การพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

นายปองลิขิต สิงห์ชัย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมซอฟต์แวร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2556
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

A DEVELOPMENT OF RETRIEVAL THAI SOFTWARE REQUIREMENTS USING SOFTWARE PROJECT HISTORICAL DATA

Mr. Ponglikit Singchai

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Software Engineering
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2013

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต
โดย	นายปองลิขิต สิงห์ชัย
สาขาวิชา	วิศวกรรมซอฟต์แวร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร. วันชัย ริ้วไพบูลย์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์

(ศาสตราจารย์ ดร. บัณฑิต เอื้ออาภรณ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(รองศาสตราจารย์ ดร. วิวัฒน์ วัฒนาวุฒิ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร. วันชัย ริ้วไพบูลย์)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. อาทิตย์ ทองทักษ์)

.....กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร. สมชาย ปราการเจริญ)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ปองลิขิต สิงห์ชัย : การพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทยโดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต. (A DEVELOPMENT OF RETRIEVAL THAI SOFTWARE REQUIREMENTS USING SOFTWARE PROJECT HISTORICAL DATA) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร. วันชัย ริ้วไพบูลย์, 84 หน้า.

วิทยานิพนธ์นี้นำเสนอระเบียบวิธีและพัฒนาเครื่องมือสำหรับการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต เพื่อช่วยให้วิศวกรความต้องการสามารถสืบค้นความต้องการของโครงการในอดีตที่มีความสัมพันธ์กับคำสืบค้น แล้วนำมาเป็นแนวทางการกำหนดความต้องการของระบบให้ครบถ้วน

ผู้วิจัยนำเอกสารกำหนดความต้องการของซอฟต์แวร์ในอดีตมาวิเคราะห์หาความสัมพันธ์ตั้งแต่ระดับโครงสร้างไปจนถึงระดับคำสำคัญในเอกสาร แล้วคำนวณหาค่าน้ำหนักของคำและความต้องการเชิงหน้าที่ หลังจากนั้นจึงนำทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมาช่วยหาค่าสนับสนุนและค่าความเชื่อมั่นเชิงน้ำหนัก เพื่อหากฎความสัมพันธ์เชิงน้ำหนักระหว่างคำ

ผู้วิจัยสร้างเครื่องมือสำหรับการสืบค้นความต้องการซอฟต์แวร์ โดยประยุกต์ใช้การประมวลผลเอกสารความต้องการเพื่อหาดัชนีคำศัพท์ของเอกสาร แล้วนำกฎความสัมพันธ์เชิงน้ำหนักมาใช้ในการตัดสินใจหาคำที่สัมพันธ์กับคำสืบค้น จนกระทั่งได้เอกสารกำหนดความต้องการของซอฟต์แวร์เบื้องต้น ซึ่งเป็นแบบร่างสำหรับวิศวกรความต้องการนำไปสอบถามกับผู้ใช้งาน เพื่อให้ได้รับความต้องการของซอฟต์แวร์ที่สมบูรณ์และรวดเร็ว ซึ่งผู้วิจัยได้ประเมินผลของเครื่องมือ พบว่าคำที่สัมพันธ์กันสามารถทำให้ได้รับความต้องการเชิงหน้าที่ที่เกี่ยวข้องกัน ซึ่งสามารถได้รับความต้องการของซอฟต์แวร์ที่ครบถ้วนมากขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต .....

สาขาวิชา วิศวกรรมซอฟต์แวร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....

ปีการศึกษา 2556

5470965921 : MAJOR SOFTWARE ENGINEERING
KEYWORDS: REQUIREMENTS SPECIFICATION / STRUCTURED DOCUMENTS RETRIEVAL / SOFTWARE ENGINEERING / SOFTWARE ENGINEERING TOOL

PONGLIKIT SINGCHAI: A DEVELOPMENT OF RETRIEVAL THAI SOFTWARE REQUIREMENTS USING SOFTWARE PROJECT HISTORICAL DATA. ADVISOR: ASSOC. PROF. DR. WANCHAI RIVEPIBOON, 84 pp.

This thesis presents a methodology and tool to assist requirement engineers searching Thai software requirements. This research aims to use keywords to classify part requirements which can be same as a guide in gathering a requirement.

The researcher used the past software requirements specification to analyze the relationship from related words in the same context of the requirement which calculates the weight of the words and functional requirements. Our method is based on a calculation of weighted support and confidence using weighted association rule theory.

Afterwards, we build tool for software requirements retrieval by applying from document processing for finding the index terms of documents, using the weighted association rule for finding the relationship of words. Requirement engineers bring software requirements specification guideline to talk with users in order to get the complete software requirements specification. We find the relationship of word association with relationship of functional requirements. If requirement engineers can find more requirements quickly, they will get software requirements specification which helps to develop software project quickly.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Department: Computer Engineering Student's Signature .....
Field of Study: Software Engineering Advisor's Signature .....
Academic Year: 2013

กิตติกรรมประกาศ

ข้าพเจ้าขอขอบพระคุณ รองศาสตราจารย์ ดร.วันชัย ริ้วไพบูลย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่เสียสละเวลาช่วยเหลือให้คำปรึกษา คำแนะนำและข้อคิดเห็นที่มีประโยชน์ ทำให้การจัดทำวิทยานิพนธ์สำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณคณาจารย์ทุกท่านที่แนะนำสั่งสอนและให้ความรู้แก่ข้าพเจ้าตลอดระยะเวลาการศึกษา ขอขอบพระคุณ รองศาสตราจารย์ ดร. วิวัฒน์ วัฒนาวุฒิ ประธานกรรมการสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.อาทิตย์ ทองทักษ์ และรองศาสตราจารย์ ดร.สมชาย ปราการเจริญ กรรมการสอบวิทยานิพนธ์ ที่กรุณาให้คำแนะนำต่าง ๆ ทำให้วิทยานิพนธ์นี้มีความถูกต้องและสมบูรณ์มากขึ้น

ขอขอบพระคุณ คุณพ่อ คุณแม่ ที่ให้ความรัก ความห่วงใย คอยให้กำลังใจ พร้อมทั้งให้ความช่วยเหลือ และสนับสนุนในด้านค่าใช้จ่ายในการศึกษาจนสำเร็จลุล่วงได้

ขอขอบคุณพี่ ๆ และเพื่อน ๆ ทุกคน สำหรับคำปรึกษาที่ดีในทุก ๆ ด้าน รวมทั้งกำลังใจและความช่วยเหลืออื่น ๆ ที่มอบให้มาโดยตลอด

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฎ
สารบัญภาพ ..... ฏ
บทที่ 1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของการวิจัย 2
1.3 ขอบเขตวิทยานิพนธ์ 2
1.4 ประโยชน์ที่คาดว่าจะได้รับ 2
1.5 โครงสร้างเนื้อหาวิทยานิพนธ์ 2
1.6 ผลงานตีพิมพ์ 3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4
2.1 แนวคิดและทฤษฎี 4
2.1.1 เอกสารกำหนดความต้องการของซอฟต์แวร์ (Software Requirements Specifications) ตามมาตรฐานไอทริปเปิลอีแปดสามศูนย์ 4
2.1.2 การประมวลผลเอกสาร (Document Preprocessing) 5
2.1.3 การหาค่าน้ำหนักคำ (Term Weighting) ตามทฤษฎีทีเอฟไอดีเอฟ 8
2.1.4 เวกเตอร์นอร์ม (Vector Norm) 9
2.1.5 การวัดความคล้ายคลึงเชิงมุมโคไซน์ (Cosine Similarity Measures) 10
2.1.6 กฎความสัมพันธ์เชิงน้ำหนัก (Weighted Association Rule) 11
2.1.7 ส่วนผลป้อนกลับของผู้ใช้งาน (User Feedback) 12
2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง 12
2.2.1 การสืบค้นความต้องการของซอฟต์แวร์จากความต้องการเชิงหน้าที่ของโครงการในอดีต 12

หน้า

2.2.2 การสืบค้นความต้องการของซอฟต์แวร์ด้วยความหมายออนโทโลจี (Semantic Ontology) และการตอบรับจากผู้ใช้งาน (User Feedback) 13
บทที่ 3 การพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต 15
3.1 การเตรียมความพร้อมระบบ 16
3.1.1 การเตรียมความพร้อมข้อมูล 16
3.1.1.1 การแปลงเอกสารกำหนดความต้องการซอฟต์แวร์ในอดีตเป็นรูปแบบไฟล์เอกซ์เอ็มแอล 16
3.1.1.2 การตัดคำจากความต้องการเชิงหน้าที่ในอดีต 16
3.1.1.3 การกำจัดคำหยุด 17
3.1.1.4 การแปลงคำพ้องความหมายเป็นรูปคำเดียวกัน 17
3.1.2 การคำนวณหาค่าน้ำหนัก 17
3.1.2.1 การนับคำ 17
3.1.2.2 การคำนวณหาค่าน้ำหนักของคำ 20
3.1.2.3 การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่ 23
3.1.3 การสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต 25
3.1.3.1 การหาความต้องการเชิงหน้าที่ที่คล้ายคลึงกัน (Similarity) 25
3.1.3.2 การคำนวณหาค่าสนับสนุนเชิงน้ำหนัก (W-Support) 26
3.1.3.3 การคำนวณหาค่าความเชื่อมั่นเชิงน้ำหนัก (W-Confidence) 27
3.2 การสืบค้นด้วยโครงสร้างความสัมพันธ์ 28
3.3 ผลป้อนกลับของผู้ใช้งาน 29
บทที่ 4 การพัฒนาเครื่องมือ 31
4.1 การออกแบบหน้าที่การทำงานของระบบ 31
4.2 การออกแบบส่วนต่อประสานของเครื่องมือ 36
4.2.1 ส่วนการเตรียมความพร้อมของระบบ 36
4.2.2 ส่วนการสืบค้น 37
4.2.3 ส่วนผลป้อนกลับของผู้ใช้งาน 39
4.3 สภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือ 39
4.3.1 ฮาร์ดแวร์ 39

