P(k_j | \bar{R})$ เป็นดังนี้

$$P(k_i | R) = \frac{|D_{r,i}| + 0.5}{|D_r| + 1} \quad \text{และ} \quad P(k_i | \bar{R}) = \frac{n_i - |D_{r,i}| + 0.5}{N - |D_r| + 1}$$

อย่างไรก็ตาม การปรับในลักษณะนี้อาจจะทำให้ผลไม่เป็นที่น่าพอใจในบางกรณี จึงมีการปรับค่าตัวประกอบอีกวิธีหนึ่งซึ่งเสนอโดย [14] คือ แทนตัวประกอบ 0.5 ด้วย n_i/N หรือ $(n_i/|D_{r,i}|)/(N-|D_{r,i}|)$ จะได้เป็นสมการใหม่ดังนี้

$$P(k_i | R) = \frac{|D_{r,i}| + \frac{n_i}{N}}{|D_r| + 1} \quad \text{และ} \quad P(k_i | \bar{R}) = \frac{n_i - |D_{r,i}| + \frac{n_i}{N}}{N - |D_r| + 1}$$

โดยในส่วนของงานวิทยานิพนธ์ฉบับนี้จะนำวิธีการปรับค่าด้วยตัวประกอบนี้มาใช้ในสูตรเพื่อคำนวณหาความคล้ายของยูสเคสกับข้อความในแบบจำลองความน่าจะเป็น

ข้อดีของวิธีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งานด้วยแบบจำลองความน่าจะเป็นคือสามารถทำการเปลี่ยนแปลงค่าน้ำหนักให้กับข้อความโดยตรง นอกจากนี้เอกสารจะถูกเรียงลำดับความน่าจะเป็นของความเกี่ยวข้องกันระหว่างข้อความกับเอกสาร แต่ข้อเสียคือ ค่าน้ำหนักของคำในชุดข้อมูลไม่ได้ถูกนำมาใช้ รวมถึงไม่มีการขยายค่าให้กับข้อความในขั้นตอนการให้ผลป้อนกลับที่ตรงประเด็นด้วย

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 Use Case Retrieval Using Terms and Use Case Structure Similarity Computation โดย Akadej Udomchaiporn และคณะ 2006 [7]

งานวิจัยนี้ นำเสนอวิธีการนำยูสเคสกลับมาใช้ใหม่ โดยพิจารณาในส่วนของคำอธิบายยูสเคส ซึ่งมีเทมเพลตดังแสดงในรูปที่ 2.10 ด้วยการคำนวณค่าความคล้ายกันของคำ และโครงสร้างระหว่างข้อความที่มีลักษณะเหมือนกับยูสเคสที่เก็บไว้ในไลบรารีหรือฐานข้อมูล อีกทั้งยังทำการพัฒนาเครื่องมือที่ใช้ในการทดลองวิธีการที่เสนอนี้ด้วย ซึ่งเครื่องมือที่พัฒนาขึ้นมาใช้ในการดำเนินงาน 2 ขั้นตอนด้วยกัน ได้แก่ ขั้นตอนการจัดเก็บ และการค้นคืนยูสเคส ในส่วนของขั้นตอนการจัดเก็บสารสนเทศนั้น เอกสารที่เก็บไว้คือ ยูสเคส และทำการแปลงยูสเคสให้เป็นดรรชนีแบบอัตโนมัติ (Automatic Indexing) โดยมีการจัดทำดรรชนี และค่าน้ำหนักคำโดยใช้ค่าความถี่ของเอกสารแบบผกผัน ในขั้นตอนการค้นคืน หลังจากที่ผู้ใช้ใส่ข้อความเข้ามา ระบบจะแปลงข้อความนั้นเป็นดรรชนี แล้วคำนวณค่าความคล้ายกันระหว่างเซตของดรรชนีที่ได้จากข้อความกับดรรชนีของยูสเคสแต่ละตัวในฐานข้อมูล ผลลัพธ์ที่ได้ในขั้นตอนการคำนวณค่าความคล้ายนั้นคือ ค่าความคล้ายกันระหว่างส่วนประกอบแต่ละส่วนของข้อความเทียบกับส่วนประกอบแต่ละส่วนของยูสเคส โดยที่แต่ละส่วนประกอบของทั้งข้อความ และยูสเคสนั้น จะถูกกำหนดค่าน้ำหนักโดยผู้ใช้ ก่อนที่จะนำค่าน้ำหนักนั้น มาใช้ในการคำนวณเพื่อหาค่าความคล้ายกันดังกล่าว แล้วจึงแสดงยูสเคส 5 รายการแรกที่มีค่าคะแนนความคล้ายกับข้อความมากที่สุดออกมา ซึ่งมีการตั้งค่าขีดแบ่งของคะแนนความคล้ายไว้ล่วงหน้าด้วย ขั้นตอนสุดท้ายของงานวิจัยชิ้นนี้คือ ขั้นตอนการประเมินประสิทธิภาพของระบบด้วยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก

ในขั้นตอนของการค้นคืนยูสเคสนั้นงานวิจัยนี้ได้ใช้การค้นคืนที่ต่างไปจากการค้นคืนโดยทั่วไป นั่นคือมีการค้นคืนโดยอาศัยลักษณะที่เป็นโครงสร้างของทั้งข้อความ และตัวเอกสารยูสเคส จึงมีข้อดีในการลดเวลาในการค้นคืนให้น้อยลงได้ เพราะการพิจารณาแต่ละส่วนนั้นจะไม่ซ้อนทับกัน นอกจากนี้ยังเสนอค่าน้ำหนักเพื่อให้ความสำคัญของแต่ละส่วนประกอบ (Element) ของยูสเคสที่แตกต่างกัน ซึ่งค่าน้ำหนักเหล่านี้จะถูกระบุโดยผู้ใช้ ดังแสดงดังรูปที่ 2.10

A Use Case Description Template		
Use Case Name:	WE ₁ =[1-5]	
Objective:	WE ₂ =[1-5]	
Actor:	WE ₃ =[1-5]	
Relationship:	Association:	WE ₄₁ =[1-5]
	Include:	WE ₄₂ =[1-5]
	Extend:	WE ₄₃ =[1-5]
	Generalization:	WE ₄₄ =[1-5]
Precondition:	WE ₅ =[1-5]	
Postcondition:	WE ₆ =[1-5]	
Normal Flow of Events:	WE ₇ =[1-5]	
Subflow:	WE ₈ =[1-5]	
Alternative or Exceptional Flow of Events:	WE ₉ =[1-5]	

รูปที่ 2.10 แสดงค่าน้ำหนักที่เป็นไปได้ในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส

โดยมีสูตรในการคำนวณผลรวมความคล้ายระหว่างข้อความ และยูสเคสดังสมการที่ 12

$$Similarity(UC_i, Query_j) = \frac{\sum_{m=1}^e [Similarity(E_{mi}, E_{mj}) WE_m]}{TotalWeight} \quad (12)$$

กำหนดให้ $Similarity(UC_i, Query_j)$ คือ ค่าคะแนนความคล้ายระหว่างข้อความกับยูสเคส

e คือ จำนวนส่วนประกอบของยูสเคสที่ i และข้อความที่ j

WE_m คือ ค่าน้ำหนักของส่วนประกอบที่ m ซึ่งระบุโดยผู้ใช้

$TotalWeight$ คือ ค่าผลรวมน้ำหนักทั้งหมดที่ผู้ใช้ระบุ

$Similarity(E_{mi}, E_{mj})$ คือ ค่าความคล้ายกันระหว่างแต่ละส่วนประกอบของยูสเคส

ที่ i กับข้อความยูสเคสที่ j โดยคำนวณได้จากสมการที่ 13

$$Similarity(E_{mi}, E_{mj}) = \frac{2 \sum_{k=1}^t (Term_{mik} W_{ik}) Term_{mjk}}{\sum_{k=1}^t Term_{mik} + \sum_{k=1}^t Term_{mjk}} \quad (13)$$

- กำหนดให้ E_{mi} คือ ส่วนประกอบส่วนที่ m ของยูสเคสที่ i
 E_{mj} คือ ส่วนประกอบส่วนที่ m ของข้อความที่ j
 $Term_{mik}$ มีค่าเป็น 1 เมื่อคำ k ปรากฏพบในส่วนประกอบที่ m ของยูสเคสที่ i และเป็น 0 หากไม่พบ
 $Term_{mjk}$ มีค่าเป็น 1 เมื่อคำ k ปรากฏพบในส่วนประกอบที่ m ของยูสเคสที่ j และเป็น 0 หากไม่พบ
 W_{ik} คือ น้ำหนักของคำ k ในยูสเคสที่ i

งานวิจัยของ อัครเดช อุดมชัยพร และคณะ ได้เสนอเฉพาะการค้นคืนยูสเคสกลับมาใช้ใหม่โดยการใช้ข้อความเริ่มต้นเท่านั้น และมีการประเมินผลเอกสารที่ตรงประเด็น หลังจากค้นคืนได้ในครั้งแรกเลยทันที ดังนั้นในงานวิทยานิพนธ์นี้จะเพิ่มเติมงานในส่วนการจัดกลุ่มเอกสารยูสเคส และการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน เพื่อช่วยในการเพิ่มประสิทธิผลของระบบให้ดีขึ้น โดยจะแสดงตารางเปรียบเทียบระหว่างงานวิจัยของอัครเดช อุดมชัยพร และคณะกับงานวิทยานิพนธ์นี้ ดังตารางที่ 2.1

ตารางที่ 2.1 เปรียบเทียบวิธีการที่ใช้ในงานวิจัย

วิธีการ	งานวิจัยของ [7]	งานวิทยานิพนธ์นี้
การจัดกลุ่มเอกสารคำอธิบายยูสเคส	-	ü
การค้นคืนโดยพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสและการให้น้ำหนักแต่ละส่วนประกอบ	ü	ü
แบบจำลองในการค้นคืนเอกสาร		
1) แบบจำลองปริภูมิเวกเตอร์	ü	ü
2) แบบจำลองความน่าจะเป็น	-	ü
การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน	-	ü

2.2.2 Improving Relevance Feedback in the Vector Space Model โดย Carol Lundquist, David A. Grossman and Ophir Frieder [15]

ในงานวิจัยนี้เห็นความสำคัญของการให้ผลป้อนกลับที่ตรงประเด็นว่าสามารถเพิ่มประสิทธิผลของระบบค้นคืนสารสนเทศได้โดยศึกษางานเกี่ยวกับการให้ผลป้อนกลับที่ตรงประเด็นหลายงานด้วยกัน โดยงานวิจัยนี้กล่าวถึงเทคนิคต่างๆ ที่ช่วยเพิ่มประสิทธิผลของการให้ผลป้อนกลับที่ตรงประเด็น ซึ่งเป็นการให้ผลป้อนกลับที่ตรงประเด็นแบบเทียม (Pseudo Relevance Feedback) อันประกอบด้วยการใช้เอกสารที่ถูกจัดอยู่ในอันดับต้นๆ (Top-Ranked Documents) หรือคำที่ใช้ในการให้ผลป้อนกลับที่ตรงประเด็น การจัดกลุ่มเอกสารที่อยู่ในอันดับต้นๆ (Clustering the Top-Ranked Documents) การเปลี่ยนแปลงค่าคงที่ หรือสูตรในการคำนวณน้ำหนักคำ และการวัดน้ำหนักของคำที่ได้จากการป้อนกลับมาจากผู้ใช้ โดยระบบค้นคืนสารสนเทศของงานวิจัยนี้ใช้แบบจำลองปริภูมิเวกเตอร์ และฐานข้อมูลเชิงสัมพันธ์

ผลที่ได้จากการทดลองแต่ละเทคนิค วัดด้วยค่าความแม่นยำ และค่าเรียกคืน โดยการใช้ชุดข้อมูลทีปสเตอร์ (Tipster Document Collection) และใช้ข้อคำถามแบบสั้นของ TREC-5 และ TREC-4 สรุปผลได้ดังนี้

- 1) การเปลี่ยนแปลงคำในข้อคำถามด้วย 10 ถึง 20 คำใหม่ทั้งแบบเป็นคำ และวลี (Words and Phrases) ให้ผลดีกว่าการเปลี่ยนแปลงข้อคำถามด้วย 50 คำใหม่ และ 20 วลีใหม่
- 2) การใช้เอกสารที่ถูกจัดอันดับให้อยู่ในอันดับ 5 ถึง 20 เอกสารแรก ให้ผลดีกว่า การใช้จำนวนเอกสารน้อยกว่า หรือมากกว่านี้
- 3) การใช้ตัวประกอบในสูตรการคำนวณเพื่อเปลี่ยนแปลงคำ และน้ำหนักคำในข้อคำถาม ให้มีค่าระหว่าง 0.4 ถึง 0.6 จะให้ค่าความแม่นยำ และค่าเรียกคืน ดีที่สุด
- 4) การพิจารณาน้ำหนักของคำ พบว่าถ้ามีการให้ผลป้อนกลับที่ตรงประเด็น ผลการพิจารณาน้ำหนักคำด้วย $n \cdot idf$ จะดีกว่า idf แต่ถ้าไม่ใช้การให้ผลป้อนกลับที่ตรงประเด็นในระบบแล้ว ใช้ idf จะได้ผลดีกว่าการใช้ $n \cdot idf$
- 5) สำหรับการจัดกลุ่มให้เอกสารที่ถูกจัดอันดับอยู่ต้นๆ (Clustering Top-Ranked Documents) สามารถเพิ่มผลการค้นคืนได้ดีในบางข้อคำถาม

งานวิจัยชิ้นนี้ทำการทดลองกับเอกสารโดยทั่วไป และแสดงให้เห็นว่าการให้ผลป้อนกลับที่ตรงประเด็นโดยอาศัยแบบจำลองปริภูมิเวกเตอร์สามารถเพิ่มประสิทธิผลการค้นคืนได้ จึงเป็นงานวิจัยชิ้นหนึ่งที่สนับสนุนได้ว่า การนำการให้ผลป้อนกลับที่ตรงประเด็นเข้ามาใช้แม้ว่าจะเป็น การให้ผลป้อนกลับที่ตรงประเด็นแบบเทียม ก็สามารถเพิ่มประสิทธิผลให้กับการค้นคืนสารสนเทศได้ ดังนั้นหากนำผลที่ได้หรือแนวทางในการวิจัย มาประยุกต์ใช้กับการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้อาศัยแบบจำลองปริภูมิเวกเตอร์เช่นเดียวกัน ก็น่าที่จะให้ผลที่ดีกับการจัดเก็บและค้นคืนเอกสารในงานวิทยานิพนธ์นี้

2.2.3 A New Algorithm to Get the Initial Centroids โดย Fang Yuan, Zeng-Hui Meng, Hong-Xia Zhang and Chun-Ru Dong [16]

งานวิจัยนี้ตรวจสอบอัลกอริทึมในการจัดกลุ่มเอกสาร ด้วยค่าเฉลี่ย k มาตรฐาน (Standard k -mean Clustering Algorithm) และได้เสนอแนวทางใหม่เพื่อปรับปรุงอัลกอริทึม โดยได้พิจารณาการเลือกเอกสารที่จะมาเป็นตัวแทนกลุ่มของเอกสาร เพื่อใช้เป็นเอกสารเริ่มแรก (Initial Centroids) ก่อนใช้อัลกอริทึมในการจัดกลุ่ม ซึ่งเดิมนั้นใช้วิธีการสุ่มเลือกเอกสาร ออกมาจำนวน k เอกสารให้เป็นตัวแทนของแต่ละกลุ่ม ซึ่งการทดลองในงานวิจัยนี้พบว่า การสุ่มเลือกเอกสารนั้น ได้ค่าความแม่นยำไม่แน่นอน มากบ้างน้อยบ้าง ส่วนการเลือกตัวแทนกลุ่มของเอกสารโดยพิจารณาจากค่าความคล้ายกันนั้น สามารถให้ผลการทดลองที่ถูกต้องมากขึ้น

โดยงานวิจัยชิ้นนี้สามารถนำมาประยุกต์ใช้ในงานวิทยานิพนธ์ได้ในบางส่วนคือ การพิจารณาเลือกตัวแทนกลุ่มของเอกสารที่มีความเหมาะสม ก่อนการดำเนินการในการจัดกลุ่มเอกสาร

2.2.4 Advancing Candidate Link Generation for Requirements Tracing: The Study of Methods โดย Jane Huffman, Alex Dekhtyar and Senthil Karthikeyan Sundaram [17]

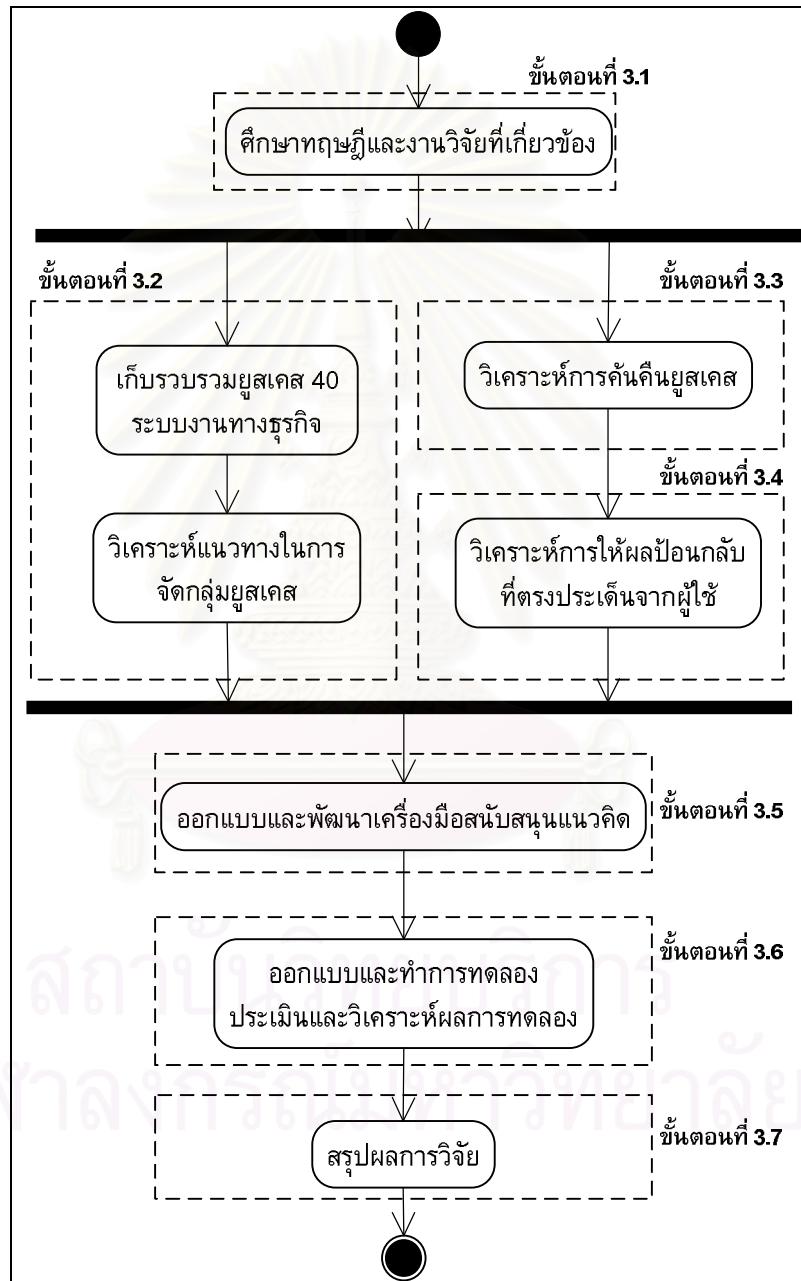
งานวิจัยนี้พิจารณาประสิทธิภาพของการค้นคืนสารสนเทศในการค้นหาเอกสารความต้องการที่เป็นข้อความ (Textual Requirements) แบบอัตโนมัติ โดยได้เสนอแนวทางที่ทำให้เห็นว่าวิธีการค้นคืนสารสนเทศจะมีทั้งประสิทธิภาพและประสิทธิผล เมื่อมีการใช้ลิงค์ลิสต์คู่แข่ง (Candidate Link Lists) ในกระบวนการติดตามความต้องการ และพัฒนาเครื่องมือที่ชื่อว่า RETRO (Requirements Tracing On-target) ซึ่งใช้หลักการของการค้นคืนสารสนเทศ เช่น การให้คำนำหน้าเอกสาร การหาค่าความคล้าย การใช้ตรรกะความหมายซ่อนเร้น (Latent Semantic Indexing) รวมทั้งการให้ผลป้อนกลับที่ตรงประเด็นด้วย นำมาประยุกต์ใช้เพื่อค้นคืนกับเอกสารความต้องการ 2 ชุดคือ MODIS และ CM-1 จากนั้นจึงประเมินผลด้วยค่าเรียกคืน ค่าความแม่นยำ และความเฉลี่ยฮาร์โมนิก และพบว่าการให้ผลป้อนกลับที่ตรงประเด็นสามารถปรับปรุงผลการติดตามความต้องการได้เป็นอย่างดี

ดังนั้นงานวิจัยชิ้นนี้จึงเป็นงานวิจัยที่สามารถสนับสนุนแนวทางในการทำวิทยานิพนธ์ได้ นั่นคือ การให้ผลป้อนกลับที่ตรงประเด็นสามารถช่วยปรับปรุงประสิทธิภาพของการค้นคืนเอกสาร ยูสเคสให้ดีขึ้นได้

บทที่ 3

วิธีการวิจัย

งานวิจัยนี้แบ่งขั้นตอนในการดำเนินงานวิจัยออกเป็น 7 ขั้นตอน สามารถแสดงได้ด้วยแผนภาพกิจกรรม ดังรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนการดำเนินการวิจัย

ขั้นตอนที่ 3.1 เป็นขั้นตอนแรกเริ่มต้นของการดำเนินการวิจัย ซึ่งเริ่มจากการศึกษาทฤษฎีและงานวิจัยต่างๆ ที่เกี่ยวข้อง ต่อมาขั้นตอนที่ 3.2 ผู้วิจัยทำการเก็บรวบรวมเอกสาร

คำอธิบายยูสเคส ทั้งสิ้น 40 ระบบงานทางธุรกิจ เพื่อเป็นชุดข้อมูลสำหรับการใช้ในการทดลอง โดยนำเอกสารคำอธิบายยูสเคส 16 ระบบงานทางธุรกิจ จากงานวิจัยเรื่อง “การนำยูสเคสกลับมาใช้ใหม่โดยการคำนวณค่าความคล้ายกันของพจน์และโครงสร้าง” [7] และเก็บรวบรวมเพิ่มเติมอีก 24 ระบบงาน รวมเป็นเอกสารคำอธิบายยูสเคสทั้งสิ้น 611 ยูสเคส แล้วจึงวิเคราะห์หาแนวทางในการจัดกลุ่มยูสเคสต่อไป ขั้นตอนที่ 3.3 เป็นขั้นตอนในการวิเคราะห์การค้นคืนยูสเคสโดยการพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส ตามแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น จากนั้นจะวิเคราะห์การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ในขั้นตอนที่ 3.4 เมื่อได้แนวทางต่างๆ ครบถ้วนแล้ว ผู้วิจัยจะทำการออกแบบและพัฒนาเครื่องมือเพื่อใช้ในการทดสอบแนวคิดที่ได้นำเสนอ จะกล่าวในส่วนของบทที่ 4 จากนั้นออกแบบและทำการทดลองโดยใช้เครื่องมือดังกล่าว แล้วจึงประเมินและวิเคราะห์ผลที่ได้จากการทดลอง ก่อนจะสรุปผลการทดลองและสรุปผลการวิจัย ซึ่งจะกล่าวในบทที่ 5 และ 6 ตามลำดับ

ดังนั้นในส่วนของบทที่ 3 นี้จะกล่าวถึงขั้นตอนที่ 3.1 ถึง 3.4 ดังรายละเอียดต่อไปนี้

3.1 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง

โดยในขั้นตอนนี้เป็นการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องต่าง ๆ ซึ่งเกี่ยวกับแผนภาพยูสเคสและลักษณะของเอกสารคำอธิบายยูสเคส รวมถึงระบบการจัดเก็บและค้นคืนสารสนเทศ เรื่องการทำดรรชนี ซึ่งใช้ดรรชนีแบบแฟ้มข้อมูลผกผัน (Inverted File) การให้ค่าน้ำหนักของคำที่ได้หลังการผ่านขั้นตอนการดำเนินการกับข้อความ ซึ่งมีทั้งการตัดแยกคำ การกำจัดคำยกเว้น และการลดรูปคำศัพท์ [12] ด้วยการใช้ค่าความถี่ของเอกสารแบบผกผันตลอดจนแนวทางในการค้นคืนเอกสารยูสเคสด้วยการพิจารณาโครงสร้างของเอกสาร การหาค่าความคล้ายกันโดยใช้ค่าสัมประสิทธิ์โคซายน์ (Cosine Efficiency) และงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้นำเครื่องมือจากงานวิจัยเรื่อง “การนำยูสเคสกลับมาใช้ใหม่โดยการคำนวณค่าความคล้ายกันของพจน์ และโครงสร้าง” [7] มาศึกษาในส่วนของการทำดรรชนีและหาค่าน้ำหนักของคำด้วย

3.2 รวบรวมยูสเคสและวิเคราะห์แนวทางในการจัดกลุ่มยูสเคส

ผู้วิจัยได้เก็บรวบรวมเอกสารคำอธิบายยูสเคสเพิ่มเติมจากงานวิจัยเดิม [7] ซึ่งเดิมมีทั้งสิ้น 315 ยูสเคสจาก 16 ระบบงานทางธุรกิจ เป็น 40 ระบบงาน รวมทั้งสิ้นในงานวิจัยนี้มีชุดข้อมูลที่เป็นเอกสารคำอธิบายยูสเคส 611 ยูสเคส ซึ่งเขียนเป็นโดยใช้ภาษาอังกฤษทั้งหมด แล้วจึงนำมาวิเคราะห์แนวทางการจัดกลุ่มยูสเคส ซึ่งในชุดข้อมูลของงานวิจัยเดิมไม่มีขั้นตอนนี้อยู่ โดยแนวทางการจัดกลุ่มยูสเคสนั้น มีขั้นตอนดังต่อไปนี้

- 1) ผู้เชี่ยวชาญเฉพาะทางหรือวิศวกรซอฟต์แวร์ ตรวจสอบเอกสารยูสเคสทั้งหมด เพื่อกำหนดว่าเอกสารยูสเคสทั้งหมดที่มีอยู่ในชุดข้อมูลนั้น ควรมีกี่กลุ่ม อะไรบ้าง โดยมีเกณฑ์ที่ใช้ในการพิจารณาจัดกลุ่มอย่างคร่าว ๆ ได้แก่ การพิจารณายูสเคสตามระบบงาน (System) ระบบงานย่อย (Subsystem) และตามหน้าที่ของยูสเคส (Function) เป็นต้น

2) หลังจากที่ได้พิจารณาแล้ว ผู้เชี่ยวชาญ ได้เสนอให้มีการจัดกลุ่มเอกสารยูสเคสออกตามโดเมนทางธุรกิจ (Business Domain) นั่นคือ ได้กลุ่มของยูสเคสตามระบบงาน ซึ่งระบบงานมีทั้งหมด 40 ระบบงาน จึงได้กลุ่มของยูสเคส 40 กลุ่มด้วยกัน เนื่องจากแต่ละระบบงานมีลักษณะที่แตกต่างกันในระดับหนึ่งแล้ว หากจัดให้มีจำนวนกลุ่มน้อยลงไปกว่านี้หรือจัดให้ความละเอียดของคุณสมบัติของยูสเคสน้อยลงแล้ว อาจจะทำให้เอกสารคำอธิบายยูสเคสที่เป็นเอกสารเป้าหมายไม่ถูกค้นคืนออกมาได้

3) หาตัวแทนของยูสเคสสำหรับแต่ละกลุ่มยูสเคสได้โดยใช้สูตรจากสมการที่ 7 ซึ่งได้อธิบายรายละเอียดไว้ในหัวข้อที่ 2.1.3 แล้ว

4) หากมีการจัดเก็บยูสเคสใหม่เข้าสู่ระบบ จะทำการจัดเก็บยูสเคสทั้งระบบงาน แล้วจึงทำการหาตัวแทนยูสเคสของระบบงานนั้นๆ ก่อนเก็บค่าลงฐานข้อมูลเพื่อใช้ในการค้นคืนยูสเคสต่อไป

3.3 วิเคราะห์การค้นคืนยูสเคส

สำหรับการค้นคืนเอกสารยูสเคสนั้น ใช้วิธีการค้นคืนเกี่ยวกับงานวิจัย [7] นั่นคือ นำข้อความที่มีลักษณะโครงสร้างเหมือนกับเอกสารคำอธิบายยูสเคส เพื่อใช้ในการเปรียบเทียบหาค่าความคล้ายระหว่างแต่ละส่วนประกอบของข้อความกับเอกสารคำอธิบายยูสเคสโดยใช้สูตรดังสมการที่ 13 และหาค่าผลรวมความคล้ายของทั้งยูสเคสด้วยสูตรดังสมการที่ 12 เพื่อค้นคืนยูสเคสที่มีค่าความคล้ายกับข้อความมากที่สุดออกมา ซึ่งต้องดำเนินการเปรียบเทียบกับทุกยูสเคสที่เก็บไว้ในฐานข้อมูล แต่สำหรับงานวิทยานิพนธ์นี้ได้เพิ่มการจัดกลุ่มยูสเคสและทำการหา ยูสเคสตัวแทนของแต่ละกลุ่มไว้แล้ว ดังนั้น ขั้นตอนการเปรียบเทียบค่าความคล้ายระหว่างข้อความกับยูสเคสจึงเริ่มจากการเปรียบเทียบกับยูสเคสตัวแทนของกลุ่มเสียก่อน แล้วจึงทำการเปรียบเทียบกับยูสเคสที่เป็นสมาชิกในกลุ่มต่อไป ดังนั้น ในส่วนของการเปรียบเทียบเพื่อหาค่าความคล้ายระหว่างข้อความกับยูสเคสตัวแทนกลุ่มนั้น สามารถใช้สูตรในการคำนวณได้เช่นเดียวกันกับที่กล่าวมา แต่สำหรับในงานวิทยานิพนธ์นี้จะใช้สูตรการคำนวณต่างออกไปจากงานวิจัยเดิม ซึ่งจะอธิบายรายละเอียดไว้ในหัวข้อที่ 3.5.2 ต่อไป

3.4 วิเคราะห์การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน

เป็นขั้นตอนที่ดำเนินการภายหลังจากที่ผู้ค้นคืนยูสเคสตรวจสอบหรือดูรายละเอียดเนื้อหาภายในเอกสารคำอธิบายของยูสเคสที่ระบบค้นคืนมาให้ในครั้งแรกเรียบร้อยแล้ว และผู้ค้นคืนระบุว่ายูสเคสใดบ้างตรงกับความต้องการของตน โดยพัฒนาส่วนต่อประสานผู้ใช้เพื่ออำนวยความสะดวกให้ผู้ใช้งานเลือกยูสเคสที่ตรงความต้องการ ซึ่งระบบจะนำข้อมูลยูสเคสที่ตรงประเด็นหรือตรงความต้องการเหล่านี้ ไปคำนวณเพื่อใช้ในการเปลี่ยนแปลงข้อความเดิมเป็นข้อความใหม่ โดยใช้วิธีการเพื่อปรับเปลี่ยนข้อความจากการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้น มี 2 วิธีการด้วยกันคือ การเปลี่ยนแปลงข้อความ หรือการขยายคำในข้อความ และ

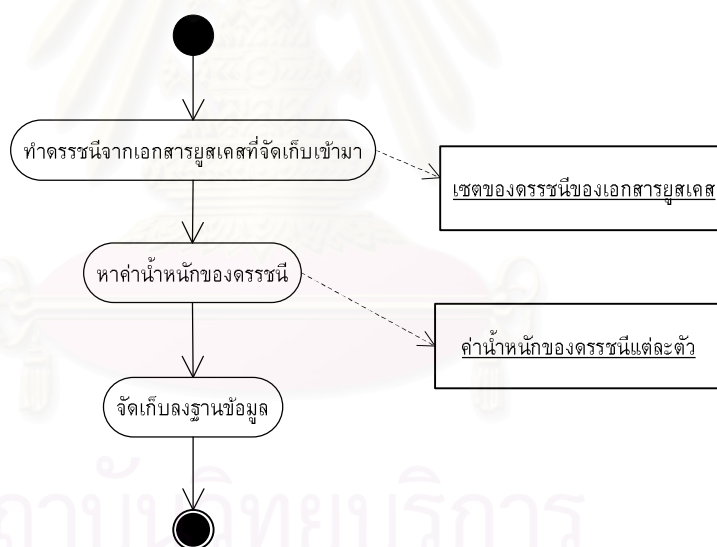
การเปลี่ยนแปลงน้ำหนักค่าในข้อคำถาม ด้วยแบบจำลองปริภูมิเวกเตอร์ และอีกวิธีคือ การเปลี่ยนแปลงค่าน้ำหนักของค่าในแบบจำลองความน่าจะเป็น

3.5 ภาพรวมการทำงานของวิธีที่นำเสนอ

หลังจากที่ได้วิเคราะห์แนวทางต่างๆ แล้วจึงพัฒนาเครื่องมือสำหรับการค้นคืนยูสเคสที่พิจารณาค่า และโครงสร้างของยูสเคส โดยมีการพัฒนาฟังก์ชันสำหรับการหาตัวแทนกลุ่มยูสเคสทั้งหมดไว้ล่วงหน้าก่อนที่จะเริ่มใช้เครื่องมือนี้ในการค้นคืนยูสเคส และพัฒนาฟังก์ชันในส่วนของการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ โดยอาศัยหลักการต่างๆ ที่ได้วิเคราะห์ไว้ข้างต้น จึงได้ระบบหรือเครื่องมือที่พัฒนาเพื่อสนับสนุนวิธีการซึ่งงานวิทยานิพนธ์นี้นำเสนอ ภาพรวมของการทำงานประกอบด้วยขั้นตอนหลัก 3 ขั้นตอน แสดงดังรูปที่ 3.3

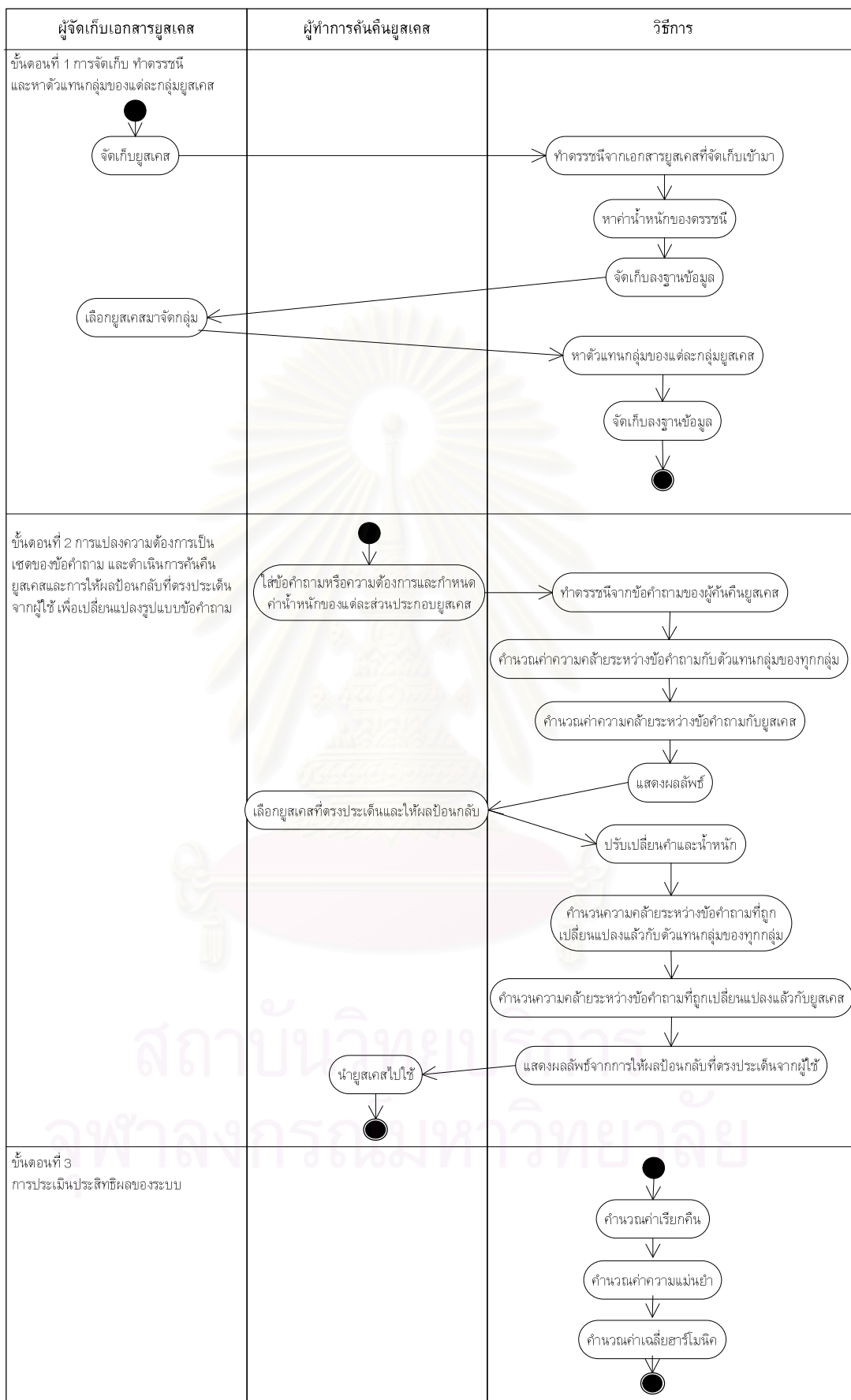
3.5.1 ขั้นตอนการจัดเก็บยูสเคส

สำหรับกิจกรรมในขั้นตอนนี้ประกอบด้วย 2 ขั้นตอนย่อยด้วยกัน ได้แก่ ขั้นตอนการจัดเก็บเอกสารคำอธิบายยูสเคส แสดงดังรูปที่ 3.2 และขั้นตอนการนำเอกสารคำอธิบายยูสเคสที่จัดเก็บมาจัดกลุ่มเพื่อหาตัวแทนยูสเคสในแต่ละกลุ่ม แสดงดังรูปที่ 3.4

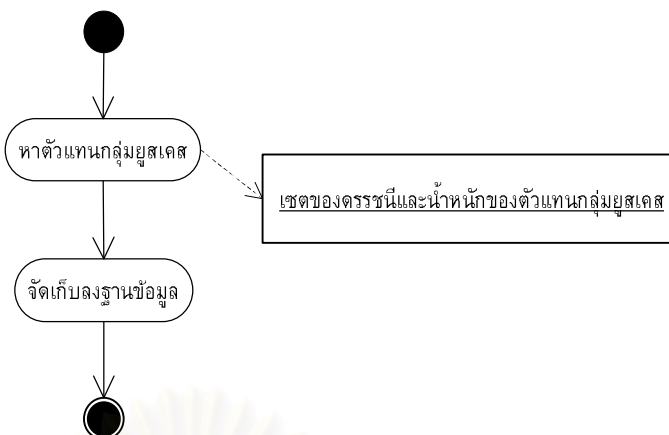


รูปที่ 3.2 แผนภาพกิจกรรมการจัดเก็บเอกสารยูสเคสเพื่อทำกรณีและน้ำหนักค่าอัตโนมัติ

หลังจากที่ผู้จัดเก็บเอกสารคำอธิบายยูสเคสทำการจัดเก็บเอกสารแต่ละชิ้นแล้ว ระบบจะใช้ฟังก์ชันในการหากรณีสำหรับเอกสารแต่ละชิ้นนั้น และหาค่าน้ำหนักให้กับกรณีเหล่านั้นแบบอัตโนมัติ โดยขั้นตอนของการทำกรณีนั้น กล่าวอธิบายไว้ในหัวข้อที่ 2.1.2 ส่วนการหาค่าน้ำหนักให้กับกรณีแต่ละตัวใช้เทคนิคของการหาค่าน้ำหนักแบบผกผัน ดังนั้นเอกสารยูสเคสที่ถูกจัดเก็บเรียบร้อยแล้วจึงมีกรณีและค่าน้ำหนักในการช้อยไปยังที่จัดเก็บเอกสารเหล่านั้นไว้



รูปที่ 3.3 แผนภาพกิจกรรมของขั้นตอนโดยรวมของวิธีที่นำเสนอ

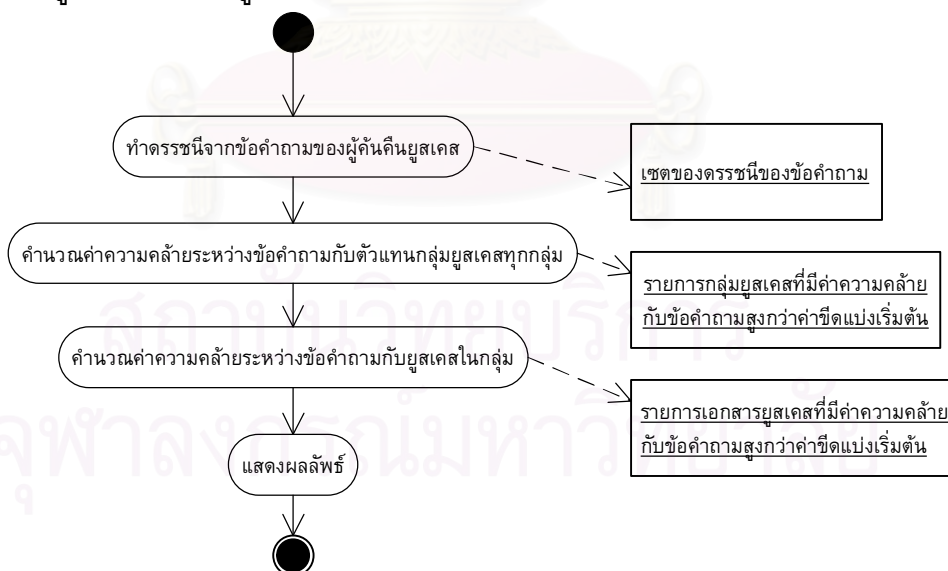


รูปที่ 3.4 แผนภาพกิจกรรมของการหาตัวแทนกลุ่มเอกสารยูสเคส

จากนั้นผู้จัดเก็บยูสเคสจะนำเอกสารยูสเคสที่อยู่ในระบบงานทางธุรกิจเดียวกัน ซึ่งถือว่าอยู่ในกลุ่มเดียวกันนั้น มาหาตัวแทนกลุ่มยูสเคส โดยการนำเซตของดรรรชนีและค่าน้ำหนักของยูสเคสทั้งหมดในกลุ่มมาคำนวณโดยสูตรในสมการที่ 7 ดังนั้น ในแต่ละกลุ่มหรือระบบงานก็จะมีเซตของดรรรชนีและน้ำหนักค่าหนึ่งซึ่งมีค่าอยู่กึ่งกลาง และทำหน้าที่เป็นตัวแทนกลุ่ม จากนั้นจึงจัดเก็บลงฐานข้อมูล เพื่อใช้ในขั้นตอนการค้นคืนยูสเคสต่อไป

3.5.2 ขั้นตอนการค้นคืนยูสเคส

สำหรับกิจกรรมในขั้นตอนนี้ประกอบด้วย 2 ขั้นตอนย่อยด้วยกัน ได้แก่ ขั้นตอนในการค้นคืนยูสเคส แสดงดังรูปที่ 3.5 และขั้นตอนการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้หรือผู้ทำการค้นคืนยูสเคส แสดงดังรูปที่ 3.6



รูปที่ 3.5 แผนภาพกิจกรรมของการค้นคืนยูสเคส

การค้นคืนยูสเคส เป็นการนำความต้องการของผู้ค้นคืนยูสเคสมาเทียบความคล้ายกับยูสเคสที่มีอยู่ โดยความต้องการนั้นต้องถูกเปลี่ยนให้เป็นดรรรชนีด้วยเนื่องจากเอกสารยูสเคสที่เก็บไว้นั้น อยู่ในรูปแบบของเซตดรรรชนีนั่นเอง ดังนั้นกิจกรรมของการค้นคืนยูสเคสจึงเริ่มจาก

ความต้องการหรือข้อความที่รับมาถูกแปลงให้เป็นเซตตรรกะของข้อความ แล้วจึงนำไปเปรียบเทียบกับเอกสารยูสเคสซึ่งในงานวิทยานิพนธ์นี้ออกแบบให้ข้อความต้องนำมาคำนวณค่าความคล้ายกับตัวแทนกลุ่มเอกสารยูสเคสก่อนที่จะคำนวณค่าความคล้ายกับเอกสารยูสเคสภายในกลุ่มที่มีค่าความคล้ายสูงพอกับค่าขีดแบ่งเริ่มต้นที่กำหนด

และเนื่องจากวิทยานิพนธ์นี้ พิจารณาลักษณะของโครงสร้างของยูสเคสดังที่กล่าวไว้ในหัวข้อที่ 3.3 และพิจารณาการให้ความสำคัญกับแต่ละส่วนประกอบของเอกสารยูสเคส ดังนั้นการหาค่าความคล้ายกันระหว่างเซตตรรกะของข้อความกับเซตตรรกะของตัวแทนกลุ่มเอกสารยูสเคสสามารถหาได้จากสูตรในสมการที่ 14

$$SIM(Centroid_j, Query_j) = \frac{\sum_{m=1}^e [SIM(ECentroid_{mj}, EQuery_{mj}) \cdot WE_m]}{TotalWeight} \quad (14)$$

กำหนดให้ $SIM(ECentroid_{mj}, EQuery_{mj})$ คือ ค่าความคล้ายระหว่างส่วนประกอบที่ m ของตัวแทนกลุ่มยูสเคสที่ i กับส่วนประกอบที่ m ของข้อความที่ j
 WE_m คือ ค่าน้ำหนักที่ผู้ค้นคืนกำหนดให้กับส่วนประกอบที่ m
 e คือ จำนวนส่วนประกอบทั้งหมดของเอกสารยูสเคส
 $TotalWeight$ คือ ค่าน้ำหนักรวมของค่าน้ำหนักทั้งหมดที่ผู้ค้นคืนยูสเคสกำหนดให้กับแต่ละส่วนประกอบ

สูตรที่ใช้ในการหาค่าความคล้ายกันระหว่างเซตตรรกะของข้อความกับเซตตรรกะของเอกสารยูสเคสแต่ละเอกสารภายในกลุ่มยูสเคส มีดังต่อไปนี้

$$SIM(UC_j, Query_j) = \frac{\sum_{m=1}^e [SIM(EUC_{mi}, EQuery_{mj}) \cdot WE_m]}{TotalWeight} \quad (15)$$

กำหนดให้ $SIM(EUC_{mi}, EQuery_{mj})$ คือ ค่าความคล้ายระหว่างส่วนประกอบที่ m ของยูสเคสที่ i กับส่วนประกอบที่ m ของข้อความที่ j

ในงานวิจัยนี้ใช้วิธีการคำนวณค่าความคล้ายด้วยแบบจำลอง 2 แบบด้วยกันได้แก่

1) กรณีใช้แบบจำลองปริภูมิเวกเตอร์ ค่าความคล้ายระหว่างข้อความและตัวแทนกลุ่มยูสเคสของแต่ละส่วนประกอบของเอกสาร เป็นดังสมการที่ 16

$$SIM(ECentroid_{mi}, EQuery_{mj}) = \frac{\sum_{k=1}^t (CTerm_{mik} \cdot W_{ik}) \cdot QTerm_{mjk}}{\sqrt{\sum_{k=1}^t (CTerm_{mik})^2 \cdot \sum_{k=1}^t (QTerm_{mjk})^2}} \quad (16)$$

กำหนดให้ $CTerm_{mik}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของตัวแทนกลุ่มยูสเคสที่ i และมีค่า 0 หากไม่ปรากฏ

$QTerm_{mjk}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของข้อความที่ j และมีค่า 0 หากไม่ปรากฏ

W_{ik} คือ น้ำหนักของคำ k ในตัวแทนกลุ่มยูสเคสที่ i

ค่าความคล้ายระหว่างข้อความกับเอกสารยูสเคสของแต่ละส่วนประกอบของเอกสารเป็นดังสูตรในสมการที่ 17

$$SIM(EUC_{mi}, EQuery_{mj}) = \frac{\sum_{k=1}^t (Term_{mik} \cdot W_{ik}) \cdot QTerm_{mjk}}{\sqrt{\sum_{k=1}^t (Term_{mik})^2 \cdot \sum_{k=1}^t (QTerm_{mjk})^2}} \quad (17)$$

กำหนดให้ $Term_{mik}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของยูสเคสที่ i และมีค่า 0 หากไม่ปรากฏ

2) กรณีใช้แบบจำลองความน่าจะเป็น เป็นดังสมการที่ 18

$$SIM(ECentroid_{mi}, EQuery_{mj}) = \sum_{k=1}^t W_{k,mi} \times W_{k,mj} \times \log \frac{N - n_k}{n_k} \quad (18)$$

กำหนดให้ $W_{k,mi}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของตัวแทนกลุ่มยูสเคสที่ i และมีค่า 0 หากไม่ปรากฏ

$W_{k,mj}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของข้อความที่ j และมีค่า 0 หากไม่ปรากฏ

N คือ จำนวนตัวแทนกลุ่มทั้งหมด

n_k คือ จำนวนตัวแทนกลุ่มเอกสารยูสเคสที่มีค่า k ปรากฏ

ค่าความคล้ายระหว่างข้อความกับเอกสารยูสเคสเป็นดังสมการที่ 19

$$SIM(EUC_{mi}, EQuery_{mj}) = \sum_{k=1}^t W_{k,mi} \times W_{k,mj} \times \log \frac{N - n_k}{n_k} \quad (19)$$

กำหนดให้ $W_{k,mi}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของยูสเคสที่ i และมีค่า 0 หากไม่ปรากฏ

$W_{k,mj}$ มีค่าเป็น 1 เมื่อปรากฏคำที่ k ในส่วนประกอบที่ m ของข้อความที่ j และมีค่า 0 หากไม่ปรากฏ

N คือ จำนวนเอกสารยูสเคสทั้งหมด

n_k คือ จำนวนเอกสารยูสเคสที่มีคำ k ปรากฏ

สูตรดังกล่าวข้างต้นนั้น ผู้ค้นคืนสามารถเลือกได้ว่าจะทำการค้นคืนด้วยวิธีการของแบบจำลองปริภูมิเวกเตอร์ หรือแบบจำลองความน่าจะเป็น จากนั้นระบบจะแสดงผลลัพธ์เป็นเอกสารยูสเคสออกมาให้กับผู้ค้นคืน

หลังจากที่ระบบแสดงผลลัพธ์แล้ว ผู้ค้นคืนจึงทำการตรวจสอบเอกสารยูสเคสว่าเอกสารใดตรงกับความต้องการของตน ซึ่งผู้ค้นคืนสามารถค้นคืนแบบใช้ผลป้อนกลับที่ตรงประเด็นได้ โดยการป้อนรายการเอกสารยูสเคสที่ตรงความต้องการกลับคืนสู่ระบบ ระบบจะนำข้อมูลเหล่านี้ไปคำนวณเพื่อแสดงผลลัพธ์อีกครั้ง โดยกิจกรรมของการค้นคืนแบบใช้ผลป้อนกลับที่ตรงประเด็นผู้ใช้ แสดงดังรูปที่ 3.5

เมื่อระบบรับรายการเอกสารยูสเคสที่ตรงความต้องการมาจากผู้ใช้แล้ว ระบบจะทำการปรับเปลี่ยนคำ และน้ำหนักของคำในข้อความ จากนั้นระบบจะทำการคำนวณค่าความคล้ายระหว่างข้อความที่ปรับเปลี่ยนแล้วกับตัวแทนกลุ่มยูสเคสแต่ละกลุ่ม และระหว่างข้อความที่ปรับเปลี่ยนกับเอกสารยูสเคสภายในกลุ่ม สำหรับกรณีของแบบจำลองปริภูมิเวกเตอร์นั้น การปรับเปลี่ยนข้อความสามารถหาได้จากสูตรในสมการที่ 8 ก่อนจะนำมาหาค่าความคล้ายโดยใช้สูตรดังที่กล่าวไว้ข้างต้น แต่สำหรับกรณีของแบบจำลองความน่าจะเป็นนั้น สูตรที่ใช้ มีการเปลี่ยนแปลงไป จากสมการที่ 18 เป็นดังสมการที่ 20 และจากสมการที่ 19 เป็นสมการที่ 21

$$SIM(ECentroid_{mi}, EQuery_{mj}) = \sum_{k=1}^t W_{k,mi} \times W_{k,mj} \times \log \frac{|D_{r,k}|}{|D_r| - |D_{r,k}|} \times \frac{N - |D_r| - (n_k - |D_{r,k}|)}{n_k - |D_{r,k}|} \quad (20)$$

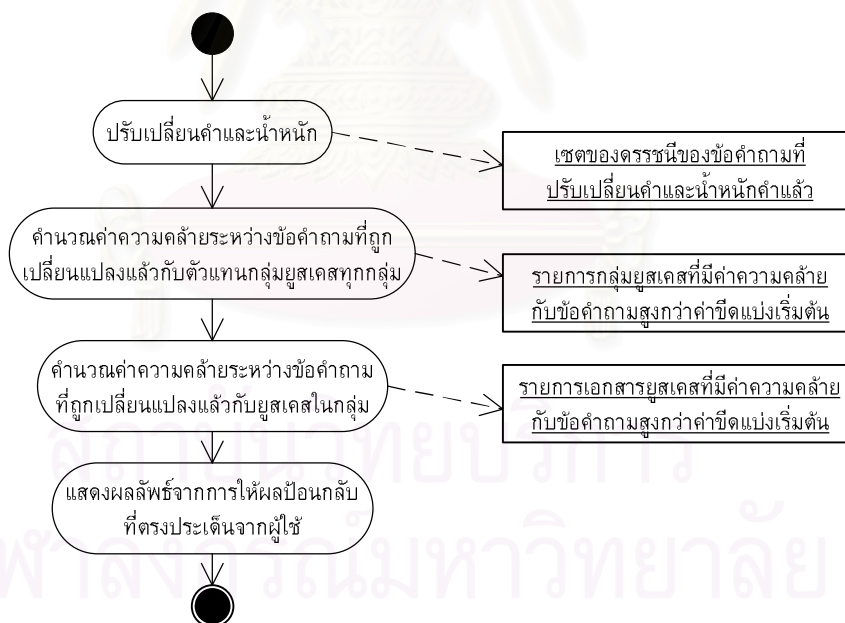
กำหนดให้ D_r คือ เซตของกลุ่มของเอกสารยูสเคสที่ตรงประเด็นที่ระบบค้นคืนได้

- $D_{r,k}$ คือ เซตย่อยของ D_r ที่ประกอบด้วยกลุ่มของเอกสารยูสเคสที่มีค่า k ปรากฏอยู่
- N คือ จำนวนตัวแทนกลุ่มทั้งหมด
- n_k คือ จำนวนตัวแทนกลุ่มเอกสารยูสเคสที่มีค่า k ปรากฏ

$$SIM(EUC_{mi}, EQuery_{mj}) = \sum_{k=1}^t w_{k,mi} \times w_{k,mj} \times \log \frac{|D_{r,k}|}{|D_r| - |D_{r,k}|} \times \frac{N - |D_r| - (n_k - |D_{r,k}|)}{n_k - |D_{r,k}|} \quad (21)$$

- กำหนดให้ D_r คือ เซตของเอกสารที่ตรงประเด็นที่ค้นคืนได้ กำหนดโดยผู้ใช้
- $D_{r,k}$ คือ สับเซตของ D_r ที่ประกอบด้วยเอกสารที่มีค่า k อยู่
- N คือ จำนวนเอกสารยูสเคสทั้งหมด
- n_k คือ จำนวนเอกสารยูสเคสที่มีค่า k ปรากฏ

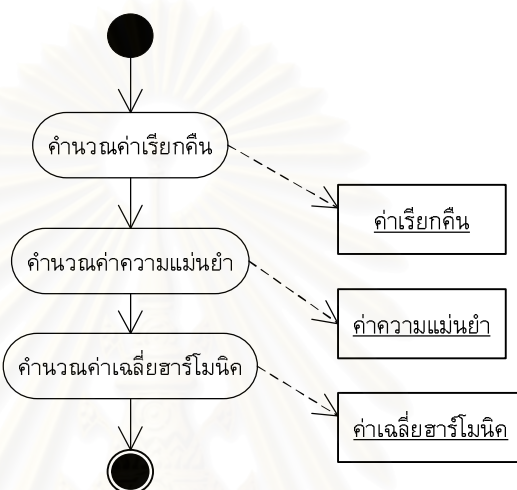
จากนั้นระบบจะแสดงผลลัพธ์เป็นเอกสารยูสเคสที่ได้จากการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ผ่านทางส่วนต่อประสานผู้ใช้ต่อไป



รูปที่ 3.6 แผนภาพกิจกรรมของการค้นคืนแบบใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

3.5.3 ขั้นตอนการประเมินประสิทธิผลของระบบการค้ำคินยूसเคส

การประเมินประสิทธิผลของระบบค้ำคินสารสนเทศนั้น นิยมประเมินจากสารสนเทศที่ค้ำคินออกมา นั้นว่ามีความถูกต้องหรือตรงกับที่ผู้ค้ำคินต้องการหรือไม่ ดังนั้นในงานวิทยานิพนธ์นี้จึงใช้ค้ำคินค่าเรียกคิน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิกเป็นมาตรวัดประสิทธิผลของการค้ำคินยूसเคส ซึ่งสามารถค้ำคินได้จากสูตรในสมการที่ 3 4 และ 5 ตามลำดับ สำหรับกิจกรรมในขั้นตอนนี้แสดงเป็นแผนภาพกิจกรรมได้ดังรูปที่ 3.7



รูปที่ 3.7 แผนภาพกิจกรรมการประเมินประสิทธิผลของระบบ

บทที่ 4

การพัฒนาเครื่องมือ

ในบทนี้จะกล่าวถึงรายละเอียดของการพัฒนาเครื่องมือสำหรับการค้นคืนยูสเคสโดยมีการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ บนแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น โดยจะกล่าวถึงสภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือ สถาปัตยกรรมของเครื่องมือ โครงสร้างของเครื่องมือ และแบบจำลองข้อมูล ซึ่งมีรายละเอียดดังต่อไปนี้

4.1 สภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือ

ฮาร์ดแวร์ (Hardware) และซอฟต์แวร์ (Software) ที่ใช้ในการพัฒนาระบบมีดังนี้

4.1.1 ฮาร์ดแวร์

- 1) เครื่องคอมพิวเตอร์ส่วนบุคคล หน่วยประมวลผล อินเทล เพนเทียมเอ็ม 1.6 กิกะเฮิร์ตซ์ โปรเซสเซอร์ 600 (Intel Pentium M 1.6 GHz Processor 600)
- 2) หน่วยความจำ (Memory) 512 เมกะไบต์
- 3) จานบันทึกแบบแข็ง (Hard disk) ความจุ 40 กิกะไบต์
- 4) จอภาพ 14.1 นิ้ว

4.1.2 ซอฟต์แวร์

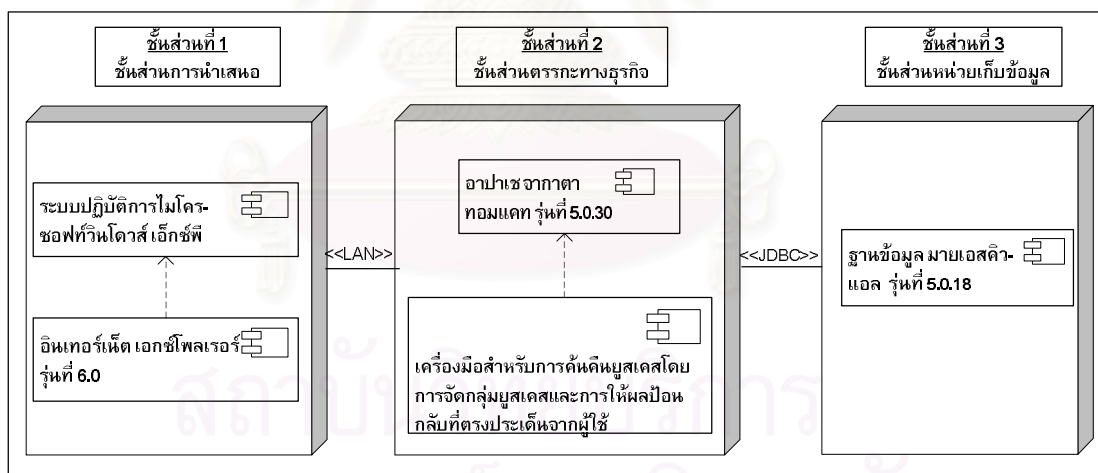
- 1) ระบบปฏิบัติการ
 - (1) ระบบปฏิบัติการไมโครซอฟท์วินโดวส์ เอ็กซ์พี โปรเฟสชันแนล (Microsoft Window XP Professional)
- 2) ซอฟต์แวร์ที่ใช้ในจัดทำเอกสาร
 - (1) ไมโครซอฟท์ออฟฟิศ 2003 (Microsoft Office 2003)
 - (2) ไมโครซอฟท์ออฟฟิศวิซิโอ โปรเฟสชันแนล 2003 (Microsoft Office Visio Professional 2003)
- 3) ซอฟต์แวร์ที่ใช้ในการพัฒนาส่วนต่อประสานกราฟิก
 - (1) เครื่องมือช่วยจัดการส่วนต่อประสานเว็บเพจ ได้แก่ มาโครมีเดียร์ ดรีมวีเวอร์ เอ็ม เอ็กซ์ 2004 รุ่นที่ 7 (Macromedia Dreamweaver: MX 2004 version 7)
 - (2) อินเทอร์เน็ต เอกซ์โพลเลอร์ รุ่นที่ 6 (Internet Explorer 6.0)
- 4) ซอฟต์แวร์ที่ใช้ในการเขียนโปรแกรมเพื่อพัฒนาเครื่องมือ
 - (1) เจครีเอเตอร์ รุ่นที่ 3.5 แอลอี (JCreator 3.5 LE)
 - (2) จาวาสแตนด์เอเลี่ยน รุ่น 1.4.2.08 สำหรับวินโดวส์ (J2SDK 1.4.2.09 for Windows)
- 5) ซอฟต์แวร์ที่ใช้ในส่วนให้บริการและส่วนสนับสนุน

- (1) เว็บเซิร์ฟเวอร์ อาปาเช จากาตา ทอมแคท รุ่นที่ 5.0.30 (Apache Jakarta Tomcat 5.0.30)
- (2) ฐานข้อมูล มายเอสคิวแอล รุ่นที่ 5.0.18 (MySQL 5.0.18)
- (3) โปรแกรมจัดการฐานข้อมูลพรีเมียมซอฟต์แวร์ นาวิแคท 2006 (PremiumSoft Navicat 2006)
- (4) ตัวเชื่อมต่อมายเอสคิวแอลรุ่นที่ J 3.0.15 (MySQL Connector/J 3.0.15)

4.2 การออกแบบสถาปัตยกรรมของเครื่องมือ

ระบบที่พัฒนามีการออกแบบสถาปัตยกรรมแบบหลายส่วนชั้น (Multi-Tiers) ดังรูปที่ 4.1 ซึ่งสามารถแบ่งออกเป็น 3 ส่วน ได้แก่

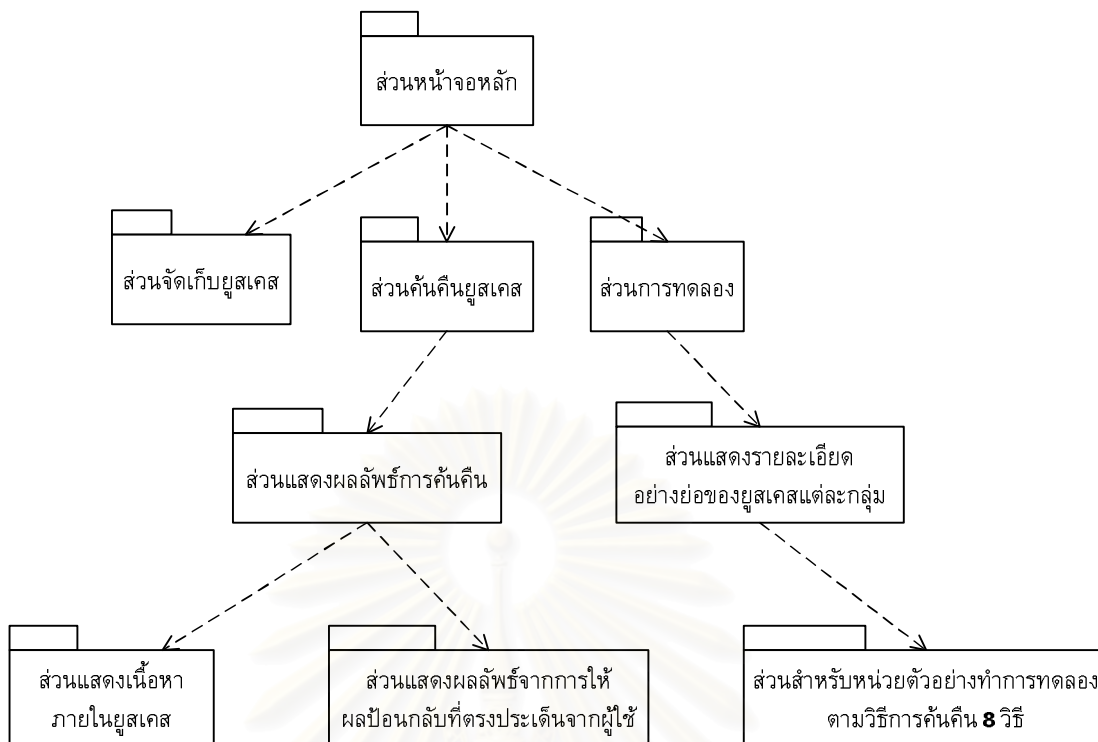
- 1) ชั้นส่วนการนำเสนอ (Presentation Tier) เป็นชั้นส่วนลูกข่ายซึ่งทำหน้าที่เป็นส่วนต่อประสานกับผู้ใช้ระบบโดยใช้โปรแกรมประยุกต์แบบเว็บเบส เพื่อแก้ปัญหาข้อจำกัดเรื่องสถานที่ทำงาน
- 2) ชั้นส่วนตรรกะทางธุรกิจ (Business Logic Tier) เป็นส่วนกลางทำหน้าที่ให้บริการแก่เครื่องลูกข่าย
- 3) ชั้นส่วนหน่วยเก็บข้อมูลของระบบ (Data Storage Tier) เป็นส่วนที่ทำหน้าที่จัดเก็บข้อมูลของระบบ โดยใช้ฐานข้อมูลมายเอสคิวแอล (MY SQL) ในการจัดเก็บข้อมูลของระบบ



รูปที่ 4.1 สถาปัตยกรรมของเครื่องมือสำหรับการค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

4.3 โครงสร้างของเครื่องมือ

โครงสร้างของเครื่องมือที่พัฒนาขึ้นสำหรับการค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ จะอธิบายด้วยแผนภาพส่วนประกอบซึ่งแสดงความสัมพันธ์ระหว่างส่วนประกอบต่างๆ ในระบบ แผนภาพแสดงส่วนประกอบโครงสร้างของเครื่องมือ แสดงดังรูปที่ 4.2



รูปที่ 4.2 แผนภาพส่วนประกอบโครงสร้างของเครื่องมือ

หน้าจอหลักแสดงรายการให้เลือกได้ 3 รายการ ดังรูปที่ 4.3 ได้แก่

- 1) Use Case Storage เป็นส่วนที่ใช้สำหรับจัดเก็บเอกสารคำอธิบายยูสเคสลงฐานข้อมูลหรือชุดข้อมูล โดยรายละเอียดของการทำงานในส่วนนี้จะอธิบายไว้ในข้อ 4.2.1
- 2) Use Case Retrieval เป็นส่วนที่ใช้สำหรับค้นคืนเอกสารคำอธิบายยูสเคสโดยรายละเอียดของการทำงานในส่วนนี้จะอธิบายไว้ในข้อ 4.2.2
- 3) The Experiment เป็นส่วนสำหรับให้หน่วยตัวอย่างทำการทดลองเพื่อวัดประสิทธิผลของระบบและวิธีการที่น่าเสนอ โดยรายละเอียดการทำงานในส่วนนี้จะอธิบายไว้ในบทที่ 5

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool	
Software Engineering Lab Center of Excellence in Software Engineering Department of Computer Engineering, Faculty of Engineering Chulalongkorn University, Bangkok, Thailand Main Menu	
Topic	Description
1. Use Case Storage	: Use Case Storage is to store use case into database/collection.
2. Use Case Retrieval	: Use Case Retrieval is to query some use cases from the database/collection.
3. The Experiment	: The Experiment is to test effectiveness of our approach with some sample units.

รูปที่ 4.3 หน้าจอแสดงรายการหลักของเครื่องมือ

4.3.1 การจัดเก็บเอกสารคำอธิบายยูสเคส

การจัดเก็บเอกสารคำอธิบายยูสเคสนี้ แสดงด้วยส่วนต่อประสานผู้ใช้ ซึ่งทำหน้าที่บันทึกข้อมูลในแต่ละส่วนของเอกสารคำอธิบายยูสเคสจากผู้จัดเก็บชุดข้อมูล ภาพหน้าจอสำหรับการจัดเก็บเอกสารคำอธิบายยูสเคส แสดงในรูปที่ 4.4

เนื่องจากในแผนภาพยูสเคสหนึ่ง ๆ ซึ่งเป็นตัวแทนของระบบงานหนึ่ง ๆ นั้น ประกอบไปด้วยหลายยูสเคส จึงสามารถนำมาเขียนเป็นเอกสารคำอธิบายยูสเคสได้หลายส่วน แผนภาพยูสเคสและเอกสารคำอธิบายยูสเคสจึงจำเป็นต้องมีรหัสของแผนภาพ และรหัสของเอกสารคำอธิบายยูสเคส ตามลำดับ เพื่อกำหนดหรือแยกแยะว่ายูสเคสเหล่านั้นเป็นยูสเคสของระบบงานใด ในหน้าจอสำหรับการจัดเก็บยูสเคสนี้ จึงออกแบบให้มีการจัดเก็บรหัสของแผนภาพยูสเคสและรหัสของเอกสารคำอธิบายยูสเคส ในช่อง Diagram ID และ Use Case ID ตามลำดับสำหรับในช่องอื่นๆ เป็นส่วนของเนื้อหาในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส หลังจากที่ผู้จัดเก็บยูสเคสใส่เนื้อหาของเอกสารคำอธิบายยูสเคสหนึ่ง ๆ เรียบร้อย และกดปุ่ม Submit แล้ว ข้อมูลในเอกสารนี้จะถูกแปลงเป็นเซตของดรรชนีโดยผ่านการทำดรรชนีอัตโนมัติ (Automatic Indexing) ซึ่งได้กล่าวไว้ใน หัวข้อที่ 2.1.2 ในส่วนการจัดเก็บสารสนเทศ

4.3.2 การคั่นคืนเอกสารคำอธิบายยูสเคส

สำหรับการคั่นคืนเอกสารคำอธิบายยูสเคส มีการออกแบบส่วนต่อประสานผู้ใช้ให้มีลักษณะเหมือนกับโครงสร้างของเอกสารคำอธิบายยูสเคส เพื่อให้ผู้ทำการคั่นคืนยูสเคสหรือวิศวกรความต้องการ ใส่คำสำคัญภาษาอังกฤษที่ต้องการคั่นลงในแต่ละช่องตามความต้องการของตน และเลือกค่าน้ำหนักหรือการให้ความสำคัญสำหรับแต่ละช่องด้วย โดยช่วงของค่าน้ำหนักอยู่ระหว่าง 0 ถึง 5 ซึ่ง 1 คือค่าน้ำหนักที่น้อยที่สุดและ 5 คือค่าน้ำหนักที่มากที่สุด หรือการให้ค่าความสำคัญกับส่วนประกอบของเอกสารคำอธิบายยูสเคสน้อยที่สุดไปจนถึงมากที่สุดตามลำดับ โดยที่ 0 คือไม่ให้ความสำคัญ หรือกรณีที่ผู้ทำการคั่นคืนไม่ใส่คาลงในส่วนประกอบนั้นนั่นเอง หน้าจอสำหรับการคั่นคืนเอกสารคำอธิบายยูสเคส แสดงดังรูปที่ 4.5

หลังจากที่ผู้ทำการคั่นคืน ใส่ข้อความลงในแต่ละช่อง และเลือกค่าน้ำหนักเรียบร้อยแล้ว จึงทำการคั่นคืนในแต่ละวิธีการซึ่งเครื่องมือจะกำหนดมาให้ที่ละวิธีการเรียงตามลำดับอย่างต่อเนื่อง ซึ่งวิธีการคั่นคืนมีทั้งหมด 4 วิธีการด้วยกัน ได้แก่

- 1) การคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์
- 2) การคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น
- 3) การคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์
- 4) การคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool

Store a Use Case

Diagram ID :

Use Case ID :

Use Case Name :

Objective :

Actor :

Relationship :

Association :

Include :

Extend :

Generalization :

Precondition :

Postcondition :

Normal Flow of Events :

Subflow :

Alternative or Exceptional Flow of Events :

[back to menu](#)

รูปที่ 4.4 หน้าจอแสดงส่วนสำหรับการจัดเก็บเอกสารคำอธิบายยูสเคส

โดยที่การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส ในงานวิจัยนี้เป็นการนำข้อคำถามในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคสในหน้าค้นคืนมาต่อกันเป็นข้อคำถามเดี่ยวก่อนนำไปเปรียบเทียบกับเอกสารคำอธิบายยูสเคสในฐานข้อมูล จึงเสมือนว่าไม่พิจารณาโครงสร้าง แต่การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสนั้น จะทำการเปรียบเทียบข้อคำถามจากแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคสในหน้าค้นคืนกับคำสำคัญในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคสในฐานข้อมูล จึงถือว่าการค้นคืนที่พิจารณาโครงสร้างของเอกสาร

เมื่อผู้ทำการค้นคืนเลือกวิธีการใดวิธีการหนึ่งแล้วกดปุ่ม Submit ข้อคำถามนั้นๆ จะถูกแปลงเป็นเซตตรรกะ และนำไปเปรียบเทียบกับค่าความคล้ายกับตัวแทนเอกสารคำอธิบายยูสเคสของแต่ละระบบงาน ก่อนที่จะนำไปเปรียบเทียบกับค่าความคล้ายกับเอกสารคำอธิบายยูสเคสแต่ละเอกสารภายในระบบงานนั้นๆ แยกตามวิธีการค้นคืนยูสเคสทั้ง 4 ที่กล่าวไปข้างต้น ผลที่ได้จากการค้นคืนเอกสารคำอธิบายยูสเคสนั้น มี 2 ส่วนด้วยกัน โดยในส่วนแรก จะเป็นผลที่ได้จากการค้นคืน ซึ่งจะแสดงชื่อของยูสเคส ดังรูปที่ 4.6 และส่วนแสดงรายละเอียดของแต่ละเอกสารคำอธิบายยูสเคส ดังรูปที่ 4.7 เมื่อผู้ใช้ทำการให้ผลป้อนกลับที่ตรงประเด็นเข้าสู่ระบบแล้ว ระบบจะค้นคืนซ้ำและแสดงชื่อของยูสเคสที่ได้จากการค้นคืนแบบให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ดังแสดงในรูปที่ 4.8

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool

Search Use Case

Use Case Name: 0 ▾

Objective: 0 ▾

Actor: 0 ▾

Relationship:

Association: 0 ▾

Include: 0 ▾

Extend: 0 ▾

Generalization: 0 ▾

Precondition: 0 ▾

Postcondition: 0 ▾

Normal Flow of Events: 0 ▾

Subflow: 0 ▾

Alternative Flow of Events: 0 ▾

Method 1 Combine elements by Vector Space Model

เลือกค่าหน้าหลัก

รูปที่ 4.5 หน้าจอแสดงส่วนสำหรับการค้นคืนเอกสารคำอธิบายยูสเคส

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool

Use Cases Results
by Element Separation on Probabilistic Model without User Relevance Feedback

Select	Use Case ID	Use Case Name	Details
<input type="checkbox"/>	1509	View course detail	details
<input type="checkbox"/>	1603	Student View All Courses	details
<input type="checkbox"/>	1510	Register course	details
<input type="checkbox"/>	1512	Study registered course	details
<input type="checkbox"/>	0205	View course operation details	details
<input type="checkbox"/>	0206	View course operation status	details
<input type="checkbox"/>	0207	View course operation report	details
<input type="checkbox"/>	0208	View course operation result	details
<input type="checkbox"/>	0219	View course information	details
<input type="checkbox"/>	0220	View course statistic information	details
<input type="checkbox"/>	1515	View history study information	details
<input type="checkbox"/>	1604	Student View Record	details
<input type="checkbox"/>	1605	Student View Schedule	details
<input type="checkbox"/>	1606	Faculty View Courses	details
<input type="checkbox"/>	0201	Save course details	details
<input type="checkbox"/>	0202	Save course operation status	details
<input type="checkbox"/>	0203	Save course operation report	details
<input type="checkbox"/>	0204	Save course operation result	details

รูปที่ 4.6 หน้าจอแสดงผลลัพธ์เอกสารคำอธิบายยูสเคสที่ค้นคืนแบบไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool	
Use Case Description	
Elements	Details
Business Domain	Course Management
Name:	View course operation details
Objective:	This use case is to view details of operations consisting of open, close, and edit course
Actor:	Department leader, lecturer, teacher
Relationship:	Association: Department leader, lecturer, teacher
	Include:
	Extend:
	Generalization:
Precondition:	A staff logged in the system
Postcondition:	Course operation details is displayed
Normal Flow:	<ol style="list-style-type: none"> 1. A staff selects a type of operation which wants to view. <ul style="list-style-type: none"> If a staff selects view course opening details, see S-1. If a staff selects view course closing details, see S-2. If a staff selects view course editing details, see S-3.
Subflow:	S-1 Save course opening result <ol style="list-style-type: none"> 1. The system displays a list of course which is operating course opening. 2. A staff selects course which wants to save course opening result. 3. A staff saves course opening result to the system.
	S-2 Save course closing result
Alternative Flow:	S-1 View course opening details <ol style="list-style-type: none"> 1. The system displays a list of course which is operating course opening. 2. A staff selects course which wants to view course opening details. 3. The system displays course opening details
	S-2 View course closing details

รูปที่ 4.7 หน้าจอแสดงรายละเอียดของเอกสารคำอธิบายยูสเคสที่ผู้ใช้เลือกเข้าไปดู

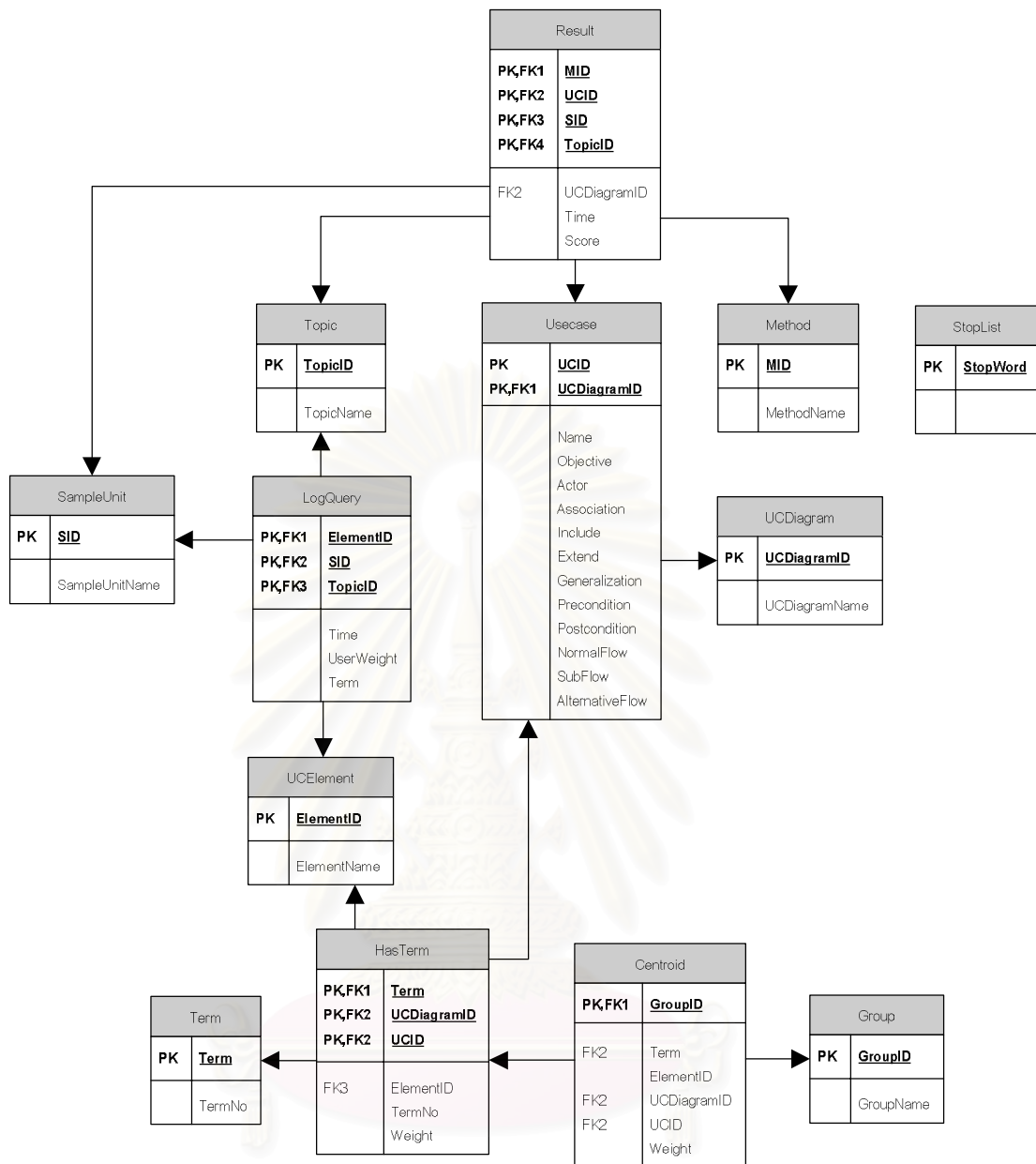
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Use Case Retrieval using Use Case Clustering and User Relevance Feedback Tool			
Use Cases Results by Element Separation on Probabilistic Model with User Relevance Feedback			
No	Use Case ID	Use Case Name	Details
1	0201	Save course details	details
2	0202	Save course operation status	details
3	0203	Save course operation report	details
4	0204	Save course operation result	details
5	0205	View course operation details	details
6	0206	View course operation status	details
7	0207	View course operation report	details
8	0208	View course operation result	details
9	0217	Save course information	details
10	0218	Edit course information	details
11	0219	View course information	details
12	0220	View course statistic information	details
13	1505	Announce course information	details
14	1509	View course detail	details
15	1510	Register course	details
16	1512	Study registered course	details
17	1603	Student View All Courses	details
18	1606	Faculty View Courses	details
19	1612	Add another course	details
20	1619	Add existing course	details

รูปที่ 4.8 หน้าจอแสดงผลลัพท์เอกสารคำอธิบายยูสเคสที่ค้นคืน
หลังใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช

4.4 แบบจำลองข้อมูล

การออกแบบแบบจำลองข้อมูลของระบบการจัดเก็บและค้นคืนยูสเคสโดยมีการจัดกลุ่มยูสเคส และการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้น เป็นการออกแบบโครงสร้างและความสัมพันธ์ระหว่างข้อมูลภายในระบบ โดยใช้ฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ซึ่งออกแบบให้รองรับทั้งข้อมูลเอกสารคำอธิบายยูสเคส ข้อมูลที่จำเป็นในการทดลอง เช่น รหัสหน่วยตัวอย่าง ข้อมูลข้อคำถาม และข้อมูลผลการทดลอง เป็นต้น โดยแบบจำลองข้อมูลทั้งหมดสามารถออกแบบเป็นแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entities Relationship Diagram: ER Diagram) ดังรูปที่ 4.9



รูปที่ 4.9 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูลของระบบการค้นคืนยูสเคสโดยการจัดกลุ่มยูสเคสและการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

จากแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูลสามารถแปลงเป็นตารางได้ทั้งหมด 13 ตาราง ดังรายละเอียดในตารางที่ 4.1

ตารางที่ 4.1 อธิบายตารางข้อมูลของระบบ

หมายเลข	ชื่อตาราง	คำอธิบาย
1	UseCase	ตารางจัดเก็บเอกสารคำอธิบายยูสเคส
2	UCDiagram	ตารางจัดเก็บชื่อโดเมนหรือแผนภาพของยูสเคส
3	UCElement	ตารางจัดเก็บชื่อส่วนประกอบของเอกสารคำอธิบายยูสเคส
4	Term	ตารางจัดเก็บคำและจำนวนของคำทั้งหมดในชุดข้อมูล
5	HasTerm	ตารางจัดเก็บคำทั้งหมดและจำนวนคำแต่ละคำในแต่ละเอกสารคำอธิบายยูสเคส
6	Centroid	ตารางจัดเก็บคำและจำนวนคำทั้งหมดในแต่ละกลุ่มของเอกสารคำอธิบายยูสเคส
7	LogQuery	ตารางจัดเก็บคำและน้ำหนักความสำคัญของแต่ละส่วนประกอบที่หน่วยทดสอบใช้ในการค้นคืนเอกสารคำอธิบายยูสเคสในแต่ละวิธีการค้นคืน
8	Method	ตารางจัดเก็บประเภทวิธีการค้นคืนเอกสารคำอธิบายยูสเคส
9	Group	ตารางจัดเก็บชื่อกลุ่มของยูสเคส
10	SampleUnit	ตารางจัดเก็บชื่อของหน่วยทดสอบ
11	Result	ตารางจัดเก็บผลการทดลองทั้งหมด
12	Topic	ตารางจัดเก็บโจทย์ปัญหาที่ใช้ในการทดลอง
13	Stoplist	ตารางจัดเก็บคำ Stop words ทั้งหมด

จากตารางอธิบายตารางข้อมูลของระบบ ผู้วิจัยได้สร้างโครงสร้างของแต่ละตารางข้อมูลดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.2 โครงสร้างตารางข้อมูล UseCase

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>UCID</u>	รหัสยูสเคส	varchar	4	PK
<u>UCDiagramID</u>	รหัสแผนภาพยูสเคส	char	2	PK,FK
Name	ชื่อยูสเคส	text		
Objective	วัตถุประสงค์	text		
Actor	ผู้กระทำหลัก	text		
Association	แอสโซซิเอชัน	text		
Include	การรวม	text		
Extend	การขยาย	text		
Generalization	เจเนอรัลไลเซชัน	text		
Precondition	เงื่อนไขก่อนทำงาน	text		
Postcondition	เงื่อนไขหลังทำงาน	text		
NormalFlow	สายงานปกติ	text		
SubFlow	สายงานย่อย	text		
AlternativeFlow	สายงานทางเลือก	text		

ตารางที่ 4.3 โครงสร้างตารางข้อมูล UCDiagram

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>UCDiagramID</u>	รหัสแผนภาพยูสเคส	char	2	PK
UCDiagramName	ชื่อแผนภาพยูสเคส	char	50	

ตารางที่ 4.4 โครงสร้างตารางข้อมูล UCElement

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>ElementID</u>	รหัสส่วนประกอบยูสเคส	char	2	PK
ElementName	ชื่อส่วนประกอบยูสเคส	char	30	

ตารางที่ 4.5 โครงสร้างตารางข้อมูล Term

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>Term</u>	คำ	char	30	PK
TermNo	จำนวนคำ	int	10	

ตารางที่ 4.6 โครงสร้างตารางข้อมูล HasTerm

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>Term</u>	คำ	char	30	PK,FK
<u>UCID</u>	รหัสยูสเคส	varchar	4	PK,FK
<u>UCDiagramID</u>	รหัสแผนภาพยูสเคส	char	2	PK,FK
ElementID	รหัสส่วนประกอบยูสเคส	char	2	FK
TermNo	จำนวนคำ	int	10	
Weight	น้ำหนักคำ	float		

ตารางที่ 4.7 โครงสร้างตารางข้อมูล Centroid

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>GroupID</u>	รหัสกลุ่มยูสเคส	char	2	PK,FK
Term	คำ	char	30	FK
ElementID	รหัสส่วนประกอบยูสเคส	char	2	
<u>UCDiagramID</u>	รหัสแผนภาพยูสเคส	char	2	FK
<u>UCID</u>	รหัสยูสเคส	varchar	4	FK
weight	น้ำหนักคำ	float		

ตารางที่ 4.8 โครงสร้างตารางข้อมูล LogQuery

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>SID</u>	รหัสหน่วยทดสอบ	char	2	PK,FK
<u>TopicID</u>	รหัสโจทย์ปัญหา	char	1	PK,FK
Time	ลำดับครั้งการค้นคืน	char	1	
<u>ElementID</u>	รหัสส่วนประกอบยูสเคส	char	2	PK,FK
Userweight	น้ำหนักจากผู้ใช้	char	1	
Term	คำ	char	30	

ตารางที่ 4.9 โครงสร้างตารางข้อมูล Method

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>MID</u>	รหัสวิธีการค้นคืน	char	1	PK
MethodName	ชื่อวิธีการค้นคืน	text		

ตารางที่ 4.10 โครงสร้างตารางข้อมูล Group

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>GroupID</u>	รหัสกลุ่มยูสเคส	char	2	PK
GroupName	ชื่อกลุ่มยูสเคส	text		

ตารางที่ 4.11 โครงสร้างตารางข้อมูล SampleUnit

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>SID</u>	รหัสหน่วยทดสอบ	char	2	PK
Name	ชื่อหน่วยทดสอบ	varchar	30	

ตารางที่ 4.12 โครงสร้างตารางข้อมูล Result

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>SID</u>	รหัสหน่วยทดสอบ	char	2	PK,FK
<u>TopicID</u>	รหัสโจทย์ปัญหา	char	1	PK,FK
Time	ลำดับครั้งการค้นคืน	char	1	
<u>MID</u>	รหัสวิธีการค้นคืน	char	1	PK,FK
<u>UCID</u>	รหัสยูสเคส	varchar	4	PK,FK
Score	คะแนนความคล้าย	float		
UCDiagramID	รหัสแผนภาพยูสเคส	char	2	FK

ตารางที่ 4.13 โครงสร้างตารางข้อมูล Topic

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>TopicID</u>	รหัสโจทย์ปัญหา	char	1	PK
TopicName	ชื่อโจทย์ปัญหา	char	50	

ตารางที่ 4.14 โครงสร้างตารางข้อมูล Stoplist

คุณลักษณะ	คำอธิบาย	ชนิด	ขนาด	คีย์
<u>stopword</u>	คำ Stop words	char	20	PK

บทที่ 5

การทดลอง

ในบทนี้จะกล่าวถึงรายละเอียดของการทดลอง เพื่อทำการทดสอบแนวคิดที่ได้นำเสนอด้วยเครื่องมือซึ่งผู้วิจัยพัฒนาขึ้นมา ประเด็นต่างๆ ในการทดลองประกอบด้วย วัตถุประสงค์ของการทดลอง วิธีการทดลอง ขั้นตอนการทดลอง สภาพแวดล้อมของการทดลอง ผลการทดลอง การวิเคราะห์ผลการทดลอง และข้ออภิปรายต่างๆ จากการทดลอง ซึ่งรายละเอียดมีดังนี้

5.1 วัตถุประสงค์ของการทดลอง

วัตถุประสงค์ของการทดลองนี้คือ เพื่อทดสอบสมมุติฐานของงานวิจัย นั่นคือ การค้นคืนยูสเคสโดยมีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้นมีประสิทธิภาพมากกว่าการค้นคืนยูสเคสที่ไม่ใช่ให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้หรือไม่ โดยเปรียบเทียบประสิทธิภาพทั้งในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น ซึ่งในทุกวิธีการของการค้นคืนนั้นได้ผ่านขั้นตอนของการจัดกลุ่มของยูสเคสเรียบร้อยแล้ว นอกจากนี้ยังเพิ่มวิธีการค้นคืนแบบไม่พิจารณาโครงสร้างของเอกสารยูสเคสขึ้นมา เพื่อเปรียบเทียบประสิทธิภาพกับการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ซึ่งเป็นวิธีการค้นคืนหลักในงานวิจัยนี้ ทำให้มีวิธีการค้นคืนยูสเคสทั้งสิ้น 8 วิธีการด้วยกัน โดยจะกล่าวถึงในหัวข้อที่ 5.2.3

ดังนั้น ในการทดลองจึงมีการตั้งเงื่อนไขและปัจจัยต่างๆ เพื่อเปรียบเทียบประสิทธิภาพของวิธีการในการค้นคืนยูสเคสแบบต่างๆ และผลลัพธ์ที่ได้จากแต่ละวิธีการค้นคืนในการทดลอง จะใช้ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก มาประเมินประสิทธิภาพของวิธีการที่นำเสนอต่อไป

5.2 วิธีการทดลอง

การออกแบบวิธีการทดลองที่ดีนั้น ปัจจัยที่ควบคุมได้สามารถช่วยลดความโน้มเอียงหรืออคติที่อาจเกิดขึ้นและส่งผลต่อการทดลอง โดยในการทดลองของงานวิทยานิพนธ์นี้ ปัจจัยควบคุมที่กำหนดขึ้นมีดังต่อไปนี้

5.2.1 ยูสเคส (Use Cases)

เอกสารคำอธิบายยูสเคสที่ใช้ในการทดลองการค้นคืนยูสเคสนี้มีจำนวนทั้งสิ้น 611 เอกสาร จาก 40 ระบบงาน เขียนอธิบายเป็นภาษาอังกฤษ รายละเอียดและจำนวนของยูสเคสในแต่ละระบบงาน แสดงไว้ในภาคผนวก ก

5.2.2 หน่วยตัวอย่าง (Sample Units)

สำหรับการทดลองนี้ใช้หน่วยตัวอย่าง 10 หน่วยตัวอย่าง มาเป็นผู้ทำการทดสอบเครื่องมือที่พัฒนาขึ้น โดยหน่วยตัวอย่างทั้งหมดเป็นนิสิตที่กำลังศึกษาระดับปริญญาโทในสาขาวิชาวิศวกรรมซอฟต์แวร์ ซึ่งมีประสบการณ์เกี่ยวกับการเก็บรวบรวม วิเคราะห์ และบันทึก

ข้อกำหนดความต้องการซอฟต์แวร์ รวมทั้งมีความสามารถในการออกแบบซอฟต์แวร์ด้วยแบบจำลองยูสเคส นอกจากนี้แล้วหน่วยตัวอย่างยังเป็นผู้มีความรู้ภาษาอังกฤษในเกณฑ์ที่ดี โดยมีผลคะแนนการสอบโทเฟล (TOEFL) หรือเทียบเท่า ไม่ต่ำกว่า 500 คะแนน

5.2.3 วิธีการค้นคืนยูสเคส (Use Case Retrieval Methods)

เพื่อสนับสนุนแนวคิดที่น่าเสนอ ในการที่จะประเมินประสิทธิผลของวิธีการค้นคืนที่มีทั้งแบบพิจารณาโครงสร้างและไม่พิจารณาโครงสร้างของยูสเคส บนแบบจำลองทั้ง 2 แบบ รวมทั้งในส่วนของการใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน จึงมีการพัฒนาเครื่องมือให้สอดคล้องกับการทดลอง โดยมีฟังก์ชันการค้นคืนยูสเคสที่รองรับการทดลอง 8 วิธี ซึ่งทุกวิธีจะค้นคืนเอกสารยูสเคสที่ได้ถูกจัดกลุ่มเอาไว้เรียบร้อยแล้ว ดังนี้

- 1) การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์
- 2) การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์
- 3) การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น
- 4) การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น
- 5) การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์
- 6) การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์
- 7) การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น
- 8) การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น

วิธีการค้นคืนยูสเคสทั้ง 8 วิธีการนั้น เมื่อนำไปใช้กับหน่วยตัวอย่างในขั้นตอนการทดลอง หน่วยทดลองจะเห็นวิธีการค้นคืนเพียง 4 วิธีเท่านั้น เนื่องจากการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้น เป็นวิธีการทดลองที่สามารถทำได้ต่อจากวิธีการที่ไม่พิจารณาผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้นเอง

ตัวอย่างเช่น กรณีที่หน่วยตัวอย่างทำการทดลองวิธีการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์ คือ วิธีการค้นคืนที่ 1 เมื่อได้ผลลัพธ์ออกมาเป็นเอกสารยูสเคสจำนวนหนึ่งแล้วนั้น หน่วยตัวอย่างทำการผลป้อนกลับที่ตรงประเด็นเข้าสู่ระบบ ก็จะได้ผลลัพธ์สำหรับวิธีการค้นคืนที่ 2 ตามมาด้วย

ดังนั้น หากหน่วยตัวอย่างทำการทดลองกับวิธีการค้นคืนที่ 1, 3, 5, และ 7 ก็จะถือว่าได้วิธีการค้นคืนที่ 2, 4, 6, และ 8 ตามลำดับ หากมีการพิจารณาผลป้อนกลับที่ตรงประเด็นจากผู้

5.2.4 โจทย์ปัญหาที่ใช้ในการทดลอง (Problem Topics used in the Experiment)

ในส่วนของโจทย์ปัญหาที่ใช้ในการทดลองนี้กำหนดขึ้นเพื่อให้หน่วยตัวอย่างป้อนข้อความในการค้นหาเอกสารคำอธิบายยูสเคสนั้น หน่วยตัวอย่างจะป้อนข้อความโดยสมมุติสถานการณ์ว่าเป็นผู้ที่กำลังต้องการออกแบบและพัฒนาซอฟต์แวร์ และต้องการค้นหาข้อมูลจากเอกสารคำอธิบายยูสเคสที่มีการออกแบบและพัฒนาไว้แล้ว มาใช้ในงานของตน ตามที่ผู้วิจัยได้กำหนดโจทย์ขึ้น หลังจากทราบรายละเอียดอย่างย่อของแต่ละระบบงานในชุดข้อมูลเรียบร้อยแล้ว สำหรับโจทย์ปัญหาที่ใช้ในการทดลองทั้งหมดในงานวิทยานิพนธ์นี้มี 5 โจทย์ ได้แก่

- 1) Educational System
- 2) Material Requirements Planning/ Manufacturing Planning
- 3) Member / User's Information Management
- 4) Product Trading Operation
- 5) Report Generation

ลักษณะของแต่ละโจทย์ปัญหานั้นมีการกำหนดเพื่อให้เกิดค่าโน้มเอียงหรืออคติ (Bias) น้อยคือ มีความหลากหลายของข้อความทั้งในมุมมองเชิงกว้าง ปานกลาง และลึก เนื่องจากหน่วยตัวอย่างแต่ละหน่วยนั้น อาจมีความรู้ในเรื่องที่ค้นคืนแตกต่างกันไป โดยในโจทย์ข้อที่ 1 จะเป็นเรื่องของระบบการศึกษา ซึ่งมีมุมมองเชิงกว้างที่สุด โจทย์ข้อที่ 2 เกี่ยวกับการวางแผนความต้องการวัสดุหรือวัตถุดิบ หรือเกี่ยวกับการวางแผนการผลิต ซึ่งเป็นเรื่องที่มีมุมมองที่ลึกกว่าโจทย์แรก ในส่วนของโจทย์ข้อที่ 3 เกี่ยวกับการจัดการสมาชิก หรือการจัดการข้อมูลผู้ใช้ ซึ่งเป็นเรื่องที่ลึกลงมาอีก โดยเน้นไปที่การจัดการข้อมูลกลุ่มคนซึ่งเป็นสมาชิกหรือผู้ใช้ทั่วไป แต่ไม่จำกัดว่าเกี่ยวข้องกับงานด้านใด ส่วนในโจทย์ข้อที่ 4 จะเกี่ยวกับการดำเนินการด้านการขายสินค้าซึ่งมีมุมมองที่ลึกขึ้นอีกเพราะเกี่ยวข้องกับงานด้านการขายสินค้าเท่านั้น และในโจทย์ข้อสุดท้ายคือ การสร้าง หรือการจัดทำรายงานนั้น เป็นโจทย์ที่เฉพาะเจาะจง เพราะเกี่ยวข้องกับฟังก์ชันหรือการทำงานหนึ่ง ๆ ในระบบงาน จึงมีมุมมองที่ลึกที่สุดในบรรดาโจทย์ทั้งหมด

เหตุผลที่เลือก 5 โจทย์ปัญหานี้มา เนื่องจาก ผู้วิจัยได้ทำการตรวจสอบเอกสารคำอธิบายยูสเคสแต่ละเอกสารในแต่ละโดเมนหรือระบบงาน ซึ่งเก็บไว้ในชุดข้อมูลทั้งหมดและพบว่า มีชุดของคำตอบที่รองรับกับข้อความจากการค้นคืนของผู้ใช้ได้

หัวข้อเรื่องที่กำหนดขึ้นมาทั้ง 5 เรื่องนี้ จะนำมาใช้ในการเปรียบเทียบประสิทธิภาพของระบบการค้นคืนกับวิธีการค้นคืนยูสเคสที่ได้ระบุไว้ใน 5.2.3 โดยในแต่ละหัวข้อเรื่องนั้น ทางผู้วิจัยได้กำหนดชุดคำตอบ (Answer Set) และจัดเก็บไว้ในฐานข้อมูล สำหรับใช้ในการคำนวณเพื่อหาค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก

5.2.5 ข้อคำถาม (Queries)

หน่วยตัวอย่างจะใช้ข้อคำถามในการค้นคืนยูสเคสทั้งแบบอาศัยแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น ซึ่งในแต่ละแบบจำลองนั้น จะค้นคืนทั้งแบบไม่พิจารณาและพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส ซึ่งจะสามารถทำการค้นคืนเอกสารคำอธิบายยูสเคสออกมาแสดงให้กับหน่วยตัวอย่างได้ทั้งการค้นคืนแบบไม่ใช้และแบบใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้อย่างนั้น หน่วยตัวอย่างหนึ่ง ๆ ต้องทำการค้นคืนด้วยข้อคำถามในโจทย์ปัญหาทั้ง 5 ข้อ และในแต่ละข้อนั้นใช้ข้อคำถาม 5 ข้อคำถาม เพื่อให้มีความหลากหลายของข้อคำถาม จึงมีข้อคำถาม 25 ข้อคำถาม สำหรับวิธีการค้นคืนทั้งหมดซึ่งหมายความว่าข้อคำถามเป็นตัวแปรหนึ่งซึ่งจะคงที่ไม่เปลี่ยนแปลงไปตามวิธีการค้นคืน