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

หน้า

4.3.2 ซอฟต์แวร์ 40
4.3.3 ฐานข้อมูล 40
4.4 การทดสอบการทำงานของเครื่องมือ 41
บทที่ 5 การประเมินผลลัพธ์ที่ได้จากวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต 42
5.1 ข้อมูลของการทำทดสอบ 44
5.2 วิธีการสืบค้นความต้องการเชิงหน้าที่ 44
5.3 วิธีประเมินผลการทดลอง 44
5.3.1 ค่าระลึก 44
5.3.2 ค่าความแม่นยำ 45
5.3.3 ค่าอัตราการรู้จำ 45
5.3.4 Precision Histogram 46
5.3.5 ผลลัพธ์ประเมินผลจากค่าระลึก ค่าความแม่นยำ และค่าอัตรารู้จำ 46
5.3.6 ผลลัพธ์ประเมินผลการทดลองด้วยการเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำ 48
5.3.7 ผลลัพธ์ประเมินผลการทดลองด้วย Precision Histogram 51
บทที่ 6 สรุปผลการวิจัย 52
6.1 สรุปผลการวิจัย 52
6.2 อภิปรายผลการวิจัย 54
6.3 ปัญหาและข้อจำกัดของงานวิจัย 54
6.4 แนวทางการวิจัยในอนาคต 55
รายการอ้างอิง 56
ภาคผนวก ก 59
คำอภิธานศัพท์ 59
ภาคผนวก ข 60
รายละเอียดโครงสร้างตารางที่ทำการจัดเก็บในฐานข้อมูล 60
ภาคผนวก ค 66
ข้อคำถามสำหรับทำการทดสอบ 66

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

หน้า

ภาคผนวก ง 68
ผลลัพธ์การเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำแต่ละข้อคำถาม 68
ประวัติผู้เขียนวิทยานิพนธ์ 84

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตารางที่ 2.1 ผลลัพธ์การตัดคำด้วยวิธีเทียบคำที่ยาวที่สุด 7
ตารางที่ 2.2 ผลลัพธ์การคำนวณหาค่าความถี่ของคำ ค่าความถี่ส่วนกลับของเอกสาร และค่าน้ำหนักของคำ 9
ตารางที่ 3.1 ตัวอย่างผลลัพธ์การนับคำ แสดงเฉพาะ 9 คำ..... 18
ตารางที่ 3.2 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของคำ แสดงเฉพาะ 9 คำ 21
ตารางที่ 3.3 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่..... 23
ตารางที่ 3.4 ตัวอย่างผลลัพธ์การคำนวณหาค่าสนับสนุนเชิงน้ำหนัก 27
ตารางที่ 3.5 ตัวอย่างผลลัพธ์การคำนวณหาค่าความเชื่อมั่นเชิงน้ำหนัก..... 28
ตารางที่ 3.6 ความหมายของการให้คะแนน สำหรับผลป้อนกลับผู้ใช้งาน 30
ตารางที่ 4.1 คำอธิบายแผนภาพยูสเคสส่วนเพิ่มเอกสารกำหนดความต้องการของซอฟต์แวร์เข้าฐานข้อมูล..... 32
ตารางที่ 4.2 คำอธิบายแผนภาพยูสเคสส่วนการสืบค้นด้วยความต้องการของผู้ใช้งาน 32
ตารางที่ 4.3 คำอธิบายแผนภาพยูสเคสส่วนผลป้อนกลับของผู้ใช้งาน..... 33
ตารางที่ 4.4 คำอธิบายแผนภาพยูสเคสส่วนคำนวณหาค่าน้ำหนัก 33
ตารางที่ 4.5 คำอธิบายแผนภาพยูสเคสส่วนวิเคราะห์โครงสร้างความสัมพันธ์ 34
ตารางที่ 4.6 คำอธิบายแผนภาพยูสเคสส่วนวิเคราะห์ผลการป้อนกลับ 34
ตารางที่ 4.7 ตัวอย่างกรณีทดสอบการสืบค้นความต้องการเชิงหน้าที่ด้วยคำสำคัญ 41
ตารางที่ 5.1 ผลลัพธ์การประเมินจากค่าระลึก ค่าความแม่นยำ และค่าอัตราการรู้จำ..... 46
ตารางที่ 5.2 ผลลัพธ์ค่าเฉลี่ยการเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำ..... 49

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญภาพ

หน้า

ภาพที่ 2.1 ขั้นตอนการทำงานของการประมวลผลเอกสาร 5
ภาพที่ 2.2 การสืบค้นและเก็บข้อมูลของความต้องการเชิงหน้าที่ [19] 13
ภาพที่ 2.3 การสืบค้นความต้องการด้วยความหมายออนโทโลจี [20] 14
ภาพที่ 3.1 ภาพรวมงานวิจัย 15
ภาพที่ 3.2 โครงสร้างความสัมพันธ์ 29
ภาพที่ 3.3 การป้อนกลับการสืบค้น 30
ภาพที่ 4.1 แผนภาพยูสเคสของระบบสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต 31
ภาพที่ 4.2 แผนภาพกิจกรรมของระบบสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต 35
ภาพที่ 4.3 หน้าจอหลักของเครื่องมือ 36
ภาพที่ 4.4 หน้าจอของส่วนการเตรียมความพร้อมของระบบ 37
ภาพที่ 4.5 หน้าจอของส่วนการสืบค้นด้วยความต้องการของผู้ใช้งาน 38
ภาพที่ 4.6 หน้าจอของส่วนการแสดงผลของการสืบค้น 38
ภาพที่ 4.7 หน้าจอของส่วนผลป้อนกลับของผู้ใช้งาน 39
ภาพที่ 5.1 ภาพรวมการประเมินผลลัพธ์ 43
ภาพที่ 5.2 กราฟแสดงความสัมพันธ์ระหว่างค่าระลึกกับค่าความแม่นยำ 50
ภาพที่ 5.3 กราฟเปรียบเทียบประสิทธิภาพระหว่างทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก และการวัดความคล้ายคลึงเชิงมุมโคไซน์ 51

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

กระบวนการความต้องการในวิศวกรรมซอฟต์แวร์ ถือเป็นกระบวนการหนึ่งที่สำคัญ ซึ่งเป็นกระบวนของการพัฒนาซอฟต์แวร์ ซึ่งหากกระบวนนี้ได้ข้อมูล หากข้อมูลที่ได้จากกระบวนการนี้ไม่ถูกต้องหรือล่าช้า ก็จะทำส่งผลให้กระบวนการต่อ ๆ ไปมีปัญหาตามมาได้อีกด้วย [1]

ปัจจุบันโครงการซอฟต์แวร์หลายโครงการประสบปัญหาไม่สามารถพัฒนาซอฟต์แวร์ได้อย่างมีประสิทธิภาพจนถึงขั้นโครงการล้มเหลว ส่วนหนึ่งมาจากการวิเคราะห์ความต้องการของซอฟต์แวร์ไม่ครบถ้วนหรือถูกต้องตรงตามความต้องการของผู้ใช้งานหรือธุรกิจอย่างแท้จริง ซึ่งถือเป็นปัญหาสำคัญที่มีมานานของโครงการซอฟต์แวร์ [2]

การวิเคราะห์ความต้องการของซอฟต์แวร์เป็นส่วนหนึ่งของขั้นตอนการดำเนินการของโครงการซอฟต์แวร์ ซึ่งนำความต้องการของผู้ใช้งานมาแปลงเป็นความต้องการของซอฟต์แวร์อยู่ในรูปแบบของเอกสารข้อกำหนดความต้องการของซอฟต์แวร์ แต่ในความเป็นจริงนักวิเคราะห์ระบบไม่สามารถรวบรวมความต้องการของผู้ใช้งานได้อย่างถูกต้อง 100% อันเนื่องมาจากหลายปัจจัย เช่น ผู้ใช้งานไม่สามารถอธิบายขั้นตอนการทำงานได้ถูกต้องครบถ้วน หรือนักวิเคราะห์ระบบมีประสบการณ์น้อย เป็นต้น อย่างไรก็ตามการกำหนดความต้องการของซอฟต์แวร์ถือเป็นเรื่องเร่งด่วนสำหรับการพัฒนาโครงการซอฟต์แวร์ ถ้าสามารถได้เอกสารข้อกำหนดความต้องการของซอฟต์แวร์ภายในเวลาที่รวดเร็วขึ้น ก็สามารถเริ่มพัฒนาโครงการได้เร็วขึ้นด้วย จะเห็นได้ว่าการสร้างเอกสารข้อกำหนดความต้องการของซอฟต์แวร์ได้อย่างถูกต้องและรวดเร็ว ถือเป็นสิ่งสำคัญอย่างหนึ่งในโครงการซอฟต์แวร์