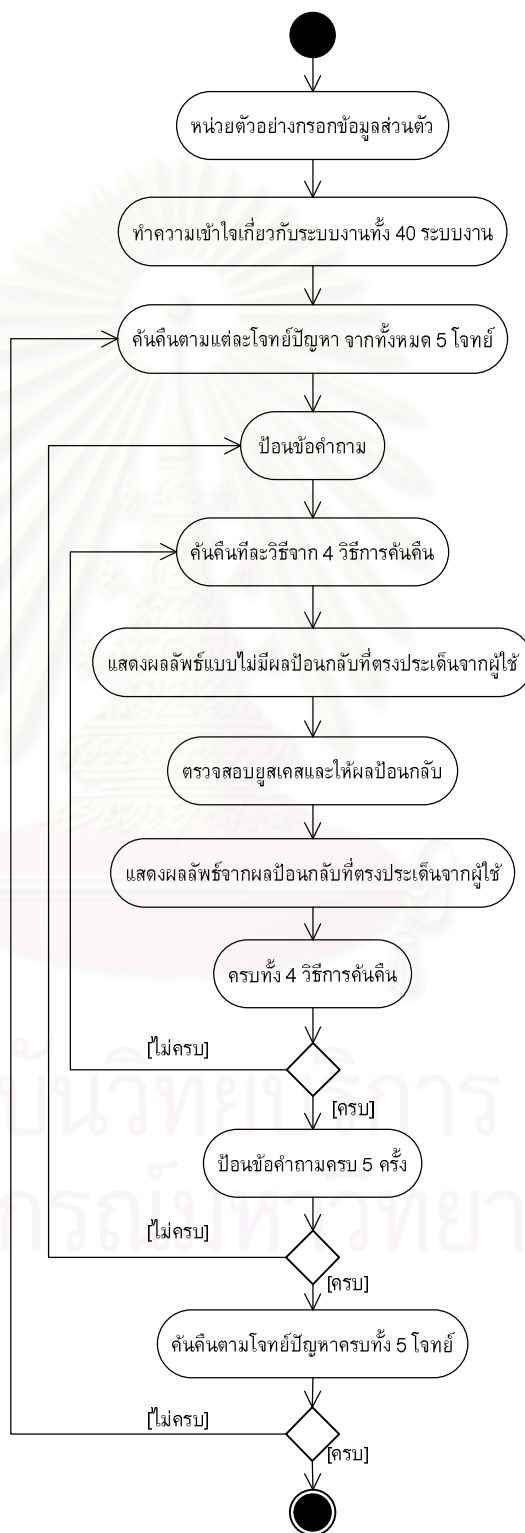
ในท้ายการทดลองนั้นจะได้ชุดข้อคำถามทั้งหมด 250 ข้อคำถามจากหน่วยตัวอย่างทั้ง 10 หน่วยตัวอย่าง โดยในแต่ละโจทย์ปัญหาจะได้ชุดข้อคำถามเรื่องละ 50 ข้อคำถาม ชุดของข้อคำถามเหล่านี้ เมื่อนำไปใช้ในขั้นตอนการค้นคืนเอกสารคำอธิบายยูสเคสออกมาแสดงผลเรียบร้อยแล้ว จะนำไปตรวจสอบกับชุดคำตอบและคำนวณหาค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก เพื่อให้สามารถเข้าใจได้ง่ายขึ้นเกี่ยวกับจำนวนข้อคำถามที่ได้จากหน่วยตัวอย่างทั้งหมดแสดงไว้ในตารางที่ 5.1

ตารางที่ 5.1 แสดงจำนวนข้อคำถามทั้งหมดที่ได้จากหน่วยตัวอย่าง

หน่วยตัวอย่าง \ โจทย์ปัญหาข้อที่	1	2	3	4	5	รวมข้อคำถามจากแต่ละหน่วยตัวอย่าง
1	5	5	5	5	5	25
2	5	5	5	5	5	25
3	5	5	5	5	5	25
4	5	5	5	5	5	25
5	5	5	5	5	5	25
6	5	5	5	5	5	25
7	5	5	5	5	5	25
8	5	5	5	5	5	25
9	5	5	5	5	5	25
10	5	5	5	5	5	25
ข้อคำถามทั้งหมดจาก แต่ละโจทย์	50	50	50	50	50	250

5.3 ขั้นตอนการทดลอง

หน่วยตัวอย่างแต่ละหน่วย จะต้องดำเนินการตามขั้นตอนการทดลองซึ่งมีลำดับการดำเนินการ 8 ขั้นตอนด้วยกัน ดังแสดงด้วยแผนภาพกิจกรรมขั้นตอนการทดลอง ในรูปที่ 5.1



รูปที่ 5.1 แผนภาพกิจกรรมขั้นตอนการทดลอง

รายละเอียดของแต่ละขั้นตอนในแผนภาพกิจกรรมนี้เริ่มต้นจาก

- 1) หน่วยตัวอย่างป้อนข้อมูลส่วนตัวของตน ซึ่งได้แก่ ชื่อหน่วยตัวอย่าง
 - 2) หน่วยตัวอย่างอ่านทำความเข้าใจเกี่ยวกับรายละเอียดอย่างย่อของระบบงานทั้ง 40 ระบบงาน จากคำอธิบายระบบ ซึ่งแสดงผ่านส่วนต่อประสานผู้ใช้
 - 3) หน่วยตัวอย่างตอบโจทยปัญหาสำหรับการค้นคืนเอกสารคำอธิบายยูสเคสที่ละโจทย จากทั้งหมด 5 โจทย ตามหัวข้อที่ 5.2.4
 - 4) หน่วยตัวอย่างป้อนข้อคำถามที่เกี่ยวข้องกับโจทยปัญหานั้นๆ โดยในแต่ละโจทย หน่วยตัวอย่างจะต้องป้อนข้อคำถามทั้งหมด 5 ครั้ง
 - 5) หน่วยตัวอย่างใช้วิธีการค้นคืนยูสเคสแต่ละวิธี ตามหัวข้อที่ 5.2.3
 - 6) ระบบแสดงผลลัพธ์เป็นชื่อระบบงานและชื่อของเอกสารคำอธิบายยูสเคสพร้อมทั้งรายละเอียดของเอกสารซึ่งมีค่าความคล้ายกับข้อคำถามจากหน่วยตัวอย่าง ออกมาแสดง
 - 7) หน่วยตัวอย่างตรวจสอบเอกสารคำอธิบายยูสเคส และเลือกเอกสารที่มีความเกี่ยวข้องหรือตรงกับความต้องการของตน แล้วจึงป้อนผลเอกสารเหล่านั้นกลับเข้าสู่ระบบให้ทำการค้นคืนอีกครั้ง
 - 8) ระบบแสดงผลลัพธ์เป็นชื่อระบบงานและชื่อของเอกสารคำอธิบายยูสเคสที่คำนวณจากผลป้อนกลับที่ตรงประเด็นจากผู้ที่ใช้หน่วยตัวอย่างป้อนกลับมา
- ทั้ง 8 ขั้นตอนดังกล่าวมาถือเป็นการสิ้นสุดขั้นตอนในการค้นคืนยูสเคส ในแต่ละครั้งที่หน่วยตัวอย่างเลือกโจทยที่จะค้นหายูสเคสข้อหนึ่ง โดยป้อนข้อคำถามเพียงครั้งเดียว และใช้วิธีการค้นคืนวิธีเดียว กระบวนการของกิจกรรมนี้จะสิ้นสุดอย่างสมบูรณ์ต่อเมื่อหน่วยตัวอย่างแต่ละหน่วยตัวอย่างใช้วิธีการค้นคืนครบทั้ง 4 วิธีการในแต่ละครั้งที่ป้อนข้อคำถาม และป้อนข้อคำถามครบทั้ง 5 ครั้งในแต่ละโจทย สุดท้ายคือตอบโจทยปัญหาครบทั้ง 5 ข้อ

5.4 การกำหนดค่าขีดแบ่งเริ่มต้นความคล้าย

ค่าขีดแบ่งเริ่มต้นความคล้าย (Similarity Threshold) เป็นตัวเลขค่าหนึ่งที่ใช้เพื่อกำหนดให้ระบบการค้นคืนทำการเปรียบเทียบค่าความคล้ายที่หาได้ระหว่างข้อคำถามกับเอกสารคำอธิบายยูสเคสกับตัวเลขนี้ หากค่าความคล้ายเหล่านั้นมีค่ามากกว่าหรือเท่ากับค่าขีดแบ่งเริ่มต้นความคล้ายแล้ว จึงจะนำเอกสารเหล่านั้นออกมาแสดงกับผู้ที่ทำการค้นคืนต่อไป ผู้วิจัยใช้ชุดข้อคำถามจากงานวิจัยของ [7] มาทดสอบกับเครื่องมือของงานวิจัยนี้ ซึ่งมีวิธีการค้นคืนดังกล่าวในหัวข้อที่ 5.2.3 และค้นคืนยูสเคสจากชุดข้อมูลยูสเคสทั้งหมด 611 เอกสารยูสเคส โดยค่าขีดแบ่งเริ่มต้นความคล้ายหาได้จากค่าเฉลี่ยที่ได้หักออกด้วยค่าเบี่ยงเบนมาตรฐาน ดังสูตรในสมการที่ 22

$$\text{ค่าขีดแบ่งเริ่มต้นความคล้าย} = \text{ค่าเฉลี่ย} - \text{ค่าเบี่ยงเบนมาตรฐาน} \quad (22)$$

ผลของค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานความคล้อยที่ได้จากวิธีการคั่นคืนตามงานวิจัยนี้แบ่งออกเป็น 2 ค่าด้วยกัน นั่นคือ ผลของค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานความคล้อยที่เกิดระหว่างข้อคำถามกับตัวแทนยูสเคสในแต่ละกลุ่ม กับค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานความคล้อยที่เกิดระหว่างข้อคำถามกับเอกสารคำอธิบายยูสเคส จึงได้ค่าขีดแบ่งเริ่มต้นความคล้อยระหว่างข้อคำถามกับตัวแทนยูสเคสในแต่ละกลุ่มและค่าขีดแบ่งเริ่มต้นความคล้อยระหว่างข้อคำถามกับเอกสารคำอธิบายยูสเคสมีค่าเป็น 0.0000629289 และ 0.000796947 ตามลำดับ

5.5 ผลการทดลอง

ผลลัพธ์ทั้งหมดที่ได้จากการทดลองทั้ง 8 วิธีการคั่นคืน ที่แสดงไว้ในภาคผนวก ค นั้น นำมาสรุปเป็นค่าเฉลี่ยของค่าเรียกคืน (R) ค่าความแม่นยำ (P) และค่าเฉลี่ยฮาร์โมนิก (F) โดยเปรียบเทียบระหว่างกรณีที่ใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ที่ได้จากแบบจำลองทั้ง 2 แบบ ในตารางที่ 5.2 ถึง 5.5 โดยในส่วนของวิธีการคั่นคืนที่ไม่พิจารณาโครงสร้างของเอกสารยูสเคส แสดงผลไว้ในตารางที่ 5.2 และ 5.3 และในส่วนของวิธีการคั่นคืนแบบพิจารณาโครงสร้างของเอกสารยูสเคส แสดงผลไว้ในตารางที่ 5.4 และ 5.5 ค่าเฉลี่ยของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิกเหล่านี้ แสดงโดยแยกตามโจทย์ปัญหาทั้ง 5 ข้อในการคั่นคืน เพื่อใช้ในการวิเคราะห์ประสิทธิผลของการคั่นคืนในวิธีต่างๆ จากข้อคำถามที่มีความหลากหลายในระดับเดียวกัน ซึ่งมีอยู่ทั้งสิ้น 5 ระดับ ดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.2 ค่าเฉลี่ยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค ของการค้นคืนแบบไม่พิจารณาโครงสร้างของเอกสารยูสเคส เปรียบเทียบการใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ของแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น

โจทย์ ข้อที่	แบบจำลองปริภูมิเวกเตอร์									แบบจำลองความน่าจะเป็น								
	ไม่ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			ใช้ผลป้อนกลับที่ตรง ประเด็นจากผู้ใช้			เพิ่มขึ้น/ลดลง ร้อยละ			ไม่ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			เพิ่มขึ้น/ลดลง ร้อยละ		
	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
1	0.447	0.612	0.479	0.608	0.458	0.452	+36.16	-25.21	-5.58	0.495	0.634	0.504	0.557	0.47	0.423	+12.68	-25.75	-16.04
2	0.372	0.833	0.358	0.472	0.338	0.355	+26.68	-59.49	-0.75	0.442	0.422	0.398	0.423	0.354	0.340	-4.23	-16.16	-14.52
3	0.437	0.136	0.200	0.741	0.099	0.172	+69.76	-27.27	-14.38	0.117	0.054	0.062	0.457	0.075	0.123	+290.6	+39.43	+99.68
4	0.419	0.300	0.338	0.515	0.241	0.307	+22.70	-19.77	-9.40	0.459	0.282	0.332	0.518	0.250	0.316	+12.87	-11.56	-4.76
5	0.719	0.423	0.513	0.808	0.318	0.428	+12.46	-24.73	-16.41	0.641	0.370	0.412	0.781	0.374	0.480	+21.86	+0.97	+16.64
ค่าเฉลี่ย	0.479	0.372	0.378	0.629	0.291	0.343	+31.34	-31.29	-9.30	0.431	0.352	0.341	0.547	0.304	0.337	+27.09	-2.61	+16.2

จากตารางที่ 5.2 พบว่าการคั่นคั้นยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ โดยมีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน จากทุกโจทย์ปัญหา มีค่าเรียกคืน มากกว่าแบบที่ไม่ใช้การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ในขณะที่ ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิกมีค่าลดลง แสดงให้เห็นว่า การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งานในการคั่นคั้นแบบนี้ นั้น ทำให้ระบบสามารถคั่นคั้นเอกสารที่ตรงกับความต้องการของผู้ใช้งานได้มากขึ้น เนื่องจากค่าเรียกคืนเพิ่มขึ้น แต่ก็ทำให้ไปถึงเอาเอกสารที่ไม่ตรงกับความต้องการออกมาด้วย โดยสังเกตจากค่าความแม่นยำมีค่าลดลง จึงเป็นผลให้ค่าเฉลี่ยฮาร์โมนิกลดลงด้วยเช่นกัน

ส่วนการคั่นคั้นยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น โดยมีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน โดยส่วนใหญ่ให้ค่าเรียกคืนมากกว่าแบบไม่ใช้การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน มีเพียงโจทย์ปัญหาข้อที่ 2 ที่มีค่าเรียกคืนลดลง คือลดลงร้อยละ 4.23 สำหรับค่าความแม่นยำ พบว่าในโจทย์ข้อที่ 3 และ 5 เท่านั้นที่เพิ่มขึ้น แต่โดยเฉลี่ยแล้ว ค่าความแม่นยำลดลงร้อยละ 2.61 แสดงให้เห็นว่าการคั่นคั้นแบบนี้ นั้น ส่วนใหญ่ทำให้คั่นคั้นยูสเคสที่ตรงกับความต้องการของผู้ใช้งานออกมาได้มากขึ้น แต่ก็ยังคั่นคั้นยูสเคสที่ไม่ตรงกับความต้องการออกมาเช่นกัน แต่โดยรวมเมื่อพิจารณาจากค่าเฉลี่ยฮาร์โมนิกมีค่าเพิ่มขึ้นถึงร้อยละ 16.2

ตารางที่ 5.3 เปรียบเทียบอัตราเพิ่มขึ้น/ลดลงของค่าเฉลี่ยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก ของการคั่นคั้นแบบไม่พิจารณาโครงสร้างของเอกสารยูสเคส ระหว่างแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น

โจทย์ข้อที่	แบบจำลองปริภูมิเวกเตอร์ เปรียบเทียบกับ แบบจำลองความน่าจะเป็น			แบบจำลองปริภูมิเวกเตอร์ เปรียบเทียบกับ แบบจำลองความน่าจะเป็น		
	ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน (เพิ่มขึ้น/ลดลง ร้อยละ)			ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน (เพิ่มขึ้น/ลดลง ร้อยละ)		
	R	P	F	R	P	F
1	-10.74	-3.59	-5.22	+8.39	-2.62	+6.42
2	-18.82	+49.34	-11.17	+10.38	-4.73	+4.23
3	+73.23	+60.29	+69.00	+38.33	24.24	+28.49
4	-9.55	+6.00	+1.78	-0.58	-3.73	-2.93
5	+10.85	+12.53	+19.69	+3.34	-17.61	-12.15
ค่าเฉลี่ย	+10.02	-15.86	+9.79	+13.04	-4.47	+1.75

หากพิจารณาระหว่าง 2 แบบจำลอง จากตารางที่ 5.3 ซึ่งเปรียบเทียบอัตราการเพิ่มขึ้น และลดลงของค่าเฉลี่ยของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค ระหว่างแบบจำลองปริภูมิเวกเตอร์ กับแบบจำลองความน่าจะเป็น ของการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคส ซึ่งนำตัวเลขจากตารางที่ 5.2 มาคำนวณ แยกระหว่างแบบไม่ใช้ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน โดยมีแบบจำลองปริภูมิเวกเตอร์เป็นตัวตั้งในการเปรียบเทียบ หมายความว่า หากตัวเลขในตารางมีค่าเป็นบวก แสดงว่าแบบจำลองปริภูมิเวกเตอร์ให้ผลลัพธ์ในหน่วยวัดนั้นๆ ได้ดีกว่า แต่หากจำนวนที่ได้มีค่าเป็นลบ แสดงว่าแบบจำลองความน่าจะเป็นให้ผลลัพธ์ในหน่วยวัดนั้นๆ ได้ดีกว่า

ในส่วนแรกแบบไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน แบบจำลองปริภูมิเวกเตอร์มีค่าเรียกคืนเฉลี่ยมากกว่าแบบจำลองความน่าจะเป็น แสดงว่าการค้นคืนแบบไม่พิจารณาโครงสร้างด้วยแบบจำลองปริภูมิเวกเตอร์และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน สามารถค้นคืนยูสเคสที่ตรงกับความต้องการของผู้ใช้ได้มากกว่าแบบจำลองความน่าจะเป็น แต่ก็ยังค้นคืนยูสเคสที่ไม่ตรงกับความต้องการของผู้ใช้ออกมาด้วยเช่นกัน โดยเห็นได้จากค่าความแม่นยำที่ติดลบในอีกแง่หนึ่งนั้นบ่งบอกว่า การค้นคืนแบบไม่พิจารณาโครงสร้างด้วยแบบจำลองความน่าจะเป็น และไม่ใช้การให้ผลกลับที่ตรงประเด็นจากผู้ใช้งาน มีความแม่นยำมากกว่าแบบจำลองปริภูมิเวกเตอร์ด้วย แต่หากพิจารณาที่ค่าเฉลี่ยฮาร์โมนิคซึ่งมีค่าเพิ่มขึ้น 9.79 จะสรุปได้ว่า การค้นคืนแบบไม่พิจารณาโครงสร้างของยูสเคส และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์มีประสิทธิผลดีกว่าแบบจำลองความน่าจะเป็น

ในส่วนที่สองแบบใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน พบว่าเป็นไปในทิศทางเดียวกันกับในส่วนแรก นั่นคือ การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างด้วยแบบจำลองปริภูมิเวกเตอร์และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ให้ค่าเรียกคืนมากกว่าแบบจำลองความน่าจะเป็น แต่แบบจำลองความน่าจะเป็น โดยการค้นคืนแบบเดียวกันนี้ มีค่าความแม่นยำที่มากกว่าแบบจำลองปริภูมิเวกเตอร์ แต่หากพิจารณาที่ค่าเฉลี่ยฮาร์โมนิคซึ่งมีค่าเพิ่มขึ้น 1.75 จะสรุปได้ว่า การค้นคืนแบบไม่พิจารณาโครงสร้างของยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์มีประสิทธิผลดีกว่าแบบจำลองความน่าจะเป็น

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.4 ค่าเฉลี่ยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค ของการค้นคืนแบบพิจารณาโครงสร้างของเอกสารยูสเคส เปรียบเทียบการใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ของแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น

โจทย์ ข้อที่	แบบจำลองปริภูมิเวกเตอร์									แบบจำลองความน่าจะเป็น								
	ไม่ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			เพิ่มขึ้น/ลดลง ร้อยละ			ไม่ใช้ผลป้อนกลับที่ ตรงประเด็นจากผู้ใช้			ใช้ผลป้อนกลับที่ตรง ประเด็นจากผู้ใช้			เพิ่มขึ้น/ลดลง ร้อยละ		
	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
1	0.316	0.702	0.409	0.317	0.8	0.433	-0.19	+13.92	+5.87	0.364	0.654	0.433	0.375	0.501	0.429	+3.03	-23.41	-0.97
2	0.195	0.451	0.251	0.201	0.553	0.280	+3.34	+22.63	+11.94	0.231	0.426	0.262	0.224	0.491	0.276	-2.94	+15.32	+5.14
3	0.279	0.178	0.183	0.257	0.217	0.214	-8.06	+21.94	+16.97	0.343	0.195	0.208	0.432	0.173	0.221	+25.80	-11.26	+6.16
4	0.29	0.459	0.336	0.265	0.511	0.328	-8.69	+11.89	-2.32	0.339	0.441	0.346	0.342	0.486	0.37	+0.91	+10.28	+7.09
5	0.579	0.571	0.543	0.54	0.599	0.543	-6.67	+4.739	0	0.21	0.262	0.17	0.571	0.446	0.457	+172.1	+70.41	+168.84
ค่าเฉลี่ย	0.332	0.472	0.344	0.316	0.536	0.36	-4.05	+13.57	+6.49	0.297	0.395	0.284	0.389	0.444	0.35	+39.78	+12.45	+37.25

จากตารางที่ 5.4 เป็นค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก ที่ได้จากการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส ซึ่งใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น พบว่าการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส โดยใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ของแบบจำลองปริภูมิเวกเตอร์นั้น โดยเฉลี่ยมีค่าเรียกคืนลดลงร้อยละ 4.05 มีเพียงโจทย์ปัญหาข้อที่ 2 เท่านั้น ที่มีค่าเรียกคืนเพิ่มขึ้นร้อยละ 3.34 แต่ในส่วนของค่าความแม่นยำพบว่า ในโจทย์ปัญหาทุกข้อนั้นมีอัตราค่าเพิ่มขึ้น แสดงว่า การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส โดยใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์นั้น ช่วยให้ยูสเคสที่ระบบค้นคืนออกมาได้นั้น ตรงกับความต้องการของผู้ใช้เพิ่มมากขึ้นกว่าแบบที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ แม้ว่าจะไม่สามารถค้นคืนยูสเคสที่ตรงกับความต้องการออกมาได้มากกว่าแบบที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ซึ่งสังเกตได้จากค่าเรียกคืนที่ลดลง แต่ก็ทำให้ค่าเฉลี่ยฮาร์โมนิกมีค่าเพิ่มขึ้นร้อยละ 6.49

ในส่วนของ การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น พบว่าการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ส่วนใหญ่ให้ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก มากกว่าแบบที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ แสดงว่าการค้นคืนแบบที่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็นและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สามารถค้นคืนที่ตรงกับความต้องการของผู้ใช้ได้มากขึ้น นอกจากนี้ยูสเคสที่ระบบค้นมาได้ทั้งหมดนั้นเป็นเอกสารที่ตรงกับความต้องการมากขึ้นด้วยเช่นกัน สังเกตได้จาก ค่าเฉลี่ยของอัตราเพิ่มขึ้นของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก

แต่หากพิจารณาระหว่าง 2 แบบจำลอง จากตารางที่ 5.5 ซึ่งเปรียบเทียบอัตราการเพิ่มขึ้นและลดลงของค่าเฉลี่ยของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก ระหว่างแบบจำลองปริภูมิเวกเตอร์ กับแบบจำลองความน่าจะเป็น ของการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส จะกล่าวถึงเป็นสองส่วนด้วยกัน ในส่วนแรกจะเป็นแบบที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ พบว่า โดยเฉลี่ยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิกในแบบจำลองปริภูมิเวกเตอร์ มีค่ามากกว่าแบบจำลองความน่าจะเป็น แสดงให้เห็นว่าการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส โดยไม่มีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์สามารถค้นคืนยูสเคสได้ดีกว่าแบบจำลองความน่าจะเป็น

ในส่วนที่สอง แบบที่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ พบว่าในแบบจำลองปริภูมิเวกเตอร์ มีค่าเรียกคืนโดยเฉลี่ยน้อยกว่าแบบจำลองความน่าจะเป็น สังเกตจากค่าที่ติดลบคือลดลงร้อยละ 23.10 แต่สำหรับค่าความแม่นยำและค่าเฉลี่ยฮาร์โมนิก มีค่ามากกว่าแบบจำลองความน่าจะเป็น สรุปได้ว่า การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคส และมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ ให้ค่าความแม่นยำในการค้นคืนเอกสารยูสเคสได้ดีกว่าแบบจำลองความน่าจะเป็น แม้ว่าค่าเรียกคืนจะต่ำกว่า แต่โดยรวมซึ่ง

พิจารณาจากค่าเฉลี่ยฮาร์โมนิกที่เพิ่มขึ้นแล้ว แบบจำลองปริภูมิเวกเตอร์มีประสิทธิภาพการค้นคืนดีกว่าแบบจำลองความน่าจะเป็น

ตารางที่ 5.5 เปรียบเทียบอัตราการเพิ่มขึ้น/ลดลงของค่าเฉลี่ยค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก ของการค้นคืนแบบพิจารณาโครงสร้างของเอกสารยูสเคส ระหว่างแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น

โจทย์ ข้อที่	แบบจำลองปริภูมิเวกเตอร์ เปรียบเทียบกับ แบบจำลองความน่าจะเป็น			แบบจำลองปริภูมิเวกเตอร์ เปรียบเทียบกับ แบบจำลองความน่าจะเป็น		
	ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ (เพิ่มขึ้น/ลดลง ร้อยละ)			ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ (เพิ่มขึ้น/ลดลง ร้อยละ)		
	R	P	F	R	P	F
1	-15.19	+6.84	-5.87	-18.30	+37.38	+0.92
2	-18.46	+5.54	-4.38	-11.44	+11.21	+1.43
3	-22.94	-9.55	-13.66	-68.10	+20.28	-3.27
4	-16.90	+3.92	-2.98	-29.06	+4.89	-12.80
5	+63.73	+45.88	+68.69	-5.74	+25.54	+15.84
ค่าเฉลี่ย	+10.54	+16.31	+17.44	-23.10	+17.16	+2.78

ผลลัพธ์ของการทดลองวิธีค้นคืนยูสเคสทั้ง 8 วิธี ข้างต้นนั้น เป็นค่าเฉลี่ยที่แยกตามโจทย์ปัญหาทั้ง 5 ข้อ เพื่อให้แสดงให้เห็นความแตกต่างของค่าเรียกคืน ค่าความแม่นยำและค่าเฉลี่ยฮาร์โมนิกอันเกิดจากข้อคำถามและชุดคำตอบที่มีมุมมองต่างระดับกัน แต่หากลองพิจารณาผลการทดลองโดยที่ไม่คำนึงถึงความแตกต่างหรือมุมมองที่ต่างระดับกันของข้อคำถามที่ถูกกำหนดไว้ด้วยโจทย์ทั้ง 5 ข้อ แล้วนั้น ผู้วิจัยนำผลของค่าเรียกคืน และค่าความแม่นยำของทั้ง 250 ข้อคำถาม โดยนำค่าความแม่นยำมาเฉลี่ยตามค่าเรียกคืน 11 ค่า จาก 0 ถึง 1 แยกตามวิธีการค้นคืนทั้ง 8 ดังกล่าวในหัวข้อที่ 5.2.3 แสดงดังตารางที่ 5.6

จุฬาลงกรณ์มหาวิทยาลัย

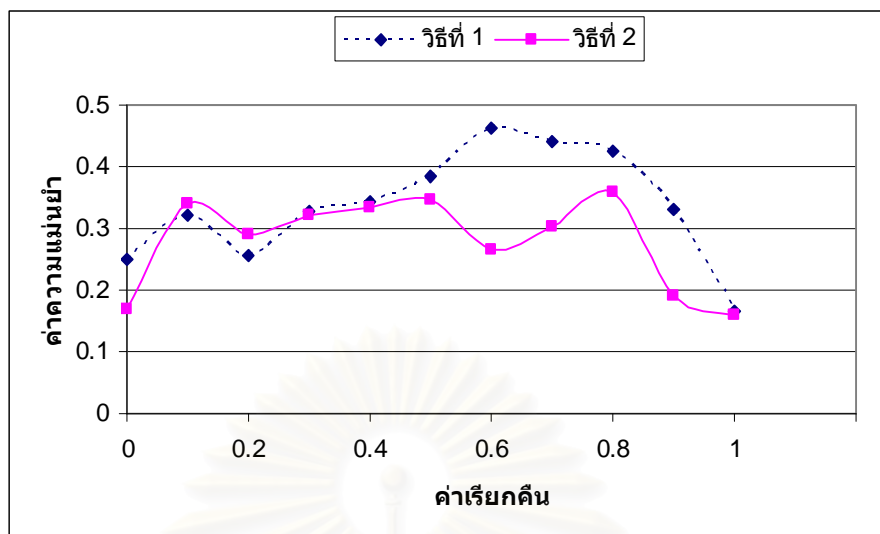
ตารางที่ 5.6 สรุปผลค่าความแม่นยำเฉลี่ยจากข้อคำถามทั้งหมด 250 ข้อคำถาม เรียงลำดับตามค่าเรียกคืนทั้ง 11 ค่า จาก 0 ถึง 1 ในแต่ละวิธีการค้นคืน

ค่าเรียกคืน	ค่าความแม่นยำ จากวิธีการค้นคืนที่							
	1	2	3	4	5	6	7	8
0.0	0.251	0.170	0.077	0.061	0.283	0.268	0.197	0.207
0.1	0.322	0.340	0.202	0.326	0.267	0.294	0.357	0.334
0.2	0.256	0.292	0.370	0.394	0.419	0.526	0.450	0.471
0.3	0.329	0.321	0.374	0.319	0.424	0.632	0.306	0.385
0.4	0.345	0.334	0.401	0.341	0.621	0.622	0.559	0.607
0.5	0.385	0.348	0.497	0.332	0.422	0.633	0.417	0.431
0.6	0.461	0.265	0.428	0.365	0.594	0.662	0.421	0.379
0.7	0.441	0.302	0.521	0.221	0.648	0.572	0.442	0.623
0.8	0.425	0.359	0.443	0.407	0.324	0.413	0.324	0.432
0.9	0.330	0.192	0.360	0.244	0.162	0.207	0.162	0.161
1.0	0.165	0.159	0.424	0.179	0.081	0.104	0.081	0.080
ค่าเฉลี่ย	0.337	0.280	0.372	0.290	0.386	0.448	0.338	0.374
ค่าเบี่ยงเบนมาตรฐาน	0.085	0.070	0.123	0.099	0.177	0.191	0.136	0.163

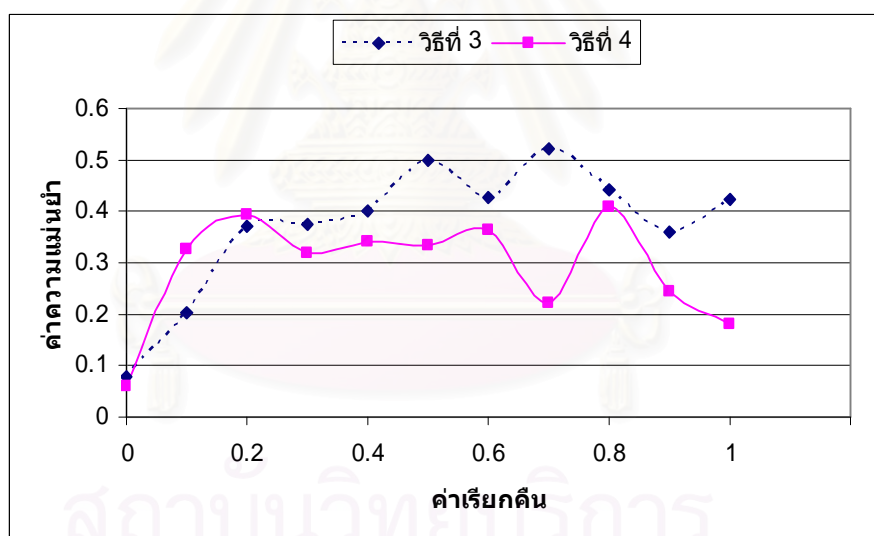
จากนั้นผู้วิจัยนำข้อมูลค่าความแม่นยำจากตารางที่ 5.6 มาแสดงในรูปแบบกราฟ เปรียบเทียบค่าความแม่นยำ ระหว่างวิธีการที่ไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ ซึ่งได้แก่คู่ของวิธีการค้นคืนที่ 1 กับ 2 ดังรูปที่ 5.2 คู่วิธีการที่ 3 กับ 4 ดังรูปที่ 5.3 คู่วิธีการที่ 5 กับ 6 ดังรูปที่ 5.4 คู่วิธีการที่ 7 กับ 8 ดังรูปที่ 5.5

จากตารางที่ 5.6 และกราฟในรูปที่ 5.2 พบว่าการค้นคืนวิธีที่ 1 มีค่าความแม่นยำโดยเฉลี่ยมากกว่าการค้นคืนวิธีที่ 2 แสดงว่าการค้นคืนโดยไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ ไม่มีผลให้การค้นคืนมีความแม่นยำที่เพิ่มขึ้นไปกว่าการค้นคืนโดยไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้

จากตารางที่ 5.6 และกราฟในรูปที่ 5.3 ซึ่งมีทิศทางเดียวกับกราฟในรูปที่ 5.2 พบว่าการค้นคืนวิธีที่ 3 มีค่าความแม่นยำโดยเฉลี่ยมากกว่าการค้นคืนวิธีที่ 4 แสดงว่าการค้นคืนโดยไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ ไม่มีผลให้การค้นคืนมีความแม่นยำเพิ่มขึ้นไปกว่าการค้นคืนโดยไม่พิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้

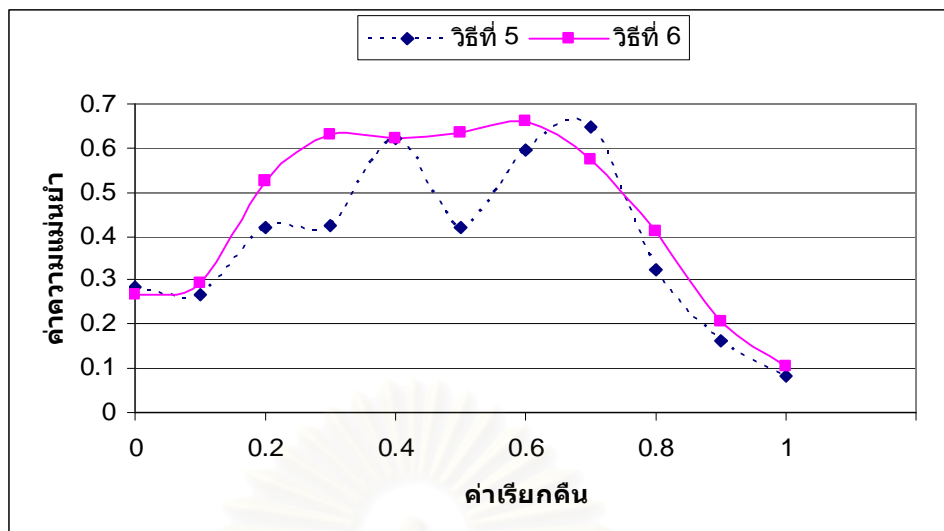


รูปที่ 5.2 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 1 และ 2 ตามลำดับ



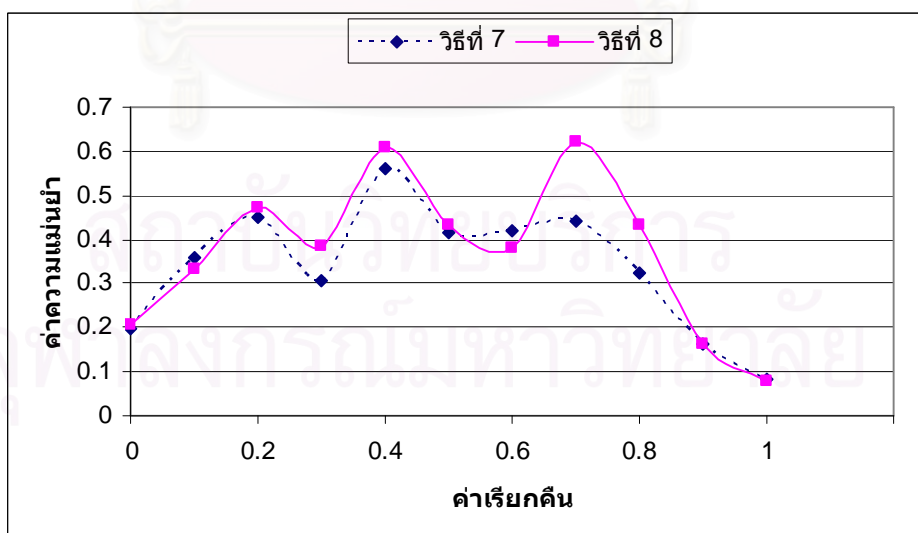
รูปที่ 5.3 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของยูสเคสด้วยแบบจำลองความน่าจะเป็น โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ วิธีค้นคืนที่ 3 และ 4 ตามลำดับ

ตารางที่ 5.6 และกราฟในรูปที่ 5.4 พบว่าการค้นคืนวิธีที่ 6 มีค่าความแม่นยำโดยเฉลี่ยมากกว่าการค้นคืนวิธีที่ 5 แสดงว่าการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีผลทำให้การค้นคืนมีความแม่นยำมากกว่าการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

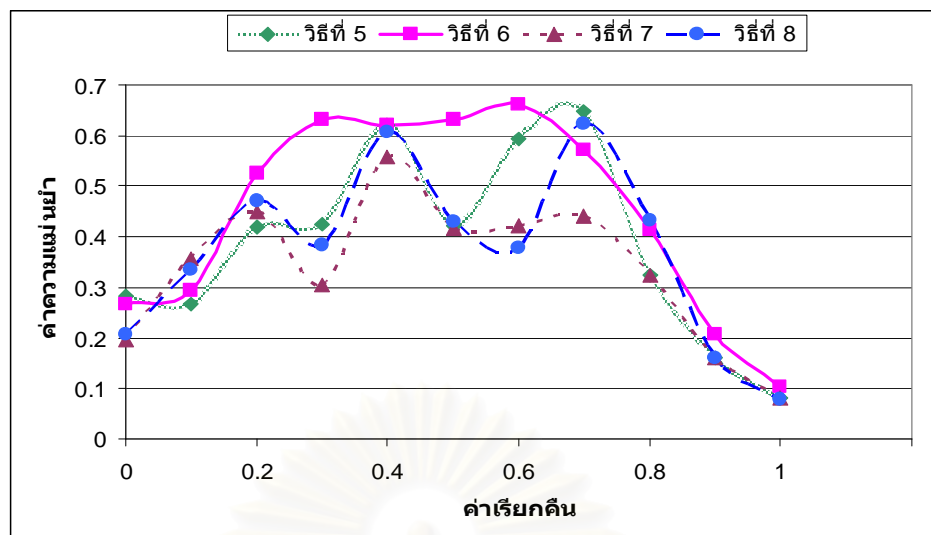


รูปที่ 5.4 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสด้วยจำลองปริภูมิเวกเตอร์ โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช่วิธีค้นคืนที่ 5 และ 6 ตามลำดับ

ตารางที่ 5.6 และกราฟในรูปที่ 5.5 ซึ่งมีทิศทางเดียวกับกราฟในรูปที่ 5.4 พบว่าการค้นคืนวิธีที่ 8 มีค่าความแม่นยำโดยเฉลี่ยมากกว่าการค้นคืนวิธีที่ 7 แสดงว่าการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีผลทำให้การค้นคืนมีความแม่นยำมากกว่าการค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้



รูปที่ 5.5 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสด้วยแบบจำลองความน่าจะเป็น โดยไม่ใช้และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช่วิธีค้นคืนที่ 7 และ 8 ตามลำดับ



รูปที่ 5.6 กราฟค่าเรียกคืน และค่าความแม่นยำของวิธีการค้นคืน 4 วิธีการ เฉพาะการค้นคืนแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยूसเคส

หลังจากนำผลของค่าความแม่นยำที่ได้จาก 250 ข้อคำถาม มาทำการเฉลี่ยเพื่อหาค่าความแม่นยำ ณ ระดับค่าเรียกคืนต่างๆ ทั้ง 11 จุด ของแต่ละวิธีการค้นคืน ซึ่งเป็นการพิจารณาผลการทดลองโดยไม่คำนึงถึงความแตกต่างกันของข้อคำถามที่ถูกกำหนดโดยโจทย์ปัญหาทั้ง 5 สรุปได้ว่า การค้นคืนแบบไม่พิจารณาโครงสร้างของยूसเคส ทั้งในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น เมื่อใช้เทคนิคการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้แล้ว ไม่สามารถทำให้ค่าความแม่นยำเพิ่มขึ้นจากที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ดังนั้นกราฟในรูปที่ 5.6 แสดงเฉพาะวิธีการค้นคืนที่พิจารณาโครงสร้างของเอกสารยूसเคส โดยเมื่อดูจากตารางที่ 5.6 และกราฟนี้แล้ว พบว่าวิธีการค้นคืนที่ 6 ให้ค่าความแม่นยำมากที่สุดคือ การค้นคืนยूसเคสแบบพิจารณาโครงสร้างของยूसเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ รองลงมาเป็นวิธีที่ 5 8 และ 7 ตามลำดับ

5.6 การวิเคราะห์ผลการทดลองโดยทางสถิติ

สำหรับผลจากการทดลองที่ได้นำเสนอไปข้างต้นแล้วนั้นยังไม่เพียงพอที่จะสรุปได้ว่าวิธีการค้นคืนใดที่เหมาะสมหรือมีประสิทธิภาพการค้นคืนเอกสารประเภทยूसเคสได้ดีที่สุด โดยงานวิจัยนี้จะวัดที่ค่าความแม่นยำเป็นหลักเนื่องจากเป็นมาตรวัดที่ใช้คำนวณเอกสารที่ระบบค้นคืนได้นั้นเป็นเอกสารที่ตรงกับความต้องการของผู้ใช้เท่าไร

ด้วยเหตุนี้จึงนำหลักการวิเคราะห์ทางสถิติมาใช้ในการสนับสนุนความน่าเชื่อถือของงานวิจัยด้วยอีกทางหนึ่ง ดังนั้นต้องมีการตั้งสมมุติฐาน และการทดสอบสมมุติฐาน ซึ่งหลักการพื้นฐานประการแรก ผู้วิจัยต้องตรวจสอบการแจกแจงของประชากรหรือข้อมูลว่ามีการแจกแจงปกติหรือไม่ เพื่อเลือกใช้วิธีการทางสถิติในการทดสอบสมมุติฐานที่เหมาะสมกับการแจกแจงประชากรนั้นๆ ต่อไป

ในงานวิจัยนี้ใช้การตรวจสอบการแจกแจงของประชากร โดยใช้สถิติทดสอบ Shapiro-Wilk Test เป็นสถิติที่ใช้ทดสอบการแจกแจงของตัวแปรเชิงปริมาณว่ามีการแจกแจงแบบปกติหรือไม่ ซึ่งเป็นสถิติทดสอบที่เหมาะสมกับปัญหาที่อาจทราบหรือไม่ทราบค่าเฉลี่ยหรือค่าความแปรปรวนของประชากรก็ได้ และขนาดประชากรไม่เกิน 50 หน่วย

สมมติฐานของการทดสอบ คือ

- $H_{0,1}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 1 มีการแจกแจงแบบปกติ
 $H_{0,2}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 2 มีการแจกแจงแบบปกติ
 $H_{0,3}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 3 มีการแจกแจงแบบปกติ
 $H_{0,4}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 4 มีการแจกแจงแบบปกติ
 $H_{0,5}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 5 มีการแจกแจงแบบปกติ
 $H_{0,6}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 6 มีการแจกแจงแบบปกติ
 $H_{0,7}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 7 มีการแจกแจงแบบปกติ
 $H_{0,8}$: ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการค้นคืนที่ 8 มีการแจกแจงแบบปกติ
 H_1 : ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากทุกวิธีการค้นคืนไม่มีการแจกแจงแบบปกติ

สถิติทดสอบ : Shapiro-Wilk

เขตปฏิเสธ H_0 : จะปฏิเสธ H_0 ถ้าค่า Sig.(Significance) ของการทดสอบน้อยกว่าระดับนัยสำคัญ (α) ที่กำหนดไว้เท่ากับ 0.05

จากนั้นทำการทดสอบสมมติฐาน โดยใช้โปรแกรม SPSS ได้ค่าสถิติทดสอบดังตารางที่ 5.7

ตารางที่ 5.7 แสดงค่าสถิติทดสอบการแจกแจงปกติของค่าความแม่นยำโดยเฉลี่ยในแต่ละวิธีการค้นคืน

วิธีการค้นคืนที่	Shapiro-Wilk Test		H_0
	df	Sig.	
1	11	0.741	ยอมรับ
2	11	0.059	ยอมรับ
3	11	0.054	ยอมรับ
4	11	0.192	ยอมรับ
5	11	0.644	ยอมรับ
6	11	0.116	ยอมรับ
7	11	0.726	ยอมรับ
8	11	0.691	ยอมรับ

จากตารางพบว่า ค่า Sig. ของทุกวิธีการคั่นคืน มีค่ามากกว่า α แสดงว่ายอมรับสมมุติฐานหลัก นั่นคือ ค่าความแม่นยำโดยเฉลี่ยตามค่าเรียกคืน 11 ค่า จากวิธีการคั่นคืนที่ 1 ถึง วิธีการคั่นคืนที่ 8 มีการแจกแจงแบบปกติ ที่ระดับนัยสำคัญ 0.05

เมื่อทราบแล้วว่าประชากรมีการแจกแจงแบบปกติ ในการทดสอบสมมุติฐานจึงเลือกใช้ การทดสอบสมมุติฐานที่ใช้พารามิเตอร์ โดยในงานวิจัยนี้ ผู้วิจัยเลือกใช้การทดสอบสมมุติฐาน เกี่ยวกับผลต่างระหว่างค่าเฉลี่ยของสองประชากรแบบจับคู่ (Paired t-test) ในการทดสอบ สมมุติฐาน เนื่องจากเป็นวิธีการที่เหมาะสมกับการทดสอบความแตกต่างของผลการใช้สิ่ง ที่ต้องการตรวจสอบกับการไม่ใช้ [18] โดยใช้ข้อมูลค่าความแม่นยำในตารางที่ 5.6 เป็นประชากร ในการทดสอบ

กำหนดให้

m_1 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของ เอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็น จากผู้ใช้

m_2 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของ เอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจาก ผู้ใช้

m_3 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของ เอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็น จากผู้ใช้

m_4 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของ เอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจาก ผู้ใช้

m_5 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

m_6 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

m_7 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

m_8 คือ ค่าความแม่นยำเฉลี่ย จากการคั่นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

ประชากรที่ใช้ทดสอบคือ ผลจากการทดลองด้วยแต่ละเทคนิควิธีในตารางที่ 5.6 นำมา เปรียบเทียบกันทีละคู่วิธีดังต่อไปนี้ โดยมีการตั้งสมมุติฐานดังต่อไปนี้

5.6.1 การคั่นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้ แบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ย

มากกว่าหรือเท่ากับการคั่นคี่นยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส โดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_1 &\geq \mu_2 \\ H_1 : \mu_1 &< \mu_2 \end{aligned} \quad \dots\dots (1)$$

5.6.2 การคั่นคี่นยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการคั่นคี่นยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_3 &\geq \mu_4 \\ H_1 : \mu_3 &< \mu_4 \end{aligned} \quad \dots\dots (2)$$

5.6.3 การคั่นคี่นยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการคั่นคี่นยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_5 &\geq \mu_6 \\ H_1 : \mu_5 &< \mu_6 \end{aligned} \quad \dots\dots (3)$$

5.6.4 การคั่นคี่นยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการคั่นคี่นยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_7 &\geq \mu_8 \\ H_1 : \mu_7 &< \mu_8 \end{aligned} \quad \dots\dots (4)$$

5.6.5 การคั่นคี่นยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการคั่นคี่นยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_2 &\geq \mu_6 \\ H_1 : \mu_2 &< \mu_6 \end{aligned} \dots\dots (5)$$

5.6.6 การค้นคืนยูสเคสแบบไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_4 &\geq \mu_8 \\ H_1 : \mu_4 &< \mu_8 \end{aligned} \dots\dots (6)$$

5.6.7 การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_7 &\geq \mu_5 \\ H_1 : \mu_7 &< \mu_5 \end{aligned} \dots\dots (7)$$

5.6.8 การค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองความน่าจะเป็น และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีความแม่นยำเฉลี่ยมากกว่าหรือเท่ากับการค้นคืนยูสเคสแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สมมุติฐานคือ

$$\begin{aligned} H_0 : \mu_8 &\geq \mu_6 \\ H_1 : \mu_8 &< \mu_6 \end{aligned} \dots\dots (8)$$

กำหนดสถิติที่ใช้ในการทดสอบคือ t_{n-1}

$$t = \frac{\bar{d} - d_0}{S_d / \sqrt{n}} \quad \text{ซึ่งจะมีการแจกแจงแบบ } t \text{ ที่องศาอิสระ } n-1$$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$ โดยจะปฏิเสธ H_0 ถ้า $t_{10} < -t_{0.95,10}$ หรือ -1.81
ผลการทดสอบสมมุติฐาน ได้ค่าสถิติ ดังตารางที่ 5.8

ตารางที่ 5.8 แสดงค่าสถิติทดสอบแบบจับคู่ระหว่างค่าเฉลี่ยสองประชากร ของการทดสอบสมมุติฐานที่ 1 ถึง 8

การทดสอบสมมุติฐานที่	t	H ₀
1) วิธีคั่นคั้นที่ 1 เปรียบเทียบกับวิธีคั่นคั้นที่ 2	2.552	ยอมรับ
2) วิธีคั่นคั้นที่ 3 เปรียบเทียบกับวิธีคั่นคั้นที่ 4	2.283	ยอมรับ
3) วิธีคั่นคั้นที่ 5 เปรียบเทียบกับวิธีคั่นคั้นที่ 6	-2.348	ปฏิเสธ
4) วิธีคั่นคั้นที่ 7 เปรียบเทียบกับวิธีคั่นคั้นที่ 8	-1.835	ปฏิเสธ
5) วิธีคั่นคั้นที่ 2 เปรียบเทียบกับวิธีคั่นคั้นที่ 6	-3.512	ปฏิเสธ
6) วิธีคั่นคั้นที่ 4 เปรียบเทียบกับวิธีคั่นคั้นที่ 8	-1.897	ปฏิเสธ
7) วิธีคั่นคั้นที่ 7 เปรียบเทียบกับวิธีคั่นคั้นที่ 5	-1.770	ยอมรับ
8) วิธีคั่นคั้นที่ 8 เปรียบเทียบกับวิธีคั่นคั้นที่ 6	-2.139	ปฏิเสธ

5.7 สรุปผลการทดลอง

ผลที่ได้จากการวิเคราะห์ผลการทดลองซึ่งเป็นการทดสอบสมมุติฐานต่างๆ ในหัวข้อที่ 5.6 นำมาสรุปได้ดังต่อไปนี้

5.7.1 การทดสอบสมมุติฐานที่ 1 ผลจากค่าสถิติทดสอบพบว่า ยอมรับสมมุติฐาน H₀ นั่นคือ ปฏิเสธสมมุติฐาน H₁ แสดงว่า การคั่นคั้นยूसเคสแบบไม่พิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคสและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์ มีความแม่นยำเฉลี่ยน้อยกว่าหรือเท่ากับ การคั่นคั้นยूसเคสแบบไม่พิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคส แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์ กล่าวได้ว่าการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ไม่สามารถทำให้มีความแม่นยำมากขึ้น

5.7.2 การทดสอบสมมุติฐานที่ 2 ผลจากค่าสถิติทดสอบพบว่า ยอมรับสมมุติฐาน H₀ นั่นคือ ปฏิเสธสมมุติฐาน H₁ แสดงว่า การคั่นคั้นยूसเคสแบบไม่พิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคสและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น มีความแม่นยำเฉลี่ยน้อยกว่าหรือเท่ากับ การคั่นคั้นยूसเคสแบบไม่พิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคส แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองความน่าจะเป็น หรือกล่าวได้ว่าการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ไม่สามารถทำให้มีความแม่นยำมากขึ้น

5.7.3 การทดสอบสมมุติฐานที่ 3 ผลจากค่าสถิติทดสอบพบว่า ปฏิเสธสมมุติฐาน H₀ นั่นคือ ยอมรับสมมุติฐาน H₁ แสดงว่า การคั่นคั้นยूसเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคสและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์ มีความแม่นยำเฉลี่ยมากกว่าการคั่นคั้นยूसเคสแบบพิจารณาโครงสร้างของเอกสาร คำอธิบายยूसเคส แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งาน ด้วยแบบจำลองปริภูมิเวกเตอร์ หรือกล่าวได้ว่าการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งานสามารถทำให้มีความแม่นยำมากขึ้น

แบบจำลองปริภูมิเวกเตอร์มีความแม่นยำมากกว่าแบบจำลองความน่าจะเป็น โดยพิจารณาถึงผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

5.8 ข้ออภิปราย

จากการทดลอง ผลการทดลอง การวิเคราะห์ผลการทดลองด้วยหลักสถิติและสรุปผลทดลอง ในหัวข้อดังกล่าวข้างต้น สามารถนำมาอภิปรายได้ดังต่อไปนี้

5.8.1 จากตารางที่ 5.2 พบว่า การค้นคืนที่ไม่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ทำให้ค่าเรียกคืนโดยเฉลี่ยมีค่าเพิ่มมากขึ้น แต่ความแม่นยำโดยเฉลี่ยไม่เพิ่มขึ้น ในแบบจำลองทั้งสองแบบ รวมทั้งข้อมูลการทดสอบสมมุติฐานข้อที่ 1 และ ข้อที่ 2 ซึ่งยอมรับสมมุติฐาน H_0 ทั้ง 2 ข้อ แสดงว่า การนำข้อคำถามจากแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคสมาต่อกันเป็นข้อคำถามเดี่ยวซึ่งเป็นการค้นคืนที่ไม่พิจารณาโครงสร้างของยูสเคส หากใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้แล้ว ทำให้พบเอกสารที่ตรงกับความต้องการของผู้ใช้เพิ่มมากขึ้น แต่ก็พบว่าเอกสารที่ค้นคืนออกมาได้นั้น เป็นเอกสารที่ไม่ตรงกับความต้องการด้วยเช่นกัน จึงทำให้ค่าความแม่นยำลดลงด้วย

5.8.2 จากตารางที่ 5.3 ซึ่งเทียบอัตราการเพิ่มขึ้นหรือลดลงโดยเฉลี่ยของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิคจากทั้ง 5 โจทย์ของการค้นคืนแบบไม่พิจารณาโครงสร้างของเอกสารยูสเคส เพื่อเปรียบเทียบประสิทธิผลระหว่าง 2 แบบจำลอง พบว่าทั้งแบบที่ใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีทิศทางเดียวกัน โดยแบบที่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีแนวโน้มของแต่ละค่าสูงกว่าแบบไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ นั่นคือแบบจำลองปริภูมิเวกเตอร์มีค่าเรียกคืนมากกว่าแบบจำลองความน่าจะเป็น แต่กลับมีค่าความแม่นยำน้อยกว่าแบบจำลองความน่าจะเป็น

5.8.3 จากตารางที่ 5.4 พบว่าการค้นคืนที่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ ทำให้ค่าความแม่นยำโดยเฉลี่ยมีค่าเพิ่มมากขึ้น แต่ค่าเรียกคืนไม่เพิ่มขึ้น แสดงว่าการค้นคืนแบบพิจารณาคำในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส ด้วยแบบจำลองนี้ ทำให้จำนวนเอกสารยูสเคสที่ระบบค้นคืนออกมาได้นั้น เป็นยูสเคสที่ตรงกับความต้องการของผู้ใช้มากขึ้นกว่าแบบที่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ แม้ว่าจะไม่สามารถค้นคืนยูสเคสที่ตรงความต้องการออกมาได้ทั้งหมดก็ตาม สำหรับการค้นคืนแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองความน่าจะเป็น ทำให้ทั้งค่าความแม่นยำโดยเฉลี่ย และค่าเรียกคืนโดยเฉลี่ยมีค่าเพิ่มมากขึ้น แสดงว่าการค้นคืนแบบพิจารณาคำในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส ในแบบจำลองนี้ ทำให้ระบบสามารถค้นคืนยูสเคสที่ตรงประเด็นกับความต้องการของผู้ใช้ได้มากขึ้น และถูกต้องแม่นยำมากขึ้นด้วย

5.8.4 จากตารางที่ 5.5 ซึ่งเทียบอัตราการเพิ่มขึ้นหรือลดลงโดยเฉลี่ยของค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิคจากทั้ง 5 โจทย์ของการค้นคืนแบบพิจารณาโครงสร้างของ

เอกสารยูสเคส เพื่อเปรียบเทียบประสิทธิภาพระหว่าง 2 แบบจำลอง พบว่าสิ่งที่เหมือนกันคือทั้งแบบที่ใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ แบบจำลองปริภูมิเวกเตอร์มีค่าความแม่นยำมากกว่าแบบจำลองความน่าจะเป็น

5.8.5 จากรูปที่ 5.6 และตารางที่ 5.6 พบว่าการค้นคืนในวิธีที่ 6 คือ การค้นคืนที่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ มีความแม่นยำมากที่สุดคือ มีค่าความแม่นยำโดยเฉลี่ย 0.448 ตามด้วยการค้นคืนในวิธีที่ 5 คือ การค้นคืนที่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ โดยมีค่าความแม่นยำโดยเฉลี่ย 0.386 ตามด้วยการค้นคืนวิธีที่ 8 และ 7 ซึ่งเป็นการค้นคืนที่พิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส ด้วยแบบจำลองความน่าจะเป็น แบบใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ โดยมีค่าความแม่นยำโดยเฉลี่ย 0.374 และ 0.338 ตามลำดับ

5.8.6 เปรียบเทียบการค้นคืนแบบพิจารณาโครงสร้างของเอกสารคำอธิบายยูสเคส และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ระหว่างแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น พบว่าแบบจำลองปริภูมิเวกเตอร์มีความแม่นยำมากกว่า 19.79 % เนื่องจากวิธีการเปลี่ยนแปลงข้อความของแบบจำลองความน่าจะเป็นนั้น มีเพียงการเปลี่ยนแปลงคำนำหน้าของคำในข้อความเท่านั้น แต่แบบจำลองปริภูมิเวกเตอร์สามารถเปลี่ยนแปลงได้ทั้งข้อความ โดยการเพิ่มคำเข้าไปในข้อความใหม่ จากยูสเคสที่ผู้ใช้เลือกกว่ามีความเกี่ยวข้องกับความต้องการ รวมทั้งสามารถเปลี่ยนแปลงคำนำหน้าของคำได้อีกด้วย จึงเป็นสาเหตุให้การค้นคืนแบบพิจารณาโครงสร้างของยูสเคสร่วมกับการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ในแบบจำลองปริภูมิเวกเตอร์ให้ค่าความแม่นยำได้ดีที่สุด และมีความเหมาะสมกับการค้นคืนเอกสารคำอธิบายยูสเคส

บทที่ 6

สรุปผลงานวิจัย

ในบทนี้กล่าวถึงส่วนสุดท้ายที่ได้จากผลงานวิจัยนั่นคือ บทสรุปของผลงานวิจัย รวมทั้งงานวิจัยในอนาคต และบทความวิชาการที่ตีพิมพ์ ซึ่งมีรายละเอียดดังต่อไปนี้

6.1 สรุปผลงานวิจัย

งานวิจัยนี้นำเสนอวิธีการค้นคืนยูสเคส โดยมีการจัดกลุ่มของเอกสารคำอธิบายยูสเคส ร่วมกับการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น ยูสเคสที่นำมาจัดเก็บในระบบ คือ เอกสารคำอธิบายยูสเคส ซึ่งเขียนเป็นภาษาอังกฤษ จำนวน 611 เอกสาร จาก 40 ระบบงานทางธุรกิจ ซึ่งรวบรวมจากเอกสารประกอบโครงงานวิศวกรรม โครงงานมหัศจรรย์ และวิทยานิพนธ์มหัศจรรย์ จากภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยนำหลักการพื้นฐานจากทฤษฎีการจัดเก็บและค้นคืนสารสนเทศ อันได้แก่ การทำดรรชนีอัตโนมัติ การหาคำนำหน้าของคำ การใช้สูตรคำนวณหาความคล้าย เป็นต้น มาประยุกต์ใช้ จากนั้นผู้วิจัยได้เพิ่มเติมวิธีการจัดกลุ่มของเอกสารคำอธิบายยูสเคส และเพิ่มเทคนิคการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้งานวิจัยของ คุณอัคเดช อุดมชัยพร และคณะ [7] ซึ่งมีเพียงการค้นคืนยูสเคสโดยการพิจารณาคำและโครงสร้างของเอกสารคำอธิบายยูสเคสเท่านั้น

ในงานวิจัยนี้จึงประกอบด้วย 5 ขั้นตอนที่สำคัญ คือ

1) การจัดกลุ่มให้กับเอกสารคำอธิบายยูสเคส

สำหรับการจัดกลุ่มให้กับยูสเคส ผู้วิจัยให้ผู้เชี่ยวชาญด้านวิศวกรรมซอฟต์แวร์ วิเคราะห์ลักษณะความเหมือนและความแตกต่างกันของยูสเคส พบว่า การจัดกลุ่มให้กับยูสเคสในชุดข้อมูลทั้งหมด ควรจัดกลุ่มยูสเคสตามระบบงานทางธุรกิจ จะมีความเหมาะสมที่สุด นั่นคือ มี 40 กลุ่มยูสเคส เนื่องจากยูสเคสตามระบบงานทางธุรกิจแต่ละระบบงาน ต่างก็มีลักษณะเฉพาะตัว หากจัดให้มีกลุ่มของยูสเคสน้อยหรือมากเกินไป อาจทำให้ระบบพลาดการค้นคืนยูสเคสที่เป็นยูสเคสเป้าหมายได้ จากนั้นคำนวณหาค่ากลางซึ่งจะใช้เป็นตัวแทนกลุ่มของเอกสารในแต่ละกลุ่ม เพื่อใช้ในขั้นตอนการค้นคืนต่อไป

2) การค้นคืนเอกสารคำอธิบายยูสเคส

มีหลักการสำคัญคือ การค้นคืนยูสเคสในแต่ละครั้ง จะพิจารณาจากคำที่ปรากฏในแต่ละส่วนประกอบของเอกสารคำอธิบายยูสเคส ร่วมกับคำนำหน้าที่ผู้ใช้ระบุให้มีความสำคัญกับแต่ละส่วนประกอบนั้นๆ ซึ่งถือเป็นการพิจารณาตามโครงสร้างของเอกสารคำอธิบายยูสเคส โดยจะได้ผลลัพธ์เป็นค่าความคล้ายกันระหว่างข้อความกับยูสเคส แต่สูตรที่ใช้ในการคำนวณหาความคล้ายนั้นจะแตกต่างกันไปตามแบบจำลองที่ใช้ ซึ่งข้อความที่นำมาใช้ค้นคืนยูสเคสในแต่ละครั้ง จะผ่านการคำนวณกับเอกสารยูสเคสที่เป็นตัวแทนของกลุ่มเสมอ ก่อนที่จะนำไปคำนวณ

กับเอกสารยูสเคสที่เป็นสมาชิกในกลุ่มนั้นๆ ส่วนการค้นคืนยูสเคสแบบมีผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้น ในแบบจำลองปริภูมิเวกเตอร์ จะใช้วิธีการเปลี่ยนแปลงข้อความ หรือการขยายคำในข้อความ และการเปลี่ยนแปลงน้ำหนักคำในข้อความเดิม เป็นข้อความใหม่ ส่วนในแบบจำลองความน่าจะเป็น จะมีเฉพาะการเปลี่ยนแปลงน้ำหนักของคำในข้อความเดิมเท่านั้น

3) การออกแบบและพัฒนาเครื่องมือ

เมื่อได้แนวคิดในงานวิจัยเรียบร้อยแล้ว จึงได้ทำการออกแบบและพัฒนาเครื่องมือที่เหมาะสมเพื่อสนับสนุนแนวคิดดังกล่าว โดยเครื่องมือที่พัฒนาขึ้น แบ่งออกเป็น 2 ส่วน ดังนี้

(1) ส่วนการจัดเก็บยูสเคส เมื่อกำหนดได้แล้วว่าจัดกลุ่มยูสเคสตามระบบงานทางธุรกิจ ในการจัดเก็บยูสเคสเข้าสู่ระบบจึงกำหนดรหัสของแต่ละกลุ่ม แล้วจึงดำเนินการกับข้อความ ค่าตรรกษีและน้ำหนักของคำในแต่ละเอกสาร แล้วจึงทำการหายูสเคสตัวแทนให้แต่ละกลุ่ม และจัดเก็บลงฐานข้อมูล

(2) ส่วนการค้นคืนยูสเคส ในขั้นตอนการค้นคืนนั้น ข้อความที่ผู้ใช้ใส่เข้ามา จะถูกนำมาเปรียบเทียบกับค่าความคล้ายกับเอกสารตัวแทนกลุ่มยูสเคสเสียก่อน หากข้อความดังกล่าวมีค่าความคล้ายกับตัวแทนกลุ่มพอเพียงกับค่าความคล้ายเริ่มต้นที่ได้กำหนดไว้ ระบบจึงจะนำข้อความนั้นไปคำนวณค่าความคล้ายกับยูสเคสที่เป็นสมาชิกของกลุ่มนั้นต่อไป การดำเนินการในลักษณะนี้ เพื่อช่วยลดเวลาในการเปรียบเทียบความคล้ายระหว่างข้อความกับยูสเคสทั้งหมดที่ละยูสเคส โดยมีการค้นคืนด้วยการพิจารณาตามโครงสร้างของเอกสาร นั่นคือ การกำหนดให้ผู้ค้นคืนทำการค้นหายูสเคสด้วยการใช้ข้อความที่มีลักษณะเดียวกับโครงสร้างของยูสเคส นอกจากนี้ผู้วิจัยใช้วิธีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้เข้ามาพร้อมด้วย โดยเสนอวิธีการค้นคืนแบบมีการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ไว้ 2 วิธีการ นั่นคือ แบบอาศัยแบบจำลองปริภูมิเวกเตอร์ ซึ่งสามารถใช้ข้อมูลจากการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มาทำการเปลี่ยนแปลงคำและน้ำหนักของคำในข้อความเดิม และแบบจำลองความน่าจะเป็น จะเปลี่ยนแปลงในส่วนของน้ำหนักของคำในข้อความเดิมเท่านั้น

4) การทดลองและประเมินผลการทดลอง

เมื่อพัฒนาเครื่องมือเรียบร้อยแล้ว ผู้วิจัยได้ออกแบบการทดลองและทำการทดลองเพื่อประเมินประสิทธิผลของวิธีการที่นำเสนอ ซึ่งเครื่องมือดังกล่าวถูกออกแบบให้ผู้ทำการทดลองหรือหน่วยตัวอย่าง สามารถค้นคืนยูสเคสได้ ทั้งตามแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น และสามารถค้นคืนโดยใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ได้ การทดลองจะทำการเปรียบเทียบประสิทธิผลระหว่างการค้นคืนโดยใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ จากแบบจำลองทั้ง 2 ข้างต้น ในการทดลองนี้มีหน่วยตัวอย่างทั้งหมด 10 หน่วยตัวอย่าง 611 เอกสารยูสเคส และ 5 โจทย์ปัญหา ซึ่งจะเป็นปัจจัยควบคุมไม่ให้ผลลัพธ์ที่ได้เกิดความเอนเอียง จากนั้นจึงใช้ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก เป็นมาตรวัดในการ

ประเมินประสิทธิผลของการคั่นคั้น จากข้อคำถามทั้งสิ้น 250 ข้อคำถามที่มีความหลากหลายต่างกันไป

5) สรุปผลการทดลอง

ผลลัพธ์ที่ได้จากการทดลอง แบ่งออกเป็น 2 กรณีด้วยกัน กรณีแรกคือ พิจารณาค่าเรียกคั้น ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก จากข้อคำถามที่แบ่งตามโจทย์ 5 ข้อ พบว่าการคั่นคั้นแบบไม่พิจารณาโครงสร้างของยูสเคสและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ทำให้ค่าเรียกคั้นดีขึ้น แต่ค่าความแม่นยำลดลง ทั้งในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น แต่สำหรับการคั่นคั้นแบบพิจารณาโครงสร้างของยูสเคสและใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ทำให้ค่าเรียกคั้นลดลง แต่ค่าความแม่นยำมากขึ้น ทั้งในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น

อีกกรณีหนึ่งคือ พิจารณาค่าเรียกคั้น ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิก โดยไม่คำนึงถึงมุมมองของข้อคำถามที่แตกต่างกัน ซึ่งนำค่าความแม่นยำจากทุกวิธีการคั่นคั้นมาเฉลี่ยตามจุดของค่าเรียกคั้นทั้ง 11 จุด ดังตารางที่ 5.6 และ กราฟรูปที่ 5.6 พบว่าการคั่นคั้นแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ และใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ มีค่าความแม่นยำโดยเฉลี่ยเป็น 0.448 ซึ่งมีค่ามากที่สุด รองลงมาคือ การคั่นคั้นแบบพิจารณาโครงสร้างของยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ แต่ไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ซึ่งมีค่าความแม่นยำโดยเฉลี่ยเป็น 0.386 จากนั้นได้ใช้หลักการทางสถิติในการวิเคราะห์ผลการทดลอง โดยได้ตั้งสมมุติฐานต่างๆ และทำการทดสอบสมมุติฐาน โดยได้ข้อสรุปว่า

1) การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ไม่สามารถทำให้มีความแม่นยำมากขึ้นหากคั่นคั้นแบบไม่พิจารณาโครงสร้างของเอกสารยูสเคส ทั้งในแบบจำลองทั้ง 2 แบบ แต่ให้ค่าเรียกคั้นมากขึ้น

2) การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ สามารถทำให้มีความแม่นยำมากขึ้นหากคั่นคั้นแบบพิจารณาโครงสร้างของเอกสารยูสเคส ในแบบจำลองทั้ง 2 แบบ แต่ให้ค่าเรียกคั้นลดลง

3) การคั่นคั้นแบบพิจารณาโครงสร้างของเอกสารยูสเคสมีความแม่นยำมากขึ้นกว่าการคั่นคั้นที่ไม่พิจารณาโครงสร้างของเอกสารยูสเคส ในแบบจำลองทั้ง 2 แบบ

4) การคั่นคั้นบนแบบจำลองปริภูมิเวกเตอร์มีความแม่นยำมากกว่าการคั่นคั้นของแบบจำลองความน่าจะเป็น ทั้งแบบที่ใช้และไม่ใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

จากผลการทดลองและการทดสอบสมมุติฐานต่างๆ สรุปได้ว่า การคั่นคั้นเอกสารที่เป็นโครงสร้างอย่างเอกสารคำอธิบายยูสเคสที่มีการจัดกลุ่มไว้แล้ว จะให้ผลลัพธ์ที่ดีที่สุด ถ้าคั่นคั้นโดยพิจารณาโครงสร้างของเอกสารยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ ร่วมกับการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้

6.2 งานวิจัยในอนาคต

1) การศึกษาและวิเคราะห์ลักษณะสำคัญของเอกสารคำอธิบายยูสเคส เป็นสิ่งที่น่าสนใจ และสามารถสานต่อจนเป็นงานวิจัยได้อีก เนื่องจากเอกสารประเภทนี้มีความสำคัญอย่างมากต่อการพัฒนาซอฟต์แวร์ประเภทต่าง ๆ ไม่ว่าจะเป็นซอฟต์แวร์ประเภทการจัดการสารสนเทศ การค้าขายอิเล็กทรอนิกส์ หรือเพื่อการวางแผนการผลิตในอุตสาหกรรมหนึ่ง ๆ ก็นิยมนำหลักการของการพัฒนาซอฟต์แวร์สมัยใหม่ที่ใช้หลักการออกแบบซอฟต์แวร์เชิงวัตถุและมีเครื่องมือจัดเก็บ และออกแบบความต้องการซอฟต์แวร์เป็นภาษายูเอ็มแอล ดังนั้นการนำเอกสารประเภทนี้มาวิเคราะห์เพื่อหาแนวทางให้สามารถนำกลับมาใช้ใหม่ได้ จะเป็นประโยชน์อย่างยิ่ง เช่น การวิเคราะห์ยูสเคสสายผลิตภัณฑ์ซอฟต์แวร์ ซึ่งสามารถนำกลับมาใช้ใหม่ได้กับการพัฒนาซอฟต์แวร์สายงานผลิตภัณฑ์ (Product Line) เป็นต้น

2) การจัดเก็บและค้นคืนเอกสารก็ยังคงเป็นงานวิจัยหนึ่งที่มีผลงานการวิจัยออกมาเผยแพร่อย่างสม่ำเสมอ เนื่องจากเอกสารในยุคสมัยนี้มีรูปแบบหลากหลาย และต้องการการจัดเก็บและค้นคืนที่ให้ทั้งประสิทธิภาพและประสิทธิผล เอกสารคำอธิบายยูสเคสก็เช่นเดียวกันสามารถนำไปเพิ่มเติมข้อมูลที่เกี่ยวข้องกับการใช้งานของเอกสารได้ เช่น วันเวลาที่ยูสเคสถูกนำไปใช้ล่าสุด จำนวนความถี่ที่ยูสเคสถูกค้นคืนไปใช้งาน หรือความนิยมที่ยูสเคสถูกนำไปใช้ เป็นต้น ซึ่งจะเป็นประโยชน์ไม่น้อยหากผู้ค้นคืนหรือวิศวกรรมซอฟต์แวร์ได้รับข้อมูลในส่วนนี้ ก่อนการตัดสินใจเลือกนำยูสเคสออกมาใช้

3) การให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้อย่างสามารถประยุกต์กับเอกสารคำอธิบายยูสเคสได้ในอีกลักษณะหนึ่งคือ ในขั้นตอนที่ผู้ทำการค้นคืนเข้าไปอ่านรายละเอียดของเอกสารนั้น ในงานวิจัยนี้กำหนดให้ผู้ค้นคืนเลือกเอกสารคำอธิบายยูสเคสที่คิดว่าตรงประเด็นโดยเลือกทั้งเอกสาร ซึ่งหากทำการเพิ่มเติมฟังก์ชันให้ผู้ค้นคืนทำการเลือกเฉพาะส่วนประกอบของเอกสารคำอธิบายยูสเคสได้ด้วย ก็น่าที่จะให้ผลลัพธ์ในการให้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ที่ดีขึ้นกว่าการเลือกทั้งเอกสารก็เป็นไปได้

4) ลักษณะเด่นของการใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้นั้น ทำให้ผลลัพธ์ของเอกสารที่ระบบค้นคืน ทำการค้นคืนออกมานั้นเป็นที่น่าพอใจกับผู้ใช้ วิธีการนี้จึงสามารถนำไปประยุกต์ใช้กับเอกสารประเภทอื่นๆ ได้ เช่น เอกสารเว็บเพจ เป็นต้น

5) การจัดกลุ่มให้เอกสารยูสเคสนั้น อาจใช้วิธีการจัดกลุ่มในรูปแบบอื่นได้ เช่น จัดให้มีกลุ่มยูสเคสที่เกี่ยวข้องกับระบบลงทะเบียนผู้ใช้งาน (Log in system) เนื่องจากระบบงานคอมพิวเตอร์โดยทั่วไปมักมีฟังก์ชันนี้ เป็นต้น

6.3 บทความวิชาการที่ตีพิมพ์

ในการวิจัยนี้ ผู้วิจัยผลงานวิชาการร่วมกับคณะผู้วิจัย เป็นบทความวิชาการระดับชาติ รวมเป็น 2 บทความ (แสดงใน ภาคผนวก จ.) ได้แก่

1) บทความวิชาการเรื่อง “การค้นคืนยูสเคสด้วยการจำแนกประเภทยูสเคสและการค้นคืนย้อนกลับจากผู้ใช้ (Use Case Retrieval using Use Case Clustering and User Relevance

Feedback)” ซึ่งได้รับการคัดเลือกเพื่อนำเสนอและตีพิมพ์ในงาน “การประชุมวิชาการร่วมสาขา
วิทยาการคอมพิวเตอร์และวิศวกรรมซอฟต์แวร์ ครั้งที่ 3 (The 3rd Joint Conference on
Computer Science and Software Engineering: JCSSE 2006)” ระหว่างวันที่ 29 – 30
มิถุนายน 2549 ณ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า วิทยาเขตพระนครเหนือ กรุงเทพฯ

2) บทความวิชาการเรื่อง “การจัดเก็บและค้นคืนยูสเคสที่มีการจัดกลุ่มยูสเคสร่วมกับการ
ค้นคืนย้อนกลับจากผู้ใช้ (Use Case Clustered Storage and Retrieval with User Relevance
Feedback)” ซึ่งได้รับการคัดเลือกเพื่อนำเสนอและตีพิมพ์ในงาน “การประชุมวิชาการทาง
วิทยาการคอมพิวเตอร์และวิศวกรรมคอมพิวเตอร์ในระดับชาติ ครั้งที่ 11 (The 11th National
Computer Science and Engineering Conference: NCSEC2007)” ระหว่างวันที่ 19 – 21
พฤศจิกายน 2550 ณ โรงแรมมิราเคิลแกรนด์คอนเวนชัน กรุงเทพฯ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

- [1] Charles. W. Krueger. Software reuse. ACM Computing Survey 24 (1992) : 133-183.
- [2] Sadahiro Isoda. Do Actors Call Use Cases?-Solving the Problems of OOSE/UML Use-Case Relationships. Proceedings of the 26th Annual International Computer Software and Applications Conference (COMPSAC'02)., 2002.
- [3] Salton, G. and Buckley, C. The SMART and SIRE experimental retrieval systems, Introduction to Modern Information Retrieval, 141-143. (n.p.) : McGraw Hill, 1983.
- [4] Rocchio , J. Relevance feedback in information retrieval, The SMART Retrieval System, 313-323. (n.p.) : Prentice-Hall, 1971.
- [5] Saeki, M. Reusing use case descriptions for requirements specification: towards use case patterns. Proceedings of Asia-pacific software engineering conference (APSEC '99)., 1999.
- [6] Woo, H. G. and Robinson, W. N. Reuse of scenario specifications using an automated relational learner: a lightweight approach. Proceedings of IEEE Joint requirements engineering international conference., 2002.
- [7] อัครเดช อุดมชัยพร. การค้นคืนยูสเคสโดยใช้การคำนวณความคล้ายกันของพจน์และโครงสร้างของยูสเคส. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาวิศวกรรมซอฟต์แวร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2548.
- [8] Fowler, M. and Scott, K. UML Distilled. Addison-Wesley, 2000.
- [9] Alan Dennis and David Tegarden. Systems analysis and design, An object-oriented approach with UML. Wiley, 2002.
- [10] Salton, G. and Buckley, C. Introduction to Modern Information Retrieval. McGraw Hill, 1983.
- [11] Recardo, B. Yates and Berthier, R. Neto. Retrieval Evaluation. Modern Information Retrieval, 73-82. Addison-Wesley, 1999.
- [12] Porter, M. F. An algorithm for suffix stripping: Program. Automated library and information system 14 (1980) : 130-137.
- [13] Salton, G. and Buckley, C. Improving retrieval performance by relevance feedback. Journal of the American Society for Information Science 41 (1990) : 288-297.
- [14] C.T. Yu, C. Buckley, K. Lam, and G. Salton. A generalized term dependence model in information retrieval. Information Technology: Research and Development,

methods. IEEE transaction on software engineering 2 (October 1983) : 129-154.

- [15] Carol Lundquist, David A. Grossman and Ophir Frieder. Improving relevance feedback in the vector space model. Proceedings of the Sixth International Conference on Information and Knowledge Management (CIKM'97)., 1997.
- [16] Fang Yuan, Zeng Hui Meng, Hong Xia Zhang and Chun Ru Dong. A new algorithm to get the initial centriods. Proceedings of the Third Conference on Machine Learning and Cybernetics., 2004.
- [17] Jane H. Hayes, Alex Dekhtyar and Senthil K. Sundaram. Advancing candidate link generation for requirements tracing: the study of methods. IEEE transaction on software engineering 32 (January 2006) : 4-19.
- [18] กัลยา วานิชย์บัญชา. การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2546.



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก
ยูสเคสของระบบงานที่เลือกใช้

ในภาคผนวก ก จะอธิบายรายละเอียดของระบบงานทั้งหมดที่นำมาใช้ในงานวิจัย ซึ่งประกอบด้วยระบบงานทั้งสิ้น 40 ระบบงาน และจัดทำเป็นเอกสารคำอธิบายยูสเคสภาษาอังกฤษ แสดงดังตารางที่ ก.1

ตารางที่ ก.1 ระบบงานที่เลือกใช้ 40 ระบบงาน

หมายเลขระบบงาน	ชื่อระบบงาน	จำนวนยูสเคสในระบบ
1	ICT Blocking and Filtering	26
2	Departmental Information System for Curriculum and Course Offering Management	24
3	Resource Management	30
4	Automatic Restaurant Lookup	15
5	Library Management	11
6	Mobile Phone Selling on the Internet	12
7	Payroll	28
8	Personal Stock Manager on PDA	24
9	Basic Product Trading	16
10	Mobile Phone Service Paying	15
11	MLM Computerize	21
12	Content Management and Platform Upgrade for the Official Website of Nation Channel	14
13	Purchase, Production, and Export	14
14	Management Information System for Insurance Company	14
15	E-Learning	21
16	Management Information System for an Academic	30
17	SME Production Planning	35
18	E-Document	18
19	Purchase and Sale Real Estate Management Online	11

ตารางที่ ก.1 ระบบงานที่เลือกใช้ 40 ระบบงาน(ต่อ)

หมายเลขระบบงาน	ชื่อระบบงาน	จำนวนยูสเคสในระบบ
20	Simulation Office	30
21	Thai Spell Checking and Suggesting	7
22	MRP-Material Resource Planning	20
23	Hided Information Security	8
24	Next Generation Requirement Management	8
25	News Operation News Reporting	9
26	Customer Follow-up and Appointment	12
27	Support and Management Information System of Online Car Insurance	11
28	Central Help Desk	11
29	Cruise Control and Monitoring	14
30	Data Converter between Bank	3
31	ATM	4
32	Public Relation Online	14
33	News Operation Staff Training	10
34	Intelligent Dictionary	6
35	Job Assignment Management	7
36	The Primary Information System to Plan Education and Check Graduation for Undergraduate Students of Engineering Faculty Chulalongkorn University	6
37	Departmental Information System for Research management	15
38	Real Time News Feeding	10
39	Staff Procurement	5
40	Software Subcontract Management	23
รวม	ยูสเคสจากระบบงานทั้ง 40 ระบบ	<u>611</u>

ตารางที่ ก.2 ระบบงาน ICT Blocking and Filtering

ชื่อระบบงาน: ICT Blocking and Filtering	
คำอธิบายระบบ: เป็นระบบเพื่อค้นหาและป้องกันสิ่งที่ไม่ดีหรือไม่เหมาะสมบนอินเทอร์เน็ต และตรวจสอบดูแลนักเล่นเกมที่มีอายุต่ำกว่า 18 ปี นอกจากนี้ระบบจะช่วยผู้ปกครอง ครู และผู้ตรวจตราไซเบอร์ ในการลดความเสี่ยง อันตราย และข้อมูลซึ่งไม่เหมาะสมที่เด็กๆ อาจได้รับทางอินเทอร์เน็ต	
รหัสยูสเคส	ชื่อยูสเคส
0101	Register for a member
0102	Edit members information
0104	Cancel member status
0105	Request report about members
0106	Manage information for users
0107	Provide other information
0108	Display information
0109	Create token
0110	Create log file
0111	Verify log file
0112	Add game online users information
0113	Manage rules
0114	Receive log file from a game online
0115	View monitoring reports
0116	Pre-Screening
0117	Manage a bad website
0118	Manage a good website
0119	Manage suspended keywords
0120	View checking blocked information report
0121	Notify a trace
0122	Initialize sending SMS
0123	Notify SMS for emergency case
0124	Check notifying a trace
0125	View a notified trace report
0126	Define a system authority for a member
0127	View a report for defining authority for a member

ตารางที่ ก.3 ระบบงาน Departmental Information System for Curriculum and Course Offering Management

ชื่อระบบงาน: Departmental Information System for Curriculum and Course Offering Management	
คำอธิบายระบบ: เป็นระบบเพื่อจัดการข้อมูลรายวิชาและหลักสูตรในสถานศึกษาระดับภาควิชา	
รหัสยูสเคส	ชื่อยูสเคส
0201	Save course details
0202	Save course operation status
0203	Save course operation report
0204	Save course operation result
0205	View course operation details
0206	View course operation status
0207	View course operation report
0208	View course operation result
0209	Save subject details
0210	Save subject operation status
0211	Store subject report
0212	Save subject operation result
0213	View subject operation details
0214	View subject operation status
0215	View subject operation result
0216	View subject operation report
0217	Save course information
0218	Edit course information
0219	View course information
0220	View course statistic information
0221	Save subject information
0222	Edit subject information
0223	View subject information
0224	View teaching schedule

ตารางที่ ก.4 ระบบงาน Resource Management

ชื่อระบบงาน: Resource Management	
คำอธิบายระบบ: เป็นระบบช่วยจัดการการปฏิบัติงานภายในองค์กรหนึ่งๆ โดยสามารถแบ่งออกเป็นระบบงานย่อยอีก 6 ระบบงานย่อย ได้แก่ ระบบงานย่อยการจัดการลูกค้า ระบบงานย่อยการจัดการการขาย ระบบงานย่อยการจัดการลูกจ้าง ระบบงานย่อยการจัดการซัพพลายเออร์ ระบบงานย่อยการจัดการสินค้า และระบบงานย่อยการจัดการรายการสั่งซื้อสินค้าจากซัพพลายเออร์	
รหัสยูสเคส	ชื่อยูสเคส
0301	Process new application
0302	Process add new customer
0303	Process customer verification
0304	Process credit customer payment
0305	Process cash sales transactions
0306	Process credit sales
0307	Process Lay-Bye sales
0308	Process sales returns
0309	Capture employees details
0310	Process employee records
0311	Process existing supplier
0312	Process new supplier
0313	Process incoming inventory
0314	Process self-manufactured products
0315	Process check inventory available
0316	Process inventory order
0317	Access control
0318	Process generate password
0319	Process order details
0320	Process receive order details
0321	Process order payment
0322	Process account payment
0323	Generate sales figures
0324	Process petty cash withdrawals
0325	Process supplier payments

ตารางที่ ก.4 ระบบงาน Resource Management (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
0326	User trail
0327	Supplier order trail
0328	Customer payment trail
0329	Customer order trail
0330	Inventory trail

ตารางที่ ก.5 ระบบงาน Automatic Restaurant lookup

ชื่อระบบงาน: Automatic Restaurant lookup	
คำอธิบายระบบ: เป็นระบบสำหรับค้นหาชื่อร้านอาหาร ภัตตาคาร และอาหารจานเด่น โดยชื่อร้านอาหาร ภัตตาคารจะสามารถถูกค้นหาได้โดยขึ้นกับสถานที่ตั้ง ประเภทของอาหาร ยิ่งไปกว่านั้น ระบบนี้สามารถทำงานได้ทั้งบนคอมพิวเตอร์ส่วนบุคคล และบนพีดีเอ	
รหัสยูสเคส	ชื่อยูสเคส
0401	New customer
0402	Manage restaurant
0403	Manage customer
0404	Find a restaurant
0405	View restaurants information
0406	Vote restaurant
0407	Manage map
0408	Post webboard
0409	Manage webboard
0410	View FAQ
0411	Manage FAQ
0412	Log in
0413	Find restaurant on PDA
0414	View restaurant's information on PDA
0415	Vote restaurant on PDA

ตารางที่ ก.6 ระบบงาน Library Management

ชื่อระบบงาน: Library Management	
คำอธิบายระบบ: เป็นระบบจัดการงานในห้องสมุดซึ่งใช้เทคโนโลยี RFID โดยที่ระบบนี้จะจัดเตรียมการจัดการทรัพยากรและบริการสำหรับเจ้าหน้าที่ห้องสมุด นักเรียน นักศึกษา และผู้ใช้บริการห้องสมุดทั่วไป	
รหัสยูสเคส	ชื่อยูสเคส
0501	Manage users information
0502	Summarize users report
0503	Manage library media information
0504	Define library media status
0505	Summarize library media
0506	Add members information
0507	Manage members information
0508	Summarize members
0509	Borrow library media
0510	Return library media
0511	Pay a fine

ตารางที่ ก.7 ระบบงาน Mobile Phone Selling on the Internet

ชื่อระบบงาน: Mobile Phone Selling on the Internet	
คำอธิบายระบบ: เป็นระบบเพื่อการขายโทรศัพท์มือถือผ่านทางอินเทอร์เน็ต โดยระบบจะสามารถให้บริการได้ทั้งการค้นหาข้อมูลโทรศัพท์มือถือ ระบบการขาย และระบบการชำระเงิน นอกจากนี้ยังสามารถให้คำแนะนำลูกค้าเกี่ยวกับการเลือกโทรศัพท์มือถืออีกด้วย	
รหัสยูสเคส	ชื่อยูสเคส
0601	Search product's information
0602	View product's information
0603	Comment
0604	Compare products
0605	Purchase products
0606	Manage a product list
0607	Check purchase order
0608	Pay money
0609	Use credit card

ตารางที่ ก.7 ระบบงาน Mobile Phone Selling on the Internet (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
0610	Transfer money
0611	Confirm transferring money
0612	Verify user's payment

ตารางที่ ก.8 ระบบงาน Payroll

ชื่อระบบงาน: Payroll	
คำอธิบายระบบ: เป็นระบบเพื่อจัดทำตารางงานของพนักงานขายป๊อปคอร์นซึ่งกระจายกันอยู่ในร้านค้าหลายสาขา นอกจากนี้ยังคำนวณเงินเดือนของพนักงานขายซึ่งเงินเดือนจะขึ้นกับการปฏิบัติงานจริงตามตารางงานของพนักงานแต่ละคน	
รหัสยูสเคส	ชื่อยูสเคส
0701	Add new selling place
0702	Edit selling places information
0703	Cancel selling place
0704	Edit sale staff information
0705	Delete sale staff information
0706	View sale staff information
0707	Display work schedule
0708	Save working time
0709	Calculate staff salary
0710	Calculate social assurance fee
0711	Calculate over time fee
0712	Fine for working late
0713	Calculate special fee
0714	Fine for absent work
0715	Decrease tax
0716	Collect additional money from selling place
0717	Calculate net salary
0718	Display net salary
0719	Create report for executive
0720	Create report for sale staff or seller
0721	View selling place information

ตารางที่ ก.8 ระบบงาน Payroll (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
0722	Add new sale staff or seller
0723	Select a sale staff or seller
0724	Calculate salary
0725	Manage salary
0726	Save taking money in advance
0727	View taking money in advance information

ตารางที่ ก.9 ระบบงาน Personal Stock Manager on PDA

ชื่อระบบงาน: Personal Stock Manager on PDA	
คำอธิบายระบบ: เป็นระบบงานเพื่อบริหารจัดการข้อมูลเกี่ยวกับหุ้นผ่านพีดีเอ โดยที่ผู้ใช้สามารถดูข้อมูลของหุ้นแบบเรียลไทม์ และสามารถซื้อขายหุ้นผ่านพีดีเอได้	
รหัสยูสเคส	ชื่อยูสเคส
0801	Log in
0802	Verify user's account
0803	Search by abbreviate
0804	Search stock
0805	View top ten
0806	View most gain
0807	View most lost
0808	View most active
0809	View portfolio
0810	View transaction
0811	View index
0812	View search results
0813	View other currencies
0814	View stock details
0815	View graph
0816	Buy stock
0817	Sell stock
0818	Verify PIN
0819	Check regulation

ตารางที่ ก.9 ระบบงาน Personal Stock Manager on PDA (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
0820	Calculate price and commission
0821	Cancel transaction
0822	Check user regulation
0823	Maintain user account
0824	Maintain regulation

ตารางที่ ก.10 ระบบงาน Basic Product Trading

ชื่อระบบงาน: Basic Product Trading	
คำอธิบายระบบ: เป็นระบบงานเพื่อจัดการการขายสินค้าของบริษัทหนึ่งๆ ผ่านทางอินเทอร์เน็ต โดยมีระบบงานย่อยคือการบริหารจัดการลูกค้า การบริหารจัดการสินค้า การบริหารจัดการการขาย และการบริหารจัดการใบแจ้งหนี้	
รหัสยูสเคส	ชื่อยูสเคส
0901	Download purchase order
0902	Maintain purchase order
0903	Maintain sale order
0904	Transfer purchase order to sale order
0905	Approve purchase order
0906	Approve price
0907	Approve credit
0908	Approve stock balance
0909	Maintain invoice
0910	Transfer sale order to invoice
0911	Confirm invoice
0912	Update stock balance
0913	Update customer credit
0914	Maintain customer credit
0915	Maintain product price
0916	Maintain product balance

ตารางที่ ก.11 ระบบงาน Mobile Phone Service Paying

ชื่อระบบงาน: Mobile Phone Service Paying	
คำอธิบายระบบ: เป็นระบบเพื่อจัดการการให้บริการชำระค่าโทรศัพท์มือถือของบริษัทหนึ่ง ซึ่งระบบนี้สามารถจัดการข้อมูลของลูกค้า คำนวณค่าธรรมเนียมการใช้บริการตามประเภทของลูกค้า และสามารถจัดทำรายงานถึงผู้บริหารระดับสูงได้	
รหัสยูสเคส	ชื่อยูสเคส
1001	Register telephone number
1002	Manage customer and telephone number information
1003	Manage customer information
1004	Manage telephone number information
1005	Manage service information
1006	Cancel service
1007	Manage service package
1008	Save service information
1009	Ask for information
1010	Check service bill
1011	Check remaining money
1012	Create reports
1013	Print paying report
1014	Print paying report by terminal network
1015	Print income summary report

ตารางที่ ก.12 ระบบงาน Basic MLM Computerize

ชื่อระบบงาน: Basic MLM Computerize	
คำอธิบายระบบ: เป็นระบบงานเพื่อจัดการข้อมูลของระบบ MLM ระบบที่จัดการข้อมูลของลูกค้า จัดการสินค้าในคลังสินค้า จัดการการขายสินค้า และคำนวณรายได้ของพนักงานแต่ละคน	
รหัสยูสเคส	ชื่อยูสเคส
1101	Sign in member
1102	Maintain member information
1103	Show child member
1104	Cancel member
1105	Adjust position

ตารางที่ ก.12 ระบบงาน Basic MLM Computerize (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1106	Purchase order
1107	Receive product
1108	Adjust product quantity
1109	Maintain information
1110	New product information
1111	Change product information
1112	New position information
1113	Change position information
1114	Order product
1115	Create invoice
1116	Save point value
1117	Print invoice
1118	Cancel invoice
1119	Print commission
1120	Calculate commission
1121	Promote position

ตารางที่ ก.13 ระบบงาน Content Management and Platform Upgrade for the Official Website of Nation Channel

ชื่อระบบงาน: Content Management and Platform Upgrade for the Official Website of Nation Channel	
คำอธิบายระบบ: เป็นระบบเพื่อการจัดการข่าวบนเว็บไซต์แบบอัตโนมัติ โดยที่เว็บไซต์ดังกล่าวนี้จะจัดเตรียมเนื้อหาของข่าว การบริการไฟล์ บริการรูปภาพ การแพร่ภาพข่าวสด เป็นต้น ซึ่งระบบจะสนับสนุนการจัดการข้อมูลข่าวเป็นหลักและจัดการการบริการอื่นๆด้วย	
รหัสยูสเคส	ชื่อยูสเคส
1201	Log in
1202	Manage photo
1203	Load photo dialogue
1204	Edit schedule table
1205	Reply messages
1206	Manage program

ตารางที่ ก.13 ระบบงาน Content Management and Platform Upgrade for the Official Website of Nation Channel (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1207	Manage today news
1208	Link video clip
1209	View today news
1210	Post today news
1211	Leave message to anchorman
1212	Vote poll
1213	Read breaking news
1214	Activate live

ตารางที่ ก.14 ระบบงาน Purchase, Producing and Export

ชื่อระบบงาน: Purchase, Producing and Export	
คำอธิบายระบบ: เป็นระบบเพื่อจัดการการสั่งซื้อวัสดุ การผลิต การนำเข้าสินค้าและการส่งออกสินค้าของบริษัทแห่งหนึ่ง	
รหัสยูสเคส	ชื่อยูสเคส
1301	Purchase system
1302	Product system
1303	Warehouse and export system
1304	Check raw material
1305	Make contract and order raw material
1306	Track inventory and generate report
1307	Configure flow, hire rate and raw material
1308	Create barcode and preparation
1309	Produce
1310	Calculate hire and create report
1311	Check inventory
1312	Contact port
1313	Contact duty
1314	Export

ตารางที่ ก.15 ระบบงาน Management Information System for Insurance Company

ชื่อระบบงาน: Management Information System for Insurance Company	
คำอธิบายระบบ: เป็นระบบงานที่สนับสนุนการขายประกันของบริษัทประกันหนึ่ง ๆ โดยระบบจะจัดเตรียมข้อมูลเกี่ยวกับรายละเอียดของประกันภัยทั้งของพนักงานและลูกค้า นอกจากนี้ระบบยังสนับสนุนการทำงานบนคอมพิวเตอร์ส่วนบุคคลและเทคโนโลยีไร้สายอีกด้วย	
รหัสยูสเคส	ชื่อยูสเคส
1401	Place a claim
1402	View claim status
1403	External approve claim
1404	Maintain company information
1405	View company information
1406	View product information
1407	Calculate an insurance application
1408	Verify underwriting rule
1409	Calculate premium
1410	Validate insurance application
1411	View agent information
1412	View agent holder information
1413	View policy holder information
1414	Authentication

ตารางที่ ก.16 ระบบงาน E-Learning

ชื่อระบบงาน: E-Learning	
คำอธิบายระบบ: เป็นระบบงานเพื่อสนับสนุนระบบการเรียนอิเล็กทรอนิกส์บนเว็บไซต์ โดยที่นักเรียนและผู้ใช้สามารถลงทะเบียนและเข้าเรียนผ่านทางเว็บไซต์ โดยที่ระบบสามารถจัดการรายวิชา ข้อมูลนักเรียน ข้อมูลอาจารย์ผู้สอน การสอบ และสามารถประเมินการเรียนการสอนได้	
รหัสยูสเคส	ชื่อยูสเคส
1501	Register a new user for E-Learning System
1502	Generate password
1503	Log in to E-Learning System
1504	Change password

ตารางที่ ก.16 ระบบงาน E-Learning (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1505	Announce course information
1506	Upload materials
1507	Create examinations and examination solutions
1508	Post answer
1509	View course detail
1510	Register course
1511	Verify condition
1512	Study registered course
1513	Take examination and get result of the examination
1514	Record time
1515	View history's study information
1516	Post question
1517	View study result
1518	Follow study result
1519	View study result
1520	Number of learner or student
1521	Maintain employee information

ตารางที่ ก.17 ระบบงาน Management Information System for an Academic

ชื่อระบบงาน: Management Information System for an Academic	
คำอธิบายระบบ: เป็นระบบเพื่อจัดการรายวิชาในสถานศึกษา ระบบแบ่งออกเป็น 3 ส่วนหลักๆ ได้แก่ นักเรียนนักศึกษา คณะต่างๆ และเจ้าหน้าที่ ซึ่งระบบสามารถบริหารจัดการข้อมูลของนักเรียนนักศึกษา ข้อมูลคณะ ตารางเรียน และข้อมูลของแต่ละวิชา	
รหัสยูสเคส	ชื่อยูสเคส
1601	Log in
1602	Student adjust schedule
1603	Student view all courses
1604	Student view record
1605	Student view schedule
1606	Faculty view course
1607	Faculty view enrollment

ตารางที่ ก.17 ระบบงาน Management Information System for an Academic (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1608	Faculty view student records
1609	Add another admin
1610	Add another building
1611	Add another class
1612	Add another course
1613	Add another department
1614	Add another faculty
1615	Add student
1616	Modify existing admin
1617	Add existing building
1618	Modify existing classroom
1619	Add existing course
1620	Modify existing department
1621	Admin modify enrollment list
1622	Modify existing faculty member
1623	Modify existing student
1624	Remove administrator
1625	Remove building
1626	Remove classroom
1627	Remove course
1628	Remove department
1629	Remove faculty member
1630	Remove student

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ก.18 ระบบงาน SME Production Planning

ชื่อระบบงาน: SME Production Planning	
คำอธิบายระบบ: เป็นระบบเพื่อการจัดการระบบการวางแผนการผลิตในอุตสาหกรรมขนาดกลางและขนาดย่อม ซึ่งได้แก่การจัดการข้อมูลวัตถุดิบ ข้อมูลเครื่องจักร ข้อมูลซัพพลายเออร์ ข้อมูลรายการสั่งซื้อสินค้า ข้อมูลสายการผลิตต่างๆ ข้อมูลการซ่อมบำรุง รวมถึงตารางการทำงานของพนักงานด้วย	
รหัสยูสเคส	ชื่อยูสเคส
1701	Record new material
1702	Record new material by ID
1703	Use material
1704	Use material by ID
1705	Edit material specification
1706	Manage supplier
1707	Add supplier
1708	Edit supplier
1709	Search supplier
1710	Search material catalogue
1711	Add machine
1712	Edit machine
1713	Search machine
1714	Record workday
1715	Edit workday
1716	Add stock
1717	Add goods detail
1718	Add material catalogue
1719	Add machine catalogue
1720	Edit goods
1721	Search goods
1722	Add material
1723	Search material
1724	Edit material
1725	Increase maintenance schedule
1726	Search maintenance schedule

ตารางที่ ก.18 ระบบงาน SME Production Planning (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1727	Edit maintenance table
1728	Add customer
1729	Search customer
1730	Edit customer
1731	Search all purchase orders
1732	Manage new purchase order
1733	Check master production schedule
1734	Check vacant machine
1735	Create manual production line

ตารางที่ ก.19 ระบบงาน E-Document

ชื่อระบบงาน: E-Document	
คำอธิบายระบบ: เป็นระบบงานสารบรรณอิเล็กทรอนิกส์ กรณีศึกษามหาวิทยาลัยราชภัฏสวนดุสิต เพื่อการบริหารจัดการเอกสารในรูปแบบของเอกสารอิเล็กทรอนิกส์ โดยมีการจัดการของข้อมูลในส่วนของผู้ใช้เอกสาร การสร้างเอกสารอิเล็กทรอนิกส์ การเรียกดู ตรวจสอบ รับ ค้นหา รวมทั้งจัดประเภทของเอกสารอิเล็กทรอนิกส์	
รหัสยูสเคส	ชื่อยูสเคส
1801	Add user data
1802	Edit user data
1803	Delete user data
1804	Create official document
1805	View output document list
1806	Edit document
1807	Check document
1808	Approve document
1809	Keep log data
1810	Receive document
1811	View input document list
1812	Search document
1813	View log data
1814	Delete document

ตารางที่ ก.19 ระบบงาน E-Document (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
1815	Specify input document type
1816	Delete input document type
1817	Specify output document type
1818	Delete output document type

ตารางที่ ก.20 ระบบงาน Purchase and Sale Real Estate Management Online

ชื่อระบบงาน: Purchase and Sale Real Estate Management Online	
คำอธิบายระบบ: เป็นระบบเพื่อการบริหารจัดการข้อมูลฝากซื้อ ฝากขายอสังหาริมทรัพย์ผ่านทางเครือข่ายอินเทอร์เน็ต โดยมีการจัดการเกี่ยวกับข้อมูลของพนักงาน การจัดการข้อมูลสินค้า ฝากซื้อ ฝากขาย จัดการข้อมูลลูกค้า จัดการรายการสินค้า และการเรียกดูสินค้า เป็นต้น	
รหัสยูสเคส	ชื่อยูสเคส
1901	Manage staff data
1902	Specify staff authority
1903	Change password
1904	Manage property entrusted purchase data
1905	Manage property entrusted sale data
1906	Manage customer data
1907	Manage recommender data
1908	Match entrusted purchase and sale list
1909	View property entrusted purchase data
1910	View property entrusted sale data
1911	Close entrusted sale property

ตารางที่ ก.21 ระบบงาน Simulation Office

ชื่อระบบงาน: Simulation Office	
คำอธิบายระบบ: เป็นระบบจำลองการทำงานของสำนักงานแห่งหนึ่ง ซึ่งมีการบริหารจัดการงานโดยทั่วไป อันได้แก่ การบริหารจัดการทรัพยากรต่างๆ พาหนะ ห้องประชุม ห้องทำงาน อื่นๆ ที่ใช้ในสำนักงาน รวมทั้งการบริหารจัดการงาน และตารางการทำงานของเจ้าหน้าที่ด้วย	
รหัสยูสเคส	ชื่อยูสเคส
2001	Reserve other resource