จากงานวิจัยนี้ [3] จึงมีแนวคิดในการนำข้อมูลของโครงการซอฟต์แวร์ในอดีตที่ประสบความสำเร็จมาใช้ให้เกิดประโยชน์ โดยนำความต้องการเชิงหน้าที่มาหาจุดเชื่อมโยงไปความต้องการเชิงหน้าที่อื่น เพื่อสืบค้นความสัมพันธ์ของข้อมูล โดยนำทฤษฎีของกฎความสัมพันธ์เชิงน้ำหนักมาช่วยในการสืบค้นความสัมพันธ์ของคำหรือกลุ่มคำ เพื่อเป็นการขยายข้อคำถามทำให้ได้ความต้องการเชิงหน้าที่ที่มีความสัมพันธ์กับข้อคำถามถึงแม้ว่าข้อคำถามจะไม่มีคำเหล่านั้นปรากฏอยู่ในความต้องการเชิงหน้าที่ที่ค้นพบก็ตาม แต่เป็นความต้องการที่เกี่ยวข้องกัน ซึ่งช่วยทำให้วิศวกรความต้องการค้นพบความต้องการของซอฟต์แวร์ได้มากขึ้นได้ นอกจากนี้ยังทำให้การรวบรวมความต้องการของซอฟต์แวร์เป็นไปอย่างสมบูรณ์และรวดเร็วมากขึ้น ส่งผลให้การเพิ่มหรือเปลี่ยนความต้องการของซอฟต์แวร์ระหว่างการพัฒนาลดน้อยลงได้

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 นำเสนอวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต

1.2.2 นำเอกสารกำหนดความต้องการของซอฟต์แวร์ของโครงการในอดีตมาช่วยสำหรับการรวบรวมความต้องการของซอฟต์แวร์อย่างรวดเร็วและตรงตามความต้องการของผู้ใช้งาน

1.3 ขอบเขตวิทยานิพนธ์

1.3.1 ใช้เอกสารข้อกำหนดความต้องการของซอฟต์แวร์ตามมาตรฐาน ไอทริปเปิลอีแปดสามศูนย์ (IEEE 830) ที่เป็นภาษาไทย ในโครงการซอฟต์แวร์ในอดีต

1.3.2 ใช้ข้อมูลบางส่วนในเอกสารข้อกำหนดความต้องการของซอฟต์แวร์ตามมาตรฐานไอทริปเปิลอีแปดสามศูนย์ ดังนี้

1. เป้าหมายของโครงการ
2. ขอบเขตของโครงการ
3. ความต้องการเชิงหน้าที่
4. ความต้องการไม่ใช่เชิงหน้าที่

1.3.3 ใช้ความต้องการของผู้ใช้งานเป็นข้อคำถามในการสืบค้น

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 เครื่องมือการทดลองสำหรับใช้สืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทยโดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต

1.5.2 นักวิเคราะห์ระบบได้รับเอกสารกำหนดความต้องการซอฟต์แวร์เบื้องต้นที่นำไปใช้เป็นแนวทาง สำหรับวิเคราะห์ร่วมกับผู้ใช้งานอย่างรวดเร็วและตรงตามความต้องการของผู้ใช้งานยิ่งขึ้น ซึ่งมีโอกาสลดระยะเวลาการรวบรวมความต้องการของซอฟต์แวร์ และการดำเนินโครงการได้

1.5 โครงสร้างเนื้อหาวิทยานิพนธ์

เนื้อหาในวิทยานิพนธ์ฉบับนี้ แบ่งออกเป็น 5 บท ดังนี้ คือ บทที่ 1 บทนำ กล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ และขอบเขตของการวิจัย วิธีดำเนินการวิจัยโดยย่อ และประโยชน์ที่คาดว่าจะได้รับจากการวิจัย รวมถึงผลงานตีพิมพ์ทางวิชาการ บทที่ 2 ในส่วนนี้

มีเนื้อหาเกี่ยวกับทฤษฎีและงานวิจัยที่เกี่ยวข้อง เพื่อชี้ให้เห็นถึงแนวคิดและมุมมองการเกิดงานวิจัยชิ้นนี้ บทที่ 3 อธิบายการวิจัยโดยละเอียด ว่ามีขั้นตอนการทำอย่างไร แต่ละขั้นตอนใช้เอกสารและข้อมูลประเภทใด ส่วนบทที่ 4 อธิบายถึงขั้นตอนการพัฒนาเครื่องมือสนับสนุนตัวแบบที่พัฒนาขึ้น บทที่ 5 กล่าวถึงการวิเคราะห์ รายงานผลการทดลอง และบทที่ 6 สรุปผลงานวิจัย อภิปรายผลการวิจัย ปัญหาและข้อจำกัด

1.6 ผลงานตีพิมพ์

การประชุมวิชาการ International Conference on Manufacturing Science and Information Engineering (ICMSIE 2013), 28 – 29 September 2013, Jinrong International Hotel, Shanghai, China ในบทความเรื่อง Framework for Fast Building Software Requirements Specification โดยผู้แต่งคือ Ponglikit Singchai และ Wanchai Rivepiboon.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

2.1.1 เอกสารกำหนดความต้องการของซอฟต์แวร์ (Software Requirements Specifications) ตามมาตรฐานไอทริปเปิลอีแปดสามศูนย์

เอกสารกำหนดความต้องการของซอฟต์แวร์เป็นเอกสารที่ใช้กำหนดข้อตกลงในการพัฒนาซอฟต์แวร์ระหว่างผู้ใช้งานและผู้พัฒนาซอฟต์แวร์ให้สามารถเข้าใจตรงกัน

ในงานวิจัยนี้จะนำเอกสารกำหนดความต้องการของซอฟต์แวร์ตามมาตรฐานไอทริปเปิลอีแปดสามศูนย์ [4] ที่ถูกรวบรวมจากโครงการซอฟต์แวร์ในอดีต นำมาจัดเก็บเป็นฐานข้อมูลในการสืบค้นความต้องการ และเป็นรูปแบบสำหรับแปลงเป็นเอกสารกำหนดความต้องการของซอฟต์แวร์เพื่อช่วยให้นักวิเคราะห์ระบบ สามารถได้เอกสารกำหนดความต้องการของซอฟต์แวร์เบื้องต้น ทำให้ลดเวลาในการรวบรวมความต้องการได้เร็วขึ้น และช่วยนักวิเคราะห์ระบบให้ได้ความต้องการที่ใกล้เคียงกับผู้ใช้งานโดยใช้ข้อมูลโครงการซอฟต์แวร์ในอดีต ซึ่งงานวิจัยนี้จะนำเฉพาะบางส่วนในเอกสารกำหนดความต้องการของซอฟต์แวร์มาใช้งาน ดังนี้

1. เป้าหมายของโครงการ (Purpose of Project) เป็นสิ่งที่ผู้ใช้งานคาดหวังให้ทางผู้พัฒนาทำให้สำเร็จ ดังนั้นนักวิเคราะห์ระบบจึงต้องเข้าใจกับเป้าหมายในการทำงาน ว่าต้องทำให้สิ่งใดบ้าง เพื่อให้สำเร็จตรงตามเป้าหมายนี้

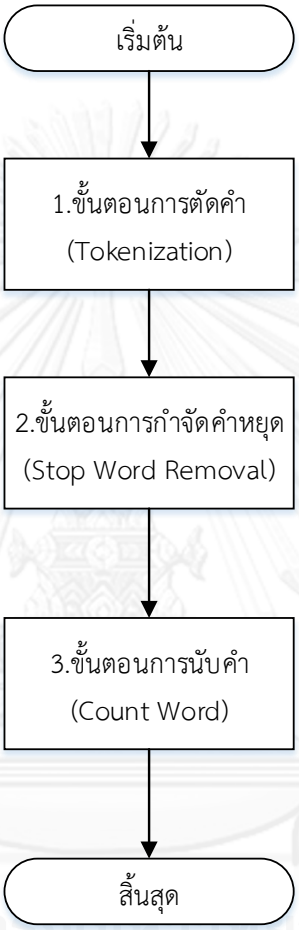
2. ขอบเขตของโครงการ (Scope of Project) เป็นสิ่งที่ผู้พัฒนากำหนดขอบเขตการพัฒนาโครงการ ซึ่งเป็นสิ่งที่ผู้พัฒนาและผู้ใช้งานตกลงร่วมกัน

3. ความต้องการเชิงหน้าที่ (Functional Requirements) เป็นความต้องการของผู้ใช้งานที่สามารถนำมาสร้างฟังก์ชันการทำงานต่าง ๆ ของซอฟต์แวร์ได้ ซึ่งถือเป็นหน้าที่หลักที่ซอฟต์แวร์ต้องคำนึงถึงเป็นอย่างแรกในการพัฒนาซอฟต์แวร์

4. ความต้องการไม่ใช่เชิงหน้าที่ (Non-Functional Requirements or Performance Requirements) ความต้องการของผู้ใช้งานที่ไม่สามารถนำมาสร้างฟังก์ชันการทำงานต่าง ๆ ของซอฟต์แวร์ได้ชัดเจน แต่เป็นความต้องการที่จะทำให้ระบบมีประสิทธิภาพสามารถทำงานได้ตรงตามการทำงานของผู้ใช้งานมากขึ้น เช่น ระบบต้องดูมีความน่าเชื่อถือ ต้องตอบสนองภายใน 10 นาที เป็นต้น

2.1.2 การประมวลผลเอกสาร (Document Preprocessing)

การประมวลผลเอกสารเป็นการสกัดคำเพื่อให้ได้คำสำคัญที่เป็นดัชนีคำศัพท์ของเอกสารได้ รวมถึงนำคำหยุดออกไป พร้อมทั้งนับความถี่ของคำในเอกสาร และจัดเก็บข้อมูล [5] ซึ่งมีขั้นตอนการทำงานดังนี้