ตารางที่ ก.21 ระบบงาน Simulation Office (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
2002	Cancel other resource reservation
2003	Check other resource reservation
2004	Approve other resource reservation list
2005	Return other resource
2006	Show reserved list
2007	Reserve vehicle
2008	Cancel vehicle reservation
2009	Check vehicle reservation
2010	Consider vehicle reservation
2011	Save vehicle usability
2012	Notify reserved result
2013	See meeting room reservation list
2014	Reserve meeting room
2015	Cancel meeting room reservation
2016	Check meeting room reservation
2017	Save information of meeting room usability
2018	See other room reservation list
2019	Reserve other room
2020	Cancel other room reservation
2021	Check other room reservation
2022	Save other room usability information
2023	Acknowledge
2024	Further assign or order
2025	Reply
2026	Create new job
2027	Create new job that send to archive
2028	Personal Calendar
2029	Name list book
2030	File form

ตารางที่ ก.22 ระบบงาน Thai Spell Checking and Suggesting

ชื่อระบบงาน: Thai Spell Checking and Suggesting	
คำอธิบายระบบ: เป็นระบบเพื่อตรวจสอบและแนะนำตัวสะกดสำหรับคำไทยที่สะกดผิดหนึ่งที โดยระบบแบ่งออกเป็นส่วนย่อยๆ ทั้งหมด 4 ส่วน ได้แก่ 1.ส่วนการตัดคำ ตรวจสอบ และแก้ไขตัวสะกด 2. ส่วนการสร้างพจนานุกรมและค้นหาคำในพจนานุกรม 3. ส่วนการจัดเรียงคำแนะนำ และส่วนสุดท้ายคือ ส่วนต่อประสานกราฟิก	
รหัสยูสเคส	ชื่อยูสเคส
2101	Create dictionary
2102	Segment word
2103	Search dictionary
2104	Correct word
2105	Generate candidate corrections
2106	Rank candidate corrections
2107	Replace word

ตารางที่ ก.23 ระบบงาน MRP-Material Resource Planning

ชื่อระบบงาน: MRP-Material Resource Planning	
คำอธิบายระบบ: เป็นระบบเพื่อการวางแผนการผลิต:กรณีศึกษาอุตสาหกรรมการผลิตตู้แช่ โดยมีเป้าหมายในการผลิตสินค้าให้ได้ตรงตามความต้องการของลูกค้า ประกอบด้วย 5 ระบบย่อย คือ ระบบการจัดการบุคลากร ระบบจัดการวัตถุดิบ ระบบจัดการเครื่องจักร ระบบจัดการสินค้า และระบบจัดตารางการทำงาน ซึ่งระบบสามารถทำหน้าที่ในการจัดการข้อมูลต่างๆ และจัดทำเอกสารตารางการผลิตหลัก เพื่อใช้ในการวางแผนการผลิต	
รหัสยูสเคส	ชื่อยูสเคส
2201	Record officer data
2202	Edit officer data
2203	Search officer data
2204	Record raw material data
2205	Edit raw material data
2206	Search raw material data
2207	Record product data
2208	Edit product data
2209	Search product data
2210	Record machine data

ตารางที่ ก.23 ระบบงาน MRP-Material Resource Planning (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
2211	Edit machine data
2212	Search machine data
2213	Manage fixing data
2214	Record master product planning
2215	Create task scheduling
2216	Search detail master production planning
2217	Create production order
2218	Check remain material
2219	Create insufficient material
2220	Look detail everyday task

ตารางที่ ก.24 ระบบงาน Hided Information Security

ชื่อระบบงาน: Hided Information Security	
คำอธิบายระบบ: เป็นระบบงานรักษาความปลอดภัยของข้อมูลที่ต้องปกปิด ซึ่งเกี่ยวข้องกับสิทธิ์การเข้าใช้ข้อมูลของผู้ใช้แต่ละคนที่มีระดับการเข้าใช้ข้อมูลที่แตกต่างกัน นั่นคือ ระบบนี้มีหน้าที่ในการบริหารจัดการข้อมูล และกำหนดสิทธิ์การเข้าใช้ให้เป็นไปตามนโยบายขององค์กร	
รหัสยูสเคส	ชื่อยูสเคส
2301	Manage MIS principle information
2302	Manage security policy
2303	Search access information
2304	MIS Logging
2305	Manage user group information
2306	Define authority
2307	Manage user information
2308	Search operation and report

ตารางที่ ก.25 ระบบงาน Next Generation Requirement Management

ชื่อระบบงาน: Next Generation Requirement Management

คำอธิบายระบบ: เป็นระบบการบริหารจัดการความต้องการต่าง ๆ ให้กับระบบงานหนึ่ง ๆ โดยสามารถทำการระบุคุณลักษณะต่าง ๆ ของความต้องการของระบบงานได้

รหัสยูสเคส	ชื่อยูสเคส
2401	Start a new system
2402	Define a list of attributes for requirements in the system
2403	Creates a sub-system for a system
2404	Defines feature for a sub-system or system
2405	Defines requirements
2406	Defines use case
2407	Defines system hierarchy
2408	Defines traceability

ตารางที่ ก.26 ระบบงาน News Operation News Reporting

ชื่อระบบงาน: News Operation News Reporting

คำอธิบายระบบ: เป็นระบบงานการรายงานข่าวการปฏิบัติการข่าวสาร ของการสื่อสารแห่งประเทศไทย ซึ่งเป็นระบบงานเพื่อการจัดการและจัดเก็บข้อมูล URL การจัดกลุ่ม URL การแปลข่าว การจัดการเอกสารต่างๆ รวมถึงการออกรายงาน

รหัสยูสเคส	ชื่อยูสเคส
2501	Store URL group data
2502	Store URL data
2503	Translate Thai into English
2504	Translate English into Thai
2505	Manage document group data
2506	Manage document file
2507	Manage operation news review
2508	Manage data encrypt-decrypt
2509	Make report

ตารางที่ ก.27 ระบบงาน Customer Follow-up and Appointment

ชื่อระบบงาน: Customer Follow-up and Appointment	
คำอธิบายระบบ: เป็นระบบที่ครอบคลุมในส่วนองงานฝ่ายลูกค้าสัมพันธ์ ซึ่งเป็นงานเกี่ยวกับการสอบถามและติดตามลูกค้า รวมถึงการนัดหมายลูกค้าล่วงหน้า โดยการทำงานบนระบบคอมพิวเตอร์นั้นประกอบด้วย ส่วนการนำเข้าข้อมูล (Import) จากฐานข้อมูลลูกค้า ส่วนการเรียกดูข้อมูลลูกค้า รวมถึงการบันทึกประวัติการติดต่อกับลูกค้าและประวัติการเข้ารับบริการของลูกค้า ส่วนการติดตามลูกค้า ส่วนการจัดการตารางนัดหมายในแต่ละวัน รวมทั้งส่วนการประมวลผลข้อมูลและแสดงในรูปแบบรายงานสรุปผลการดำเนินการ	
รหัสยูสเคส	ชื่อยูสเคส
2601	Import data
2602	Receive data from sell department
2603	Receive data from service center
2604	Follow customer
2605	Follow new customer
2606	Follow old customer
2607	Follow after service customer
2608	Make appointment
2609	Make appointment by follow-up
2610	Make appointment by customer
2611	Make follow-up report
2612	Make appointment report

ตารางที่ ก.28 ระบบงาน Support and Management Information System of Online Car Insurance

ชื่อระบบงาน: Support and Management Information System of Online Car Insurance	
คำอธิบายระบบ: เป็นระบบงานสนับสนุนและบริหารจัดการข้อมูลการทำประกันภัยรถยนต์ออนไลน์ให้มีประสิทธิภาพ เพื่อรองรับกับความต้องการของบริษัทผู้ค้า โดยการนำเทคโนโลยีเว็บเซอร์วิสเข้ามาช่วยในการพัฒนา	
รหัสยูสเคส	ชื่อยูสเคส
2701	Manage user data
2702	Manage insurance data
2703	Inform insurance data
2704	Manage quotation data

ตารางที่ ก.28 ระบบงาน Support and Management Information System of Online Car Insurance (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
2705	Print quotation document
2706	Print acceptance letter
2707	Print letter of regret
2708	Search quotation data
2709	Manage CMI data
2710	Search CMI data
2711	Conclude insurance contract

ตารางที่ ก.29 ระบบงาน Central Help Desk

ชื่อระบบงาน: Central Help Desk	
คำอธิบายระบบ: เป็นระบบงานเพื่อช่วยแก้ไขปัญหาด้านฮาร์ดแวร์และซอฟต์แวร์ โดยผู้ใช้ซึ่งใช้งานคอมพิวเตอร์และประสบกับปัญหาต่างๆ จะแจ้งปัญหาเข้ามายังระบบนี้ จากนั้นเจ้าหน้าที่ผู้ช่วยแก้ไขปัญหาจึงสามารถดำเนินการแก้ไขต่อไป	
รหัสยูสเคส	ชื่อยูสเคส
2801	Log in
2802	Display ticket list
2803	Add new ticket
2804	Response ticket
2805	Escalate ticket
2806	Close ticket
2807	Cancel ticket
2808	Look up user information
2809	Look up hardware information
2810	Look up software information
2811	Look up solution

ตารางที่ ก.30 ระบบงาน Cruise Control and Monitoring System Case Study

ชื่อระบบงาน: Cruise Control and Monitoring System Case Study	
คำอธิบายระบบ: เป็นระบบการควบคุมและตรวจสอบการเดินเรือ ซึ่งเป็นระบบงานคอมพิวเตอร์แบบเรียลไทม์ชนิดฝังในอุปกรณ์ ซึ่งแบ่งการทำงานออกเป็น 2 ส่วนคือ ส่วนควบคุมการเดินเรือ เช่นการควบคุมความเร็ว การกำหนดระยะทาง การแสดงการวัดต่างๆ และ ส่วนการตรวจสอบดูแล ซึ่งจะทำหน้าที่ ตรวจสอบการบำรุงรักษาน้ำมัน แอร์ รวมทั้งบริการต่างๆ การคำนวณการใช้เชื้อเพลิงในการเดินเรือแต่ละครั้ง เป็นต้น	
รหัสยูสเคส	ชื่อยูสเคส
2901	Control speed
2902	Perform calibration
2903	Update shaft rotation count
2904	Determine distance and speed
2905	Reset trip speed
2906	Reset trip fuel consumption
2907	Calculate trip speed
2908	Calculate trip fuel consumption
2909	Reset oil maintenance
2910	Reset Air Filter maintenance
2911	Reset major service maintenance
2912	Check oil maintenance
2913	Check Air Filter maintenance
2914	Check Major Service maintenance

ตารางที่ ก.31 ระบบงาน Data converter between bank

ชื่อระบบงาน: Data converter between bank	
คำอธิบายระบบ: เป็นระบบโปรแกรมปรับเปลี่ยนรูปแบบของแฟ้มข้อมูลที่ใช้ระหว่างธนาคารกรณีศึกษา ธนาคาร เอบีเอ็น แอมโร เพื่อรองรับการปรับเปลี่ยนรูปแบบของแฟ้มข้อมูลให้สอดคล้องกับระบบการพิมพ์เช็คของธนาคารหุ้นส่วนได้ โดยผู้ดูแลระบบจะสามารถกำหนดรูปแบบแฟ้มข้อมูลให้ตรงกับความต้องการของ ธนาคารหุ้นส่วนได้เอง	
รหัสยูสเคส	ชื่อยูสเคส
3001	Authenticate
3002	Manage Project
3003	Process Transaction File

ตารางที่ ก.32 ระบบงาน ATM

ชื่อระบบงาน: ATM	
คำอธิบายระบบ: เป็นระบบการทำงานของเครื่องกดเงินอัตโนมัติ หรือ เอทีเอ็ม โดยแสดงเฉพาะหน้าที่ในส่วนการทำงานของลูกค้าที่มาใช้บริการเท่านั้น	
รหัสยูสเคส	ชื่อยูสเคส
3101	Validate PIN
3102	Withdraw Funds
3103	Query Account
3104	Transfer Funds

ตารางที่ ก.33 ระบบงาน Public Relation Online

ชื่อระบบงาน: Public Relation Online	
คำอธิบายระบบ: เป็นระบบงานปฏิบัติการจิตวิทยาประชาสัมพันธ์ข้อมูลบนเครือข่ายสากล ซึ่งเกี่ยวข้องกับข่าวสารต่างๆ การบริหารจัดการข่าวสาร วิเคราะห์และเผยแพร่ข่าวสาร นอกจากนี้ยังเกี่ยวข้องกับการรับเรื่องราวร้องทุกข์ ตอบปัญหาเรื่องทุกข์ รวมถึงการจัดทำรายงานต่างๆ	
รหัสยูสเคส	ชื่อยูสเคส
3201	Manage data news broadcast issue
3202	Manage news
3203	Analyze news
3204	Manage RSS news
3205	Approve news broadcast
3206	Gather complaint
3207	Receive complaint online
3208	Receive complaints cross Call Center
3209	Answer problem
3210	Manage FAQ
3211	Make report
3212	Make Public information usage Report
3213	Make Works data Report
3214	Make Influence Report

ตารางที่ ก.34 ระบบงาน News Operation Staff Training

ชื่อระบบงาน: News Operation Staff Training	
คำอธิบายระบบ: เป็นระบบงานเพื่อบริหารจัดการการฝึกอบรมบุคลากรการปฏิบัติการข่าวสาร อันได้แก่ การจัดการกับข้อมูลต่างๆ เช่น ข้อมูลผู้เข้ารับการฝึก ข้อมูลบทเรียน ข้อมูลแบบทดสอบ เป็นต้น การประเมินการเข้าฝึกอบรม และการจัดทำรายงาน	
รหัสยูสเคส	ชื่อยูสเคส
3301	Manage Program Testing data
3302	Manage trainee data
3303	Manage Lesson Topic
3304	Manage lesson content
3305	Manage practical lesson content
3306	Manage theoretical lesson content
3307	Manage theoretical test
3308	Manage trainee data
3309	Manage training evaluation information
3310	Make report

ตารางที่ ก.35 ระบบงาน Intelligent Dictionary

ชื่อระบบงาน: Intelligent Dictionary	
คำอธิบายระบบ: โปรแกรมพจนานุกรมอัจฉริยะ เป็นโปรแกรมค้นหาคำศัพท์ซึ่งมีความสามารถในการช่วยแก้ปัญหากรณีที่ผู้ใช้พิมพ์ผิด โดยมีการแจ้งเตือนผู้ใช้งานว่าไม่พบคำในระบบโดยอาจเกิดจากการพิมพ์ผิดของผู้ใช้ หรืออาจเกิดจากคำที่ไม่มีในพจนานุกรมและสามารถประเมินคำที่ใกล้เคียงกับความต้องการของผู้ใช้ แสดงเป็นรายการให้ผู้ใช้เลือกได้	
รหัสยูสเคส	ชื่อยูสเคส
3401	Search
3402	Refine query
3403	Manage database
3404	Rank
3405	Find synonym
3406	Make Index

ตารางที่ ก.36 ระบบงาน Job Assignment Management

ชื่อระบบงาน: Job Assignment Management	
คำอธิบายระบบ: เป็นระบบเพื่อการจัดการงานของพนักงานผ่านเครือข่ายอินเทอร์เน็ต โดยระบบงานนี้สนับสนุนการจัดการปัญหาการใช้งานระบบสารสนเทศ การตรวจสอบสิทธิ์เข้าใช้ระบบ การจัดการงานที่ได้รับมอบหมาย การกำหนดสถานะของงาน เป็นต้น	
รหัสยูสเคส	ชื่อยูสเคส
3501	Request service
3502	Verify user
3503	Assign service
3504	Re-Assign service
3505	Management service
3506	Alarm Request
3507	Report Service

ตารางที่ ก.37 ระบบงาน The Primary Information System to Plan Education and Check Graduation for Undergraduate Students of Engineering Faculty, Chulalongkorn University

ชื่อระบบงาน: The Primary Information System to Plan Education and Check Graduation for Undergraduate Students of Engineering Faculty, Chulalongkorn University.	
คำอธิบายระบบ: เป็นระบบจัดการการวางแผนการศึกษาและตรวจสอบสถานะภาพการสำเร็จการศึกษาสำหรับนิสิตปริญญาตรี คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งสามารถเรียกดูรายละเอียดเงื่อนไขข้อบังคับของแต่ละวิชาและเรียกดูรายชื่อวิชาที่ตัวเองได้ลงทะเบียนไปแล้วผ่านทางอินเทอร์เน็ตช่วยให้นิสิตสามารถวางแผนจัดตารางเวลาการลงทะเบียนวิชาต่างๆ ของตัวเองได้ เป็นการลดงานของเจ้าหน้าที่งานทะเบียนคณะและสามารถตรวจสอบผลการจบศึกษาของนิสิตแต่ละคนได้ โดยแสดงเป็นรายงานต่างๆ	
รหัสยูสเคส	ชื่อยูสเคส
3601	Manage curriculum
3602	Manage course
3603	Plan education
3604	Search course information
3605	Check graduation status
3606	Create report graduate

ตารางที่ ก.38 ระบบงาน Departmental Information System for Research Management

ชื่อระบบงาน: Departmental Information System for Research Management	
คำอธิบายระบบ: เป็นระบบงานเพื่อใช้ในการบริหารจัดการด้านงานวิจัยในระดับภาควิชา ตามกระบวนการทางวิศวกรรมซอฟต์แวร์ โดยใช้ข้อมูลขั้นตอนการดำเนินงานด้านงานวิจัยของภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เป็นกรณีศึกษา	
รหัสยูสเคส	ชื่อยูสเคส
3701	Add thesis data
3702	Edit thesis data
3703	Add - Edit thesis theme data
3704	Show thesis data
3705	Show thesis theme data
3706	Add academic data
3707	Add conference data
3708	Add conference fund data
3709	Add - Edit organization data
3710	Add - Delete instructor of laboratory
3711	Edit academic data
3712	Show academic data
3713	Summary instructor conference
3714	Show organization data
3715	Report thesis proposal list

ตารางที่ ก.39 ระบบงาน Real Time News Feeding

ชื่อระบบงาน: Real Time News Feeding	
คำอธิบายระบบ: เป็นระบบการรับส่งข่าวแบบเรียลไทม์ที่สามารถใช้งานบนระบบปฏิบัติการวินโดวส์ โดยแก้ไขปัญหาระบบการรับส่งข่าวในปัจจุบัน ทั้งในด้านบริหารจัดการ การพัฒนาระบบและดูแลรักษาระบบ ซึ่งสามารถช่วยลดกำลังบุคลากรได้	
รหัสยูสเคส	ชื่อยูสเคส
3801	Control NFS Windows Service
3802	Maintain Configuration file
3803	Send Command

ตารางที่ ก.39 ระบบงาน Real Time News Feeding (ต่อ)

รหัสยูสเคส	ชื่อยูสเคส
3804	View Process Log
3805	Control News Transfer
3806	Load Configuration file
3807	Manipulate News
3808	Log Working Process
3809	Receive News
3810	Send News

ตารางที่ ก.40 ระบบงาน Staff Procurement

ชื่อระบบงาน: Staff Procurement	
คำอธิบายระบบ: เป็นระบบการจัดการพนักงานของบริษัทจัดหาพนักงานแห่งหนึ่งซึ่งทำการจัดหาพนักงานให้ตรงกับความต้องการของบริษัทลูกค้า เริ่มจากบริษัทลูกค้าระบุคุณสมบัติของพนักงานหรือความต้องการมา จากนั้นทางระบบจะตรวจสอบหลักฐานและเงื่อนไขต่างๆ หากถูกต้องจึงทำการตรวจสอบคุณสมบัติของพนักงานในฐานข้อมูล จองตัวพนักงาน ติดต่อพนักงาน และทำการปิดงานการจัดการพนักงานนั้นๆ	
รหัสยูสเคส	ชื่อยูสเคส
3901	Request staff
3902	Determine Contract
3903	Place available staff
3904	Arrange staff
3905	Close staffing request

ตารางที่ ก.41 ระบบงาน Software Subcontract Management

ชื่อระบบงาน: Software Subcontract Management	
คำอธิบายระบบ: เพื่อสนับสนุนการดำเนินการกระบวนการจัดการงานรับจ้างช่วงพัฒนาซอฟต์แวร์ มีหน้าที่การทำงานพื้นฐานที่สอดคล้องกับข้อกำหนดของกระบวนการหลักเรื่องจัดการงานรับจ้างช่วงพัฒนาซอฟต์แวร์ จากแบบจำลองวุฒิภาวะความสามารถการพัฒนาซอฟต์แวร์ระดับที่ 2 โดยจัดการฐานข้อมูลผู้รับจ้างพัฒนาซอฟต์แวร์ การจัดลำดับความสามารถผู้รับจ้างช่วงพัฒนาซอฟต์แวร์ การสืบค้นข้อมูลผู้รับจ้างช่วงพัฒนาซอฟต์แวร์ การบันทึกและรายงานสถานะสิ่งส่งมอบของโครงการ การปรับปรุงข้อมูลโครงการ การแสดงข้อความเตือนเมื่อมีกำหนดต้องส่งมอบงาน เพื่อติดตามการดำเนินงานของผู้รับจ้างช่วงพัฒนาซอฟต์แวร์ และการเรียกแสดงเอกสารที่สนับสนุนการดำเนินกิจกรรมของกระบวนการ	
รหัสยูสเคส	ชื่อยูสเคส
4001	Input data
4002	Authenticate
4003	Record data
4004	Present data
4005	Show data by letter order
4006	Show data by district order
4007	Identify weight
4008	Authenticate
4009	Evaluate capability
4010	Present capability rank
4011	Identify search condition
4012	Search by Subcontractor Name
4013	Search by District
4014	Search by Experience year
4015	Present searching data
4016	Record deliverable
4017	Authenticate
4018	Report deliverable status
4019	Update data
4020	Authenticate
4021	Identify alarm time
4022	Alarm at defined time
4023	Show electronic document

ภาคผนวก ข
ข้อคำถามจากหน่วยตัวอย่าง

สำหรับชุดของข้อคำถามทั้งหมดที่หน่วยตัวอย่างทั้ง 10 หน่วยตัวอย่างใช้ในแต่ละโจทย์
ปัญหา แสดงไว้ในตารางที่ ข.1-ข.5

ตารางที่ ข.1 แสดงชุดของข้อคำถามที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Educational System"

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อคำถาม	ส่วนประกอบที่ (Element)	หน้าหลัก ที่ให้
01	1	course	1	5
		check course	2	2
		student	3	5
	2	save course	1	5
		save course and report result	2	3
		staff	3	3
	3	subject	1	4
		display	2	2
		student	3	3
	4	check number of student	1	3
		check number of student in each course	2	3
		officer	3	3
	5	edit student detail	1	4
		student can edit our detail	2	2
		student	3	5
02	1	view curriculum	1	5
		Faculty Student	3	4
	2	Manage Grade	1	5
		Student Teacher	3	5
		Evaluate Grade	5	3
		Print Grade	6	2

ตารางที่ ข.1 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Educational System" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
02	2	Student completes the exam in all subjects and teacher scores them all	8	2
	3	Search Impact Factor	1	5
		To search for impact factor for the conferences and journals.	2	3
		Faculty Professor	3	5
	4	View Thesis Data	1	5
		Student Faculty	3	4
	5	study E-learning material	1	5
		Student Faculty	3	4
03	1	Educational System course management teacher information	1	5
		Teacher student	3	4
	2	Education course information, academic	1	5
		Teacher student	3	5
	3	Educational System course management	1	5
		Student	3	4
	4	Education course schedule study	1	4
		Student	3	5
	5	Education course	1	5
		student	3	5
04	1	education	1	5
		student	3	3
	2	education	1	5
		detail of course planning grade	2	5
		student	3	3

ตารางที่ ข.1 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Educational System" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
04	3	curriculum education	1	5
		planning course of detail view display	2	5
	4	education academic curriculum	1	5
		detail of course planning course information	2	5
		learner student	3	5
	5	education course information, curriculum	1	5
		student learner detail of course faculty	2	5
	05	1	research	1
student			3	4
2		course	1	5
		student	3	4
3		schedule	1	3
		student	3	5
4		subject	1	5
		teacher	3	5
5		thesis	1	5
		student teacher	3	5
06	1	education	1	5
		registration course graduation student teacher	2	5
		student teacher staff	3	5
	2	curriculum	1	5
		thesis	1	4
	3	status report	2	4
		scholarship	1	5
	4	student status grade GPA	2	5

ตารางที่ ข.1 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Educational System" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	หน้าหลัก ที่ให้
06	5	course	1	4
		teacher	2	3
07	1	login	1	4
		student	3	5
	2	add student	1	5
		administrator	3	5
	3	view	1	3
		student	3	3
	4	view	1	3
		teacher	3	3
	5	add teacher	1	4
		admin	3	5
08	1	Education System	1	5
	2	Education	1	4
		manage course	2	4
		teacher student	3	4
		students register	4	2
		login	5	2
	3	1. create course 2. remove course 3. summary course 4. print and public course into web	10	3
		1.1 create course name 1.2 create course objective and description 1.3 create course outline	11	2
		student register	1	5
	3	register to course study	2	2
student		3	5	

ตารางที่ ข.1 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Educational System" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
08	4	manage course online material	1	4
		manager outline lecture and content on internet	2	3
		teacher	3	3
		manager course online	4	2
		create the course subject before add any content	8	2
	5	view study result	1	4
09	1	Open curriculum	1	5
		To open new curriculum	2	5
	2	Track research progress	1	4
		To track research progress	2	2
		advisor	3	2
	3	Search thesis	1	5
		To search for thesis	2	2
		Student	3	2
	4	Add student data	1	4
		To add new student data into system	2	2
		Officer	3	1
	5	Edit study plan	1	4
	10	1	Select course	1
Course description			2	5
Teacher student			3	5
2		View research topic	1	5
		View research topic of student	2	5
		Teacher student	3	5
3		View grade	1	5
		student	3	3

ตารางที่ ข.1 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Educational System" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
10	4	Search curriculum or program	1	5
		Description of curriculum	2	4
		Student teacher	3	3
	5	Register course	1	5
		Student learner	3	3

ตารางที่ ข.2 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Material Requirements Planning/Manufacturing Planning"

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
01	1	planning project	1	4
	2	view machine	2	3
		officer	3	4
	3	view material	1	3
		officer	3	3
	4	check material	1	4
		manager	3	3
	5	edit detail	1	4
manager		3	4	
02	1	Manage inventory	1	5
		Officer, Maintainer	3	3
	2	check available material	1	5
		manufacturer, stock manager, engineer	3	5
		print available material	6	3
	3	View out of stock product	1	5
		Stock Manager	3	5
Manufacturing plan		8	3	

ตารางที่ ข.2 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Material Requirements Planning/Manufacturing Planning" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
02	4	Monitor manufacturing schedule and effort	1	4
		Manufacturing Manager	3	4
	5	manage requirement	1	4
		Manufacturer	3	4
03	1	Material Requirements Plan, Manufacturing Plan	1	5
	2	Material Requirements, Manufacturing	1	5
		Plan	2	5
	3	Material, Manufacture	1	5
		Planning	2	4
	4	Material Requirements Planning, Manufacturing Planning	1	4
		Planning	2	5
		Planning	1	4
	5	Material Requirements Planning, Manufacturing Planning	2	5
04	1	management, planning	1	5
		Requirements Planning, Manufacturing Planning	2	5
		manager	3	3
	2	requirement analysis, production planning, check material	1	5
		planning, Manufacturing	2	5
	3	Material, Planning, stock	1	5
		Planning, manage	2	5
	4	Requirements Planning, Manufacturing Planning, check stock	2	5

ตารางที่ ข.2 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Material Requirements Planning/Manufacturing Planning" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
04	4	Planning, analysis	3	5
	5	Material Requirements Planning	1	5
		Material Requirements Planning	2	5
		requirement analysis, production planning	3	3
05	1	product	1	5
	2	pricelist	1	5
		customer	3	4
	3	company detail	1	4
		customer buyer	3	4
	4	payment	1	4
	5	add supplier	1	5
		manager	3	5
06	1	planning	1	5
		material manufacturing	2	5
	2	Requirement	1	3
		Material Manufacturing Requirement	2	5
		staff employee	3	2
	3	planning	1	3
		demand supply	2	5
		store	3	5
	4	process	1	5
		warehouse store production	2	3
	5	planning	1	2
		manufactory requirements plan	2	3
07	1	View	1	3
		admin	3	5

ตารางที่ ข.2 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Material Requirements Planning/Manufacturing Planning" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
07	2	Product planning	1	3
		user	3	4
	3	Add material	1	4
		admin	3	5
	4	Log in	1	5
		User stock	3	3
	5	Order material	1	4
		stock	3	4
08	1	Material Requirements Planning	1	5
	2	order material	1	4
		planning to purchase material and request order	2	3
		officer	3	2
	3	production planning	1	5
		planning production line by the billing order	2	3
		industrial engineer	3	3
	4	inventory support	1	5
	5	production assurance	2	4
	09	1	Add material	1
To add new material			2	2
2		Show material	1	5
		Show detail of material	2	2
3		Use material	1	4
4		Check material remain	1	4
		Check material remain in stock	2	2
5	Plan product schedule	1	4	
10	1	Search material	1	5
		Staff	3	5

ตารางที่ ข.2 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Material Requirements Planning/Manufacturing Planning" (ต่อ)

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
10	2	Record material	1	5
		Store staff	3	5
	3	Process planning	1	5
		Process planning	2	4
	4	Edit stock goods material	1	5
		Edit stock goods material	2	4
	5	Check order	1	5
		Check order of stock goods material	2	4
		store	3	3

ตารางที่ ข.3 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Member/ User's Information Management"

รหัสหน่วยตัวอย่าง	ค้นคืนครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
01	1	search user	1	3
		user	3	4
	2	edit detail	1	4
		member	3	4
	3	view information	1	5
		manager	3	5
	4	view information log	1	3
		manager	3	4
	5	verify member data	1	5
		officer	3	3
02	1	view member privilege details	1	5
		User, Member	3	5
	2	check valid user	1	5
		administrator	3	5

ตารางที่ ข.3 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Member/ User's Information Management" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
02	3	manage user/member profile	1	5
		administrator, user, member	3	5
	4	remove inactive user/member	1	5
		Administrator	3	5
	5	assign premium member	1	5
		Administrator	3	3
03	1	user, member, information management	1	5
	2	information management	1	5
		user, member	3	5
	3	user, member, information	1	5
		user, member	3	4
	4	user, member, information	1	5
		manage user, member information	2	5
		user, member	3	3
	5	information management	1	3
		user, member	3	5
04	1	User Member credit	1	5
		check balance	2	4
		check credit	10	3
	2	User Member credit	1	5
		check credit, check balance	2	4
		customer, manager	3	3
	3	check credit	1	5
		check customer credit	2	3
		customer, manager	3	3
	4	manage customer credit	1	5
check customer credit		2	4	

ตารางที่ ข.3 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Member/ User's Information Management" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
04	4	customer, manager	3	3
	5	credit customer	1	5
		payment, check credit	2	4
		customer	3	3
05	1	edit username	1	4
		administrator	3	5
	2	remove member	1	3
		member	3	4
	3	edit member ,edit user ,edit account	1	4
		manager, administrator	3	4
	4	history edit remove add	1	5
	5	member information	1	5
06	1	management	1	5
	2	management	1	4
		user member information	2	4
		User member	3	4
	3	salary position	1	3
		user member information	2	5
		User member	3	5
	4	training	1	3
		experience knowledge	2	5
		member user officer	3	3
	5	event	1	2
		member event	2	2
		member user officer	3	5
07	1	add user	1	4
		admin	3	5

ตารางที่ ข.3 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Member/ User's Information Management" (ต่อ)

รหัสหน่วย ตัวอย่าง	ค้นคืน ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้	
07	2	edit user data	1	3	
		user	3	4	
	3	user member	1	3	
		admin	3	4	
	4	View user information	1	3	
		admin	3	5	
	5	register	1	5	
		user	3	4	
	08	1	User and member information management	1	5
		2	user login	1	4
sign-in to authenticate			2	4	
member			3	3	
user is a member			8	2	
3		Delete member data	1	5	
		Delete member data from database	2	4	
		Administration	3	5	
		Maintain member data	4	3	
		Remove all data and related field from database and keep operation log into file	9	4	
4		View profile	1	4	
		View member information	2	4	
		Member user administration	3	4	
5		Find member	1	3	
		Search matching member data	2	3	
		Member	3	3	
		User must be a member	8	2	

ตารางที่ ข.3 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Member/ User's Information Management" (ต่อ)

รหัสหน่วยตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
09	1	Add user	1	5
		Add new user	2	4
		administrator	3	5
	2	Edit member data	1	4
		Edit member data	2	2
	3	Log in	1	5
	3	user	3	5
	4	Verify user	1	5
		Verify user log in	2	1
	5	Delete user	1	4
		Delete user information	2	2
		Officer administrator	3	4
10	1	Register a new member or user	1	5
		Member user	3	5
	2	Update member or user profile	1	5
		Administrator or staff	3	5
	3	View member or user details	1	5
		View member or user details	2	5
		Member staff	3	5
	4	Manage member or user information	1	5
		Manage member or user information	2	5
	5	Member user information management	1	5
		Administrator staff manager	3	5

ตารางที่ ข.4 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Product Trading Operation"

รหัสหน่วย ตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
01	1	view product	1	5
		user	3	3
	2	check status of product	1	5
		manager	3	3
	3	Trading product	1	5
		user	3	3
	4	verify product	1	5
		for verify product list	2	2
		manager	3	2
	5	operate product	1	3
		user	3	3
	02	1	perform trade	1
staff			3	5
2		estimate product cost	1	5
		business analyst	3	5
3		establish product	1	5
		acquirer	3	5
4		evaluate product	1	5
		Acquirer	3	4
5		purchase product	1	5
		customer	3	4
03	1	Product trading	1	5
	2	Product Trading Operation	1	5
	3	Product Operation	1	5
	4	Product	1	5
	5	Trading	1	5
		Product	3	5

ตารางที่ ข.4 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Product Trading Operation" (ต่อ)

รหัสหน่วย ตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
04	1	Product Trading Operation	1	5
	2	Transaction, sell product, purchase product	1	5
	3	transaction status of trade	1	5
		transaction status of trade	2	5
	4	sell purchase trading product	1	5
	5	list of product to purchase	1	5
		display list of product	2	5
05	1	import product	1	3
	2	Export product	1	5
	3	Calculate price	1	3
	4	Product trading	1	5
	5	order product	1	4
06	1	Product	1	5
		Trading information	2	3
	2	Trading	1	5
		Product Information	2	5
	3	trading operation	1	4
		product demand price	2	4
		customer	3	5
	4	product claim	1	4
	5	sale promotion	1	3
		commission	2	4
07	1	order product	1	3
		user	3	5
	2	add product	1	3
		admin	3	4
	3	Add goods	1	4
		admin	3	5

ตารางที่ ข.4 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Product Trading Operation" (ต่อ)

รหัสหน่วย ตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
07	4	customer	1	4
		admin	3	5
	5	Edit good product	1	3
		admin	3	5
08	1	Trading product	1	5
	2	Product operation	1	5
	3	Control product trading to inventory	2	5
		manager	3	3
	4	Trading report	1	5
		Generate trading product report	2	4
	5	Operate product for trading	1	5
		inventory	2	1
09	1	Search product	1	4
	2	Add product order	1	5
		Add new product order	2	1
	3	Add supplier	1	5
		Add new supplier data	2	5
	4	Add customer	1	5
	5	Search supplier	1	5
		Search for supplier data	2	2
10	1	Search product	1	5
		Search product or goods	2	4
		Customer seller	3	3
	2	Order product	1	5
		Customer order product	2	5
		customer	3	5
	3	Add new product	1	5
		Add new product	2	5

ตารางที่ ข.4 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Product Trading Operation" (ต่อ)

รหัสหน่วยตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
10	4	View product detail	1	5
		View product detail	2	4
		Customer staff salesman	3	3
	5	View order detail	1	5
		View order detail	2	4
		Staff salesman	3	3

ตารางที่ ข.5 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Report Generation"

รหัสหน่วยตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
01	1	view report	1	5
		manager	3	4
	2	create report	1	5
		company	3	2
	3	report information	1	4
		report information to manager	2	3
		user	3	3
	4	save report	1	4
		user	3	3
	5	summary report	1	5
02	1	present in graph representation	1	5
		Management, Executive	3	4
	2	design report	1	5
		operator	3	5
	3	print report	1	5
		officer, operator	3	5
	4	create annual report	1	5
		Manager	3	3

ตารางที่ ข.5 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Report Generation" (ต่อ)

รหัสหน่วย ตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	น้ำหนัก ที่ให้
02	5	create progress report	1	5
		staff	3	4
03	1	Report Generation	1	5
	2	Report	1	5
		Generation	2	5
	3	Report	1	5
		Generate	10	4
	4	report	1	5
Report Generation		2	4	
03	5	Report Generation	1	3
		Report Generation	2	5
04	1	report generation for error	1	5
	2	purchase order report	1	5
	3	report about member or user	1	5
	4	curriculum report	1	5
	5	appointment report	1	5
05	1	report status	1	5
	2	view report	1	4
	3	print report	1	3
	4	remove report	1	5
	5	make report	1	3
06	1	Report	1	5
		administrator	3	5
	2	report	1	5
		version control	2	4
	3	report service	1	5
	4	curriculum report generation	1	5
5	sale trading report	1	5	

ตารางที่ ข.5 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์ปัญหาเรื่อง "Report Generation" (ต่อ)

รหัสหน่วยตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบที่(Element)	น้ำหนักที่ให้
07	1	view	1	5
		administrator	3	5
	2	Print	1	4
		user	3	5
	3	edit	1	4
		administrator	3	5
	4	Make report	1	5
		officer	3	5
	5	report	1	4
		user	3	5
08	1	Daily report	1	5
		Static report system for every day	2	3
		user	3	3
	2	Operation report	1	5
		View operation summary	2	2
		officer	3	1
	3	Summary report	1	5
		View in education field	2	5
		Teacher student	3	1
	4	Purchase report	1	5
	5	Create system status report	1	5
		Generate report for monitor system status and help	2	5
	09	1	Generate report	1
Generate report			2	1
2		View report	1	5
3		Print income report	1	5

ตารางที่ ข.5 แสดงชุดของข้อความที่หน่วยตัวอย่างใช้สำหรับการค้นคืนยูสเคสในโจทย์
ปัญหาเรื่อง "Report Generation" (ต่อ)

รหัสหน่วย ตัวอย่าง	ครั้งที่	เซตข้อความ	ส่วนประกอบ ที่(Element)	หน้าหลัก ที่ให้
09	4	Print sale summary report	1	5
		To print a sale summary report	2	2
	5	Information report	1	4
		Report information	2	4
10	1	Create report	1	4
		Create report	2	3
		Manager staff	3	3
	2	Print summary report	1	5
		User manager	3	5
	3	Make customer report	1	5
		Make customer report	2	5
		staff	3	5
	4	View product report	1	5
		Producer staff	3	5
	5	View progress report	1	5
		Teacher advisor	3	5

ภาคผนวก ค
ค่าเรียกคืน ค่าความแม่นยำ
และค่าเฉลี่ยฮาร์โมนิค ที่ได้จากการทดลอง

ค่าเรียกคืน (R) ค่าความแม่นยำ (P) และค่าเฉลี่ยฮาร์โมนิค (F) ทั้งหมดที่ได้จากการทดลองแยกตามโจทย์ปัญหาที่ใช้ในการทดลองได้แก่

1. Educational System
2. Material Requirement Planning/ Manufacturing Planning
3. Member / User's information Management
4. Product Trading Operation
5. Report Generation

และแยกวิธีการค้นคืนที่ใช้ในการทดลองดังที่กล่าวไว้ในหัวข้อที่ 5.2.3 โดยในส่วนแรกจะเป็นผลการทดลองที่ได้จากการวิธีค้นคืนแบบไม่พิจารณาโครงสร้างของยูสเคส และวิธีค้นคืนแบบพิจารณาโครงสร้างของยูสเคส ตามลำดับ ซึ่งสรุปเป็นตารางได้ดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ค.1 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4

วิธีค้นคืนที่	1			2			3			4		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.573	0.640	0.604	0.719	0.489	0.582	0.677	0.970	0.798	0.583	0.966	0.727
2	0.542	0.319	0.402	0.698	0.253	0.371	0.708	0.333	0.453	0.708	0.238	0.356
3	0.490	0.959	0.648	0.771	0.301	0.433	0.479	0.979	0.643	0.927	0.255	0.400
4	0.604	0.365	0.455	0.740	0.302	0.429	0.719	0.645	0.680	0.719	0.267	0.390
5	0.573	0.437	0.496	0.656	0.255	0.367	0.438	1.000	0.609	0.646	0.236	0.345
6	0.542	0.693	0.608	0.615	0.578	0.596	0.583	0.496	0.536	0.583	0.496	0.536
7	0.563	0.482	0.519	0.625	0.233	0.339	0.615	0.317	0.418	0.594	0.479	0.530
8	0.167	0.232	0.194	0.313	0.170	0.221	0.260	0.198	0.225	0.260	0.139	0.181
9	0.542	0.464	0.500	0.583	0.283	0.381	0.583	0.253	0.353	0.583	0.253	0.353
10	0.417	0.465	0.440	0.479	0.451	0.465	0.479	0.451	0.465	0.479	0.451	0.465
11	0.583	0.519	0.549	1.000	0.179	0.304	0.406	1.000	0.578	1.000	0.167	0.287
12	0.583	0.767	0.663	0.917	0.313	0.467	0.917	0.854	0.884	0.917	0.300	0.451
13	0.563	0.523	0.545	1.000	0.184	0.311	0.198	1.000	0.330	1.000	0.171	0.291
14	0.510	0.731	0.601	0.740	0.755	0.747	0.625	0.723	0.670	0.625	0.723	0.670
15	0.563	1.000	0.720	0.750	0.973	0.847	0.604	0.967	0.744	0.604	0.967	0.744
16	0.365	1.000	0.534	0.375	0.900	0.529	0.365	1.000	0.534	0.365	1.000	0.534
17	0.479	0.622	0.541	0.708	0.391	0.504	0.698	0.588	0.638	0.698	0.358	0.474
18	0.552	0.473	0.510	0.917	0.230	0.368	0.542	0.520	0.531	0.917	0.221	0.356
19	0.552	0.609	0.579	0.958	0.292	0.448	0.958	0.564	0.710	0.958	0.262	0.412
20	0.625	0.759	0.686	0.958	0.287	0.441	0.958	0.814	0.880	0.958	0.263	0.413
21	0.344	1.000	0.512	0.344	1.000	0.512	0.344	1.000	0.512	0.344	1.000	0.512
22	0.573	1.000	0.728	0.729	0.921	0.814	0.583	0.966	0.727	0.583	0.966	0.727
23	0.344	0.611	0.440	0.333	0.582	0.424	0.344	0.611	0.440	0.344	0.611	0.440
24	0.208	0.870	0.336	0.271	0.867	0.413	0.240	0.852	0.374	0.240	0.852	0.374
25	0.458	0.936	0.615	0.500	0.873	0.636	0.469	0.938	0.625	0.469	0.938	0.625
26	0.667	0.780	0.719	0.896	0.462	0.610	0.792	0.411	0.541	0.792	0.411	0.541
27	0.052	1.000	0.099	0.063	0.857	0.117	0.052	1.000	0.099	0.052	1.000	0.099
28	0.177	0.250	0.207	0.271	0.260	0.265	0.188	0.184	0.186	0.115	0.177	0.139

ตารางที่ ค.1 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4 (ต่อ)

วิธีค้นคืนที่	1			2			3			4		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.313	0.714	0.435	0.521	0.602	0.559	0.427	0.547	0.480	0.427	0.547	0.480
30	0.354	0.919	0.511	0.406	0.886	0.557	0.396	0.884	0.547	0.396	0.884	0.547
31	0.344	0.418	0.377	0.385	0.440	0.411	0.344	0.413	0.375	0.344	0.413	0.375
32	0.469	0.398	0.431	0.646	0.300	0.409	0.646	0.827	0.725	0.646	0.279	0.390
33	0.479	0.554	0.514	0.552	0.564	0.558	0.500	0.457	0.478	0.500	0.457	0.478
34	0.333	0.444	0.381	0.458	0.500	0.478	0.354	0.366	0.360	0.354	0.366	0.360
35	0.313	0.297	0.305	0.521	0.259	0.346	0.208	0.357	0.263	0.510	0.234	0.321
36	0.052	1.000	0.099	0.896	0.214	0.345	0.042	1.000	0.080	0.042	1.000	0.080
37	0.792	0.392	0.524	0.885	0.230	0.366	0.875	0.356	0.506	0.875	0.305	0.453
38	0.563	0.900	0.692	0.625	0.896	0.736	0.615	0.894	0.728	0.615	0.894	0.728
39	0.438	0.235	0.305	0.635	0.189	0.292	0.583	0.602	0.593	0.656	0.188	0.292
40	0.354	0.425	0.386	0.438	0.404	0.420	0.177	0.548	0.268	0.427	0.304	0.355
41	0.219	0.600	0.321	0.229	0.195	0.211	0.230	0.177	0.200	0.229	0.177	0.200
42	0.031	0.300	0.057	0.240	0.225	0.232	0.094	0.077	0.085	0.094	0.077	0.085
43	0.313	0.330	0.321	0.448	0.254	0.325	0.417	0.284	0.338	0.417	0.205	0.275
44	0.479	0.243	0.323	0.854	0.196	0.318	0.542	0.981	0.698	0.500	0.308	0.381
45	0.354	0.442	0.393	0.469	0.285	0.354	0.135	0.295	0.186	0.448	0.267	0.335
46	0.740	0.855	0.793	0.854	0.343	0.490	0.906	0.659	0.763	0.906	0.275	0.422
47	0.531	0.554	0.543	0.542	0.403	0.462	0.552	0.457	0.500	0.552	0.457	0.500
48	0.469	0.616	0.533	0.510	0.557	0.533	0.500	0.457	0.478	0.500	0.457	0.478
49	0.625	0.531	0.574	0.708	0.412	0.521	0.719	0.523	0.605	0.719	0.377	0.495
50	0.552	0.869	0.675	0.646	0.886	0.747	0.646	0.886	0.747	0.646	0.886	0.747

ตารางที่ ค.2 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4

วิธีค้นคืนที่	1			2			3			4		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.224	0.528	0.314	0.224	0.528	0.314	0.224	0.528	0.314	0.224	0.528	0.314
2	0.247	0.153	0.189	0.318	0.157	0.210	0.365	0.135	0.197	0.365	0.135	0.197
3	0.294	0.170	0.216	0.376	0.183	0.246	0.400	0.141	0.209	0.259	0.195	0.222
4	0.388	0.268	0.317	0.506	0.189	0.276	0.000	0.000	0.000	0.000	0.000	0.000
5	0.424	0.238	0.305	0.729	0.224	0.343	0.000	0.000	0.000	0.000	0.000	0.000
6	0.165	0.113	0.134	0.388	0.149	0.215	0.329	0.301	0.315	0.200	0.090	0.124
7	0.494	0.316	0.385	0.706	0.204	0.317	0.706	0.400	0.511	0.706	0.186	0.294
8	0.494	0.296	0.370	0.659	0.207	0.315	0.659	0.389	0.489	0.659	0.176	0.278
9	0.294	0.362	0.325	0.435	0.178	0.253	0.435	0.468	0.451	0.282	0.511	0.364
10	0.294	0.225	0.255	0.247	0.160	0.194	0.353	0.789	0.488	0.365	0.143	0.205
11	0.482	0.494	0.488	0.482	0.446	0.463	0.482	0.446	0.463	0.482	0.446	0.463
12	0.482	0.494	0.488	0.482	0.436	0.458	0.482	0.446	0.463	0.482	0.446	0.463
13	0.471	0.833	0.602	0.471	0.769	0.584	0.471	0.800	0.593	0.471	0.800	0.593
14	0.471	0.519	0.494	0.494	0.452	0.472	0.482	0.446	0.463	0.482	0.446	0.463
15	0.471	0.519	0.494	0.494	0.452	0.472	0.482	0.446	0.463	0.482	0.446	0.463
16	0.482	0.275	0.350	0.541	0.200	0.292	0.541	0.793	0.643	0.542	0.197	0.288
17	0.506	0.422	0.460	0.612	0.250	0.355	0.612	0.380	0.468	0.612	0.241	0.346
18	0.471	0.417	0.442	0.588	0.219	0.319	0.600	0.486	0.537	0.600	0.215	0.317
19	0.412	0.380	0.395	0.565	0.259	0.356	0.565	0.440	0.495	0.565	0.239	0.336
20	0.447	0.369	0.404	0.482	0.328	0.390	0.482	0.328	0.390	0.482	0.328	0.390
21	0.235	0.385	0.292	0.341	0.439	0.384	0.329	0.431	0.373	0.329	0.431	0.373
22	0.176	0.254	0.208	0.165	0.200	0.181	0.224	0.224	0.224	0.224	0.224	0.224
23	0.318	0.303	0.310	0.506	0.259	0.343	0.529	0.333	0.409	0.224	0.200	0.211
24	0.094	0.800	0.168	0.247	0.778	0.375	0.188	0.800	0.305	0.188	0.800	0.305
25	0.247	0.167	0.199	0.329	0.151	0.207	0.341	0.604	0.436	0.247	0.176	0.206
26	0.471	0.833	0.602	0.471	0.800	0.593	0.471	0.800	0.593	0.471	0.800	0.593
27	0.435	0.298	0.354	0.435	0.161	0.235	0.435	0.160	0.234	0.435	0.389	0.411
28	0.341	0.492	0.403	0.400	0.523	0.453	0.523	0.453	0.400	0.400	0.523	0.453

ตารางที่ ค.2 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4 (ต่อ)

วิธีค้นคืนที่	1			2			3			4		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.671	0.479	0.559	0.918	0.373	0.531	0.929	0.373	0.532	0.929	0.373	0.532
30	0.259	0.379	0.308	0.306	0.377	0.338	0.306	0.377	0.338	0.306	0.377	0.338
31	0.521	0.538	0.529	0.510	0.441	0.473	0.531	0.447	0.486	0.531	0.447	0.486
32	0.306	0.394	0.344	0.671	0.213	0.324	0.694	0.557	0.618	0.694	0.211	0.324
33	0.259	0.268	0.263	0.329	0.197	0.247	0.318	0.397	0.353	0.329	0.190	0.241
34	0.153	0.123	0.136	0.647	0.167	0.265	0.224	0.373	0.279	0.635	0.160	0.255
35	0.341	0.483	0.400	0.388	0.508	0.440	0.341	0.475	0.397	0.341	0.475	0.397
36	0.424	0.439	0.431	0.424	0.414	0.419	0.424	0.414	0.419	0.424	0.414	0.419
37	0.459	0.305	0.366	0.576	0.232	0.331	0.576	0.229	0.328	0.576	0.229	0.328
38	0.459	0.481	0.470	0.494	0.467	0.480	0.494	0.467	0.480	0.494	0.467	0.480
39	0.082	0.500	0.141	0.082	0.412	0.137	0.082	0.500	0.141	0.082	0.500	0.141
40	0.188	0.320	0.237	0.247	0.269	0.258	0.329	0.412	0.366	0.329	0.412	0.366
41	0.318	0.262	0.287	0.518	0.250	0.337	0.506	0.336	0.404	0.518	0.224	0.313
42	0.318	0.429	0.365	0.600	0.381	0.466	0.506	0.410	0.453	0.671	0.350	0.460
43	0.247	0.750	0.372	0.247	0.724	0.368	0.247	0.724	0.368	0.247	0.724	0.368
44	0.365	0.290	0.323	0.435	0.243	0.312	0.447	0.413	0.429	0.447	0.241	0.313
45	0.282	0.364	0.318	0.412	0.422	0.417	0.365	0.369	0.367	0.365	0.369	0.367
46	0.400	0.266	0.319	0.471	0.221	0.301	0.459	0.199	0.278	0.471	0.360	0.408
47	0.553	0.336	0.418	0.741	0.297	0.424	0.753	0.242	0.367	0.753	0.496	0.598
48	0.776	0.537	0.635	0.835	0.493	0.620	0.835	0.480	0.609	0.224	0.905	0.358
49	0.435	0.325	0.372	0.518	0.226	0.314	0.482	0.519	0.500	0.482	0.214	0.296
50	0.471	0.260	0.335	0.518	0.250	0.337	0.529	0.417	0.466	0.529	0.238	0.328

ตารางที่ ค.3 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4

วิธีค้นคืนที่	1			2			3			4		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.475	0.153	0.232	0.675	0.101	0.176	0.000	0.000	0.000	0.000	0.000	0.000
2	0.150	0.068	0.094	0.400	0.075	0.127	0.000	0.000	0.000	0.000	0.000	0.000
3	0.375	0.094	0.150	0.750	0.088	0.158	0.000	0.000	0.000	0.000	0.000	0.000
4	0.500	0.106	0.175	0.625	0.084	0.148	0.000	0.000	0.000	0.000	0.000	0.000
5	0.400	0.096	0.155	0.550	0.106	0.177	0.575	0.086	0.150	0.400	0.088	0.144
6	0.250	0.092	0.134	0.800	0.099	0.176	0.350	0.200	0.255	0.750	0.163	0.268
7	0.625	0.195	0.298	0.875	0.109	0.194	0.025	0.037	0.030	0.875	0.097	0.175
8	0.675	0.160	0.258	0.925	0.097	0.176	0.025	0.014	0.018	0.850	0.115	0.203
9	0.525	0.276	0.362	0.800	0.132	0.227	0.075	0.071	0.073	0.850	0.117	0.205
10	0.325	0.194	0.243	0.350	0.161	0.220	0.625	0.181	0.281	0.625	0.181	0.281
11	0.575	0.192	0.288	0.900	0.081	0.148	0.025	0.033	0.029	0.950	0.099	0.180
12	0.475	0.142	0.218	0.900	0.082	0.151	0.000	0.000	0.000	0.000	0.000	0.000
13	0.775	0.238	0.365	0.950	0.100	0.181	0.100	0.133	0.114	0.400	0.100	0.160
14	0.800	0.158	0.264	0.950	0.082	0.150	0.000	0.000	0.000	0.000	0.000	0.000
15	0.475	0.142	0.218	0.925	0.082	0.151	0.000	0.000	0.000	0.000	0.000	0.000
16	0.375	0.139	0.203	0.775	0.100	0.178	0.100	0.077	0.087	0.775	0.095	0.169
17	0.375	0.125	0.188	0.900	0.093	0.168	0.150	0.056	0.081	0.900	0.080	0.147
18	0.200	0.060	0.092	0.525	0.074	0.130	0.050	0.033	0.040	0.525	0.072	0.126
19	0.250	0.063	0.101	0.500	0.070	0.123	0.025	0.017	0.020	0.525	0.072	0.126
20	0.125	0.057	0.079	0.250	0.065	0.103	0.125	0.045	0.066	0.275	0.067	0.108
21	0.625	0.202	0.305	0.875	0.157	0.266	0.350	0.163	0.222	0.875	0.148	0.253
22	0.025	0.143	0.043	0.050	0.074	0.060	0.025	0.143	0.043	0.025	0.143	0.043
23	0.625	0.130	0.216	0.925	0.090	0.165	0.000	0.000	0.000	0.800	0.104	0.184
24	0.300	0.115	0.167	0.525	0.111	0.183	0.100	0.129	0.113	0.575	0.120	0.199
25	0.400	0.184	0.252	0.675	0.118	0.201	0.000	0.000	0.000	0.000	0.000	0.000
26	0.225	0.095	0.133	0.400	0.094	0.152	0.000	0.000	0.000	0.000	0.000	0.000
27	0.675	0.152	0.248	0.925	0.081	0.149	0.000	0.000	0.000	0.000	0.000	0.000
28	0.625	0.189	0.291	0.925	0.095	0.173	0.100	0.074	0.085	0.275	0.067	0.108

ตารางที่ ค.3 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการคิดคืนที่ 1 2 3 และ 4 (ต่อ)

วิธีคืนคืนที่	1			2			3			4		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.075	0.047	0.058	0.800	0.103	0.182	0.200	0.061	0.094	0.800	0.092	0.165
30	0.125	0.068	0.088	0.775	0.095	0.170	0.150	0.054	0.079	0.275	0.081	0.125
31	0.300	0.148	0.198	0.725	0.128	0.217	0.025	0.024	0.025	0.650	0.111	0.190
32	0.550	0.131	0.212	0.850	0.099	0.177	0.000	0.000	0.000	0.375	0.083	0.136
33	0.175	0.130	0.149	0.575	0.122	0.202	0.025	0.024	0.025	0.650	0.111	0.190
34	0.450	0.130	0.202	0.925	0.096	0.175	0.025	0.023	0.024	0.925	0.097	0.175
35	0.225	0.220	0.222	0.675	0.142	0.235	0.175	0.189	0.182	0.675	0.114	0.195
36	0.475	0.142	0.218	0.925	0.084	0.155	0.000	0.000	0.000	0.000	0.000	0.000
37	0.450	0.137	0.211	0.700	0.106	0.184	0.300	0.068	0.111	0.700	0.104	0.181
38	0.550	0.096	0.163	0.925	0.090	0.163	0.925	0.089	0.163	0.775	0.105	0.186
39	0.600	0.155	0.246	0.850	0.113	0.199	0.025	0.016	0.019	0.950	0.092	0.168
40	0.225	0.077	0.115	0.600	0.095	0.164	0.225	0.065	0.101	0.750	0.090	0.160
41	0.700	0.155	0.253	0.875	0.113	0.200	0.025	0.100	0.040	0.700	0.160	0.260
42	0.275	0.079	0.122	0.600	0.114	0.192	0.050	0.063	0.056	0.550	0.107	0.179
43	0.375	0.167	0.231	0.625	0.103	0.177	0.000	0.000	0.000	0.000	0.000	0.000
44	0.525	0.145	0.227	0.725	0.110	0.191	0.025	0.067	0.036	0.350	0.086	0.139
45	0.775	0.164	0.271	0.925	0.091	0.166	0.000	0.000	0.000	0.450	0.088	0.147
46	0.475	0.202	0.284	0.675	0.119	0.202	0.175	0.189	0.182	0.675	0.114	0.195
47	0.500	0.124	0.199	0.925	0.093	0.169	0.300	0.051	0.087	0.925	0.082	0.150
48	0.400	0.089	0.145	0.850	0.084	0.154	0.375	0.099	0.156	0.425	0.084	0.141
49	0.650	0.116	0.196	0.925	0.080	0.147	0.000	0.000	0.000	0.000	0.000	0.000
50	0.725	0.139	0.233	0.975	0.077	0.142	0.000	0.000	0.000	0.000	0.000	0.000

ตารางที่ ค.4 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4

วิธีค้นคืนที่	1			2			3			4		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.483	0.226	0.308	0.603	0.135	0.221	0.552	0.260	0.354	0.638	0.116	0.196
2	0.483	0.204	0.287	0.466	0.104	0.170	0.000	0.000	0.000	0.517	0.166	0.251
3	0.414	0.364	0.387	0.569	0.150	0.237	0.569	0.258	0.355	0.483	0.400	0.438
4	0.569	0.194	0.289	0.672	0.127	0.213	0.552	0.291	0.381	0.690	0.113	0.194
5	0.466	0.265	0.338	0.638	0.148	0.240	0.603	0.261	0.365	0.586	0.262	0.362
6	0.121	0.086	0.101	0.172	0.068	0.098	0.362	0.120	0.180	0.362	0.120	0.180
7	0.397	0.418	0.407	0.448	0.361	0.400	0.466	0.342	0.394	0.466	0.342	0.394
8	0.000	0.000	0.000	0.138	0.320	0.193	0.017	0.077	0.028	0.017	0.077	0.028
9	0.379	0.468	0.419	0.448	0.356	0.397	0.448	0.371	0.406	0.448	0.371	0.406
10	0.655	0.576	0.613	0.724	0.483	0.579	0.724	0.483	0.579	0.724	0.483	0.579
11	0.431	0.481	0.455	0.466	0.386	0.422	0.483	0.400	0.438	0.483	0.400	0.438
12	0.500	0.349	0.411	0.603	0.259	0.363	0.621	0.267	0.373	0.621	0.267	0.373
13	0.483	0.311	0.378	0.569	0.254	0.351	0.586	0.262	0.362	0.586	0.262	0.362
14	0.397	0.442	0.418	0.431	0.385	0.407	0.448	0.400	0.423	0.448	0.400	0.423
15	0.431	0.481	0.455	0.466	0.386	0.422	0.483	0.400	0.438	0.483	0.400	0.438
16	0.500	0.349	0.411	0.552	0.252	0.346	0.621	0.267	0.373	0.621	0.267	0.373
17	0.828	0.527	0.644	0.897	0.433	0.584	0.879	0.443	0.590	0.879	0.443	0.590
18	0.155	0.148	0.151	0.190	0.149	0.167	0.172	0.152	0.161	0.172	0.152	0.161
19	0.776	0.556	0.647	0.828	0.462	0.593	0.845	0.467	0.601	0.845	0.467	0.601
20	0.672	0.273	0.388	0.914	0.141	0.244	0.552	0.421	0.478	0.948	0.132	0.231
21	0.397	0.469	0.430	0.431	0.391	0.410	0.448	0.388	0.416	0.448	0.388	0.416
22	0.397	0.442	0.418	0.448	0.400	0.423	0.448	0.377	0.409	0.448	0.377	0.409
23	0.190	0.239	0.212	0.207	0.245	0.224	0.207	0.245	0.224	0.245	0.224	0.245
24	0.431	0.481	0.455	0.483	0.378	0.424	0.483	0.400	0.438	0.483	0.400	0.438
25	0.397	0.442	0.418	0.431	0.397	0.413	0.448	0.400	0.423	0.448	0.400	0.423
26	0.500	0.345	0.408	0.655	0.136	0.226	0.034	0.286	0.062	0.466	0.104	0.170
27	0.500	0.345	0.408	0.655	0.132	0.220	0.069	0.286	0.111	0.466	0.104	0.170
28	0.603	0.297	0.398	0.810	0.239	0.369	0.810	0.239	0.369	0.810	0.239	0.369

ตารางที่ ค.4 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4 (ต่อ)

วิธีค้นคืนที่	1			2			3			4		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.397	0.460	0.426	0.431	0.352	0.388	0.448	0.382	0.413	0.448	0.382	0.413
30	0.293	0.370	0.327	0.345	0.328	0.336	0.345	0.364	0.354	0.345	0.364	0.354
31	0.397	0.288	0.333	0.552	0.129	0.209	0.517	0.288	0.370	0.603	0.126	0.209
32	0.379	0.220	0.278	0.500	0.167	0.250	0.448	0.255	0.325	0.500	0.162	0.245
33	0.017	0.014	0.016	0.069	0.035	0.047	0.017	0.036	0.023	0.103	0.048	0.066
34	0.259	0.165	0.201	0.414	0.212	0.281	0.414	0.212	0.281	0.414	0.282	0.336
35	0.414	0.209	0.277	0.603	0.164	0.257	0.448	0.248	0.319	0.569	0.154	0.243
36	0.431	0.481	0.455	0.431	0.403	0.417	0.483	0.400	0.438	0.483	0.400	0.438
37	0.483	0.311	0.378	0.586	0.264	0.364	0.586	0.262	0.362	0.586	0.262	0.362
38	0.500	0.305	0.379	0.621	0.181	0.280	0.241	0.177	0.204	0.569	0.142	0.227
39	0.483	0.286	0.359	0.552	0.230	0.325	0.552	0.246	0.340	0.552	0.246	0.340
40	0.517	0.341	0.411	0.603	0.261	0.365	0.621	0.255	0.362	0.483	0.368	0.418
41	0.397	0.240	0.299	0.431	0.200	0.273	0.466	0.314	0.375	0.466	0.186	0.266
42	0.414	0.205	0.274	0.500	0.184	0.269	0.448	0.351	0.394	0.500	0.178	0.262
43	0.138	0.068	0.091	0.276	0.085	0.130	0.138	0.133	0.136	0.276	0.076	0.119
44	0.293	0.155	0.202	0.397	0.147	0.215	0.448	0.232	0.306	0.466	0.148	0.224
45	0.138	0.074	0.096	0.224	0.069	0.105	0.138	0.121	0.129	0.259	0.075	0.117
46	0.517	0.210	0.299	0.690	0.197	0.307	0.759	0.268	0.396	0.759	0.208	0.326
47	0.603	0.333	0.429	0.724	0.318	0.442	0.724	0.321	0.444	0.724	0.321	0.444
48	0.414	0.205	0.274	0.500	0.178	0.262	0.448	0.351	0.394	0.500	0.178	0.262
49	0.638	0.190	0.292	0.879	0.155	0.264	0.897	0.182	0.302	0.810	0.170	0.281
50	0.293	0.112	0.162	0.483	0.117	0.188	0.414	0.144	0.213	0.638	0.128	0.213

ตารางที่ ค.5 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4

วิธีค้นคืนที่	1			2			3			4		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.773	0.246	0.374	0.818	0.173	0.286	0.136	0.207	0.164	0.886	0.289	0.436
2	0.750	0.359	0.485	0.864	0.330	0.478	0.432	0.352	0.388	0.864	0.311	0.458
3	0.750	0.209	0.327	0.864	0.106	0.188	0.000	0.000	0.000	0.000	0.000	0.000
4	0.727	0.286	0.410	0.818	0.132	0.227	0.136	1.000	0.240	0.818	0.632	0.713
5	0.773	0.872	0.819	0.795	0.660	0.722	0.841	0.627	0.718	0.841	0.627	0.718
6	0.205	0.125	0.155	0.273	0.080	0.124	0.023	0.500	0.043	0.114	0.152	0.130
7	0.727	0.464	0.566	0.909	0.320	0.473	0.909	0.333	0.488	0.909	0.333	0.488
8	0.818	0.259	0.393	0.909	0.202	0.331	0.909	0.182	0.303	0.909	0.182	0.303
9	0.773	0.298	0.430	0.886	0.163	0.276	0.227	0.833	0.357	0.864	0.352	0.500
10	0.750	0.277	0.405	0.886	0.218	0.350	0.614	0.169	0.265	0.932	0.195	0.323
11	0.750	0.623	0.680	0.841	0.500	0.627	0.841	0.507	0.632	0.841	0.507	0.632
12	0.750	0.623	0.680	0.841	0.500	0.627	0.841	0.507	0.632	0.841	0.507	0.632
13	0.750	0.623	0.680	0.841	0.500	0.627	0.841	0.507	0.632	0.841	0.507	0.632
14	0.795	0.507	0.619	0.841	0.500	0.627	0.841	0.507	0.632	0.841	0.507	0.632
15	0.750	0.508	0.606	0.841	0.500	0.627	0.841	0.507	0.632	0.841	0.507	0.632
16	0.750	0.303	0.431	0.818	0.313	0.453	0.841	0.285	0.425	0.841	0.285	0.425
17	0.727	0.492	0.587	0.818	0.419	0.554	0.818	0.419	0.554	0.818	0.419	0.554
18	0.750	0.516	0.611	0.818	0.205	0.327	0.432	0.514	0.469	0.818	0.632	0.713
19	0.705	0.795	0.747	0.750	0.493	0.595	0.818	0.590	0.686	0.818	0.590	0.686
20	0.705	0.738	0.721	0.705	0.564	0.626	0.818	0.563	0.667	0.818	0.563	0.667
21	0.750	0.458	0.569	0.818	0.409	0.545	0.841	0.398	0.540	0.841	0.398	0.540
22	0.750	0.351	0.478	0.864	0.358	0.507	0.886	0.289	0.436	0.886	0.289	0.436
23	0.773	0.694	0.731	0.841	0.587	0.692	0.864	0.543	0.667	0.864	0.543	0.667
24	0.705	0.738	0.721	0.727	0.762	0.744	0.818	0.563	0.667	0.818	0.563	0.667
25	0.705	0.646	0.674	0.795	0.593	0.680	0.818	0.632	0.713	0.818	0.632	0.713
26	0.727	0.364	0.485	0.841	0.222	0.351	0.841	0.220	0.349	0.841	0.220	0.349
27	0.705	0.544	0.614	0.795	0.417	0.547	0.818	0.450	0.581	0.818	0.450	0.581
28	0.705	0.470	0.564	0.750	0.446	0.559	0.818	0.375	0.514	0.818	0.375	0.514

ตารางที่ ค.5 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 1 2 3 และ 4 (ต่อ)

วิธีค้นคืนที่	1			2			3			4		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.750	0.579	0.653	0.841	0.481	0.612	0.841	0.481	0.612	0.841	0.481	0.612
30	0.750	0.452	0.564	0.818	0.319	0.459	0.841	0.359	0.503	0.841	0.359	0.503
31	0.250	0.092	0.135	0.364	0.086	0.139	0.341	0.077	0.125	0.341	0.077	0.125
32	0.318	0.275	0.295	0.500	0.242	0.326	0.273	0.316	0.293	0.273	0.522	0.358
33	0.227	0.091	0.130	0.250	0.055	0.091	0.068	0.086	0.076	0.182	0.066	0.096
34	0.727	0.281	0.405	0.864	0.244	0.380	0.886	0.209	0.338	0.886	0.209	0.338
35	0.750	0.516	0.611	0.841	0.152	0.258	0.432	0.514	0.469	0.818	0.632	0.713
36	0.818	0.456	0.585	0.909	0.107	0.191	0.045	0.200	0.074	0.818	0.474	0.600
37	0.864	0.235	0.369	0.955	0.166	0.283	0.977	0.155	0.267	0.977	0.155	0.267
38	0.773	0.330	0.463	0.886	0.235	0.371	0.886	0.234	0.370	0.886	0.234	0.370
39	0.727	0.492	0.587	0.773	0.400	0.527	0.818	0.419	0.554	0.818	0.419	0.554
40	0.864	0.262	0.402	0.977	0.078	0.145	0.000	0.000	0.000	0.250	0.306	0.275
41	0.750	0.508	0.606	0.841	0.493	0.622	0.841	0.507	0.632	0.841	0.507	0.632
42	0.750	0.351	0.478	0.818	0.295	0.434	0.886	0.289	0.436	0.886	0.289	0.436
43	0.773	0.680	0.723	0.864	0.585	0.697	0.864	0.535	0.661	0.864	0.535	0.661
44	0.886	0.443	0.591	0.909	0.255	0.398	0.909	0.190	0.314	0.909	0.190	0.314
45	0.727	0.244	0.366	0.864	0.138	0.238	0.023	0.333	0.043	0.818	0.632	0.713
46	0.818	0.243	0.375	0.932	0.137	0.239	0.136	0.162	0.148	0.932	0.196	0.324
47	0.864	0.380	0.528	0.909	0.147	0.252	0.295	0.260	0.277	0.909	0.175	0.293
48	0.773	0.241	0.370	0.932	0.180	0.301	0.932	0.178	0.299	0.932	0.178	0.299
49	0.750	0.220	0.340	0.977	0.172	0.293	0.977	0.162	0.277	0.955	0.226	0.365
50	0.705	0.369	0.484	0.818	0.263	0.398	0.886	0.271	0.415	0.886	0.271	0.415

ตารางที่ ค.6 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.438	1.000	0.609	0.448	1.000	0.619	0.458	1.000	0.629	0.458	0.587	0.515
2	0.396	0.481	0.434	0.333	0.627	0.435	0.521	0.407	0.457	0.521	0.360	0.426
3	0.365	1.000	0.534	0.313	1.000	0.476	0.375	0.667	0.480	0.375	0.667	0.480
4	0.323	0.295	0.308	0.313	0.366	0.337	0.438	0.429	0.433	0.438	0.368	0.400
5	0.375	0.766	0.504	0.354	0.756	0.482	0.521	0.485	0.503	0.521	0.543	0.532
6	0.427	0.641	0.513	0.365	0.897	0.519	0.427	0.539	0.477	0.427	0.539	0.477
7	0.052	0.556	0.095	0.365	0.921	0.522	0.396	0.731	0.514	0.396	0.826	0.535
8	0.063	0.333	0.105	0.073	0.259	0.114	0.125	0.090	0.105	0.125	0.090	0.105
9	0.448	0.422	0.434	0.396	0.603	0.478	0.448	0.344	0.389	0.302	0.367	0.331
10	0.323	0.660	0.434	0.271	0.743	0.397	0.333	0.640	0.438	0.333	0.640	0.438
11	0.542	0.658	0.594	0.510	0.731	0.601	0.594	0.731	0.655	0.573	0.362	0.444
12	0.542	0.929	0.684	0.500	0.906	0.644	0.563	1.000	0.720	0.563	0.514	0.537
13	0.438	0.646	0.522	0.438	0.933	0.596	0.479	0.742	0.582	0.458	0.427	0.442
14	0.479	0.902	0.626	0.458	0.917	0.611	0.490	0.870	0.627	0.490	0.870	0.627
15	0.438	1.000	0.609	0.438	1.000	0.609	0.448	1.000	0.619	0.448	1.000	0.619
16	0.250	1.000	0.400	0.219	1.000	0.359	0.250	1.000	0.400	0.250	1.000	0.400
17	0.385	0.974	0.552	0.365	0.897	0.519	0.500	0.727	0.593	0.500	0.727	0.593
18	0.448	0.896	0.597	0.365	0.897	0.519	0.510	0.458	0.483	0.510	0.458	0.483
19	0.396	0.905	0.551	0.458	0.863	0.599	0.208	0.625	0.313	0.563	0.730	0.635
20	0.521	0.926	0.667	0.500	0.941	0.653	0.604	0.795	0.686	0.573	0.786	0.663
21	0.240	1.000	0.387	0.250	1.000	0.400	0.240	1.000	0.387	0.240	1.000	0.387
22	0.427	1.000	0.599	0.385	1.000	0.556	0.438	1.000	0.609	0.438	1.000	0.609
23	0.250	0.800	0.381	0.260	0.806	0.394	0.250	0.774	0.378	0.250	0.774	0.378
24	0.198	1.000	0.330	0.375	1.000	0.545	0.198	1.000	0.330	0.198	1.000	0.330
25	0.375	1.000	0.545	0.375	1.000	0.545	0.375	1.000	0.545	0.375	1.000	0.545
26	0.521	0.714	0.602	0.646	0.939	0.765	0.698	0.545	0.612	0.698	0.545	0.612
27	0.010	1.000	0.021	0.010	1.000	0.021	0.010	1.000	0.021	0.010	1.000	0.021
28	0.167	0.410	0.237	0.156	0.682	0.254	0.167	0.444	0.242	0.167	0.327	0.221

ตารางที่ ค.6 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 1 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8 (ต่อ)

วิธีค้นคืนที่	5			6			7			8		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.125	1.000	0.222	0.177	1.000	0.301	0.167	0.615	0.262	0.167	0.615	0.262
30	0.240	1.000	0.387	0.240	1.000	0.387	0.240	0.958	0.383	0.240	0.958	0.383
31	0.240	0.958	0.383	0.219	0.955	0.356	0.240	0.958	0.383	0.240	0.958	0.383
32	0.417	0.417	0.417	0.365	0.515	0.427	0.417	0.460	0.437	0.417	0.317	0.360
33	0.417	0.563	0.479	0.344	1.000	0.512	0.417	0.533	0.468	0.417	0.533	0.468
34	0.271	0.456	0.340	0.260	0.862	0.400	0.271	0.426	0.331	0.271	0.426	0.331
35	0.354	0.378	0.366	0.188	0.277	0.224	0.302	0.382	0.337	0.302	0.382	0.337
36	0.010	0.200	0.020	0.042	1.000	0.080	0.031	0.214	0.055	0.031	0.214	0.055
37	0.000	0.000	0.000	0.000	0.000	0.000	0.479	0.648	0.551	0.500	0.873	0.636
38	0.406	0.951	0.569	0.375	0.947	0.537	0.510	0.925	0.658	0.510	0.925	0.658
39	0.021	0.143	0.036	0.271	0.413	0.327	0.281	0.587	0.380	0.406	0.520	0.456
40	0.281	0.458	0.348	0.292	0.778	0.424	0.281	0.429	0.340	0.281	0.429	0.340
41	0.021	1.000	0.041	0.021	1.000	0.041	0.021	1.000	0.041	0.021	1.000	0.041
42	0.021	0.333	0.039	0.073	0.636	0.131	0.083	0.070	0.076	0.083	0.070	0.076
43	0.240	0.605	0.343	0.167	0.444	0.242	0.313	0.199	0.243	0.250	0.358	0.294
44	0.302	0.218	0.253	0.365	0.368	0.366	0.198	0.174	0.185	0.531	0.188	0.278
45	0.125	0.400	0.190	0.135	0.500	0.213	0.125	0.375	0.188	0.125	0.375	0.188
46	0.656	1.000	0.792	0.667	1.000	0.800	0.719	0.986	0.831	0.719	0.986	0.831
47	0.479	0.767	0.590	0.500	0.889	0.640	0.521	0.562	0.541	0.521	0.568	0.543
48	0.417	0.588	0.488	0.344	0.892	0.496	0.417	0.533	0.468	0.417	0.533	0.468
49	0.563	0.806	0.663	0.500	0.889	0.640	0.635	0.663	0.649	0.635	0.663	0.649
50	0.438	0.955	0.600	0.365	0.946	0.526	0.448	0.956	0.610	0.448	0.956	0.610

ตารางที่ ค.7 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.024	0.500	0.045	0.024	0.400	0.044	0.024	0.500	0.045	0.024	0.500	0.045
2	0.106	0.078	0.090	0.106	0.118	0.112	0.106	0.113	0.109	0.106	1.000	0.191
3	0.165	0.109	0.131	0.188	0.168	0.178	0.188	0.176	0.182	0.188	0.667	0.294
4	0.294	0.352	0.321	0.294	0.714	0.417	0.329	0.277	0.301	0.329	0.277	0.301
5	0.294	0.362	0.325	0.247	0.368	0.296	0.329	0.286	0.306	0.329	0.452	0.381
6	0.047	0.045	0.046	0.047	0.114	0.067	0.047	0.400	0.084	0.047	0.667	0.088
7	0.106	0.220	0.143	0.424	0.632	0.507	0.471	0.336	0.392	0.471	0.336	0.392
8	0.141	0.167	0.153	0.176	0.313	0.226	0.224	0.141	0.173	0.224	0.141	0.173
9	0.329	0.452	0.381	0.259	0.688	0.376	0.341	0.326	0.333	0.341	0.326	0.333
10	0.176	0.341	0.233	0.176	0.455	0.254	0.176	0.882	0.294	0.188	0.211	0.199
11	0.212	0.750	0.330	0.200	0.739	0.315	0.212	0.720	0.327	0.212	0.720	0.327
12	0.224	0.760	0.345	0.212	0.621	0.316	0.224	0.731	0.342	0.224	0.731	0.342
13	0.224	0.826	0.352	0.259	1.000	0.411	0.224	0.792	0.349	0.224	0.792	0.349
14	0.235	0.870	0.370	0.235	0.870	0.370	0.235	0.741	0.357	0.235	0.741	0.357
15	0.271	0.639	0.380	0.271	0.719	0.393	0.271	0.548	0.362	0.271	0.548	0.362
16	0.141	0.182	0.159	0.118	0.417	0.183	0.282	0.233	0.255	0.165	0.161	0.163
17	0.306	0.565	0.397	0.294	0.556	0.385	0.318	0.415	0.360	0.318	0.415	0.360
18	0.224	0.311	0.260	0.235	0.541	0.328	0.000	0.000	0.000	0.247	0.500	0.331
19	0.153	0.464	0.230	0.200	0.586	0.298	0.200	0.386	0.264	0.200	0.436	0.274
20	0.271	0.575	0.368	0.318	0.614	0.419	0.353	0.556	0.432	0.353	0.556	0.432
21	0.106	0.409	0.168	0.094	0.444	0.155	0.106	0.391	0.167	0.106	0.391	0.167
22	0.082	0.206	0.118	0.082	0.194	0.116	0.082	0.171	0.111	0.082	0.171	0.111
23	0.106	0.209	0.141	0.082	0.159	0.109	0.141	0.226	0.174	0.141	0.226	0.174
24	0.047	0.800	0.089	0.071	0.857	0.130	0.059	0.833	0.110	0.059	0.833	0.110
25	0.259	0.314	0.284	0.247	0.356	0.292	0.294	0.248	0.269	0.294	0.248	0.269
26	0.212	0.783	0.333	0.212	1.000	0.350	0.212	0.783	0.333	0.212	0.783	0.333
27	0.224	0.333	0.268	0.224	0.576	0.322	0.247	0.176	0.206	0.247	0.525	0.336
28	0.094	0.889	0.170	0.141	1.000	0.247	0.141	0.923	0.245	0.141	0.923	0.245

ตารางที่ ค.7 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 2 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8 (ต่อ)

วิธีค้นคืนที่	5			6			7			8		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.682	0.547	0.607	0.329	0.651	0.438	0.376	0.500	0.430	0.376	0.500	0.430
30	0.071	0.333	0.117	0.071	0.400	0.120	0.118	0.400	0.182	0.118	0.400	0.182
31	0.229	0.537	0.321	0.083	0.571	0.145	0.229	0.386	0.288	0.229	0.386	0.288
32	0.059	0.104	0.075	0.094	0.533	0.160	0.106	0.088	0.096	0.106	0.375	0.165
33	0.200	0.500	0.286	0.212	0.600	0.313	0.259	0.458	0.331	0.259	0.458	0.331
34	0.047	0.148	0.071	0.047	0.125	0.068	0.047	0.043	0.045	0.047	0.043	0.045
35	0.200	0.850	0.324	0.212	1.000	0.350	0.224	0.826	0.352	0.224	0.826	0.352
36	0.200	0.739	0.315	0.200	0.773	0.318	0.200	0.708	0.312	0.200	0.708	0.312
37	0.200	0.181	0.190	0.247	0.412	0.309	0.271	0.343	0.303	0.271	0.160	0.201
38	0.282	0.545	0.372	0.353	0.638	0.455	0.353	0.484	0.408	0.353	0.484	0.408
39	0.047	0.667	0.088	0.047	0.667	0.088	0.047	0.667	0.088	0.047	0.667	0.088
40	0.118	0.263	0.163	0.129	0.297	0.180	0.129	0.282	0.177	0.129	0.282	0.177
41	0.259	0.468	0.333	0.247	0.553	0.341	0.494	0.263	0.343	0.494	0.280	0.357
42	0.200	0.500	0.286	0.188	0.615	0.288	0.247	0.404	0.307	0.247	0.404	0.307
43	0.176	0.833	0.291	0.176	1.000	0.300	0.176	0.789	0.288	0.176	0.789	0.288
44	0.224	0.352	0.273	0.259	0.415	0.319	0.271	0.338	0.301	0.271	0.338	0.301
45	0.141	0.387	0.207	0.165	0.538	0.252	0.153	0.394	0.220	0.153	0.394	0.220
46	0.247	0.253	0.250	0.282	0.490	0.358	0.294	0.203	0.240	0.294	0.543	0.382
47	0.224	0.253	0.238	0.271	0.434	0.333	0.294	0.221	0.253	0.294	0.735	0.420
48	0.376	0.889	0.529	0.435	0.822	0.569	0.812	0.511	0.627	0.329	0.875	0.479
49	0.294	0.333	0.313	0.294	0.463	0.360	0.306	0.271	0.287	0.306	0.271	0.287
50	0.200	0.333	0.250	0.294	0.410	0.342	0.318	0.386	0.348	0.318	0.342	0.329

ตารางที่ ค.8 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.275	0.190	0.224	0.100	0.121	0.110	0.575	0.204	0.301	0.575	0.261	0.359
2	0.025	0.036	0.029	0.025	0.033	0.029	0.050	0.056	0.053	0.050	0.056	0.053
3	0.325	0.114	0.169	0.325	0.130	0.186	0.250	0.072	0.112	0.425	0.108	0.173
4	0.325	0.105	0.159	0.200	0.091	0.125	0.325	0.083	0.132	0.325	0.083	0.132
5	0.250	0.081	0.123	0.200	0.093	0.127	0.225	0.095	0.133	0.075	0.429	0.128
6	0.125	0.071	0.091	0.125	0.075	0.093	0.275	0.090	0.136	0.275	0.090	0.136
7	0.475	0.221	0.302	0.400	0.348	0.372	0.525	0.266	0.353	0.625	0.194	0.296
8	0.650	0.213	0.321	0.500	0.278	0.357	0.725	0.165	0.269	0.850	0.156	0.264
9	0.425	0.227	0.296	0.400	0.205	0.271	0.550	0.272	0.364	0.625	0.229	0.336
10	0.400	0.208	0.274	0.250	0.270	0.260	0.000	0.000	0.000	0.450	0.173	0.250
11	0.400	0.186	0.254	0.425	0.405	0.415	0.500	0.444	0.471	0.650	0.215	0.323
12	0.475	0.186	0.268	0.375	0.231	0.286	0.375	0.139	0.203	0.625	0.140	0.228
13	0.400	0.250	0.308	0.300	0.279	0.289	0.625	0.216	0.321	0.725	0.207	0.322
14	0.450	0.247	0.319	0.600	0.348	0.440	0.125	0.132	0.128	0.325	0.141	0.197
15	0.450	0.198	0.275	0.325	0.283	0.302	0.375	0.139	0.203	0.625	0.140	0.228
16	0.050	0.667	0.093	0.100	0.118	0.108	0.125	1.000	0.222	0.425	0.333	0.374
17	0.100	0.058	0.073	0.100	0.121	0.110	0.525	0.158	0.243	0.525	0.146	0.228
18	0.075	0.034	0.047	0.100	0.087	0.093	0.175	0.047	0.074	0.200	0.053	0.084
19	0.150	0.061	0.086	0.100	0.087	0.093	0.200	0.053	0.083	0.175	0.051	0.079
20	0.100	0.068	0.081	0.100	0.103	0.101	0.125	0.068	0.089	0.175	0.081	0.111
21	0.450	0.200	0.277	0.325	0.241	0.277	0.250	0.167	0.200	0.475	0.157	0.236
22	0.025	0.143	0.043	0.025	0.083	0.038	0.025	0.143	0.043	0.025	0.143	0.043
23	0.500	0.153	0.234	0.450	0.300	0.360	0.600	0.161	0.254	0.675	0.147	0.241
24	0.075	0.054	0.063	0.125	0.098	0.110	0.150	0.098	0.119	0.150	0.098	0.119
25	0.350	0.264	0.301	0.250	0.435	0.317	0.000	0.000	0.000	0.000	0.000	0.000
26	0.225	0.188	0.205	0.250	0.345	0.290	0.000	0.000	0.000	0.000	0.000	0.000
27	0.400	0.205	0.271	0.450	0.409	0.429	0.150	0.079	0.103	0.225	0.111	0.149
28	0.225	0.164	0.189	0.300	0.343	0.320	0.150	0.082	0.106	0.425	0.150	0.222

ตารางที่ ค.8 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 3 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8 (ต่อ)

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
29	0.050	0.032	0.039	0.050	0.039	0.044	0.275	0.110	0.157	0.400	0.088	0.144
30	0.150	0.066	0.092	0.225	0.100	0.138	0.250	0.109	0.152	0.400	0.087	0.143
31	0.075	0.125	0.094	0.125	0.250	0.167	0.425	0.362	0.391	0.425	0.362	0.391
32	0.375	0.163	0.227	0.275	0.256	0.265	0.625	0.152	0.244	0.625	0.152	0.244
33	0.025	0.500	0.048	0.125	0.556	0.204	0.400	0.889	0.552	0.400	0.889	0.552
34	0.200	0.133	0.160	0.100	0.133	0.114	0.525	0.216	0.307	0.625	0.217	0.323
35	0.175	0.219	0.194	0.075	0.273	0.118	0.275	0.134	0.180	0.275	0.134	0.180
36	0.400	0.186	0.254	0.450	0.321	0.375	0.500	0.444	0.471	0.650	0.215	0.323
37	0.050	0.333	0.087	0.050	0.333	0.087	0.500	0.290	0.367	0.500	0.290	0.367
38	0.050	0.071	0.059	0.375	0.160	0.224	0.425	0.104	0.167	0.575	0.113	0.189
39	0.525	0.143	0.225	0.400	0.176	0.244	0.250	0.086	0.128	0.625	0.117	0.198
40	0.050	0.083	0.063	0.000	0.000	0.000	0.250	0.073	0.113	0.175	0.070	0.100
41	0.500	0.215	0.301	0.475	0.204	0.286	0.700	0.204	0.316	0.700	0.175	0.280
42	0.150	0.066	0.092	0.150	0.087	0.110	0.275	0.096	0.142	0.275	0.096	0.142
43	0.150	0.150	0.150	0.100	0.121	0.110	0.250	0.114	0.156	0.250	0.114	0.156
44	0.300	0.706	0.421	0.250	0.667	0.364	0.450	0.391	0.419	0.475	0.244	0.322
45	0.550	0.148	0.233	0.475	0.151	0.229	0.175	0.292	0.219	0.775	0.183	0.297
46	0.275	0.262	0.268	0.125	0.185	0.149	0.575	0.240	0.338	0.575	0.240	0.338
47	0.475	0.150	0.228	0.375	0.214	0.273	0.625	0.172	0.270	0.625	0.225	0.331
48	0.150	0.059	0.085	0.200	0.099	0.132	0.550	0.110	0.183	0.425	0.183	0.256
49	0.425	0.170	0.243	0.550	0.314	0.400	0.100	0.571	0.170	0.250	0.227	0.238
50	0.675	0.142	0.235	0.650	0.268	0.380	0.775	0.137	0.233	0.850	0.115	0.203

ตารางที่ ค.9 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.328	0.218	0.262	0.379	0.361	0.370	0.500	0.215	0.301	0.500	0.215	0.301
2	0.362	0.263	0.304	0.397	0.377	0.387	0.431	0.229	0.299	0.190	0.131	0.155
3	0.259	0.417	0.319	0.259	0.938	0.405	0.431	0.248	0.314	0.379	0.917	0.537
4	0.362	0.375	0.368	0.379	0.512	0.436	0.483	0.228	0.309	0.483	0.311	0.378
5	0.259	0.250	0.254	0.293	0.354	0.321	0.431	0.216	0.287	0.379	0.846	0.524
6	0.086	0.081	0.083	0.069	0.108	0.084	0.190	0.129	0.154	0.000	0.000	0.000
7	0.259	0.652	0.370	0.310	0.667	0.424	0.379	0.710	0.494	0.379	0.710	0.494
8	0.000	0.000	0.000	0.017	0.091	0.029	0.000	0.000	0.000	0.000	0.000	0.000
9	0.000	0.000	0.000	0.017	0.091	0.029	0.000	0.000	0.000	0.000	0.000	0.000
10	0.483	0.933	0.636	0.500	0.906	0.644	0.552	0.914	0.688	0.552	0.914	0.688
11	0.362	1.000	0.532	0.328	1.000	0.494	0.379	0.957	0.543	0.379	0.957	0.543
12	0.362	0.600	0.452	0.328	0.576	0.418	0.379	0.579	0.458	0.379	0.579	0.458
13	0.362	0.600	0.452	0.310	0.563	0.400	0.379	0.579	0.458	0.379	0.579	0.458
14	0.362	0.955	0.525	0.328	0.950	0.487	0.379	0.957	0.543	0.379	0.957	0.543
15	0.069	0.235	0.107	0.103	0.286	0.152	0.103	0.286	0.152	0.103	0.286	0.152
16	0.362	0.600	0.452	0.345	0.588	0.435	0.379	0.579	0.458	0.379	0.579	0.458
17	0.621	0.818	0.706	0.621	0.857	0.720	0.655	0.826	0.731	0.655	0.826	0.731
18	0.034	0.143	0.056	0.034	0.118	0.053	0.034	0.083	0.049	0.034	0.083	0.049
19	0.603	0.897	0.722	0.621	0.923	0.742	0.638	0.881	0.740	0.638	0.881	0.740
20	0.517	0.526	0.522	0.552	0.582	0.566	0.603	0.407	0.486	0.603	0.427	0.500
21	0.362	0.955	0.525	0.310	0.947	0.468	0.379	0.917	0.537	0.379	0.917	0.537
22	0.379	0.957	0.543	0.328	0.864	0.475	0.397	0.920	0.554	0.397	0.920	0.554
23	0.034	0.125	0.054	0.034	0.154	0.056	0.034	0.125	0.054	0.034	0.125	0.054
24	0.362	1.000	0.532	0.293	0.944	0.447	0.379	0.957	0.543	0.379	0.957	0.543
25	0.362	0.955	0.525	0.310	0.947	0.468	0.379	0.957	0.543	0.379	0.957	0.543
26	0.431	0.641	0.515	0.310	0.692	0.429	0.121	0.700	0.206	0.379	0.846	0.524
27	0.431	0.463	0.446	0.328	0.487	0.392	0.138	0.471	0.213	0.362	0.553	0.438
28	0.362	0.296	0.326	0.241	0.333	0.280	0.483	0.308	0.376	0.483	0.368	0.418

ตารางที่ ค.9 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 4 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8 (ต่อ)

วิธีค้นคืนที่	5			6			7			8		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.379	0.880	0.530	0.293	0.810	0.430	0.379	0.846	0.524	0.379	0.846	0.524
30	0.224	0.650	0.333	0.155	0.563	0.243	0.259	0.682	0.375	0.259	0.682	0.375
31	0.259	0.306	0.280	0.155	1.000	0.269	0.431	0.248	0.314	0.379	0.957	0.543
32	0.259	0.441	0.326	0.103	0.750	0.182	0.397	0.426	0.411	0.397	0.426	0.411
33	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.032	0.022	0.017	0.032	0.022
34	0.034	0.118	0.053	0.034	0.133	0.055	0.034	0.100	0.051	0.034	0.100	0.051
35	0.259	0.536	0.349	0.293	0.895	0.442	0.414	0.522	0.462	0.414	0.522	0.462
36	0.362	1.000	0.532	0.310	1.000	0.474	0.379	0.957	0.543	0.379	0.957	0.543
37	0.362	0.600	0.452	0.293	0.531	0.378	0.379	0.579	0.458	0.379	0.579	0.458
38	0.121	0.175	0.143	0.103	0.250	0.146	0.121	0.108	0.114	0.414	0.240	0.304
39	0.310	0.295	0.303	0.155	0.180	0.167	0.379	0.449	0.411	0.379	0.278	0.321
40	0.328	0.500	0.396	0.310	0.500	0.383	0.379	0.524	0.440	0.379	0.524	0.440
41	0.362	0.477	0.412	0.293	0.459	0.358	0.379	0.440	0.407	0.379	0.440	0.407
42	0.362	0.404	0.382	0.310	0.474	0.375	0.431	0.284	0.342	0.414	0.414	0.414
43	0.069	0.056	0.062	0.086	0.069	0.077	0.138	0.067	0.090	0.121	0.065	0.084
44	0.034	0.049	0.040	0.034	0.051	0.041	0.052	0.061	0.056	0.052	0.061	0.056
45	0.086	0.093	0.089	0.034	0.050	0.041	0.138	0.078	0.099	0.138	0.078	0.099
46	0.431	0.329	0.373	0.397	0.397	0.397	0.552	0.294	0.383	0.552	0.294	0.383
47	0.414	0.338	0.372	0.483	0.452	0.467	0.621	0.364	0.459	0.621	0.364	0.459
48	0.379	0.319	0.346	0.345	0.339	0.342	0.431	0.284	0.342	0.431	0.284	0.342
49	0.534	0.217	0.308	0.534	0.272	0.360	0.707	0.193	0.304	0.707	0.196	0.307
50	0.155	0.105	0.125	0.172	0.167	0.169	0.310	0.118	0.171	0.310	0.118	0.171

ตารางที่ ค.10 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8

วิธีค้นคืนที่	5			6			7			8		
ข้อคำถาม	R	P	F	R	P	F	R	P	F	R	P	F
1	0.591	0.252	0.354	0.591	0.289	0.388	0.455	0.161	0.238	0.750	0.306	0.434
2	0.591	0.765	0.667	0.636	0.757	0.691	0.364	0.533	0.432	0.750	0.702	0.725
3	0.614	0.300	0.403	0.659	0.296	0.408	0.000	0.000	0.000	0.045	0.100	0.063
4	0.523	0.383	0.442	0.545	0.600	0.571	0.477	0.194	0.276	0.341	0.259	0.294
5	0.659	1.000	0.795	0.523	1.000	0.687	0.136	1.000	0.240	0.727	1.000	0.842
6	0.182	0.157	0.168	0.114	0.217	0.149	0.182	0.116	0.142	0.182	0.116	0.142
7	0.545	0.632	0.585	0.614	0.675	0.643	0.091	0.222	0.129	0.091	0.222	0.129
8	0.636	0.262	0.371	0.659	0.266	0.379	0.182	0.267	0.216	0.841	0.276	0.416
9	0.591	0.377	0.460	0.636	0.509	0.566	0.409	0.202	0.271	0.773	0.324	0.456
10	0.568	0.333	0.420	0.591	0.400	0.477	0.636	0.259	0.368	0.841	0.346	0.490
11	0.659	0.906	0.763	0.614	0.964	0.750	0.068	0.500	0.120	0.727	0.914	0.810
12	0.614	0.900	0.730	0.591	0.929	0.722	0.023	0.250	0.042	0.023	0.250	0.042
13	0.136	0.857	0.235	0.545	0.750	0.632	0.045	0.118	0.066	0.727	0.681	0.703
14	0.705	0.912	0.795	0.705	0.912	0.795	0.023	0.250	0.042	0.023	0.250	0.042
15	0.614	0.900	0.730	0.705	0.912	0.795	0.068	0.500	0.120	0.773	0.872	0.819
16	0.614	0.900	0.730	0.568	0.893	0.694	0.068	0.500	0.120	0.727	0.914	0.810
17	0.636	0.683	0.659	0.545	0.649	0.593	0.045	0.133	0.068	0.705	0.705	0.705
18	0.682	0.857	0.759	0.545	0.923	0.686	0.182	0.320	0.232	0.705	0.646	0.674
19	0.636	0.966	0.767	0.386	0.630	0.479	0.023	0.500	0.043	0.705	0.969	0.816
20	0.636	0.903	0.747	0.545	0.889	0.676	0.045	0.400	0.082	0.705	0.912	0.795
21	0.659	0.806	0.725	0.568	0.781	0.658	0.273	0.600	0.375	0.705	0.795	0.747
22	0.682	0.395	0.500	0.568	0.431	0.490	0.295	0.210	0.245	0.750	0.402	0.524
23	0.682	0.909	0.779	0.455	0.800	0.580	0.136	0.667	0.226	0.750	0.917	0.825
24	0.636	0.800	0.709	0.545	0.727	0.623	0.000	0.000	0.000	0.000	0.000	0.000
25	0.682	1.000	0.811	0.545	1.000	0.706	0.000	0.000	0.000	0.000	0.000	0.000
26	0.591	0.271	0.371	0.614	0.391	0.478	0.000	0.000	0.000	0.000	0.000	0.000
27	0.591	0.897	0.712	0.568	0.862	0.685	0.000	0.000	0.000	0.000	0.000	0.000
28	0.636	0.667	0.651	0.591	0.650	0.619	0.159	0.333	0.215	0.705	0.689	0.697

ตารางที่ ค.10 ค่าเรียกคืน ค่าความแม่นยำ และค่าเฉลี่ยฮาร์โมนิค จากโจทย์ปัญหาข้อที่ 5 ด้วยวิธีการค้นคืนที่ 5 6 7 และ 8 (ต่อ)

วิธีค้นคืนที่	5			6			7			8		
	R	P	F	R	P	F	R	P	F	R	P	F
29	0.636	0.875	0.737	0.568	0.862	0.685	0.091	0.500	0.154	0.727	0.889	0.800
30	0.614	0.628	0.621	0.477	0.677	0.560	0.068	0.158	0.095	0.727	0.667	0.696
31	0.205	0.071	0.105	0.182	0.133	0.154	0.182	0.063	0.093	0.227	0.067	0.104
32	0.159	0.175	0.167	0.159	0.219	0.184	0.250	0.128	0.169	0.114	0.625	0.192
33	0.136	0.061	0.084	0.114	0.082	0.095	0.091	0.067	0.077	0.136	0.050	0.073
34	0.636	0.292	0.400	0.545	0.286	0.375	0.227	0.256	0.241	0.841	0.274	0.413
35	0.591	0.448	0.510	0.341	0.750	0.469	0.341	0.172	0.229	0.795	0.327	0.464
36	0.568	0.658	0.610	0.636	0.667	0.651	0.091	0.222	0.129	0.773	0.343	0.476
37	0.591	0.306	0.403	0.568	0.658	0.610	0.477	0.189	0.271	0.864	0.211	0.339
38	0.500	0.379	0.431	0.523	0.479	0.500	0.409	0.220	0.286	0.727	0.333	0.457
39	0.636	0.683	0.659	0.477	0.656	0.553	0.045	0.133	0.068	0.705	0.705	0.705
40	0.682	0.441	0.536	0.750	0.413	0.532	0.250	0.423	0.314	0.795	0.276	0.409
41	0.659	0.906	0.763	0.659	0.906	0.763	0.068	0.500	0.120	0.773	0.872	0.819
42	0.682	0.395	0.500	0.545	0.453	0.495	0.295	0.210	0.245	0.750	0.402	0.524
43	0.636	0.875	0.737	0.568	0.862	0.685	0.136	0.600	0.222	0.750	0.892	0.815
44	0.727	0.604	0.660	0.727	0.604	0.660	0.432	0.134	0.204	0.818	0.226	0.355
45	0.682	0.361	0.472	0.659	0.315	0.426	0.000	0.000	0.000	0.000	0.000	0.000
46	0.659	0.271	0.384	0.727	0.427	0.538	0.705	0.173	0.278	0.818	0.293	0.431
47	0.636	0.315	0.421	0.591	0.338	0.430	0.636	0.172	0.271	0.795	0.357	0.493
48	0.614	0.276	0.380	0.659	0.387	0.487	0.477	0.171	0.252	0.841	0.264	0.402
49	0.591	0.208	0.308	0.500	0.306	0.379	0.545	0.154	0.240	0.773	0.264	0.393
50	0.500	0.355	0.415	0.250	0.344	0.289	0.295	0.197	0.236	0.750	0.384	0.508

ภาคผนวก ง
สรุปสูตรที่ใช้ในงานวิจัย

ในงานวิจัยนี้มีสูตรที่ใช้ทั้งหมด 16 สูตร ด้วยกัน แสดงไว้ดังตารางที่ ง.1

ตารางที่ ง.1 สรุปสูตรและคำอธิบายสูตรทั้งหมดที่ใช้ในงานวิจัย

ที่	สูตร
1	$Weight_{ik} = \frac{Freq_{ik}}{TotFreq_k}$
	เป็นสูตรสำหรับหาค่าน้ำหนักของคำแต่ละคำในแต่ละเอกสาร โดยใช้ค่าน้ำหนักของเอกสารแบบผกผัน (IDF)
2	$Recall (R) = \frac{ Ra }{R}$
	สำหรับคำนวณหาค่าเรียกคืน เพื่อประเมินประสิทธิผลของระบบ
3	$Precision (P) = \frac{ Ra }{A}$
	สำหรับคำนวณหาค่าความแม่นยำ เพื่อประเมินประสิทธิผลของระบบ
4	$Harmonic\ Mean\ (F) = \frac{2}{\frac{1}{(R)} + \frac{1}{(P)}}$
	เป็นสูตรคำนวณค่าเฉลี่ยฮาร์โมนิก ซึ่งเป็นค่าเฉลี่ยระหว่างค่าเรียกคืนและค่าความแม่นยำ
5	$CTerm_{pk} = \frac{1}{m} \sum_{i=1}^m Term_{ik}$
	สำหรับคำนวณน้ำหนักโดยเฉลี่ยของคำ ในเอกสารกลุ่มเดียวกันเพื่อหาตัวแทนของเอกสารกลุ่มนั้นๆ
6	$Q' = \alpha Q + \beta \frac{1}{R'} \sum_{i \in D_{R'}} DOC_i - \gamma \frac{1}{N'} \sum_{i \in D_{N'}} DOC_i$
	สำหรับเปลี่ยนแปลงค่าน้ำหนักของคำจากข้อความเดิมเป็นข้อความใหม่ ในแบบจำลองปริภูมิเวกเตอร์

ตารางที่ ง.1 สรุปสูตรทั้งหมดที่ใช้ในงานวิทยานิพนธ์ (ต่อ)