ภาพที่ 2.1 ขั้นตอนการทำงานของการประมวลผลเอกสาร

จากภาพที่ 2.1 เป็นขั้นตอนการทำงานของการประมวลผลเอกสาร เริ่มจากการนำเอาประโยคต่าง ๆ ในเอกสาร มาทำงานการตัดแยกคำที่คาดว่าจะมีความหมายออกมา ซึ่งคำเหล่านี้จะเป็นคำสำคัญหรือดัชนีคำศัพท์ของเอกสาร เพื่อให้แน่ใจว่าคำที่ได้มาเป็นคำที่ไม่มีความหมายหรือไม่มีผลต่อเอกสาร ซึ่งไม่ใช่คำสำคัญหรือดัชนีคำศัพท์ของเอกสาร ดังนั้นจึงต้องกำจัดคำหยุดหรือคำที่ไม่มีความหมายออกจากเอกสารด้วย หลังจากนั้นจึงทำการนับคำ เพื่อทราบความถี่หรือน้ำหนักของคำในเอกสาร

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1. ขั้นตอนการตัดคำ (Tokenization)

ขั้นตอนการตัดคำเป็นการอ่านประโยคต่าง ๆ ในเอกสารในรูปแบบไฟล์ข้อความ โดยจะทำการแยกข้อความออกให้เป็นคำ ซึ่งเป็นหน่วยย่อยที่มีความหมายได้ [6] เพื่อให้สามารถได้คำสำคัญหรือดัชนีคำศัพท์ของเอกสารต่อไปได้ ซึ่งจะได้คำที่มีใจความสำคัญของเนื้อหาจริง ๆ วิธีการตัดคำของแต่ละภาษาก็มีวิธีการและเทคนิคที่แตกต่างกันไป เนื่องจากรูปแบบของแต่ละภาษาไม่เหมือนกัน เช่น ภาษาจะมีช่องว่างคำเสมอ แต่ภาษาไทยส่วนใหญ่ไม่มีช่องว่างคำ เป็นต้น นอกจากนั้นในแต่ละวิธีการมีข้อดี และข้อเสียที่แตกต่างกันไป

ในงานวิจัยนี้จะใช้เอกสารกำหนดความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย ซึ่งการตัดคำภาษาไทยจะทำได้ยากกว่าภาษาอังกฤษเนื่องจากภาษาไทยไม่ได้มีช่องว่างคำเหมือนในภาษาอังกฤษที่ง่ายต่อการตัดคำได้ถูกต้อง นอกจากนั้นภาษาไทยยังมีปัญหาด้านคำกำกวม เช่น ตากลม สามารถตัดคำที่มีความหมายได้ 2 แบบ คือ ตา กลม และ ตาก ลม ซึ่งทำให้การตัดคำผิดไม่ตรงตามความหมายที่ต้องการสื่อได้ รวมถึงคำบางคำสามารถมีได้หลายความหมาย

วิธีการตัดคำภาษาไทยมีหลายวิธี ในแต่ละวิธีการมีข้อดี และข้อเสียที่แตกต่างกัน จากงานวิจัยนี้ [7] พบว่าวิธีการเทียบคำที่ยาวที่สุด (Longest Matching) จากคำในพจนานุกรมที่รวบรวมมา เป็นวิธีการตัดคำที่ได้รับความนิยมและมีความแม่นยำที่ยอมรับได้

วิธีการเทียบคำที่ยาวที่สุดเป็นวิธีการตัดคำทางวิทยาการศึกษาสำนึก (Heuristic) วิธีหนึ่งซึ่งต้องใช้พจนานุกรมช่วยในการรู้จักคำภาษาไทย โดยจะทำการตรวจสอบหรือสแกนข้อความในเอกสารจากซ้ายไปขวา นำไปเทียบกับพจนานุกรมดูว่าเป็นหนึ่งคำหรือไม่ หากเทียบแล้วไม่พบคำในพจนานุกรม จะทำการลดความยาวลงทีละตัว จนกว่าจะสามารถเทียบเป็นคำในพจนานุกรมได้ ก็จะทำเครื่องหมายเพื่อเป็นจุดย้อนกลับ จากนั้นจะเริ่มทำงานจากจุดย้อนกลับนั้นเพื่อตรวจสอบส่วนที่เหลือว่าจะสามารถตัดให้เป็นคำได้หรือไม่ หากตัวเลือกในตอนแรกนี้สามารถสืบค้นคำที่เหลือได้ตัวเลือกนี้ก็จะเป็นคำแรกของข้อความได้จริง ไม่เช่นนั้นก็จะกลับไปยังจุดย้อนกลับที่ทำเครื่องหมายไว้เพื่อแก้ไขคำแรกใหม่ จากนั้นจะเริ่มทำงานต่อไปโดยเริ่มจากจุดย้อนกลับ หากยังไม่สามารถเทียบกับคำในพจนานุกรมได้ก็จะทำการลดตัวอักษรลงทีละตัวจนกว่าจะเทียบคำในพจนานุกรมได้ และทำงานในรูปแบบนี้ไปจนจบข้อความ [8][9] ตัวอย่างเช่น ข้อความ “ความต้องการซอฟต์แวร์ของโครงการในอดีต” เมื่อไม่สามารถเทียบคำกับพจนานุกรมได้ก็จะลดเหลือ “ความต้องการซอฟต์แวร์ของโครงการในอดี” จนได้คำที่มีในพจนานุกรม “ความต้องการ” ซึ่งสามารถเทียบคำในพจนานุกรมได้จึงตัดเป็นคำแรกของข้อความ และทำเครื่องหมายไว้เป็นจุดย้อนกลับ ผลการตัดคำทั้งหมดจะเป็นดังตารางที่ 2.1

ตารางที่ 2.1 ผลลัพธ์การตัดคำด้วยวิธีเทียบคำที่ยาวที่สุด

ส่วนของคำที่ยาวที่สุดที่ตัดได้	ส่วนหลังที่เหลือจากจุดย้อนกลับ
ควมต้องการ	ซอฟต์แวร์ของโครงการในอดีต
ซอฟต์แวร์	ของโครงการในอดีต
ของ	โครงการในอดีต
โครงการ	ในอดีต
ใน	อดีต
อดีต	

โดยในงานวิจัยนี้จะนำประโยคต่างๆในเอกสารกำหนดความต้องการซอฟต์แวร์นำมาตัดให้เป็นคำสำคัญที่มีความหมาย

2. ขั้นตอนการกำจัดคำหยุด (Stop Word Removal)

การกำจัดคำหยุดเป็นการนำเอาคำที่ไม่มีความสำคัญหรือคำที่ไม่มีความหมายหรือคำทั่วไปที่พบได้บ่อย ๆ ออกไป รวมถึงคำที่เป็นสัญลักษณ์ต่าง ๆ ให้เหลือเฉพาะคำที่อธิบายถึงความหมายที่แท้จริงหรือคำสำคัญเท่านั้น ซึ่งส่วนใหญ่มักจะเป็นคำกริยาและคำนาม เพื่อให้ลดเวลาที่ไปสืบค้นคำที่มีประโยชน์ไป ให้สืบค้นเฉพาะคำที่ควรเป็นดัชนีคำศัพท์ของเอกสารจริง ๆ [10]

ในงานวิจัยนี้จะใช้รายการที่เป็นคำหยุดที่ได้รวบรวมไว้ เช่น ก็ กับ และได้ เป็นต้น นำเข้าฐานข้อมูล เพื่อนำมากำจัดคำที่ไม่มีความหมายเหล่านี้ให้เหลือแต่คำสำคัญหรือดัชนีคำศัพท์ของเอกสารเท่านั้น ทำให้ลดคำที่ไม่สำคัญหรือไม่ใช่ดัชนีคำศัพท์ของเอกสารที่แท้จริง และใช้เวลาได้น้อยลง

Chulalongkorn University

3. ขั้นตอนการนับคำ (Count Word)

การนับคำเป็นการนับความถี่ของแต่ละคำสำคัญหรือดัชนีคำศัพท์ของเอกสาร เพื่อให้ทราบว่าในเอกสารมีคำสำคัญเหล่านั้นปรากฏอยู่มากน้อยเพียงใด และจำนวนเอกสารที่คำสำคัญเหล่านั้นปรากฏอยู่ เพื่อใช้สำหรับคำนวณค่าน้ำหนักของคำในเอกสารต่อไปได้

ในงานวิจัยนี้จะนับจำนวนคำสำคัญหรือดัชนีคำศัพท์ของเอกสาร สำหรับคำนวณค่าน้ำหนักของคำ และสร้างโครงสร้างความสัมพันธ์ต่อไป

2.1.3 การหาค่าน้ำหนักคำ (Term Weighting) ตามทฤษฎีทีเอฟไอดีเอฟ

การหาค่าน้ำหนักคำเป็นการนำค่าที่ได้จากการนับคำ คือ จำนวนคำสำคัญหรือดัชนีคำศัพท์ของเอกสารที่มีอยู่แต่ละเอกสาร และจำนวนเอกสารที่คำเหล่านั้นปรากฏอยู่ ซึ่งนำค่าเหล่านี้มาคำนวณหาค่าความถี่ของคำสำคัญหรือดัชนีคำศัพท์ของเอกสารว่าปรากฏในเอกสารมากน้อยเพียงใด และค่าความถี่ส่วนกลับของเอกสารว่าคำเหล่านั้นปรากฏในเอกสารต่าง ๆ มากน้อยเพียงใด เพื่อให้ได้ค่าน้ำหนักคำตามทฤษฎีทีเอฟไอดีเอฟ [11] มีสูตรการคำนวณดังนี้

1. การหาค่าความถี่ของคำ (Term Frequency – TF)

การหาค่าความถี่ของคำเป็นการนำจำนวนคำสำคัญหรือดัชนีคำศัพท์ของเอกสารที่ปรากฏในเอกสารหรือกลุ่มคำหรือประโยคว่ามีความถี่ของแต่ละคำเท่าไร ซึ่งหากคำเหล่านั้นปรากฏในเอกสารนั้นเยอะ แสดงว่าคำเหล่านั้นน่าจะมีความสำคัญกับเอกสารนั้น ก็จะทำให้มีค่าความถี่ของคำมาก ส่งผลให้ค่าน้ำหนักคำมากด้วย มีสูตรการคำนวณดังสมการที่ (2.1)

$$tf_{i,j} = \begin{cases} 1+\log_2 f_{i,j} & ; f_{i,j} > 0 \\ 0 & ; otherwise \end{cases} \quad (2.1)$$

$f_{i,j}$ คือ จำนวนคำที่ปรากฏ

2. การหาค่าความถี่ส่วนกลับของเอกสาร (Inverse Document Frequency - IDF)

การหาค่าความถี่ส่วนกลับของเอกสารเป็นการนำเอาจำนวนเอกสารที่คำเหล่านั้นไปปรากฏมาคำนวณเพื่อหาค่าความถี่ส่วนกลับของเอกสาร ซึ่งหากคำเหล่านั้นปรากฏในหลายเอกสารแสดงว่าคำเหล่านั้นไม่น่าจะเป็นคำสำคัญหรือดัชนีคำศัพท์ของเอกสารได้ อาจจะเป็นแค่คำทั่วไปที่มักต้องใช้กันบ่อย ๆ ไม่มีความหมาย ก็จะทำให้มีค่าความถี่ส่วนกลับของเอกสารน้อย ส่งผลให้ค่าน้ำหนักของคำน้อยด้วย มีสูตรการคำนวณดังสมการที่ (2.2)

$$idf_i = \log_2 \frac{N}{n_i} \quad (2.2)$$

N คือ จำนวนเอกสารทั้งหมด

n_i คือ จำนวนเอกสารที่มีคำนั้นปรากฏ

หลังจากได้รับค่าความถี่ของแต่ละคำและค่าความถี่ส่วนกลับของเอกสาร แล้วนำทั้ง 2 ค่ามาคูณกัน เพื่อได้รับค่าน้ำหนักคำเหล่านั้นในเอกสาร มีสูตรการคำนวณค่าน้ำหนักคำตามทฤษฎีทีเอฟไอดีเอฟ ดังสมการที่ (2.3)

$$w_{i,j} = \begin{cases} tf_{i,j} \times idf_i & ; f_{i,j} > 0 \\ 0 & ; otherwise \end{cases} \quad (2.3)$$

ตัวอย่างสมมุติว่ามี 2 เอกสาร ดังนี้

เอกสารที่ 1: ความต้องการของซอฟต์แวร์ของโครงการในอดีต

เอกสารที่ 2: การสืบค้นความต้องการของซอฟต์แวร์

หลังจากประมวลผลเอกสาร โดยการตัดคำ กำจัดคำหยุด และนับคำ ทั้งสองเอกสารแล้วนำมาคำนวณหาค่าความถี่ของคำ และค่าความถี่ส่วนกลับของเอกสารจะได้ค่าน้ำหนักของคำ ผลลัพธ์จากตัวอย่างเป็นดังตาราง 2.2

ตารางที่ 2.2 ผลลัพธ์การคำนวณหาค่าความถี่ของคำ ค่าความถี่ส่วนกลับของเอกสาร และค่าน้ำหนักของคำ

Term	TF (D1)	TF (D2)	IDF	W (D1)	W (D2)
ความต้องการ	1	1	1	1	1
ซอฟต์แวร์	1	1	1	1	1
โครงการ	1	0	2	2	0
สืบค้น	0	1	2	0	2

ในงานวิจัยนี้คำนวณหาค่าน้ำหนักคำตามทฤษฎี TF-IDF เพื่อให้ทราบน้ำหนักคำแล้วนำน้ำหนักคำทั้งหมดของแต่ละความต้องการเชิงหน้าที่มารวมคะแนนเพื่อให้ได้ค่าน้ำหนักของความต้องการเชิงหน้าที่ สำหรับนำไปสร้างโครงสร้างความสัมพันธ์ต่อไป

2.1.4 เวกเตอร์นอร์ม (Vector Norm)

เวกเตอร์นอร์มเป็นวิธีการหนึ่งที่ทำวัดความยาวของเอกสารอย่างเป็นมาตรฐาน ซึ่งเป็นเหมือนตัวแทนของเวกเตอร์ของน้ำหนักของคำ [12] ดังสมการที่ (2.4)

$$| \overrightarrow{d_j} | = \sqrt{\sum_{i}^{t} w_{i,j}^2} \quad (2.4)$$

w คือ น้ำหนักของคำ

t คือ คำที่ปรากฏในเอกสาร

จากสมการที่ (2.4) สามารถนำมาคำนวณหาค่าเวกเตอร์นอร์มจากตารางที่ 2.2 ได้ผลลัพธ์ของเอกสารที่ 1 และ 2 เป็น 2.4495

ในงานวิจัยนี้จะนำเอาค่าของน้ำหนักคำที่อยู่ในความต้องการเชิงหน้าที่เดียวกันมาหาค่าเวกเตอร์นอร์ม เพื่อเป็นค่าน้ำหนักของความต้องการเชิงหน้าที่ สำหรับนำไปหาโครงสร้างความสัมพันธ์ต่อไป

2.1.5 การวัดความคล้ายคลึงเชิงมุมโคไซน์ (Cosine Similarity Measures)

เป็นวิธีการวัดความคล้ายคลึงเชิงมุมโคไซน์เป็นวิธีการเปรียบเทียบความคล้ายคลึงของเอกสารสองเอกสาร โดยแต่ละเอกสารจะถูกแทนด้วยเวกเตอร์ขนาดเอ็น (N-Dimensional Vector) ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในเอกสารนั้น (N-Dimensional Vector in Term Space) [12] ดังสมการที่ (2.5)

$$Sim(S_i, S_j) = \sum_{t_i \in S_i} \sum_{t_j \in S_j} \frac{sim(t_i, t_j)}{\sqrt{|S_i \| S_j|}} \qquad (2.5)$$

t_i คือ คำที่ปรากฏอยู่ในประโยค S_i

t_j คือ คำที่ปรากฏอยู่ในประโยค S_j

S_i คือ ประโยคที่ i

S_j คือ ประโยคที่ j

การเปรียบเทียบความคล้ายคลึงของเอกสารจะเปรียบเทียบโดยดูจากมุมโคไซน์ของมุมระหว่าง 2 เวกเตอร์ของเอกสาร หากเอกสารทั้งสองเอกสารคล้ายคลึงกันมาก เวกเตอร์ของเอกสารทั้ง 2 จะทำมุมระหว่างกัน 0 องศา มุมจึงมีค่าน้อย ค่าโคไซน์ที่ได้จะมีค่าเท่ากับ 1 แต่ถ้าเอกสารทั้งสองเอกสารคล้ายคลึงกันน้อย เวกเตอร์ของเอกสารทั้ง 2 จะทำมุมระหว่างกัน 90 องศา มุมจึงมีค่ามาก ค่าโคไซน์ที่ได้จะมีค่าเท่ากับ 0

ในงานวิจัยนี้จะใช้การวัดความคล้ายคลึงเชิงมุมโคไซน์สำหรับเปรียบเทียบกันแต่ละประโยคในกลุ่มของเป้าหมายของโครงการ ขอบเขตของโครงการ ความต้องการของซอฟต์แวร์ ความต้องการเชิงหน้าที่ และความต้องการไม่ใช่เชิงหน้าที่ เพื่อให้ทราบว่าประโยคไหนมีความคล้ายคลึงกันบ้าง เพื่อไม่ให้แสดงข้อมูลซ้ำซ้อนกัน เมื่อทำการแสดงผลลัพธ์จากการค้นหา ซึ่งจะทำการคำนวณหาค่าความคล้ายคลึงกัน ในขั้นตอนเตรียมความพร้อมของระบบ

2.1.6 กฎความสัมพันธ์เชิงน้ำหนัก (Weighted Association Rule)

กฎความสัมพันธ์เชิงน้ำหนักเป็นการนำความสัมพันธ์หรือเชื่อมโยงของข้อมูล 2 ชุดหรือมากกว่า ที่มีรูปแบบคล้ายกันมาสร้างเป็นกฎความสัมพันธ์ เพื่อให้ทราบว่าถ้ามีข้อมูลนี้แล้วจะเกิดข้อมูลนั้นตามมา ซึ่งทำให้สามารถคาดการณ์สิ่งที่จะเกิดล่วงหน้าได้ง่ายขึ้น การสร้างกฎความสัมพันธ์จะทำให้รูปแบบเป็นสากลง่ายต่อการที่คนอื่นจะนำไปใช้ได้ง่าย [13]

กฎความสัมพันธ์เป็นเทคนิคที่สำคัญของ Data Mining และสามารถนำไปประยุกต์ใช้ได้กับงานต่าง ๆ มักใช้ในการสืบค้นความสัมพันธ์ที่แฝงอยู่ของข้อมูลหรือเพื่อหารูปแบบที่เกิดบ่อย (Frequent Pattern) ความเชื่อมโยง (Association) สหสัมพันธ์ (Correlation) หรือโครงสร้างอื่น ๆ ในกลุ่มของไอเท็ม (Item) หรือ อ็อบเจกต์ (Object) ในฐานข้อมูลแบบทรานเซ็คชั่น (Transaction) ฐานข้อมูลเชิงสัมพันธ์และคลังข้อมูลอื่น ๆ โดยผลลัพธ์จะปรากฏอยู่ในรูปพิจารณากฎ A -> B เพื่อนำไปใช้ในการวิเคราะห์หรือทำนายปรากฏการณ์ต่าง ๆ จากข้อมูลในตารางที่รวบรวมไว้ ผลการวิเคราะห์ที่ได้จะเป็นคำตอบของปัญหา ซึ่งการวิเคราะห์แบบนี้เป็นการใช้กฎความสัมพันธ์ (Association Rule) เพื่อหาความสัมพันธ์ของข้อมูลการสืบค้นกฎความสัมพันธ์เป็นการวิเคราะห์ความสัมพันธ์ที่แฝงอยู่ในข้อมูลจากฐานข้อมูล เพื่อบอกว่าไอเท็มตัวหนึ่ง ๆ มีความสัมพันธ์กับไอเท็มตัวใดบ้าง โดยแสดงความสัมพันธ์ที่ได้ในรูปของกฎความสัมพันธ์ [14]

กฎความสัมพันธ์เชิงน้ำหนัก เป็นกฎความสัมพันธ์รูปแบบหนึ่งที่มองความน่าสนใจจากค่าน้ำหนักที่มีเพียงพอที่จะถูกนำมาสร้างเป็นกฎความสัมพันธ์เชิงน้ำหนัก [15]

ค่าสนับสนุนเชิงน้ำหนัก (Weighted Support) คือความน่าจะเป็นที่จะมีการเกิดสิ่งเหล่านั้นในทรานเซคชั่นทั้งหมด [16] ดังสมการที่ (2.6)

$$W_{\sup}(X \rightarrow Y) = (X \cup Y) \tag{2.6}$$

X คือ ข้อมูลตั้งต้น

Y คือ ข้อมูลที่จะเกิดขึ้นภายหลัง

ค่าความเชื่อมั่นเชิงน้ำหนัก (Weighted Confidence) คือความน่าจะเป็นหากเกิดข้อมูล A แล้วจะมีโอกาสเกิดข้อมูล B ตามมา [15] ดังสมการที่ (2.7)

$$W_{conf}(X \rightarrow Y) = \frac{W_{\sup}(X \cup Y)}{W_{\sup}(X)} \tag{2.7}$$

กฎความสัมพันธ์จะถูกพิจารณาในการสมควรที่จะเป็นกฎจากค่าสนับสนุนและค่าความเชื่อมั่น โดยต้องมีค่าสนับสนุนไม่น้อยกว่าค่าสนับสนุนขั้นต่ำ (Minimum Support) และมีค่าความเชื่อมั่นไม่น้อยกว่าค่าความเชื่อมั่นขั้นต่ำ (Minimum Confidence) ที่ผู้ใช้กำหนดไว้ [17]

ในงานวิจัยนี้จะนำทฤษฎีของกฎความสัมพันธ์เชิงน้ำหนัก มาช่วยในการสร้างโครงสร้างความสัมพันธ์ สำหรับการรวบรวมความต้องการของซอฟต์แวร์ด้านทรัพยากรบุคคล เพื่อสามารถเก็บรวบรวมความต้องการของซอฟต์แวร์ได้สมบูรณ์และรวดเร็วมากขึ้น โดยจะนำค่าน้ำหนักของความต้องการเชิงหน้าที่มาใช้ สำหรับเป็นค่าน้ำหนักในการหาค่าสนับสนุนเชิงน้ำหนัก

2.1.7 ส่วนผลป้อนกลับของผู้ใช้งาน (User Feedback)

ส่วนผลป้อนกลับของผู้ใช้งานเป็นการให้ข้อมูลความคิดเห็นจากผู้ใช้งานต่อระบบ [18] เพื่อช่วยในการเป็นข้อเสนอแนะให้กับการปรับปรุงระบบให้ดีขึ้น ซึ่งจะระบุเป็นตัวเลขหรือคะแนนที่ออกมาจากความรู้สึกของผู้ใช้งาน ว่าสิ่งที่ผู้ใช้งานได้รับผลลัพธ์จากการสืบค้นว่าตรงตามความต้องการมากน้อยเพียงใด เพื่อนำกลับไปเป็นค่าสำหรับคำนวณความคล้ายคลึงระหว่างข้อคำถามและข้อมูลให้มีน้ำหนักมากขึ้นหรือน้อยลง ซึ่งจะสามารถทำให้ผลลัพธ์ที่จะได้ในคราวต่อไปตรงตามความต้องการของผู้ใช้งานมากขึ้น

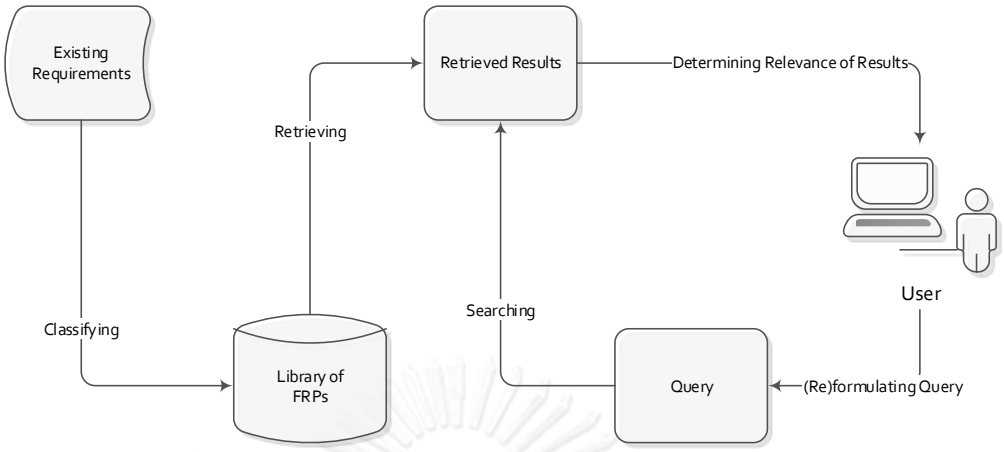
ในงานวิจัยนี้จะใช้ความคิดเห็นของผู้ใช้งาน สำหรับระบุแต่ละกฎที่ถูกสร้างและได้รับความต้องการเชิงหน้าที่ตรงตามความต้องการหรือไม่

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

2.2.1 การสืบค้นความต้องการของซอฟต์แวร์จากความต้องการเชิงหน้าที่ของโครงการในอดีต

ในปี 2010 Anas Mahmoud กล่าวว่าการสืบค้นความต้องการของซอฟต์แวร์จากความต้องการเชิงหน้าที่ของโครงการในอดีต เป็นการนำความต้องการจากโครงการที่ผ่านมาทำการจัดกลุ่มเป็นคลัง (Library) ของความต้องการเชิงหน้าที่ เพื่อเป็นฐานข้อมูล ซึ่งผู้ใช้งานสามารถใส่ข้อคำถามที่ต้องการ นำมาสืบค้นข้อมูลความต้องการเชิงหน้าที่ ดังภาพที่ 2.2 [19]

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



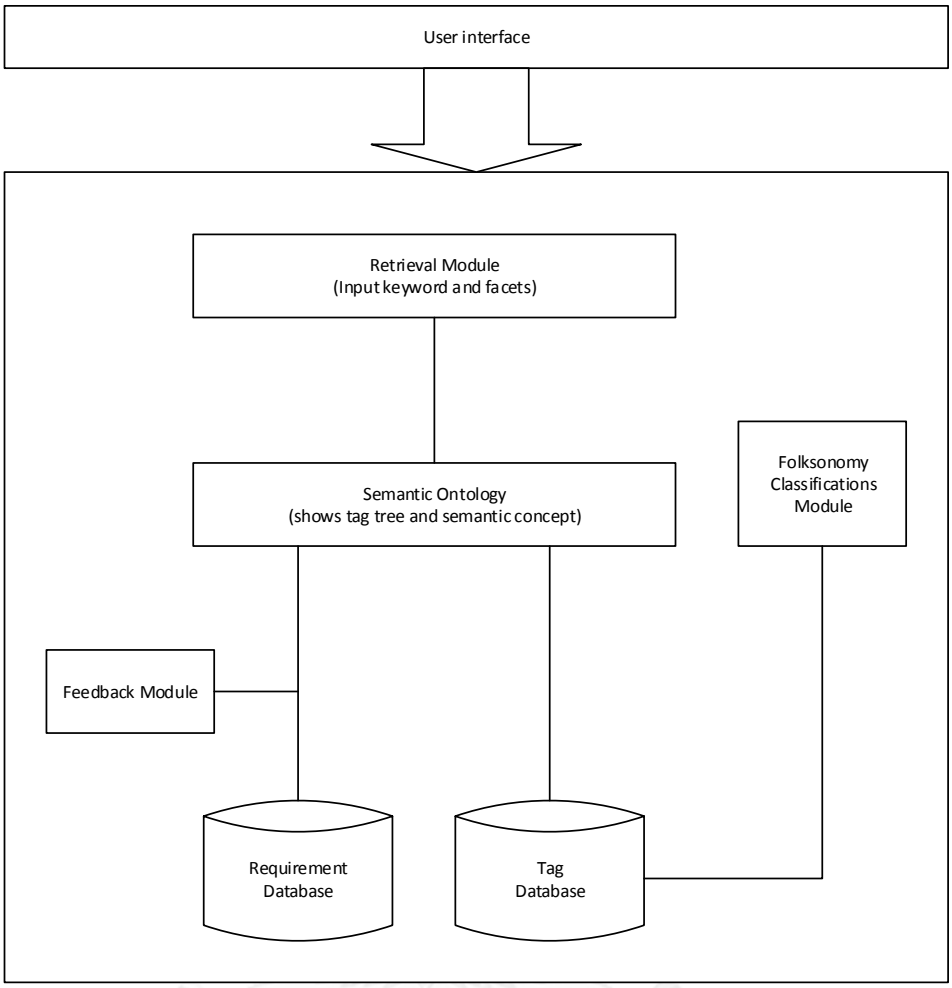
ภาพที่ 2.2 การสืบค้นและเก็บข้อมูลของความต้องการเชิงหน้าที่ [19]

จากภาพที่ 2.2 เมื่อผู้ใช้งานป้อนข้อคำถามแล้ว ข้อคำถามจะสืบค้นความต้องการเชิงหน้าที่ที่มีอยู่ในฐานข้อมูล เพื่อได้ผลลัพธ์จากการสืบค้นสำหรับนำมาให้ผู้ใช้งานตัดสินใจ

จากงานวิจัยนี้จึงเกิดแนวคิดในการใช้ข้อมูลเพิ่มเติมนอกเหนือจากความต้องการเชิงหน้าที่ เช่น เป้าหมายของโครงการ ขอบเขตของโครงการ ความต้องการของซอฟต์แวร์ ความต้องการเชิงหน้าที่ และความต้องการไม่ใช่เชิงหน้าที่ นำมาสร้างความสัมพันธ์ของแต่ละส่วน เพื่อประกอบกันเป็นเอกสารข้อกำหนดความต้องการของซอฟต์แวร์เบื้องต้นได้ จะสามารถทำให้ขั้นตอนการรวบรวมความต้องการสามารถทำได้รวดเร็วมากขึ้น

2.2.2 การสืบค้นความต้องการของซอฟต์แวร์ด้วยความหมายออนโทโลจี (Semantic Ontology) และการตอบรับจากผู้ใช้งาน (User Feedback)

ในปี 2010 Tao Zhang กล่าวว่าการสืบค้นความต้องการของซอฟต์แวร์ด้วยความหมายออนโทโลจีและการตอบรับจากผู้ใช้งานเป็นการสืบค้นที่หาความต้องการในหลายโดเมนและใช้ผลตอบรับของผู้ใช้งานในการปรับน้ำหนักในการสืบค้น เพื่อให้ได้ข้อมูลที่แม่นยำขึ้น ดังภาพที่ 2.3 [20]



ภาพที่ 2.3 การสืบค้นความต้องการด้วยความหมายออนโทโลจี [20]

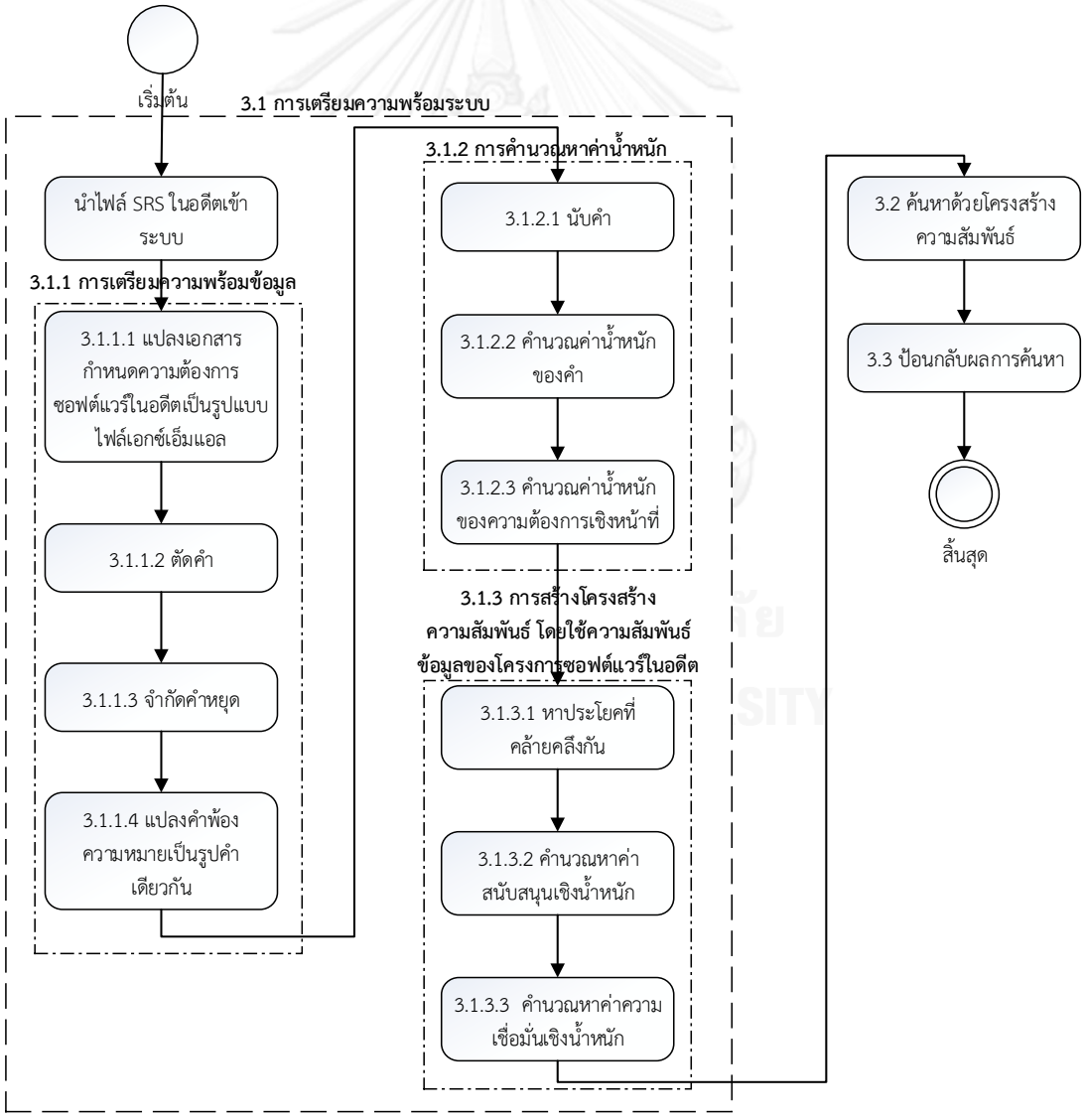
จากภาพที่ 2.3 เมื่อผู้ใช้งานใส่คำสำคัญและปัจจัยแล้วความหมายออนโทโลจีจะช่วยในการระบุความต้องการเฉพาะเจาะจงหรือโดเมนที่เกี่ยวข้องให้ชัดเจนขึ้น หลังจากระบบแสดงผลลัพธ์แล้วผู้ใช้งานสามารถให้คะแนนกับความต้องการเหล่านั้น

จากงานวิจัยนี้จึงเกิดแนวคิดที่จะนำการตอบรับจากผู้ใช้งานมาช่วยในการเพิ่มความแม่นยำให้สืบค้นตรงตามความต้องการมากขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 3
การพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

ผู้วิจัยมีแนวคิดที่จะพัฒนาวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ข้อมูลโครงการซอฟต์แวร์ในอดีต ซึ่งจะนำเอาความต้องการเชิงหน้าที่ในอดีต มาสืบค้นความสัมพันธ์ของระหว่างคำต่าง ๆ เพื่อนำมาขยายวลีหรือกลุ่มคำที่เกี่ยวข้องในการสืบค้นได้อย่างมีประสิทธิภาพมากขึ้น โดยการนำเอาโครงสร้างความสัมพันธ์ด้วยทฤษฎีกฎความสัมพันธ์ มาช่วยเป็นคำแนะนำให้กับนักวิเคราะห์ระบบเบื้องต้นได้ ซึ่งสามารถอธิบายแนวคิดและวิธีดำเนินการวิจัย ดังภาพที่ 3.1



ภาพที่ 3.1 ภาพรวมงานวิจัย

จากภาพที่ 3.1 เป็นภาพรวมงานวิจัย ซึ่งจะถูกแบ่งออกเป็น 3 ส่วนหลัก คือ การเตรียมความพร้อมระบบ การสืบค้นด้วยโครงสร้างความสัมพันธ์ และผลป้อนกลับของผู้ใช้งาน โดยจะนำเอกสารความต้องการซอฟต์แวร์ในอดีตมาเตรียมความพร้อมระบบ ด้วยการเตรียมความพร้อมข้อมูล การคำนวณค่าน้ำหนัก และการสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต หลังจากระบบถูกเตรียมความพร้อมเรียบร้อยแล้ว วิศวกรความต้องการสามารถนำความต้องการของผู้ใช้งานมาสืบค้น เพื่อได้รับเอกสารกำหนดความต้องการซอฟต์แวร์เบื้องต้น ซึ่งเป็นการแนะนำความต้องการซอฟต์แวร์ที่คาดว่าจะเกี่ยวข้อง สำหรับให้วิศวกรความต้องการสามารถรวบรวมความต้องการซอฟต์แวร์ได้รวดเร็ว สำหรับนำไปวิเคราะห์และนำไปคุยกับผู้ใช้งาน เพื่อให้ได้เอกสารกำหนดความต้องการซอฟต์แวร์ที่สมบูรณ์ หลังจากสิ้นสุดโครงการวิศวกรซอฟต์แวร์กลับมาให้คะแนนด้วยการผลป้อนกลับของผู้ใช้งาน เพื่อปรับปรุงโครงสร้างความสัมพันธ์ให้มีประสิทธิภาพที่ดีขึ้น

วิธีดำเนินการวิจัยประกอบด้วย 3 ขั้นตอน ดังต่อไปนี้

3.1 การเตรียมความพร้อมระบบ

การเตรียมความพร้อมระบบ ประกอบด้วย 3 ขั้นตอน ดังต่อไปนี้

3.1.1 การเตรียมความพร้อมข้อมูล

การเตรียมความพร้อมข้อมูลความต้องการเชิงหน้าที่ในอดีต แบ่งเป็น 4 ขั้นตอน

3.1.1.1 การแปลงเอกสารกำหนดความต้องการซอฟต์แวร์ในอดีตเป็นรูปแบบไฟล์เอกซ์เอ็มแอล

การแปลงเอกสารกำหนดความต้องการซอฟต์แวร์ในอดีตเป็นรูปแบบไฟล์เอกซ์เอ็มแอล เป็นการนำเอกสารกำหนดความต้องการซอฟต์แวร์ในอดีตที่รวบรวมไว้ในรูปแบบไฟล์เวิร์ด มาจัดเป็นสัดส่วนในรูปแบบไฟล์เอกซ์เอ็มแอล ซึ่งสามารถนำรูปแบบไฟล์ชนิดต่าง ๆ มาแปลงเพิ่มได้ง่ายกว่า และมีรูปแบบที่ดูเป็นมาตรฐานและเข้าใจง่าย

3.1.1.2 การตัดคำจากความต้องการเชิงหน้าที่ในอดีต

การตัดคำจากความต้องการเชิงหน้าทีในอดีตในรูปแบบไฟล์ข้อความ โดยผู้วิจัยใช้วิธีการตัดคำ โดยการใช้โปรแกรมเล็กซ์โตของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ซึ่งมีประสบการณ์การวิจัยด้านการตัดคำภาษาไทยมาเป็นเวลานานและได้รับความเชื่อถือและเป็นที่ยอมรับอย่างแพร่หลาย จากการประกวดการตัดคำไทย หลังจากนั้นมีบุคคลที่นำมาประยุกต์เชื่อมต่อเข้ากับโปรแกรมที่พัฒนาขึ้นเองอยู่บ่อย ด้วยวิธีพัฒนาด้วยโปรแกรมจาวา หรือเว็บเซอร์วิช

เพื่อเชื่อมต่อกับภาษาอื่น ๆ ได้ นอกจากนั้นเมื่อถูกนำไปใช้งานบ่อย ๆ และสามารถตัดคำได้ใกล้เคียงกับที่ผู้ใช้งานต้องการจึงได้รับความนิยมในการตัดคำ ซึ่งผู้วิจัยนำโปรแกรมมาเชื่อมต่อด้วยวิธีการใช้เว็บเซอร์วิช โปรแกรมนี้ใช้วิธีการตัดคำด้วยการเทียบคำที่ยาวที่สุดจากคำที่มีในพจนานุกรม โดยจะทำการตรวจสอบหรือสแกนข้อความในเอกสารจากซ้ายไปขวา นำไปเทียบกับพจนานุกรมดูว่าเป็นหนึ่งคำหรือไม่ ซึ่งจะวนทำซ้ำจนครบข้อความ ในงานวิจัยนี้จะทำการนำความต้องการเชิงหน้าที่ที่มีมาตัดคำด้วยโปรแกรมเล็กซ์โตทีละอัน แล้วเก็บข้อมูลลงฐานข้อมูลจนครบ

3.1.1.3 การกำจัดคำหยุด

การกำจัดคำหยุดเป็นการนำเอาคำที่ไม่มีความสำคัญหรือคำที่ไม่มีความหมายหรือคำทั่วไปที่พบได้บ่อย ๆ ออกไป รวมถึงคำที่เป็นสัญลักษณ์ต่าง ๆ ผู้วิจัยจะนำเอารายการคำหยุดที่รวบรวมมาจากอินเทอร์เน็ตตัดออกให้เหลือแต่คำที่ไม่ใช่คำหยุดเหล่านี้ ช่วยลดเวลาประมวลผลเกี่ยวกับคำที่ไม่สำคัญและไม่จำเป็นต้องใช้ในสืบค้นออกไป ทำให้ได้ระยะเวลาการประมวลผลที่เร็วขึ้น เนื่องจากลดจำนวนคำที่ใช้ในการประมวลผลลง นอกจากนั้นยังทำให้คำที่ได้มาเป็นคำสำคัญที่แท้จริงมากขึ้น

3.1.1.4 การแปลงคำพ้องความหมายเป็นรูปคำเดียวกัน

การแปลงคำพ้องความหมายเป็นรูปคำเดียวกันเป็นการนำเอาคำที่มีความหมายเหมือนกัน จัดให้เป็นรูปคำ เพื่อลดความซ้ำซ้อนของคำและยังสามารถนับคำได้ถูกต้องมากขึ้น จากการที่มองว่าคำพ้องความหมายที่มีรูปแบบคำไม่เหมือนกันเป็นมากกว่า 1 คำ ให้กลายเป็นคำเดียวกันได้ ซึ่งผู้วิจัยจะนำเอารายการคำพ้องความหมายที่รวบรวมมาช่วยแปลงคำพ้องความหมายให้มีรูปคำเดียวกัน เพื่อให้ทราบจำนวนคำที่ตัดออกมาที่แท้จริงได้

3.1.2 การคำนวณหาค่าน้ำหนัก

การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่ แบ่งเป็น 3 ขั้นตอนดังนี้

3.1.2.1 การนับคำ

การนับคำเป็นการนับความถี่ของแต่ละคำสำคัญหรือดัชนีคำศัพท์ของเอกสาร เพื่อให้ทราบว่าในเอกสารมีคำสำคัญเหล่านั้นปรากฏอยู่มากน้อยเพียงใด และจำนวนเอกสารที่คำสำคัญเหล่านั้นปรากฏอยู่ ผู้วิจัยจะนำคำที่ตัดออกมาได้ แล้วทำการกำจัดคำหยุดและแปลงคำพ้องความหมายให้เป็นรูปคำเดียวกัน นำมานับจำนวนความถี่ของแต่ละคำที่ตัดออกมาจากความต้องการเชิงหน้าที่ และจำนวนความต้องการเชิงหน้าที่ที่คำเหล่านั้นปรากฏอยู่ สำหรับนำไปคำนวณหาค่าน้ำหนักของคำต่อไป ในงานวิจัยนี้นำคำสำคัญที่ถูกตัดคำมาแล้ว มานับคำ โดยใช้เครื่องมือที่ถูกพัฒนาขึ้นมาสำหรับนับคำที่ซ้ำ ๆ กันว่าปรากฏในประโยคนั้นทั้งหมดมีคำสำคัญอยู่กี่คำบ้าง ซึ่งจะได้ตัวอย่างผลลัพธ์การนับคำแสดงเฉพาะ 9 คำ ตามตารางที่ 3.1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 3.1 ตัวอย่างผลลัพธ์การนับคำ แสดงเฉพาะ 9 คำ

ความต้องการเชิงหน้าที่ \ คำ	ฟิลด์	รองรับ	รหัส	พนักงาน	ตัวอักษร	ตัวเลข	จำนวน	บัตร	รูด
มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว	2	1	1	1	1	1	1	0	0
สามารถกำหนดรหัสพนักงาน และ หมายเลขบัตรพนักงาน ได้ทั้งตัวเลขและตัวอักษร เหมือนหรือแตกต่างกันได้	0	0	1	2	1	1	0	1	0
สามารถกำหนดรหัสพนักงานและรหัสบัตรรูดหรือแสกนนิ้ว ได้เหมือนหรือต่างกันก็ได้	0	0	2	1	0	0	0	1	1
บันทึกพนักงานยกเว้นการรูดบัตร และ บันทึกการออกบัตรแทน และสามารถบันทึกเหตุผลการไม่รูดบัตร	0	0	0	1	0	0	0	3	2
รองรับการรูดบัตรเข้าออก 2 คู่เวลา หรือ 4 ครั้งต่อวัน (เวลาปกติ 1 คู่และช่วงล่วงเวลา 1 คู่)	0	1	0	0	0	0	0	1	1
สามารถบันทึกพนักงานยกเว้นการรูดบัตร และบันทึกการออกบัตรแทนได้	0	0	0	1	0	0	0	2	1
สามารถรองรับกับเครื่องรูดบัตรได้ทุกรุ่น ทุกยี่ห้อ	0	1	0	0	0	0	0	1	1
กรณีไม่มีข้อมูลรหัสพนักงานในทะเบียนประวัติ สามารถกำหนดให้สร้างรหัสพนักงานขึ้น	0	0	2	2	0	0	0	0	0
การตั้งสิทธิในการลาพักร้อน การลาอื่นๆ โดยสามารถเข้ากับเครื่องรูดบัตรได้ทุกรุ่น	0	0	0	0	0	0	0	1	