ที่	สูตร
7	$sim = (d_j, q) = \sum_i^t w_{i,q} \times w_{i,j} \times \log \frac{N - n_i}{n_i}$
	สูตรหาค่าความคล้ายระหว่างเอกสารกับข้อความ เพื่อการค้นหาเอกสารในครั้งแรก ด้วยแบบจำลองความน่าจะเป็น
8	$sim(d_j, q) = \sum_{i=1}^t w_{i,q} \times w_{i,j} \times \log \frac{ D_{r,i} }{ D_r - D_{r,i} } \times \frac{N - D_r - (n_i - D_{r,i})}{n_i - D_{r,i} }$
	สูตรหาค่าความคล้ายระหว่างเอกสารกับข้อความ เพื่อการค้นหาเอกสารโดยใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใช้ ด้วยแบบจำลองความน่าจะเป็น
9	$sim(Centriod_j, Query_j) = \frac{\sum_m^e [sim(ECentroid_{mj}, EQuery_{mj})] \cdot WE_m}{TotWeight}$
	สูตรคำนวณหาค่าความคล้ายระหว่างตัวแทนเอกสารในกลุ่มกับข้อความ
10	$sim(UC_j, Query_j) = \frac{\sum [sim(EUC_{mj}, EQuery_{mj})] \cdot WE_m}{TotWeight}$
	สูตรคำนวณหาค่าความคล้ายระหว่างเอกสารยูสเคสกับข้อความ
11	$sim(ECentroid_{mj}, Equery_{mj}) = \frac{\sum [(CTerm_{mk} \cdot W_{ik}) \cdot QTerm_{mjk}]}{\sqrt{\sum_{k=1}^t (CTerm_{mik})^2 \cdot \sum_{k=1}^t (QTerm_{mjk})^2}}$
	สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารตัวแทนกลุ่มกับข้อความ สำหรับแบบจำลองปริภูมิเวกเตอร์
12	$sim(ECentroid_{mj}, Equery_{mj}) = \sum_{k=1}^t w_{k,mj} \times w_{k,mj} \times \log \frac{N - n_k}{n_k}$
	สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารตัวแทนกลุ่มกับข้อความ สำหรับแบบจำลองความน่าจะเป็น

ตารางที่ ง.1 สรุปสูตรทั้งหมดที่ใช้ในงานวิทยานิพนธ์ (ต่อ)

ที่	สูตร
13	$sim(EUC_{mi}, Equery_{mj}) = \frac{\sum (Term_{mk} \cdot w_{ik}) \cdot QTerm_{mjk}}{\sqrt{\sum_{k=1}^t (Term_{mik})^2 \cdot \sum_{k=1}^t (QTerm_{mjk})^2}}$
	<p>สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารยูสเคสกับข้อความ สำหรับแบบจำลองปริภูมิเวกเตอร์</p>
14	$sim(EUC_{mi}, Equery_{mj}) = \sum_{k=1}^t w_{k,mi} \times w_{k,mj} \times \log \frac{N - n_k}{n_k}$
	<p>สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารยูสเคสกับข้อความ สำหรับแบบจำลองความน่าจะเป็น สำหรับการค้นหาเอกสารในครั้งแรก</p>
15	$sim(ES_{mi}, Equery_{mj}) = \sum_{k=1}^t w_{k,mi} \times w_{k,mj} \times \log \frac{ D_{r,k} }{ D_r - D_{r,k} } \times \frac{N - D_r - (n_k - D_{r,k})}{n_k - D_{r,k} }$
	<p>สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารตัวแทนกลุ่มกับข้อความ สำหรับแบบจำลองความน่าจะเป็น สำหรับการค้นหาเอกสารในครั้งถัดไปซึ่งใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใ้</p>
16	$sim(EUC_{mi}, EQuery_{mj}) = \sum_{k=1}^t w_{k,mi} \times w_{k,mj} \times \log \frac{ D_{r,k} }{ D_r - D_{r,k} } \times \frac{N - D_r - (n_k - D_{r,k})}{n_k - D_{r,k} }$
	<p>สูตรคำนวณค่าความคล้ายในแต่ละส่วนระหว่างเอกสารยูสเคสกับข้อความ สำหรับแบบจำลองความน่าจะเป็น สำหรับการค้นหาเอกสารในครั้งถัดไปซึ่งใช้ผลป้อนกลับที่ตรงประเด็นจากผู้ใ้</p>

ภาคผนวก จ

บทความวิชาการที่ตีพิมพ์

ในการวิจัยนี้ ผู้วิจัยมีผลงานวิชาการร่วมกับคณะผู้วิจัย เป็นบทความวิชาการระดับชาติ รวมเป็น 2 บทความ ได้แก่

จ.1 บทความวิชาการเรื่อง “การค้นคืนยูสเคสด้วยการจำแนกประเภทยูสเคสและการค้นคืนย้อนกลับจากผู้ใช้ (Use Case Retrieval using Use Case Clustering and User Relevance Feedback)” ซึ่งได้รับการคัดเลือกเพื่อนำเสนอและตีพิมพ์ในงาน “การประชุมวิชาการร่วมสาขาวิทยาการคอมพิวเตอร์และวิศวกรรมซอฟต์แวร์ ครั้งที่ 3 (The 3rd Joint Conference on Computer Science and Software Engineering: JCSSE 2006)” ระหว่างวันที่ 29 – 30 มิถุนายน 2549 ณ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า วิทยาเขตพระนครเหนือ กรุงเทพฯ

จ.2 บทความวิชาการเรื่อง “การจัดเก็บและค้นคืนยูสเคสที่มีการจัดกลุ่มยูสเคสร่วมกับการค้นคืนย้อนกลับจากผู้ใช้ (Use Case Clustered Storage and Retrieval with User Relevance Feedback)” ซึ่งได้รับการคัดเลือกเพื่อนำเสนอและตีพิมพ์ในงาน “การประชุมวิชาการทางวิทยาการคอมพิวเตอร์และวิศวกรรมคอมพิวเตอร์ในระดับชาติ ครั้งที่ 11 (The 11th National Computer Science and Engineering Conference: NCSEC 2007)” ระหว่างวันที่ 19 – 21 พฤศจิกายน 2550 ณ โรงแรมมิราเคิลแกรนด์คอนเวนชั่น กรุงเทพฯ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การค้นคืนยูสเคสด้วยการจำแนกประเภทยูสเคสและการค้นคืนย้อนกลับจากผู้ใช้ Use Case Retrieval Using Use Case Clustering and User Relevance Feedback

สุดาทิพย์ สุขสอาด และ นครทิพย์ พร้อมพูล

ห้องปฏิบัติการวิศวกรรมซอฟต์แวร์ ศูนย์เชี่ยวชาญเฉพาะทางด้านวิศวกรรมซอฟต์แวร์

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล: kartoon_s@hotmail.com และ nakornthip.s@chula.ac.th

บทคัดย่อ

ยูสเคสเป็นเครื่องมือที่สำคัญในการบันทึกข้อกำหนดความต้องการซอฟต์แวร์ของระบบซึ่งเป็นขั้นตอนที่สำคัญมากของการพัฒนาซอฟต์แวร์ เนื่องจากเป็นขั้นตอนที่ต้องรวบรวมความต้องการของผู้ใช้ระบบ ดังนั้นหากผู้พัฒนาระบบสามารถนำยูสเคสที่ได้มีการบันทึกจากระบบเดิมที่มีความคล้ายกันกับยูสเคสที่ต้องการสร้างนั้นกลับมาใช้ใหม่ได้จะช่วยลดเวลาและค่าใช้จ่ายในการพัฒนาซอฟต์แวร์ งานวิจัยนี้เสนอการค้นคืนยูสเคสเพื่อการนำกลับมาใช้ใหม่ โดยมีการจัดเก็บยูสเคสด้วยการจำแนกประเภทของยูสเคส และให้มีการคิวรีโดยมีการค้นคืนย้อนกลับจากผู้ใช้ การประเมินประสิทธิภาพของผลลัพธ์จากวิธีการที่นำเสนอใช้มาตรวัดคือค่าความแม่นยำค่าเรียกคืน และเฉลี่ยฮาร์โมนิก

คำสำคัญ การค้นคืนยูสเคส ยูสเคส เอกสารการบรรยาย ยูสเคส การจำแนกประเภทยูสเคส การค้นคืนย้อนกลับจากผู้ใช้

Abstract

Use case is an important tool for keeping the software requirements specification which is a significant phase of software development because the user requirements are collected in this phase. Therefore, if the system developers can reuse some use cases which are recorded and similar to some relevant use cases from the previous system, they will

reduce some times and costs in software development. This research proposes the use case retrieval method for reusing which is applied the use case clustering for use case storage and user relevance feedback in querying process. The performance evaluations of the results from this proposed method are precision, recall and Harmonic mean.

Keywords: Use Case Retrieval, Use Case, Use Case Description, Use Case Clustering, User Relevance Feedback

1. บทนำ

ในกระบวนการพัฒนาซอฟต์แวร์หนึ่งๆ ประกอบด้วยหลายขั้นตอน โดยแต่ละขั้นตอนใช้เวลาและค่าใช้จ่ายต่างกัน การลดเวลาและค่าใช้จ่ายเท่าที่จะเป็นไปได้คือเป้าหมายที่สำคัญของทุกองค์กร และถือว่าเป็นความท้าทายในกระบวนการดังกล่าวอีกด้วย วิธีการที่นิยมใช้ในการลดเวลาและค่าใช้จ่ายได้อย่างมีประสิทธิภาพ คือการนำซอฟต์แวร์กลับมาใช้ใหม่ [2] โดยขั้นตอนการวิเคราะห์ความต้องการเป็นขั้นตอนหนึ่งที่สำคัญมากเนื่องจากใช้เวลาและค่าใช้จ่ายสูง เพื่อให้ได้ข้อกำหนดความต้องการซอฟต์แวร์ (Software Requirements Specification) มาจากผู้ใช้ระบบ มีแผนภาพยูสเคส (Use Case Diagram) และเอกสารการบรรยายยูสเคส (Use Case Description) เป็น

เครื่องมือที่นิยมใช้เพื่อบันทึกข้อกำหนดความต้องการต่างๆ ของระบบ ดังนั้นจะเป็นประโยชน์อย่างยิ่งหากผู้บันทึกข้อกำหนดความต้องการหรือผู้พัฒนาระบบ ใช้ข้อมูลจากยูสเคสของระบบเดิมที่มีความคล้ายกันและได้ระบุถึงฟังก์ชันต่างๆอย่างครบถ้วนกลับมาใช้ แล้วปรับแต่งเพิ่มเติมให้เหมาะสมกับความต้องการ ไม่ต้องเริ่มต้นใหม่โดยปราศจากข้อมูล ซึ่งจะช่วยลดเวลาและค่าใช้จ่ายได้เป็นอย่างดี

ระบบจัดเก็บและค้นคืนสารสนเทศ เป็นระบบที่สนับสนุนการนำยูสเคสในไลบรารีกลับมาใช้ใหม่ โดยจำเป็นต้องมีองค์ประกอบที่สำคัญ 4 ส่วน ได้แก่ 1) ตัวแทนของยูสเคส (Use Case Representation) 2) ภาษาคิวรี(Query Language) เพื่อเป็นตัวแทนความต้องการจากผู้ใช้ 3) ฟังก์ชันการเปรียบเทียบความคล้าย (Similarity Function) ระหว่างคิวรีกับยูสเคส และ 4) ฟังก์ชันการจัดอันดับผล (Ranking Function) ของยูสเคสที่ค้นคืนได้ ที่เหมาะสมด้วยเพื่อให้ได้ยูสเคสที่ผู้พัฒนาระบบต้องการมากที่สุด แต่เนื่องจากยูสเคสมีจำนวนมากและอาจมีมากขึ้นอีก จึงควรเพิ่มเติมองค์ประกอบในการจัดกลุ่มให้กับยูสเคสเพื่อให้ง่ายต่อการเข้าถึงและประหยัดเวลาในการค้นคืน นอกจากนี้ความต้องการยูสเคสที่ต่างกันของผู้พัฒนาระบบแต่ละคน เป็นปัญหาหนึ่งที่ทำให้ระบบค้นคืนให้ผลลัพธ์ที่ไม่ตรงกับความต้องการ แนวทางแก้ไขที่เหมาะสมคือการใช้การค้นคืนย้อนกลับจากผู้ใช้ (User Relevance Feedback) [4] เพราะเป็นวิธีนำผลป้อนกลับจากการประเมินผลลัพธ์ที่ได้จากการค้นคืนในครั้งแรกจากผู้ใช้มาช่วยในการปรับเปลี่ยนคิวรีใหม่ที่เหมาะสมกว่าคิวรีเดิม

ในบทความวิชาการนี้ประกอบด้วย งานวิจัยที่เกี่ยวข้องในส่วนที่ 2 ระบบการจัดเก็บและค้นคืนสารสนเทศในส่วนที่ 3 แผนภาพยูสเคสและเอกสารการบรรยายยูสเคสในส่วนที่ 4 วิธีการค้นคืนยูสเคสกลับมาใช้ใหม่ในส่วนที่ 5 โดยส่วนที่ 6 และ 7 อธิบายเกี่ยวกับการประเมินผลประสิทธิภาพของระบบและกรณีตัวอย่าง ตามลำดับ และ

ในส่วนสุดท้าย คือ สรุปผลการวิจัยและงานที่จะดำเนินการในอนาคต

2. งานวิจัยที่เกี่ยวข้อง

Saeki [8] ได้เสนอรูปแบบ (Pattern) ต่างๆที่สามารถนำกลับมาใช้ใหม่ เพื่อให้ได้มาซึ่งข้อกำหนดความต้องการ โดยที่ข้อกำหนดความต้องการนั้น สามารถนำมาแปลงเป็นเอกสารการบรรยายยูสเคส (Use Case Description) ได้แต่ข้อกำหนดของงานวิจัยนี้คือจะไม่สามารถนำข้อมูลภายในคำอธิบายยูสเคสที่มีอยู่เดิมกลับมาใช้ใหม่ได้ หรือในงานวิจัยของWoo และ Robinson [6] ได้เสนอวิธีการนำแผนภาพยูสเคสกลับมาใช้ใหม่ ด้วยการพัฒนาเครื่องมือที่ชื่อว่า "ScenAsst" เพื่อแปลงยูสเคสให้อยู่ในรูปแบบของกราฟ แล้วจัดกลุ่มเก็บไว้ในไลบรารี หลังจากนั้นทำการค้นคืนด้วยคิวรีในรูปแบบของกราฟเช่นกัน มาเปรียบเทียบค่าความคล้ายกันกับยูสเคสที่เก็บไว้ซึ่งเทคนิคในงานวิจัยนี้ต้องใช้เวลาในการประมวลผลค่อนข้างมากเนื่องจากมีความซับซ้อนของโครงสร้างสูง แต่โดยภาพรวมจากผลงานวิจัยทั้งสองข้างต้นถือว่าให้ผลลัพธ์ที่ดีในการนำยูสเคสกลับมาใช้ใหม่

Salton [5] ได้ออกแบบระบบและพัฒนาระบบสมาร์ต (The SMART System) ระบบซึ่งเป็นที่รู้จักกันดีในการศึกษาและทดลองเกี่ยวกับการจัดเก็บและค้นคืนสารสนเทศ ซึ่งมีวิธีการตั้งแต่การทำดัชนีอัตโนมัติ การจัดกลุ่มเอกสารอัตโนมัติ รวมถึงการเปลี่ยนแปลงคิวรีด้วยข้อมูลจากผู้ใช้ เป็นต้น โดยที่เอกสารในระบบมีรูปแบบเวกเตอร์ ผลการทดลองเกี่ยวกับการจัดกลุ่มเอกสารนั้นพบว่าช่วยลดเวลาในการเข้าถึงเอกสารได้ ส่วนการเปลี่ยนแปลงคิวรีการค้นคืนย้อนกลับจากผู้ใช้ พบว่าทำให้ผลการค้นคืนเอกสารมีความถูกต้องมากขึ้นด้วย Lundquist และคณะ [3] เห็นความสำคัญของการค้นคืนย้อนกลับ ได้เสนอการปรับปรุงประสิทธิภาพการค้นคืน

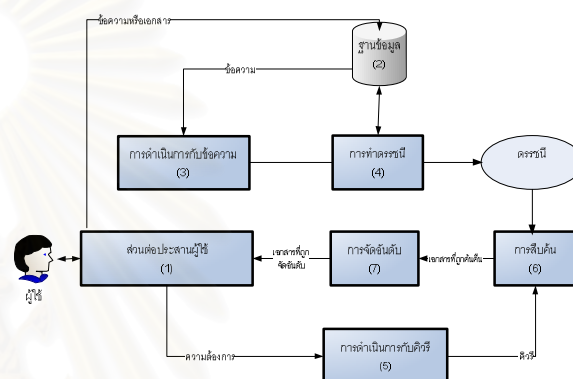
ย้อนกลับแบบเทียม (Pseudo Relevance Feedback) คือไม่มีผู้ชมาระบุความเกี่ยวข้องของเอกสาร โดยงานวิจัยนี้ศึกษาและทดลองด้วยเทคนิคต่างๆ และใช้แบบจำลองเวกเตอร์เป็นตัวแทนเอกสารซึ่งพบว่าให้ผลการค้นคืนที่ดีขึ้น

Akadej [1] เสนอวิธีนำยูสเคสกลับมาใช้ใหม่ ด้วยการคำนวณค่าความคล้ายกันของพจน์และโครงสร้างระหว่างเอกสารการบรรยายยูสเคส (Use Case Description) ที่เก็บไว้ในไลบรารีกับคิวรีที่มีลักษณะเป็นเอกสารการบรรยายยูสเคสอย่างง่าย และสร้างเครื่องมือทดลองวิธีที่นำเสนอโดยมีการดำเนินงาน 2 ส่วนคือ การจัดเก็บและค้นคืนยูสเคส ในส่วนแรกทำการจัดเก็บยูสเคสที่เป็นเอกสารการบรรยายยูสเคสลงฐานข้อมูล และทำการแปลงเอกสารดังกล่าวด้วยกรรมวิธีแบบอัตโนมัติ (Automatic Indexing) จากนั้นกำหนดค่าน้ำหนักเทอมให้กับกรรมวิธีด้วยความถี่ของเอกสารแบบผกผัน (Inverse Document Frequency :IDF) ส่วนการค้นคืนนั้น คิวรีที่ผู้ใช้ใส่เข้ามาจะถูกแปลงเป็นกรรมวิธี จากนั้นระบบจะคำนวณค่าผลรวมความคล้ายกันระหว่างกรรมวิธีของคิวรีกับกรรมวิธีของเอกสารการบรรยายยูสเคสแต่ละชิ้นในไลบรารี ซึ่งได้จากการเปรียบเทียบความคล้ายกันระหว่างส่วนประกอบแต่ละส่วนของเอกสารการบรรยายยูสเคส โดยส่วนประกอบแต่ละส่วนจะถูกกำหนดค่าน้ำหนักจากผู้ใช้ก่อนเริ่มการสืบค้น กล่าวได้ว่าวิธีการค้นคืนยูสเคสของงานวิจัยนี้มีการค้นคืนโดยอาศัยลักษณะที่เป็นโครงสร้างของทั้งคิวรีและเอกสาร นอกจากนี้ยังมีส่วนที่ให้ผู้ชมาระบุค่าน้ำหนักเพื่อให้ความสำคัญกับแต่ละส่วนประกอบของเอกสารการบรรยายยูสเคสที่แตกต่างกันด้วย จึงมีข้อดีในเรื่องความแม่นยำในการค้นคืนมากขึ้น แต่งานวิจัยนี้ไม่ได้พิจารณาการจำแนกประเภทยูสเคส และการค้นคืนย้อนกลับจากผู้ใช้

ดังนั้นงานวิจัยนี้ได้นำหลักการในการจำแนกประเภทและการค้นคืนย้อนกลับจากผู้ใช้ทั้งจากแบบจำลองเวกเตอร์และแบบจำลองความน่าจะเป็นมาใช้ในการพัฒนาระบบการจัดเก็บและค้นคืนยูสเคสต่อไป

3. การจัดเก็บและการค้นคืนสารสนเทศ

ระบบการจัดเก็บและค้นคืนสารสนเทศ เป็นส่วนงานสำคัญที่ถูกนำมาใช้ในระบบสืบค้นสารสนเทศ โดยผู้ใช้จะค้นหาข้อมูลด้วยการป้อนคิวรีที่ประกอบด้วยคำสำคัญต่างๆ (Keywords) เข้าสู่ระบบ จากนั้นระบบจะค้นคืนสารสนเทศออกมาตามวิธีกระบวนการค้นคืน โดยโครงสร้างของระบบการจัดเก็บและค้นคืนสารสนเทศ แสดงดังรูปที่ 1 [10]



รูปที่ 1 แสดงกระบวนการจัดเก็บและค้นคืนสารสนเทศ

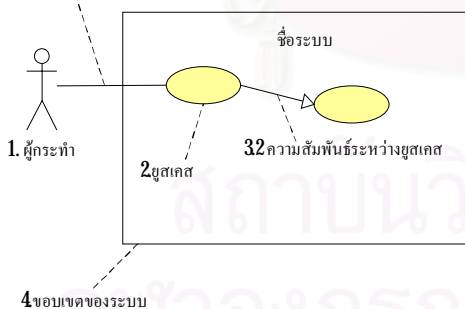
- 1) ส่วนต่อประสานผู้ใช้ (User Interface) คือส่วนที่เชื่อมต่อระหว่างผู้ใช้กับระบบ เพื่อให้ผู้ใช้ใส่คิวรีเข้าสู่ระบบและเป็นส่วนแสดงรายการเอกสารที่ค้นคืนได้จากระบบ นอกจากนี้เป็นส่วนให้เจ้าหน้าที่ควบคุมระบบใส่ข้อมูลต่างๆ ไปเก็บในฐานข้อมูลอีกด้วย
- 2) ฐานข้อมูล (Database) คือส่วนที่เก็บเอกสารทั้งหมดในระบบ
- 3) การดำเนินการกับข้อความ (Text Operation) คือขั้นตอนในการดำเนินการกับข้อความ (Text) ที่เก็บไว้ในฐานข้อมูล เป็นขั้นตอนในการเตรียมเอกสารก่อนนำไปทำกรรมวิธี เช่น การตัดคำ การกำจัดคำยกเว้น การลดรูปคำศัพท์ให้มีรากศัพท์เดียวกันและการสร้างพจนานุกรมคำศัพท์ที่มีความหมายคล้าย เป็นต้น

- 4) การทำดัชนี (Indexing) คือขั้นตอนที่เริ่มต้นจากการนำข้อความในเอกสารหนึ่งๆ ผ่านการดำเนินการกับข้อความ เพื่อให้ได้คำสำคัญในแต่ละเอกสารออกมาใช้เป็นดัชนีชี้โยงไปยังเอกสารนั้นๆ
- 5) การดำเนินการกับคิวรี (Query Operation) คือขั้นตอนในการแปลงความต้องการของผู้ใช้ให้เป็นคิวรีเพื่อป้อนเข้าสู่ระบบโดยจะผ่านขั้นตอนเดียวกันกับการดำเนินการกับข้อความเพื่อให้ได้คำสำคัญที่จะใช้การขั้นตอนการสืบค้นต่อไป
- 6) การสืบค้น (Searching) คือขั้นตอนการค้นหาเอกสารด้วยการใช้ดัชนีที่ได้ออกแบบไว้
- 7) การจัดอันดับ (Ranking) คือขั้นตอนในการจัดอันดับเอกสารที่ค้นคืน

4 แผนภาพยูสเคสและเอกสารการบรรยายยูสเคส

แผนภาพยูสเคส [7] (Use Case Diagram) คือเครื่องมือที่ใช้นำมาอธิบายปฏิสัมพันธ์ระหว่างระบบงานกับผู้ที่เกี่ยวข้องกับระบบงาน

31 ความสัมพันธ์ระหว่างผู้กระทำและยูสเคส

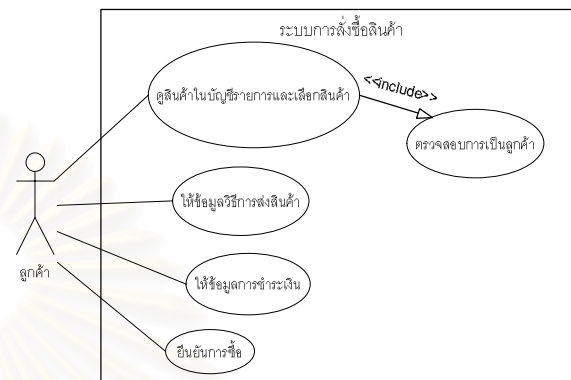


รูปที่ 2 แสดงองค์ประกอบของแผนภาพยูสเคส

โดยมีองค์ประกอบสำคัญ 4 ส่วนดังแสดงในรูปที่ 2 ได้แก่
 1) ชุดของผู้ที่เกี่ยวข้องกับระบบ 2) ชุดของยูสเคส 3) ความสัมพันธ์ต่างๆระหว่างชุดของยูสเคสกับชุดของผู้ที่เกี่ยวข้องกับระบบ 4) ขอบเขตของระบบ โดยชุดของยูส

เคสประกอบด้วยยูสเคสต่างๆที่อธิบายการทำงานของระบบในแต่ละหน้าที่หลัก

ตัวอย่างแผนภาพยูสเคสในระบบการสั่งซื้อสินค้า (Ordering System) ดังรูปที่ 3



รูปที่ 3 ระบบการสั่งซื้อสินค้า

ซึ่งยูสเคสแต่ละยูสเคสสามารถบันทึกรายละเอียดการทำงานด้วยเอกสารการบรรยายยูสเคส (Use Case Description) ดังรูปที่ 4

ระบบสั่งซื้อสินค้า		
ชื่อยูสเคส :	ดูสินค้าในบัญชีรายการและเลือกสินค้า	
วัตถุประสงค์ :	ยูสเคสนี้ถูกใช้เพื่อการดูสินค้าในบัญชีรายการสินค้าและทำการเลือกสินค้าที่ต้องการสั่งซื้อโดยลูกค้า	
ผู้กระทำ :	ลูกค้า	
ความสัมพันธ์ :	ความเกี่ยวข้อง :	-
	ความสัมพันธ์แบบรวม :	ตรวจสอบการเป็นลูกค้า
	ความสัมพันธ์แบบขยาย :	-
	เงื่อนไขลัดเล็งขั้น :	-
เงื่อนไขก่อนการทำงาน :	-	
เงื่อนไขหลังการทำงาน :	-	
ลำดับเหตุการณ์การทำงานปกติ	1. ลูกค้าล็อกอินเข้าสู่ระบบการสั่งซื้อสินค้า 2. ลูกค้าดูสินค้าที่มีในบัญชีรายการสินค้า 3. ลูกค้าเลือกสินค้าที่ต้องการสั่งซื้อ 4. บันทึกรายการสินค้าที่ลูกค้าต้องการสั่งซื้อ	
เหตุการณ์ย่อย	-	
ลำดับทางเลือกการทำงาน	-	

รูปที่ 4 รูปแบบของเอกสารการบรรยายยูสเคส

รายละเอียดแต่ละส่วนประกอบของเอกสารการบรรยายยูสเคส ประกอบไปด้วยโครงสร้างหลัก 9 ส่วน ได้แก่

- 1) ชื่อยูสเคส (Use Case Name) คือชื่อที่ใช้เรียกยูสเคส

- 2) วัตถุประสงค์ (Objective) คือวัตถุประสงค์หลักในการทำงานของยูสเคส
- 3) ผู้ใช้งานยูสเคส (Actor) คือผู้ที่เกี่ยวข้องในระบบซึ่งมีบทบาทกับการทำงานของยูสเคส
- 4) ความสัมพันธ์ (Relationship) คือคำอธิบายความสัมพันธ์ระหว่างแต่ละส่วนภายในแผนภาพ
- 5) เงื่อนไขก่อนการทำงาน (Precondition) คือเงื่อนไขหรือสถานะของระบบที่จะทำให้ยูสเคสเริ่มทำงาน
- 6) เงื่อนไขหลังการทำงาน (Postcondition) คือเงื่อนไขหรือสถานะของระบบที่ต้องเป็นจริงเสมอหลังจบการทำงานยูสเคส
- 7) ลำดับเหตุการณ์การทำงานปกติ (Normal Flow of Events) คือลำดับการทำงานของยูสเคส จะบอกถึงการทำงานที่เป็นขั้นตอนเริ่มตั้งแต่ต้นจนจบการทำงาน
- 8) เหตุการณ์ย่อย (Subflow) คือลำดับการทำงานย่อยที่เกิดขึ้นภายในลำดับเหตุการณ์หลัก
- 9) ลำดับทางเลือกการทำงาน (Alternative or Exceptional Flow of Events) คือลำดับการทำงานที่มีขั้นตอนการทำงานที่ต่างจากลำดับเหตุการณ์ทำงานปกติ

5. วิธีการค้นคืนยูสเคสกลับมาใช้ใหม่

การค้นคืนยูสเคสเพื่อนำกลับมาใช้ใหม่ในงานวิจัยนี้จะนำเสนอ 2 วิธีการอันได้แก่ วิธีการในการจำแนกประเภทยูสเคสและวิธีการในการค้นคืนย้อนกลับจากผู้ใช้ ซึ่งเป็นวิธีการที่จะนำมาเพิ่มเติมเข้าไปในระบบการจัดเก็บและค้นคืนสารสนเทศแบบทั่วไป ดังนั้นจึงขอกล่าวถึงตั้งแต่วิธีการเริ่มต้นของระบบงานซึ่งมีขั้นตอนได้แก่ การจัดเก็บยูสเคส การจัดกลุ่มหรือจำแนกยูสเคส การค้นคืนยูสเคส และการค้นคืนยูสเคสแบบย้อนกลับจากผู้ใช้ ดังนี้

5.1 การจัดเก็บยูสเคส

การจัดเก็บแต่ละยูสเคสเข้าสู่ฐานข้อมูลนั้นต้องมีการสร้างดัชนีเพื่อเป็นตัวแทนของยูสเคส และกำหนดค่าน้ำหนักให้กับดัชนีด้วยโดยในงานวิจัยนี้ใช้ดัชนีแบบเพิ่มข้อมูลผกผัน (Inverted File) ซึ่งการสร้างดัชนีคือการนำข้อความในเอกสารมาผ่านการดำเนินการกับข้อความ แล้วจึงกำหนดน้ำหนักให้กับดัชนีด้วยค่าความถี่ของเอกสารแบบผกผัน วิธีการกำหนดเป็นดังสมการที่ 1 นั่นคือค่าน้ำหนักของคำหนึ่งในเอกสารใดๆ จะได้จากสัดส่วนความถี่ของคำนั้นทั้งหมดที่ปรากฏในเอกสารต่อความถี่ของคำนั้นทั้งหมดที่ปรากฏอยู่ในฐานข้อมูล

$$W_{ik} = \frac{Freq_{ik}}{TotFreq_k} \quad (1)$$

เมื่อ W_{ik} คือค่าน้ำหนักของคำ k ในเอกสาร i
 $Freq_{ik}$ คือความถี่ของคำ k ทั้งหมดที่ปรากฏในเอกสาร i
 $TotFreq_k$ คือความถี่ของคำ k ทั้งหมดที่ปรากฏในฐานข้อมูล

5.2 การจำแนกประเภทยูสเคส

หลังจากจัดเก็บยูสเคสทั้งหมดเข้าสู่ระบบเรียบร้อยแล้วจะดำเนินการจำแนกประเภทยูสเคสเหล่านั้นเพื่อจัดกลุ่มให้กับยูสเคสแบบอัตโนมัติดังนี้

- 1) ตั้งค่าขีดแบ่งเริ่มต้น (Threshold) เพื่อใช้กำหนดว่ายูสเคสในกลุ่มเดียวกันจะต้องมีค่าความคล้ายกันอยู่ในช่วงใด
- 2) เปรียบเทียบความคล้ายระหว่างดัชนีของยูสเคสทีละคู่จนครบทุกยูสเคส โดยใช้สูตรของค่าสัมประสิทธิ์โคไซน์ (Cosine Coefficient) ดังนี้

$$COSIN(UC_i, UC_j) = \frac{\sum_{k=1}^l (W_{ik} \cdot W_{jk})}{\sqrt{\sum_{k=1}^l (W_{ik})^2 \cdot \sum_{k=1}^l (W_{jk})^2}} \quad (2)$$

เมื่อ UC_i คือดัชนียูสเคส i
 UC_j คือดัชนียูสเคส j

W_{ik} คือน้ำหนักของคำ k ในยูสเคส i

W_{jk} คือน้ำหนักของคำ k ในยูสเคส j

- 3) นำค่าความคล้ายที่ได้มาจัดกลุ่มให้กับยูสเคส
- 4) หายูสเคสตัวแทนกลุ่ม (Centroid) ของแต่ละกลุ่ม โดยยูสเคสตัวแทนกลุ่มคือยูสเคสที่มีค่าครรชนีเฉลี่ยจากครรชนีของยูสเคสในกลุ่ม
- 5) กรณีมีการใส่ยูสเคสใหม่เข้ามาในระบบ สามารถจัดกลุ่มให้ยูสเคสนั้นได้โดยการเปรียบเทียบความคล้ายระหว่างยูสเคสนั้นกับยูสเคสตัวแทนกลุ่มที่ละกลุ่ม หากพบว่าค่าความคล้ายของยูสเคสใหม่ดังกล่าวอยู่ในช่วงค่าขีดแบ่งเริ่มต้นกับกลุ่มใดจึงจัดให้อยู่ในกลุ่มนั้นได้

5.3 การค้นคืนยูสเคส

ในการค้นคืนยูสเคส ผู้ใช้จะป้อนความต้องการของตน ซึ่งจะถูกระบุการเพื่อสร้างเป็นครรชนี โดยเฉพาะงานวิจัยนี้จะใช้รูปแบบการค้นคืนเช่นเดียวกับ Akdej [1] คือใช้คิวรีที่มีลักษณะเป็นเอกสารการบรรยายยูสเคสอย่างง่าย โดยจะเริ่มที่การเปรียบเทียบความคล้ายระหว่างแต่ละส่วนประกอบของเอกสารการบรรยายยูสเคสกับแต่ละส่วนประกอบของคิวรี โดยใช้สูตรดังสมการที่ 3

$$Sim(E_{mi}, E_{mj}) = \frac{2 \cdot \sum_{k=1}^l (TERM_{mik} \cdot W_{ik}) \cdot TERM_{mjk}}{\sum_{k=1}^l TERM_{mik} + \sum_{k=1}^l TERM_{mjk}} \quad (3)$$

เมื่อ E_{mi} คือส่วนประกอบที่ m ของยูสเคสที่ i

E_{mj} คือส่วนประกอบที่ m ของคิวรีที่ j

$TERM_{mik}$ เป็น 1 เมื่อปรากฏคำ k ในส่วนประกอบที่ m ของยูสเคสที่ i และเป็น 0 ในทางกลับกัน

$TERM_{mjk}$ เป็น 1 เมื่อปรากฏคำ k ในส่วนประกอบที่ m ของคิวรีที่ j และเป็น 0 ในทางกลับกัน

W_{ik} คือน้ำหนักของคำ k ในยูสเคสที่ i

แล้วจึงรวมผลค่าความคล้ายกันจากแต่ละส่วนประกอบของยูสเคสและคิวรีด้วยสูตรดังสมการที่ 4

$$Sim(UC_i, Query_j) = \frac{\sum_{m=1}^e [Sim(E_{mi}, E_{mj}) \cdot WE_m]}{TotalWeight} \quad (4)$$

เมื่อ WE_m คือน้ำหนักของส่วนประกอบที่ m ซึ่งระบุโดยผู้ใช้

TotalWeight คือผลรวมน้ำหนักทั้งหมดที่ผู้ใช้ระบุ

e คือจำนวนส่วนประกอบในยูสเคสที่ i และคิวรี j

เพื่อค้นคืนยูสเคสที่มีค่าความคล้ายกับคิวรีมากที่สุดออกมา โดยงานวิจัยนี้ได้เพิ่มงานในการจำแนกประเภทยูสเคส และหายูสเคสตัวแทนของแต่ละกลุ่ม ดังนั้นการเปรียบเทียบคิวรีเพื่อหายูสเคสที่มีความคล้ายกันมากที่สุดจึงไม่ต้องเปรียบเทียบกับทุกยูสเคส สามารถเริ่มการเปรียบเทียบกับยูสเคสที่เป็นตัวแทนกลุ่มได้เลย แล้วจึงทำการเปรียบเทียบกับยูสเคสที่เป็นสมาชิกในกลุ่มต่อไป ดังนั้นในส่วนของการเปรียบเทียบเพื่อหาค่าความคล้ายระหว่างคิวรีกับยูสเคสตัวแทนกลุ่มนั้น ใช้สูตรในการคำนวณดังสมการที่ 3 และ 4 ซึ่งยูสเคสที่ถูกนำมาคำนวณความคล้ายกับคิวรีนั้นจะเริ่มที่ยูสเคสตัวแทนกลุ่มที่ละกลุ่ม หลังจากพบว่ายูสเคสตัวแทนกลุ่มใดมีความคล้ายกับคิวรีสูงพอตามค่าขีดแบ่งที่ตั้งไว้ จึงเปรียบเทียบความคล้ายกับยูสเคสที่เป็นสมาชิกในกลุ่มต่อไป ก่อนแสดงรายการยูสเคสทางส่วนต่อประสานผู้ใช้

5.4 การค้นคืนย้อนกลับจากผู้ใช้

เป็นวิธีหลังจากที่ผู้ใช้ตรวจสอบยูสเคสที่ระบบค้นคืนมาให้ในครั้งแรก เพื่อให้ผู้ใช้ระบุว่ายูสเคสที่ระบบค้นคืนได้นั้น ยูสเคสใดบ้างที่ตรงกับความต้องการผ่านทางส่วนต่อประสานผู้ใช้ ระบบจะนำข้อมูลนี้ไปคำนวณเพื่อใช้เปลี่ยนแปลงเป็นคิวรีใหม่ [4 10] โดยมี 2 แนวทางที่จะใช้ในงานวิจัยนี้ได้แก่

- 1) การเปลี่ยนแปลงเทอมในคิวรีและการเปลี่ยนแปลงน้ำหนักคำในคิวรีในแบบจำลองเวกเตอร์ (Query Expansion and Term Reweighting for the Vector Model)

2) การเปลี่ยนแปลงค่าน้ำหนักของคำในควิรีในแบบจำลองความน่าจะเป็น (Term Reweighting for the Probabilistic Model)

โดยในแนวทางแรกจะได้ควิรีที่เปลี่ยนแปลงทั้งคำและน้ำหนัก เพื่อป้อนกลับเข้าสู่ระบบค้นคืน จากสมการที่ 5

$$Q' = aQ + b \left(\frac{1}{R'} \sum_{i \in D_{R'}} DOC_i \right) - g \left(\frac{1}{N'} \sum_{i \in D_{N'}} DOC_i \right) \quad (5)$$

เมื่อ Q' คือควิรีใหม่

Q คือควิรีเริ่มต้น

R' คือจำนวนเอกสารที่ตรงความต้องการของผู้ใช้ โดยจะขึ้นอยู่กับผู้ใช้แต่ละคน

N' คือจำนวนเอกสารที่ไม่ตรงความต้องการของผู้ใช้

DOC_i คือเวกเตอร์ของเอกสารที่ i

$D_{R'}$ คือเซตเอกสารที่เกี่ยวข้องในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด

$D_{N'}$ คือเซตเอกสารที่ไม่เกี่ยวข้องในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด

α, β, γ คือค่าคงที่สำหรับการปรับค่า นิยมให้ $\alpha=1$ และกำหนดให้ β มีค่ามากกว่า γ

หลังจากระบบปรับควิรีใหม่แล้ว จะทำการค้นคืนยูสเคสด้วยวิธีการเดียวกันกับการค้นคืนยูสเคสดังที่กล่าวในข้อ 5.3 อีกแนวทางหนึ่งในการเปลี่ยนแปลงน้ำหนักคำในควิรีเดิมด้วยแบบจำลองความน่าจะเป็นนั้น พิจารณาหาความน่าจะเป็นได้จากแนวคิดที่ว่า มีความน่าจะเป็นเท่าไรที่เอกสารชิ้นหนึ่งตรงความต้องการของผู้ใช้

ดังนั้นสูตรที่ใช้ในการค้นคืนย้อนกลับจากผู้ใช้จะสามารถปรับเปลี่ยนค่าน้ำหนักของคำในควิรีด้วยแบบจำลองความน่าจะเป็นได้ โดยหาค่าของความคล้ายระหว่างยูสเคสกับควิรีดังนี้

$$Sim(E_{mi}, E_{mj}) = \sum_{k=1}^l W_{mik} \times W_{mjk} \times \log \left[\frac{|D_{r,k}|}{|D_r| - |D_{r,k}|} \cdot \frac{N - |D_r| - n_k + |D_k|}{n_k - |D_{r,k}|} \right] \quad (6)$$

เมื่อ E_{mi} คือส่วนประกอบที่ m ของยูสเคสที่ i

E_{mj} คือส่วนประกอบที่ m ของควิรีที่ j

W_{mik} มีค่าเป็น 1 เมื่อคำ k ปรากฏในส่วนประกอบที่ m ของยูสเคสที่ i และเป็น 0 หากไม่พบ

W_{mjk} มีค่าเป็น 1 เมื่อคำ k ปรากฏในส่วนประกอบที่ m ของควิรีที่ j และเป็น 0 หากไม่พบ

n_k คือจำนวนยูสเคสในชุดเอกสารทั้งหมดที่มีคำ k อยู่ภายใน

N คือจำนวนยูสเคสทั้งหมด

D_r คือเซตของยูสเคสที่เกี่ยวข้องที่ค้นคืนได้

$D_{r,k}$ คือสับเซตของ D_r ที่ประกอบด้วยยูสเคสที่มีคำ k

โดยสมการนี้ใช้ในการเปรียบเทียบความคล้ายระหว่างส่วนประกอบแต่ละส่วนของเอกสารการบรรยายยูสเคสกับแต่ละส่วนประกอบของควิรีซึ่งมีการเปลี่ยนแปลงค่าน้ำหนักคำในควิรีไปพร้อมกัน หลังจากนั้นจึงสามารถนำค่าความคล้ายที่ได้จากแต่ละส่วนประกอบไปคำนวณด้วยผลรวมความคล้ายดังสมการที่ 4 ต่อไป

ในการเปลี่ยนแปลงควิรีใหม่ทั้ง 2 วิธีที่กล่าวมานี้คือการเปลี่ยนแปลงผลป้อนกลับจากผู้ใช้ เป็นความต้องการใหม่ที่ระบบจะนำไปค้นคืนเอกสารอีกครั้งเพื่อให้ตรงกับความต้องการของผู้ใช้มากขึ้นกว่าการค้นคืนในครั้งแรกเพียงครั้งเดียว ซึ่งผู้วิจัยจะทำการพัฒนาเครื่องมือเพื่อทดลองและหาผลลัพธ์ที่ได้จากทั้ง 2 แนวทางต่อไป

6 การประเมินประสิทธิภาพของระบบ

สำหรับวิธีในการประเมินผลการค้นคืน ใช้วิธีการวัดด้วยค่าเรียกคืน (Recall: r) และค่าความแม่นยำ (Precision: p) เพื่อวัดประสิทธิภาพของระบบ [9] กล่าวคือวัดจากความเกี่ยวข้องกับควิรีหรือข้อมูลความต้องการของผู้ใช้หนึ่งๆ โดยในการวัดค่าเรียกคืนจะคิดสัดส่วนของยูสเคสที่มีความเกี่ยวข้องกับที่ค้นคืนได้ต่อยูสเคสที่เกี่ยวข้องทั้งหมดในฐานข้อมูล ดังสมการที่ 7

$$Recall(r) = \frac{ReIUC}{TotReIUC} \quad (7)$$

เมื่อ $ReIUC$ คือจำนวนยูสเคสที่มีความเกี่ยวข้องที่ ถูกค้นคืน

$TotReIUC$ คือจำนวนยูสเคสที่เกี่ยวข้องทั้งหมดใน ฐานข้อมูล

และค่าความแม่นยำนั้นจะคิดสัดส่วนของยูสเคสที่ เกี่ยวข้องที่ค้นคืนได้ต่อยูสเคสที่ค้นคืนได้ทั้งหมด ดัง สมการที่ 8

$$Precision(p) = \frac{ReIUC}{TotRetUC} \quad (8)$$

เมื่อ $ReIUC$ คือจำนวนยูสเคสที่มีความเกี่ยวข้องที่ ถูกค้นคืน

$TotRetUC$ คือจำนวนยูสเคสที่ค้นคืนได้ทั้งหมด

ค่าเรียกคืนและค่าความแม่นยำจะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าค่าเรียกคืนหรือค่าความถูกต้องเข้าใกล้ 1 มากเท่าใด แสดงว่าระบบมีประสิทธิภาพมากขึ้นด้วย

มาตรวัดอีกหนึ่งวิธีที่จะใช้ในงานวิจัยนี้ได้แก่ การวัด ด้วยค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) ซึ่งเป็นการวัด จากค่าเรียกคืนและค่าความแม่นยำไว้รวมกัน มีค่าอยู่ ระหว่าง 0 ถึง 1 เช่นกัน โดยมีค่าเป็น 1 เมื่อเอกสารทุก เอกสารที่ค้นคืนได้เป็นเอกสารที่ตรงความต้องการ และจะ เป็น 0 เมื่อเอกสารที่ค้นคืนมาไม่มีเอกสารใดที่ตรงความ ต้องการเลย มีสูตรการคำนวณดังสมการที่ 9

$$F = \frac{2}{\frac{1}{r} + \frac{1}{p}} \quad (9)$$

เมื่อ F คือค่าเฉลี่ยฮาร์โมนิก

r คือค่าเรียกคืน

p คือค่าความแม่นยำ

ดังนั้นในงานวิจัยนี้จะทำการประเมินประสิทธิภาพของ ระบบด้วยมาตรวัด 3 ค่า ดังที่กล่าวมานี้

7. กรณีตัวอย่าง

เพื่อให้มีความเข้าใจมากยิ่งขึ้นในวิธีการที่งานวิจัยนี้ เสนอ ในหัวข้อนี้จึงขอนำกรณีตัวอย่างในการค้นคืนยูสเคส โดยเริ่มต้นจากการจำแนกเอกสารการบรรยายยูสเคส

ทั้งหมดในฐานข้อมูลก่อน แล้วตามด้วยการค้นคืนยูสเคสที่ ทำการเปรียบเทียบจากเอกสารทีละกลุ่ม และการใช้การค้น คืนย้อนกลับจากผู้ใช้ โดยข้อมูลสมมุติของกรณีตัวอย่างมี ดังนี้

- สมมุติให้ยูสเคสทั้งหมดในฐานข้อมูลมี 7 ยูสเคส โดยหลังจากที่ระบบทำการเปรียบเทียบยูสเคสที ละคู่แล้วพบว่าได้ค่าความคล้ายระหว่างแต่ละคู่ ยูสเคสเป็นดังตารางที่ 1
- สมมุติค่าขีดเริ่มต้น (Threshold) เป็น 0.6

ตารางที่ 1 ค่าความคล้ายกันระหว่างยูสเคสแต่ละคู่

Similarity	U1	U2	U3	U4	U5	U6	U7
U1	-	0.3	0.5	0.6	0.8	0.9	0.5
U2	0.3	-	0.4	0.5	0.7	0.8	0.3
U3	0.5	0.4	-	0.3	0.5	0.2	0.7
U4	0.6	0.5	0.3	-	0.4	0.1	0.2
U5	0.8	0.7	0.5	0.4	-	0.3	0.1
U6	0.9	0.8	0.2	0.1	0.3	-	0.5
U7	0.5	0.3	0.7	0.2	0.1	0.5	-

ดังนั้นจะจัดกลุ่มให้ยูสเคสทั้ง 7 ได้เป็น 2 กลุ่มตามค่า ขีดแบ่งเริ่มต้นที่กำหนดไว้ โดยยูสเคสในกลุ่มแรกได้แก่ U1, U2, U4, U5 และ U6 ส่วนกลุ่มที่ 2 ได้แก่ U3 และ U7

หลังจากที่สามารถแบ่งกลุ่มยูสเคสได้แล้ว จะทำการ หาตัวแทนยูสเคสในแต่ละกลุ่ม โดยจะสมมุติค่าบรรชนีของ แต่ละยูสเคสไว้ดังตารางที่ 2 โดยขอยกตัวอย่างการหา ตัวแทนยูสเคสของกลุ่มแรกเท่านั้น

ตารางที่ 2 บรรชนีของยูสเคสในกลุ่มที่ 1

Group 1	Term1	Term2	Term3	Term4
U1	2	3	6	0
U2	4	2	4	2
U4	2	4	6	1
U5	1	5	7	1
U6	2	3	8	5
CentroidGroup1	1.85	2.85	5.71	1.43

ดังนั้นในการค้นหาเอกสารยูสเคสด้วยคิวรีหนึ่งๆที่
ป้อนจากผู้ใช้ คิวรีนั้นจะถูกนำมาเปรียบเทียบกับบรรณานุกรม
เป็นตัวแทนยูสเคสทีละกลุ่มจนครบทุกกลุ่มเสียก่อน หาก
พบว่าค่าใกล้เคียงหรืออยู่ในช่วงค่าขีดแบ่งที่กำหนดไว้กับกลุ่ม
ใด ก็จะดำเนินการเปรียบเทียบกับยูสเคสที่เป็นสมาชิกใน
กลุ่มนั้นๆต่อไป

ต่อมาขอเสนอตัวอย่างการค้นคืนย้อนกลับจากผู้ใช้
โดยสมมติให้คิวรีจากผู้ใช้และยูสเคสที่ค้นคืนได้เป็นดัง
ตารางที่ 3

ตารางที่ 3 บรรณานุกรมเริ่มต้นและเอกสารที่ค้นคืนได้

	Term1	Term2	Term3	Term4	Term5
Q	5	0	3	0	1
U1	2	1	2	0	0
U2	1	0	0	0	2

และสมมติว่าผู้ใช้ระบุว่ายูสเคสที่ตรงความต้องการคือ
ยูสเคสที่ 1 จะสามารถนำข้อมูลจากผู้ใช้นี้ ไปเปลี่ยนแปลง
คิวรีใหม่จากสมการที่ 5 ได้เป็นบรรณานุกรมของคิวรีใหม่ดัง
ตารางที่ 4

ตารางที่ 4 เวกเตอร์ของคิวรีใหม่

	Term1	Term2	Term3	Term4	Term5
Q'	7	1	5	0	1

หมายเหตุ กำหนดค่า $\alpha, \beta=1$ และ $\gamma=1$

เมื่อทดลองเปรียบเทียบความคล้ายระหว่างคิวรีใหม่กับ
ยูสเคสดังกล่าวจะพบว่า คิวรีใหม่ที่ได้มีค่าความคล้ายกับ
ยูสเคสที่ผู้ใช้ระบุว่า เป็นยูสเคสที่ตรงความต้องการมากขึ้น
และมีความคล้ายกับยูสเคสที่ผู้ใช้ระบุว่าไม่ตรงความ
ต้องการน้อยลงด้วย ดังแสดงในตารางที่ 5

ตารางที่ 5 ค่าความคล้ายระหว่างแต่ละเอกสารกับคิวรีเดิม
และคิวรีใหม่

sim	Q	Qc
U1	$(5*2)+(2*3)/5+9=1.14$	$(1*5)+(2*1)/3+9=0.58$
U2	$(2*7)+(1*1)+(2*5)/5+14=1.31$	$(1*7)+(2*1)/3+14=0.52$

จากนั้นนำคิวรีใหม่ที่ได้นี้ไปเปรียบเทียบกับบรรณานุกรม
ของยูสเคสในการค้นคืนอีกครั้ง ซึ่งผลการค้นคืนยูสเคสใน
แต่ละครั้งสามารถประเมินผลได้จากค่าความแม่นยำ ค่า
เรียกคืน และค่าเฉลี่ยฮาร์โมนิก โดยสมมุติค่าต่างๆ ดังนี้

- 1) สำหรับการค้นคืนครั้งแรกด้วยคิวรีเริ่มต้น
 - สมมติให้จำนวนยูสเคสที่เกี่ยวข้องที่ค้นคืน
ได้มี 6 ยูสเคส
- 2) สำหรับการค้นคืนย้อนกลับด้วยคิวรีใหม่
 - สมมติให้จำนวนยูสเคสที่เกี่ยวข้องที่ค้นคืน
ได้มี 8 ยูสเคส

โดยจำนวนยูสเคสทั้งหมดที่ค้นคืนได้มี 10 ยูสเคส
จำนวนยูสเคสทั้งหมดที่เกี่ยวข้องในฐานข้อมูลมี 15 ยูสเคส
แล้วใช้สมการที่ 7, 8 และ 9 ในการวัดประสิทธิภาพของ
ระบบได้ดังตารางที่ 6

ตารางที่ 6 ค่าความแม่นยำค่าเรียกคืนและค่าเฉลี่ยฮาร์โมนิก
จากคิวรีเริ่มต้นและคิวรีใหม่

มาตรวัด	Q (คิวรีเริ่มต้น)	Q' (คิวรีใหม่)	% ประสิทธิภาพ ที่เพิ่มขึ้น
ค่าความแม่นยำ (p)	$\frac{6}{10}=0.6$	$\frac{8}{10}=0.8$	33.33
ค่าเรียกคืน (r)	$\frac{6}{15}=0.4$	$\frac{8}{15}=0.53$	32.5
ค่าเฉลี่ยฮาร์โมนิก (F)	0.48	0.64	33.33

จากตารางที่ 6 พบว่าการใช้คิวรีใหม่ในการค้นคืนทำให้
ทั้งค่าความถูกต้อง ค่าเรียกคืน และค่าเฉลี่ยฮาร์โมนิก เพิ่ม
มากขึ้นกว่าคิวรีเริ่มต้น โดยเปอร์เซ็นต์ประสิทธิภาพที่
เพิ่มขึ้นของค่าความแม่นยำ ค่าเรียกคืน และค่าเฉลี่ยฮาร์โมนิก
เป็น 33.33%, 32.5% และ 33.33% ตามลำดับ

8. สรุปผลการวิจัยและงานที่จะดำเนินการในอนาคต

งานวิจัยนี้นำเสนอวิธีการค้นคืนยูสเคสเพื่อนำกลับมา
ใช้ใหม่ โดยนำระบบการจัดเก็บและค้นคืนสารสนเทศเข้า

มาใช้ ยูสเคสที่เก็บในฐานข้อมูลเป็นเอกสารการบรรยายยูสเคสที่ถูกจำแนกออกเป็นกลุ่มๆและมีการหาค่าครรชนีของเอกสารที่เป็นตัวแทนกลุ่มไว้ด้วย นอกจากนี้ยังเพิ่มเติมความสามารถให้กับระบบด้วยการนำการค้นคืนย้อนกลับจากผู้ใช้ ด้วยการเปลี่ยนแปลงคำและน้ำหนักในควิรีเดิมเป็นควิรีใหม่จากผลป้อนกลับมาในแบบจำลองเวกเตอร์ และเปลี่ยนแปลงเฉพาะน้ำหนักคำในควิรีเดิมในแบบจำลองความน่าจะเป็น ข้อดีของงานวิจัยนี้คือเป็นการนำยูสเคสกลับมาใช้ใหม่ซึ่งมีการดำเนินงานแบบอัตโนมัติเพื่ออำนวยความสะดวกต่อผู้ใช้งาน แต่อย่างไรก็ตามงานวิจัยนี้จะมีการพัฒนาต่อไปในส่วนของกรทดลองโดยต้องสร้างเครื่องมือขึ้นมาดำเนินการส่วนนี้ด้วย

9 บรรณานุกรม

- [1] Akadej Udomchaiporn, *Use Case Retrieval Using Terms and Use Case Structures Similarity Computation*, Thesis proposal, Department of Computer Engineering, Faculty of Engineering, Chulalongkorn university, 2006
- [2] C.W. Krueger, *Software reuse*, ACM Computing Survey, vol. 24, 1992, 131-183
- [3] Carol Lundquist, David A. Grossman and Ophir Frieder, *Improving Relevance Feedback in the Vector Space Model*, CIKM 97 LasVegas Nevada, 1997
- [4] G. Salton and C. Buckley, *Improving retrieval performance by relevance feedback*, Journal of the American Society for Information Science, 41, page 288-297, 1990.
- [5] G. Salton and C. Buckley., *The SMART and SIRE experimental retrieval systems*, Introduction to Modern Information Retrieval, page 141-143.
- [6] H. G. Woo, and Robinson, W.N., *Reuse of scenario specifications using an automated relational learner: a lightweight approach*, presented at Requirements Engineering, 2002, Proceedings, IEEE Joint International Conference on, 2002
- [7] M. Fowler, K. Scott, *UML Distilled*, 2000, Addison-Wesley.Inc
- [8] M. Saeki, *Reusing use case descriptions for requirements specification: towards use case patterns*, presented at Software Engineering Conference, 1999, (APSEC '99) Proceedings. Sixth Asia Pacific, 1999
- [9] R.B.Yates, Retrieval Evaluation, Modern Information Retrieval, page 9-11,1999, Addison-Wesley,Inc.
- [10] R.B.Yates, Query Operation, Modern Information Retrieval, page 117-121,1999, Addison-Wesley,Inc.

การจัดเก็บและค้นคืนยูสเคสที่มีการจัดกลุ่มยูสเคสร่วมกับการค้นคืนย้อนกลับจากผู้ใช้ Use Case Clustered Storage and Retrieval with User Relevance Feedback

สุดาทิพย์ สุขสอาด และ นครทิพย์ พร้อมพล

ห้องปฏิบัติการวิศวกรรมซอฟต์แวร์ ศูนย์เชี่ยวชาญเฉพาะทางด้านวิศวกรรมซอฟต์แวร์

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล: sudathip.s@student.chula.ac.th และ nakornthip.s@chula.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์และออกแบบระบบเพื่อการจัดเก็บและค้นคืนข้อกำหนดความต้องการซอฟต์แวร์ในรูปแบบของเอกสารคำอธิบายยูสเคส โดยการพิจารณาค่าความคล้ายของคำที่ใช้ในแต่ละส่วนประกอบของเอกสารกับข้อความ ซึ่งถือว่าการพิจารณาค่าตามโครงสร้างของเอกสาร และนำวิธีการจำแนกประเภทเพื่อจัดกลุ่มให้กับเอกสารคำอธิบายยูสเคส พร้อมทั้งวิธีการค้นคืนย้อนกลับจากผู้ใช้ ทั้งในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น มาปรับปรุงการค้นคืนเอกสารคำอธิบายยูสเคส นอกจากนี้ได้พัฒนาเครื่องมือเพื่อทดสอบแนวคิดที่นำเสนอ โดยอาศัยหลักการพื้นฐานจากทฤษฎีการจัดเก็บและค้นคืนสารสนเทศ และใช้การประเมินประสิทธิภาพของระบบด้วย ค่าเรียกคืนและค่าความแม่นยำ

คำสำคัญ การค้นคืนยูสเคส ยูสเคส คำอธิบายยูสเคส การจัดกลุ่มยูสเคส การค้นคืนย้อนกลับจากผู้ใช้

Abstract

The objective of this research is to analyze and design the system for storing and retrieving software requirements specifications which are in a use case description format. We consider the similarity of terms in each element of use case collections and use case query which is regarded as the structure consideration. The use case clustering is applied to

classify all use case descriptions in collection. Moreover, user relevance feedback on the vector space and probabilistic models are also applied to improve use case retrieval. A supporting tool is developed to validate our proposed method based on the fundamental of the information storage and retrieval theory. Recall and precision are two metrics used to evaluate our approach.

Keywords: Use Case Retrieval, Use Case, Use Case Description, Use Case Clustering, User Relevance Feedback

1. บทนำ

ด้วยการเติบโตและการแข่งขันที่มากขึ้นในอุตสาหกรรมซอฟต์แวร์ในปัจจุบัน และขั้นตอนต่างๆ ก่อนได้ซอฟต์แวร์มานั้น ยุ่งยากและซับซ้อน วิศวกรซอฟต์แวร์ส่วนใหญ่จึงเล็งเห็นความสำคัญของการนำซอฟต์แวร์ที่มีคุณภาพกลับมาใช้ใหม่ เพื่อลดปัญหาและความเสี่ยงต่างๆ ที่มักเกิดขึ้นได้ในช่วงของการพัฒนาซอฟต์แวร์ และช่วยจัดการเวลาและงบประมาณอย่างมีประสิทธิภาพได้ [1] แผนภาพยูสเคส และเอกสารคำอธิบายยูสเคสเป็นเอกสารสำคัญที่ใช้ในการบันทึกข้อกำหนดความต้องการซอฟต์แวร์ หากวิศวกรซอฟต์แวร์ นำเนื้อหาของข้อกำหนดความต้องการจากการพัฒนาซอฟต์แวร์เดิมที่มีความคล้ายกันกับความต้องการซอฟต์แวร์ที่กำลังจะพัฒนา กลับมาใช้ใหม่ได้ จะเป็นประโยชน์อย่างยิ่งต่อครั้งนั้นๆ ด้วยจุดเด่นของเอกสารคำอธิบายยูสเคสที่มีลักษณะเป็นโครงสร้าง จึงใช้ข้อดีนี้ประยุกต์เข้ากับบางวิธีการในระบบ

จัดเก็บและค้นคืนสารสนเทศ อย่างไรก็ตาม ความต้องการของผู้ใช้ที่แตกต่างกัน เป็นปัญหาหนึ่งที่ทำให้ผลการค้นคืนเอกสารไม่ตรงกับความต้องการ ซึ่งวิธีที่นำมาแก้ไขได้คือ การใช้การค้นคืนย้อนกลับจากผู้ช่วยแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น สำหรับการทดลอง ได้พัฒนาเครื่องมือรองรับวิธีการจะที่น่าสนใจ ส่วนการประเมินผล ใช้มาตรวัดคือ ค่าเรียกคืน (Recall) และค่าความแม่นยำ (Precision)

ในส่วนที่ 2 ของงานวิจัย คือ งานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 คือ เป้าหมายและแนวความคิด ส่วนที่ 4 คือ ลักษณะของแผนภาพและคำอธิบายยูสเคส ส่วนที่ 5 คือ การทดลองและผลการทดลอง และส่วนสุดท้ายคือ สรุปผลการทำวิจัย

2 งานวิจัยที่เกี่ยวข้อง

Saeki [2] เสนอแบบรูปต่างๆ (Patterns) ที่สามารถนำกลับมาใช้ใหม่ เพื่อให้ได้มาซึ่งข้อกำหนดความต้องการโดยข้อกำหนดความต้องการนั้น สามารถแปลงเป็นเอกสารคำอธิบายยูสเคสได้ แต่ข้อจำกัดของงานวิจัยนี้คือไม่สามารถนำข้อมูลภายในคำอธิบายยูสเคสที่มีอยู่เดิมกลับมาใช้ใหม่ได้ Woo และ Robinson [3] เสนอวิธีนำแผนภาพยูสเคสกลับมาใช้ใหม่ ด้วยเครื่องมือ “ScenAsst” เพื่อแปลงยูสเคสให้เป็นกราฟ แล้วจัดกลุ่มเก็บไว้ แล้วจึงค้นคืนด้วยข้อความที่แปลงเป็นกราฟเช่นกัน และเปรียบเทียบค่าความคล้ายกันกับยูสเคสที่เก็บไว้ แม้วิธีการนี้จะให้ผลลัพธ์ที่ดีในการนำยูสเคสกลับมาใช้ใหม่ แต่ใช้เวลาประมวลผลค่อนข้างมาก เนื่องจากมีความซับซ้อนทางโครงสร้างสูง

Salton [4] ออกแบบและพัฒนาระบบสมาร์ท (The SMART System) ซึ่งรู้จักกันดีในการศึกษาและทดลองเกี่ยวกับการจัดเก็บและค้นคืนสารสนเทศ ซึ่งมีทั้งการทำดัชนีอัตโนมัติ การจัดกลุ่มเอกสารอัตโนมัติ การเปลี่ยนแปลงข้อความด้วยข้อมูลจากผู้ใช้เป็นต้น โดยเอกสารในระบบนี้อยู่ในรูปแบบเวกเตอร์ และผลการทดลองการจัดกลุ่มเอกสารพบว่าช่วยลดเวลาในการเข้าถึงเอกสารได้ ส่วนการเปลี่ยนแปลงข้อความด้วยการค้นคืน

ย้อนกลับจากผู้ช่วย พบว่าทำให้ผลการค้นคืนเอกสารมีความแม่นยำมากขึ้น Lundquist และคณะ [5] เสนอการปรับปรุงประสิทธิภาพการค้นคืนย้อนกลับแบบเทียม (Pseudo Relevance Feedback) คือ ไม่มีผู้ช่วยมาระบุความเกี่ยวข้องของเอกสาร แต่ทดลองด้วยเทคนิคอื่น และใช้เวกเตอร์เป็นตัวแทนเอกสารซึ่งพบว่าให้ผลลัพธ์การค้นคืนที่แม่นยำขึ้น

Udomchaiporn และคณะ [6] เสนอวิธีนำยูสเคสกลับมาใช้ใหม่ ด้วยการคำนวณค่าความคล้ายกันของคำและโครงสร้างระหว่างเอกสารคำอธิบายยูสเคสในชุดข้อมูลกับข้อความที่มีลักษณะเป็นเอกสารยูสเคสอย่างง่าย และสร้างเครื่องมือทดสอบ ซึ่งประกอบด้วยส่วนจัดเก็บและส่วนค้นคืนยูสเคส ส่วนแรกจัดเก็บยูสเคสที่เป็นเอกสารคำอธิบายยูสเคสลงฐานข้อมูล และจัดการเอกสารด้วยการทำดัชนีแบบอัตโนมัติ และกำหนดค่าน้ำหนักคำดัชนีด้วยความถี่ของเอกสารแบบผกผัน ส่วนการค้นคืนนำข้อความจากแต่ละส่วนประกอบมาเปรียบเทียบความคล้ายกันระหว่างส่วนประกอบแต่ละส่วนของยูสเคสที่แปลงเป็นดัชนี แล้วจึงคำนวณค่าผลรวมความคล้ายกันระหว่างข้อความกับยูสเคสในชุดข้อมูล ร่วมกับจากที่ผู้ใช้ระบุให้กับแต่ละส่วนประกอบก่อนเริ่มการค้นคืน งานวิจัยนี้จึงมีข้อดีเรื่องการค้นคืนที่ความแม่นยำมากขึ้น แต่ไม่มีการจัดกลุ่มของยูสเคส และการค้นคืนย้อนกลับจากผู้ช่วย

3 เป้าหมายและแนวความคิด

เป้าหมายสำคัญของงานวิจัยนี้คือ การนำลักษณะโครงสร้างของยูสเคสมาใช้ในกระบวนการค้นคืนเพื่อสร้างโอกาสในการค้นคืนเอกสารความต้องการซอฟต์แวร์ที่ตรงความต้องการได้แม่นยำมากขึ้น กลับมาใช้ใหม่ และเพื่อเปรียบเทียบประสิทธิภาพของระบบการจัดเก็บและค้นคืนยูสเคสที่ได้ทำการจัดกลุ่มของยูสเคส ระหว่างแบบที่ใช้และไม่ใช้การค้นคืนย้อนกลับจากผู้ช่วย ในแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น โดยมีแนวความคิดสำคัญดังต่อไปนี้

3.1 การจัดกลุ่มเอกสารยูสเคส

การจัดเอกสารเป็นกลุ่มนั้นถูกนำมาใช้มากในการทำเหมืองข้อมูล (Data Mining) และการค้นคืนเอกสารที่มีเนื้อหาแตกต่างกันจำนวนมาก และพบว่าช่วยเพิ่มค่าเรียกคืน และค่าความแม่นยำให้กับระบบค้นคืนสารสนเทศ [7] แต่การจำแนกเอกสารมีหลายเทคนิคและวิธีการ การเลือกใช้วิธีใดจึงขึ้นอยู่กับความเหมาะสมและลักษณะของเอกสารด้วย งานวิจัยนี้ เอกสารที่ใช้ในการทดลองคือเอกสารคำอธิบายยูสเคส ซึ่งเป็นเอกสารข้อกำหนดความต้องการซอฟต์แวร์ จึงต้องคำนึงถึงลักษณะอื่นๆ เช่น วิศวกรรมซอฟต์แวร์ด้วย ดังนั้นการวิเคราะห์แนวทางการจัดกลุ่มเอกสารยูสเคส ซึ่งในชุดข้อมูลของงานวิจัยเดิม [6] ไม่มีขั้นตอนนี้รวมอยู่ มีขั้นตอนดังต่อไปนี้

1) ผู้เชี่ยวชาญเฉพาะทางด้านวิศวกรรมซอฟต์แวร์ตรวจสอบยูสเคสทั้งหมดในชุดข้อมูล เพื่อจำแนกว่า ควรมีจำนวนกลุ่มเท่าใด ะไรบ้าง โดยเกณฑ์ที่ใช้ในการจัดกลุ่มอย่างคร่าวๆ ได้แก่ การพิจารณายูสเคสตามระบบงาน (System) ระบบงานย่อย (Subsystem) และตามหน้าที่ของยูสเคส (Function) เป็นต้น

2) จากนั้น ผู้เชี่ยวชาญฯ เสนอให้จัดเอกสารยูสเคสตามระบบงานทางธุรกิจ (Business Domain) ในงานวิจัยนี้มีทั้งหมด 40 ระบบงานทางธุรกิจ จึงได้ 40 กลุ่มยูสเคสด้วย เนื่องจากแต่ละระบบงานมีลักษณะที่แตกต่างกันในระดับหนึ่งแล้ว หากจัดให้มีจำนวนกลุ่มน้อยลง หรือจัดให้ความละเอียดของคุณสมบัติของยูสเคสน้อยกว่านี้ อาจทำให้เอกสารยูสเคสเป้าหมาย ไม่ถูกค้นคืนออกมา

3) คำนวณหาตัวแทนกลุ่มยูสเคส ซึ่งเป็นค่าเฉลี่ยของคำสำคัญ หรือเทอมทั้งหมดในกลุ่ม ด้วยสูตรดังสมการที่ 1

$$CTerm_{pk} = \frac{1}{m} \sum_{i=1}^m Term_{ik} \quad (1)$$

เมื่อ $CTerm_{pk}$ คือ คำนวณน้ำหนักของเทอม k เฉลี่ยในเอกสารของกลุ่ม p , m คือ จำนวนเอกสารยูสเคสทั้งหมดในกลุ่ม p , $Term_{ik}$ คือ คำนวณน้ำหนักของเทอม k ในเอกสารยูสเคส i

4) หากมีการจัดเก็บยูสเคสใหม่เข้าสู่ระบบจะจัดเก็บยูสเคสทั้งระบบงาน ก่อนหาตัวแทนกลุ่มยูสเคสของระบบงานนั้นๆ และเก็บลงฐานข้อมูลเพื่อใช้ค้นคืนยูสเคส

3.2 การจัดเก็บและค้นคืน

การค้นคืนสารสนเทศเกี่ยวข้องกับการกำหนดตัวแทนสารสนเทศ จัดเก็บ จัดการ และเข้าถึงรายการของสารสนเทศ [8] จึงประยุกต์บางทฤษฎีของระบบการจัดเก็บและค้นคืนสารสนเทศกับงานวิจัยนี้ เช่น การทำดัชนีอัตโนมัติ การกำหนดน้ำหนักค่า การคำนวณความคล้าย และการประเมินผลการค้นคืน เป็นต้น ซึ่งอธิบายไว้ดังนี้

1) การทำดัชนีอัตโนมัติคือ ขั้นตอนการกำหนดและจัดเตรียมคำดัชนีสำหรับข้อความที่เป็นภาษาธรรมชาติ หลังการจัดเก็บ เนื้อหาจะถูกวิเคราะห์ และเปลี่ยนรูปเป็นชุดดัชนี 3 ขั้นตอนที่นิยมใช้ในการทำดัชนีอัตโนมัติ คือ การวิเคราะห์เนื้อหาหรือโครงสร้างในเอกสาร (Parsing Contents from Documents) การตัดคำที่ไม่มีผลต่อการค้นหาทิ้งไป (Eliminating the Stop Words) และการลดรูปคำให้เป็นรากศัพท์ตามไวยากรณ์ (Stemming Words) [9]

2) การกำหนดน้ำหนักค่าคือ ขั้นตอนการกำหนดน้ำหนักหรือค่าให้กับคำดัชนีแต่ละคำในแต่ละเอกสาร โดยในงานวิจัยนี้ใช้ค่าน้ำหนักของความถี่ของเอกสารแบบผกผัน (Inverse Document Frequency: IDF) [8]

3) การคำนวณความคล้าย ฟังก์ชันที่นิยมได้แก่ การเปรียบเทียบความคล้ายกันของเอกสารด้วยค่าสัมประสิทธิ์โคซายน์ เนื่องจากมีช่วงค่าความคล้ายอยู่ระหว่าง 0 ถึง 1 เหมาะแก่การกำหนดค่าขีดแบ่งเริ่มต้น ในการเปรียบเทียบความคล้ายของเอกสารแต่ละคู่ซึ่งอยู่ในรูปแบบเวกเตอร์

4) การประเมินผลการค้นคืนของระบบ นิยมใช้ค่าเรียกคืน และค่าความแม่นยำเป็นมาตรวัด โดยค่าเรียกคืน คิดสัดส่วนของเอกสารที่เกี่ยวข้องที่ค้นคืนมาได้ต่อเอกสารที่เกี่ยวข้องทั้งหมดที่ในชุดข้อมูล และค่าความแม่นยำ คิดสัดส่วนของเอกสารที่เกี่ยวข้องที่ค้นคืนมาได้ต่อเอกสารที่ค้นคืนออกมาทั้งหมด

3.3 การวัดความคล้ายของข้อความและเอกสารยูสเคส

ด้วยลักษณะเอกสารแบบมีโครงสร้าง สามารถคำนวณค่าความคล้ายของคำอธิบายยูสเคสกับข้อความแบบแยกตามแต่ละส่วนประกอบทางโครงสร้าง รวมกับน้ำหนักที่ผู้ใช้ระบุ ด้วยสูตรดังสมการที่ 2

$$\text{SIM}(UC_i, \text{Query}_j) = \frac{\sum_m [\text{SIM}(EUC_{mi}, E\text{Query}_{mj}) \cdot WE_m]}{\text{TotalWeight}} \quad (2)$$

เมื่อ $\text{SIM}(EUC_{mi}, E\text{Query}_{mj})$ คือ ค่าความคล้ายระหว่างส่วนประกอบ m ของยูสเคส i กับส่วนประกอบ m ของข้อความ j , WE_m คือ น้ำหนักในส่วนประกอบ m และ TotalWeight คือ น้ำหนักรวมทั้งหมดที่ได้จากผู้ใช้

3.4 การค้นคืนย้อนกลับจากผู้ใช้

การค้นคืนย้อนกลับจากผู้ใช้เป็นวิธีในการเปลี่ยนแปลงรูปแบบข้อความที่นิยมมากที่สุด นักวิจัยหลายท่านพบว่าการค้นคืนย้อนกลับเป็นการค้นคืนสารสนเทศซ้ำที่ให้ผลลัพธ์ที่ดี [10] งานวิจัยนี้ใช้ 2 วิธีการค้นคืนย้อนกลับจากผู้ใช้ ได้แก่

1) การขยายคำและเปลี่ยนแปลงน้ำหนักของคำในข้อความในแบบจำลองปริภูมิเวกเตอร์ เพื่อเปลี่ยนเป็นข้อความให้ ใช้สูตรดังสมการที่ 3

$$Q' = \alpha Q + \beta \left(\frac{1}{R'} \sum_{i \in DR'} \text{DOC}_i \right) - \gamma \left(\frac{1}{N'} \sum_{i \in DN'} \text{DOC}_i \right) \quad (3)$$

Q คือ ข้อความเริ่มต้น, R' คือ จำนวนเอกสารที่ตรงความต้องการ, N' คือ จำนวนเอกสารที่ไม่ตรงความต้องการ, DOC_i คือ เวกเตอร์เอกสาร i , $D_{R'}$ คือ ชุดของเอกสารที่เกี่ยวข้องในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด และ $D_{N'}$ คือ ชุดของเอกสารที่ไม่เกี่ยวข้องในจำนวนเอกสารที่ค้นคืนได้ทั้งหมด α, β, γ คือ ค่าคงที่สำหรับการปรับค่า ในงานวิจัยนี้กำหนดให้ $\alpha=1, \beta=0.5$ และ $\gamma=0$

2) การเปลี่ยนแปลงน้ำหนักของคำในข้อความในแบบจำลองความน่าจะเป็น ใช้สูตรดังสมการที่ 4

$$\text{sim}(d_i, q) = \sum_{k=1}^n w_{i,q} \times w_{i,j} \times \left[\log \frac{P(k_i | R)}{1 - P(k_i | R)} + \log \frac{1 - P(k_i | \bar{R})}{P(k_i | \bar{R})} \right] \quad (4)$$

เมื่อ $P(k_i | R)$ คือ ความน่าจะเป็นที่พบเทอม k_i ในชุดเอกสารที่เกี่ยวข้อง (R) ระบุโดยผู้ใช้, $P(k_i | \bar{R})$ คือ ความน่าจะเป็นที่พบเทอม k_i ในชุดเอกสารที่ไม่เกี่ยวข้อง (\bar{R}), $w_{i,j}$ คือ ค่าน้ำหนักเทอม k_i ในเอกสาร d_j , $w_{i,q}$ คือ ค่าน้ำหนักเทอม k_i ในข้อความ q ซึ่ง $w \in \{0,1\}$

4 ลักษณะของแผนภาพและคำอธิบายยูสเคส

ส่วนประกอบของแผนภาพยูสเคสซึ่งเป็นเครื่องมือช่วยอธิบายปฏิสัมพันธ์ระหว่างผู้ใช้กับระบบ ได้แก่ 1) ผู้กระทำแสดงโดยคนหรือสิ่งของที่ใช้งานระบบ 2) ยูสเคสคือ สิ่งผู้กระทำดำเนินการกับระบบ 3) ความสัมพันธ์ระหว่างผู้กระทำกับยูสเคส และ 4) ขอบเขตของระบบ ซึ่งแผนภาพยูสเคสสามารถนำมาเขียนเป็นเอกสารคำอธิบายยูสเคสประกอบด้วยส่วนประกอบพื้นฐานสำคัญ 9 ส่วน [11] ได้แก่ (1) ชื่อยูสเคสคือ ชื่อที่ใช้เรียกยูสเคส ซึ่งต้องไม่ซ้ำกัน และเหมาะสมกับวัตถุประสงค์ (2) วัตถุประสงค์อธิบายเป้าหมายทางธุรกิจหรือการทำงานหลักของยูสเคสนั้นๆ (3) ผู้กระทำคือ คน สิ่งของ หรือระบบภายนอก ที่มีบทบาทกับการทำงานของยูสเคส (4) ความสัมพันธ์คือ การเชื่อมโยงกันแบบต่างๆระหว่างส่วนต่างๆ ภายในระบบ (5) เงื่อนไขก่อนทำงานคือ ข้อกำหนดที่ต้องเป็นจริงก่อนที่ยูสเคสจะถูกเรียกใช้งาน (6) เงื่อนไขหลังการทำงานคือ ข้อกำหนดที่ต้องเป็นจริงหลังจากจบการทำงานของยูสเคส (7) เหตุการณ์ปกติใช้อธิบายลำดับการทำงานทั่วไปของยูสเคส (8) เหตุการณ์ย่อย อธิบายลำดับเหตุการณ์ที่อาจแยกออกมาจากเหตุการณ์ปกติ (9) เหตุการณ์ทางเลือก อธิบายการทำงานพิเศษตามเงื่อนไขของเหตุการณ์ปกติ

ลักษณะของเอกสารคำอธิบายยูสเคสสามารถแสดงเป็นรูปแบบเดียวกับเครื่องมือในส่วนค้นคืนยูสเคส ดังรูปที่ 1

5 การทดลองและผลการทดลอง

วัตถุประสงค์ของการทดลอง คือ เพื่อทดสอบว่าการค้นคืนยูสเคสด้วยการจัดกลุ่มยูสเคส และ ใช้การค้นคืน

ย้อนกลับจากผู้ใช้นั้นมีประสิทธิภาพมากกว่าการค้นคืนยูส-
เคสโดยไม่ใช้วิธีการค้นคืนย้อนกลับจากผู้ใช้หรือไม่

ดังนั้นในการทดลองจึงมีปัจจัยต่างๆ เพื่อเปรียบเทียบ
ประสิทธิภาพของวิธีการในการค้นคืนยูสเคสได้แก่

5.1 ปัจจัยควบคุม

ปัจจัยที่ควบคุมได้จะช่วยลดความโน้มเอียงหรืออคติ
(Bias) ที่ส่งผลต่อการทดลอง ในงานวิจัยนี้มีดังต่อไปนี้

1) ยูสเคส (Use Cases) เอกสารคำอธิบายยูสเคสที่ใช้
ในการทดลองมีจำนวนทั้งหมด 61 เอกสาร จาก 40
ระบบงาน ซึ่งเขียนอธิบายเป็นภาษาอังกฤษ โดยรวบรวม
จากเอกสารโครงการวิศวกรรมความต้องการ โครงการงาน
มหาบัณฑิต และวิทยานิพนธ์ ของนิสิตภาควิชาวิศวกรรม
คอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

2) หน่วยตัวอย่าง (Sample Units) ใช้ 10 หน่วยตัวอย่าง
ทดสอบเครื่องมือที่พัฒนาขึ้น ซึ่งหน่วยตัวอย่างเป็นนิสิตที่
กำลังศึกษาระดับปริญญาโท สาขาวิชาวิศวกรรมซอฟต์แวร์
ผู้มีประสบการณ์ในการรวบรวม วิเคราะห์และบันทึก
ข้อกำหนดความต้องการซอฟต์แวร์ สามารถออกแบบ
ซอฟต์แวร์ด้วยแผนภาพยูสเคส นอกจากนี้ หน่วยตัวอย่าง
เป็นผู้มีความรู้ภาษาอังกฤษในเกณฑ์ดี โดยมีผลคะแนน โท-
เฟล (TOEFL) หรือเทียบเท่า ไม่ต่ำกว่า 500 คะแนน

รูปที่ 1 รูปแบบเครื่องมือค้นคืนยูสเคส

3) โจทย์ปัญหา (Problems) กำหนดขึ้นเพื่อให้หน่วย
ตัวอย่างป้อนข้อความคำถามในการค้นหายูสเคสนั้น โดยจะป้อน
ข้อความที่มีความใกล้เคียงหรือเกี่ยวข้องกับโจทย์ที่
กำหนดขึ้น เสมือนกับกำลังค้นหายูสเคสที่จะนำไปใช้เพื่อ
พัฒนาซอฟต์แวร์ โจทย์ปัญหาดังกล่าวได้แก่

1. Educational System
2. Material Requirements Planning
3. Member/User's Information Management
4. Product Trading Operation
5. Report Generation

โจทย์ที่กำหนดขึ้นทั้ง 5 นี้ นำมาใช้เปรียบเทียบ
ประสิทธิภาพของการค้นคืนกับสมมุติฐานทุกข้อ โดยแต่ละ
โจทย์นั้น ผู้วิจัยได้กำหนดชุดคำตอบ (Answer Set) และ
เก็บไว้ในฐานข้อมูล เพื่อใช้ในการคำนวณค่าเรียกคืน และ
ค่าความแม่นยำต่อไป

4) ข้อคำถาม (Queries) หน่วยตัวอย่างจะป้อนข้อ
คำถามเพื่อค้นหายูสเคสในแบบจำลองทั้งสองโดยในแต่ละ
แบบจำลอง แบ่งการพิจารณาโครงสร้างของเอกสารยูสเคส
เป็นการรวมข้อคำถามจากแต่ละส่วนของเอกสาร และ
พิจารณาข้อคำถามแยกกัน ก่อนค้นคืนด้วยแบบจำลอง
ข้างต้น ซึ่งสามารถแสดงผลค้นคืนเอกสารยูสเคสให้กับ
หน่วยตัวอย่างได้ ทั้งแบบไม่ใช้และใช้การค้นคืนย้อนกลับ
จากผู้ใช้ ดังนั้นหน่วยตัวอย่างหนึ่ง ต้องค้นคืนด้วยข้อ
คำถามจากโจทย์ทั้ง 5 ข้อ และแต่ละโจทย์ใช้ข้อคำถาม 5
ข้อคำถาม จึงมีข้อคำถาม 25 ข้อคำถามสำหรับสมมุติฐาน
ทั้งหมด และแต่ละโจทย์จะได้ชุดข้อคำถาม โจทย์ละ 50 ข้อ
คำถาม รวมเป็น 250 ข้อคำถามจาก 10 หน่วยตัวอย่าง แต่
ละข้อคำถามเมื่อใช้ค้นคืนยูสเคสแล้ว ผลที่ได้จะนำมา
ตรวจสอบกับชุดคำตอบเพื่อหาค่าเรียกคืน และค่าความ
แม่นยำ ข้อคำถามจึงเป็นตัวแทนที่ตัวหนึ่งด้วย

เพื่อให้เข้าใจง่ายขึ้น จำนวนข้อคำถามจากหน่วย
ตัวอย่างทั้งหมด แสดงดังตารางที่ 1

5) ค่าขีดแบ่งเริ่มต้นความคล้าย (Similarity
Threshold) เป็นเลขค่าหนึ่งที่ตั้งขึ้นให้ระบบการค้นคืน
เปรียบเทียบค่าความคล้ายที่หาได้ระหว่างข้อคำถามกับ
เอกสารยูสเคสด้วยตัวเลขนี้ หากค่าความคล้ายมีค่ามากกว่า

ตารางที่ 1 จำนวนข้อคำถามทั้งหมดจากหน่วยตัวอย่าง

โจทย์ปัญหา หน่วยตัวอย่าง	1	2	3	4	5	ข้อคำถาม จากแต่ละ หน่วย ตัวอย่าง
1	5	5	5	5	5	25
2	5	5	5	5	5	25
3	5	5	5	5	5	25
4	5	5	5	5	5	25
5	5	5	5	5	5	25
6	5	5	5	5	5	25
7	5	5	5	5	5	25
8	5	5	5	5	5	25
9	5	5	5	5	5	25
10	5	5	5	5	5	25
ข้อคำถามจาก แต่ละโจทย์	50	50	50	50	50	250

หรือเท่ากับค่าขีดแบ่งเริ่มต้นแล้ว จึงจะนำเอกสารเหล่านั้นออกมาแสดงกับผู้ใช้ โดยใช้ชุดข้อคำถามจาก [6] มาทดสอบกับเครื่องมือของงานวิจัยนี้ และค้นคืนยูสเคสจากชุดข้อมูลยูสเคสทั้งหมด 611 ยูสเคส โดยค่าขีดแบ่งเริ่มต้นความคล้ายหาได้จากค่าเฉลี่ยความคล้ายทั้งหมดที่ได้ลบออกด้วยค่าเบี่ยงเบนมาตรฐาน ดังสูตรในสมการที่ 5

$$\text{Threshold value} = \text{Mean} - \text{S.D.} \quad (5)$$

ผลของค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐานความคล้ายที่ได้จากวิธีการค้นคืนตามงานวิจัยนี้แบ่งออกเป็น 2 ค่าด้วยกัน คือค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐานของค่าความคล้ายระหว่างข้อคำถามกับตัวแทนกลุ่มยูสเคสทั้งหมด กับค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐานของค่าความคล้ายระหว่างข้อคำถามกับเอกสารยูสเคสทั้งหมด มีค่าเท่ากับ 0.0000629289 และ 0.000796947 ตามลำดับ

5.2 สมมุติฐานของการทดลอง

เครื่องมือที่พัฒนาขึ้นเพื่อสนับสนุนวิธีวิจัยที่น่าเสนอ มีฟังก์ชันการทำงานที่รองรับสมมุติฐานการทดลองดังนี้

1) ประสิทธิภาพการค้นคืนยูสเคสแบบพิจารณาโครงสร้างร่วมกันของเอกสารคำอธิบายยูสเคส ด้วย

แบบจำลองปริภูมิเวกเตอร์ โดยใช้การค้นคืนย้อนกลับจากผู้ใช้ดีกว่าแบบไม่ใช้หรือไม่

2) ประสิทธิภาพการค้นคืนยูสเคสแบบพิจารณาโครงสร้างร่วมกันของเอกสารคำอธิบายยูสเคส ด้วยแบบจำลองความน่าจะเป็น โดยใช้การค้นคืนย้อนกลับจากผู้ใช้ดีกว่าแบบไม่ใช้หรือไม่

3) ประสิทธิภาพการค้นคืนยูสเคสแบบพิจารณาโครงสร้างแยกกันของเอกสารคำอธิบายยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ โดยใช้การค้นคืนย้อนกลับจากผู้ใช้ดีกว่าแบบไม่ใช้หรือไม่

4) ประสิทธิภาพการค้นคืนยูสเคสแบบพิจารณาโครงสร้างแยกกันของเอกสารคำอธิบายยูสเคส ด้วยแบบจำลองความน่าจะเป็น โดยใช้การค้นคืนย้อนกลับจากผู้ใช้ดีกว่าแบบไม่ใช้หรือไม่

งานวิจัยนี้ทดสอบการค้นคืนแบบพิจารณาโครงสร้างร่วมกันของเอกสารคำอธิบายยูสเคส เพื่อเป็นข้อสังเกตอีกทางหนึ่งว่าการค้นคืนที่อาศัยโครงสร้างจะให้ประสิทธิภาพดีกว่า และเนื่องจากยูสเคสทั้งหมดถูกจัดกลุ่มไว้ 40 กลุ่มก่อนเริ่มการค้นคืน ดังนั้นในการค้นคืนยูสเคสแบบต่างๆ ที่กล่าวมานั้น ต้องผ่านฟังก์ชันในการเปรียบเทียบความคล้ายกับตัวแทนกลุ่มยูสเคสทั้ง 40 กลุ่ม ก่อนเสมอ

5.3 ขั้นตอนของการทดลอง

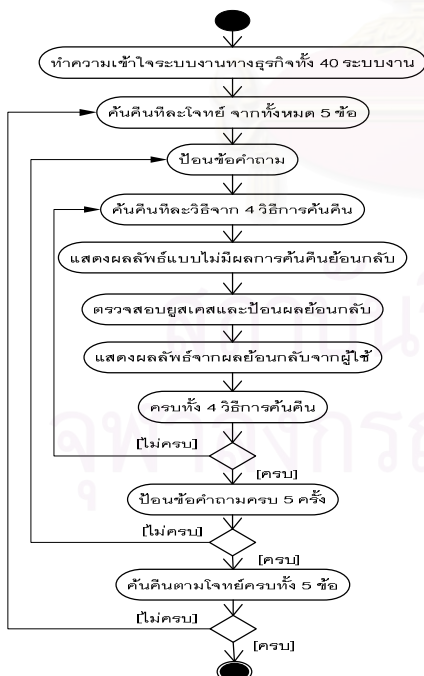
รายละเอียดของขั้นตอนการทดลองดังรูปที่ 2 เริ่มจาก (1) หน่วยตัวอย่างศึกษาระบบงานทั้ง 40 ระบบงาน (2) หน่วยตัวอย่างค้นคืนยูสเคสที่ละโจทย์ปัญหา ซึ่งมีทั้งหมด 5 ข้อ (3) หน่วยตัวอย่างป้อนข้อคำถามที่เกี่ยวข้องกับโจทย์ปัญหานั้นๆ โดยแต่ละโจทย์ หน่วยตัวอย่างจะต้องป้อนข้อคำถามทั้งหมด 5 ข้อคำถาม (4) หน่วยตัวอย่างใช้วิธีการค้นคืนยูสเคสทีละวิธีจาก 4 วิธีการค้นคืน (5) ระบบแสดงผลลัพธ์แบบไม่ใช้ผลการค้นคืนย้อนกลับ เป็นชื่อยูสเคสและเนื้อหาในคำอธิบายยูสเคส (6) หน่วยตัวอย่างตรวจสอบยูสเคสและเลือกยูสเคสที่เกี่ยวข้องกับความต้องการของตนแล้วป้อนผลการค้นคืนย้อนกลับ เข้าสู่

ระบบเพื่อดำเนินการค้นคืนซ้ำอีกครั้ง (7) ระบบแสดงชื่อ ยูสเคสที่ได้จากผลย้อนกลับจากผู้ (8) หน่วยตัวอย่างค้นคืน ยูสเคส จนครบทั้ง 4 วิธี (9) หน่วยตัวอย่างป้อนชื่อ คำถามจนครบทั้ง 5 ครั้ง (10) หน่วยตัวอย่างค้นคืน ยูสเคส ครบทั้ง 5 โจทย์

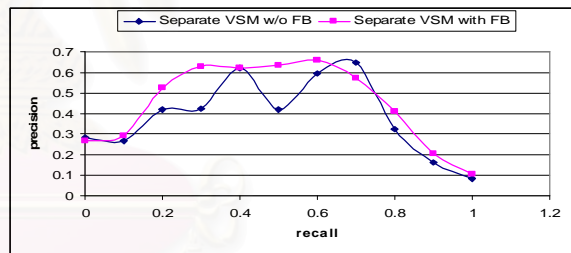
5.4 ผลการทดลอง

ผลการทดลองทั้งหมดจาก 250 ข้อคำถาม ประเมินด้วย ค่าเรียกคืน และค่าความแม่นยำ จัดลำดับตามค่าเรียกคืน 11 ค่า ทดสอบด้วยสมมุติฐานทั้ง 4 ดังตารางที่ 2 พบว่าวิธีที่ 1 และ 2 ซึ่งใช้การค้นคืนแบบพิจารณาโครงสร้างรวมกันไม่ ช่วยให้ค่าความแม่นยำเพิ่มขึ้นแม้ว่าจะใช้การค้นคืน ย้อนกลับจากผู้ ทั้งแบบจำลองปริภูมิเวกเตอร์และ แบบจำลองความน่าจะเป็น แต่วิธีที่ 3 และ 4 ซึ่งใช้การค้น คืนแบบพิจารณาโครงสร้างของยูสเคสแยกกัน และเพิ่มการ ค้นคืนย้อนกลับจากผู้เข้าไปด้วย ความแม่นยำมีค่ามาก ขึ้นจากวิธีที่ไม่ใช้การค้นคืนย้อนกลับ และมีค่าความ แม่นยำมากกว่าการค้นคืนแบบพิจารณาโครงสร้างรวมกัน ด้วย รูปที่ 3 และ 4 แสดงให้เห็นว่า การใช้การค้นคืนย้อน-

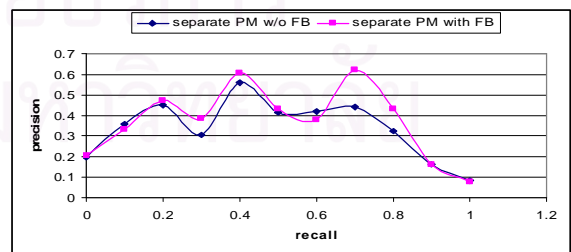
กลับในการค้นคืนแบบพิจารณาโครงสร้างแยกกันมีค่า ความแม่นยำมากกว่าไม่ใช้การค้นคืนย้อนกลับ ทั้ง 2 แบบจำลอง แต่ถ้าเทียบระหว่างวิธีที่ใช้การค้นคืนแบบ พิจารณาโครงสร้างของยูสเคสแยกกัน และใช้การค้นคืน ย้อนกลับ ดังตารางที่ 2 กับ รูปที่ 5 พบว่าแบบจำลองปริภูมิ- เวกเตอร์มีค่าความแม่นยำ โดยเฉลี่ย 0.448 ในขณะที่ แบบจำลองความน่าจะเป็น ให้ค่าความแม่นยำเฉลี่ยเท่ากับ 0.374 ซึ่งแบบจำลองปริภูมิเวกเตอร์มีความแม่นยำมากกว่า 19.79 % เนื่องจากการเปลี่ยนแปลงข้อคำถามของ แบบจำลองความน่าจะเป็นนั้น มีเพียงการเปลี่ยนแปลงค่า น้ำหนักเทอมในข้อคำถามเท่านั้น แต่แบบจำลองปริภูมิ- เวกเตอร์สามารถเปลี่ยนแปลงได้ทั้งเทอมและน้ำหนักเทอม ในข้อคำถาม จากยูสเคสที่ผู้ใช้เลือกที่มีความเกี่ยวข้องกับ ความต้องการ จึงเป็นสาเหตุให้การค้นคืนแบบพิจารณา โครงสร้างแยกกันของยูสเคสร่วมกับการค้นคืนย้อนกลับ จากผู้ใช้ในแบบจำลองปริภูมิเวกเตอร์ให้ค่าความแม่นยำได้ ดีที่สุด และมีความเหมาะสมกับการค้นคืนเอกสารยูสเคส



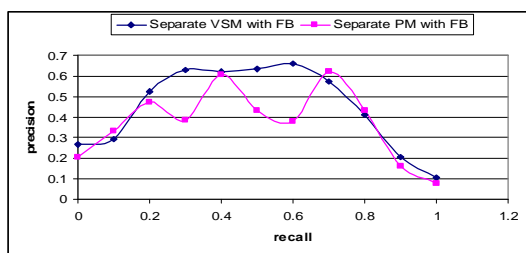
รูปที่ 2 ขั้นตอนกิจกรรมการทดลอง



รูปที่ 3 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบพิจารณาโครงสร้างแยกกันของยูสเคสด้วยจำลอง ปริภูมิเวกเตอร์ โดยใช้และไม่ใช้การค้นคืนย้อนกลับจากผู้



รูปที่ 4 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยูสเคสแบบพิจารณาโครงสร้างแยกกันของยูสเคสด้วยแบบจำลอง ความน่าจะเป็น โดยใช้และไม่ใช้การค้นคืนย้อนกลับจากผู้



รูปที่ 5 กราฟค่าเรียกคืน และค่าความแม่นยำระหว่างการค้นคืนยુสเคสแบบพิจารณาโครงสร้างแยกกันของยูสเคสโดยใช้ค้นคืนย้อนกลับจากผู้ใช้ ทั้งแบบจำลองปริภูมิเวกเตอร์และแบบจำลองความน่าจะเป็น

6 สรุปผลการทำวิจัย

งานวิจัยนี้เสนอวิธีค้นคืนยูสเคสเพื่อนำกลับมาใช้ใหม่ โดยนำระบบจัดเก็บและค้นคืนสารสนเทศเข้ามาใช้ ชุดข้อมูลเป็นเอกสารคำอธิบายยูสเคสที่ถูกจัดกลุ่ม และมีการหาคะรขนี้เอกสารตัวแทนกลุ่ม และเพิ่มเติมความสามารถให้กับระบบโดยนำการค้นคืนย้อนกลับจากผู้ใช้ มาเปลี่ยนแปลงคำ และนำหน้าคำในข้อความเดิมเป็นข้อความใหม่จากผลป้อนกลับมาจากผู้ใช้ ในแบบจำลองปริภูมิเวกเตอร์ และเปลี่ยนแปลงเฉพาะหน้าคำในข้อความเดิมด้วยแบบจำลองความน่าจะเป็น และสร้างเครื่องมือรองรับวิธีการที่นำเสนอ ผลการทดลองพบว่า การค้นคืนยูสเคสโดยพิจารณาโครงสร้างแยกกันร่วมกับการค้นคืนย้อนกลับจากผู้ใช้ ทั้งในแบบจำลองปริภูมิเวกเตอร์ และแบบจำลองความน่าจะเป็น ให้ค่าความแม่นยำเพิ่มขึ้นจากการไม่ใช้การค้นคืนย้อนกลับแต่สำหรับผลที่ได้จาก

แบบจำลองปริภูมิเวกเตอร์ให้ค่าความแม่นยำมากกว่าแบบจำลองความน่าจะเป็น จึงสรุปว่า การค้นคืนเอกสารที่เป็นโครงสร้างอย่างเอกสารคำอธิบายยูสเคสที่มีการจัดกลุ่มไว้แล้ว จะให้ผลลัพธ์ที่ดีที่สุด ถ้าค้นคืนโดยพิจารณาโครงสร้างแยกกันของเอกสารยูสเคส ด้วยแบบจำลองปริภูมิเวกเตอร์ร่วมกับการใช้ค้นคืนย้อนกลับจากผู้ใช้

7. บรรณานุกรม

- [1] C.W. Krueger, "Software reuse," ACM Computing Surveys, vol. 24, pp. 131-183, 1992.
- [2] M.Saeki, "Reusing use case descriptions for requirements specification: towards use case patterns," the 6th APSEC, 1999.
- [3] H.G. Woo, and W.N. Robinson, "Reuse of scenario specifications using an automated relational learner: a lightweight approach," RE Conference, 2002.
- [4] G. Salton and C. Buckley., "The SMART and SIRE experimental retrieval systems, introduction to modern information retrieval," page 141-143.
- [5] C. Lundquist, D. A.Grossman and O. Frieder., "Improving relevance feedback in the vector space model," CIKM, 1997.
- [6] A.Udomchaiporn, N. Prompoon and P. Kanongchaiyos, "Software requirements retrieval using use case terms and structure similarity computation," the 13th APSEC, 2006.
- [7] M. Steinbach, G. Karypis and V. Kuman, "A comparison of document clustering techniques," the 6th ACM SIGKDD, 2000.
- [8] M.J. McGill and G. Salton, "Introduction to modern information retrieval," McGraw-Hill, 1983.
- [9] B.R. Neto and R.B. Yates, "Modern information retrieval," Addison-Wesley, 1999.
- [10] D. Harman, "Relevance feedback revisited," the 15th Ann Int'l SIGIR, 1992.
- [11] G. Winters, "Use case terminology," Software, IEEE, vol. 22, pp. 67, 2005.

ตารางที่ 2 สรุปผลค่าความแม่นยำเฉลี่ยจาก 250 ข้อคำถามเรียงลำดับตามค่าเรียกคืน 11 ค่า จาก 0 ถึง 1 ในแต่ละวิธีการ

Method	(1) Combine VSM		(2) Combine PM		(3) Separate VSM		(4) Separate PM	
	w/o FB	with FB	w/o FB	with FB	w/o FB	with FB	w/o FB	with FB
Recall	Precision							
0.0	0.251	0.17	0.077	0.061	0.283	0.268	0.197	0.207
0.1	0.322	0.34	0.202	0.326	0.267	0.294	0.357	0.334
0.2	0.256	0.292	0.37	0.394	0.419	0.526	0.45	0.471
0.3	0.329	0.321	0.374	0.319	0.424	0.632	0.306	0.385
0.4	0.345	0.334	0.401	0.341	0.621	0.622	0.559	0.607
0.5	0.385	0.348	0.497	0.332	0.422	0.633	0.417	0.431
0.6	0.461	0.265	0.428	0.365	0.594	0.662	0.421	0.379
0.7	0.441	0.302	0.521	0.221	0.648	0.572	0.442	0.623
0.8	0.425	0.359	0.443	0.407	0.324	0.413	0.324	0.432
0.9	0.33	0.192	0.36	0.244	0.162	0.207	0.162	0.161
1.0	0.165	0.159	0.424	0.179	0.081	0.104	0.081	0.080
Mean	0.337	0.280	0.372	0.29	0.386	0.448	0.338	0.374
S.D.	0.085	0.070	0.123	0.099	0.177	0.191	0.136	0.163

หมายเหตุ: Combine VSM คือ วิธีค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสร่วมกันด้วยแบบจำลองปริภูมิเวกเตอร์
 Combine PM คือ วิธีค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสร่วมกันด้วยแบบจำลองความน่าจะเป็น
 Separate VSM คือ วิธีค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสแยกกันด้วยแบบจำลองปริภูมิเวกเตอร์
 Separate PM คือ วิธีค้นคืนยูสเคสแบบพิจารณาโครงสร้างของยูสเคสแยกกันด้วยแบบจำลองความน่าจะเป็น
 FB คือ การค้นคืนย้อนกลับจากผู้ใช้

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวสุดาทิพย์ สุขสะอาด เกิดวันที่ 12 สิงหาคม พ.ศ. 2524 ที่จังหวัดชลบุรี สำเร็จการศึกษาระดับปริญญาวิทยาศาสตรบัณฑิต จากภาควิชาคณิตศาสตร์ สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2546 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2547



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย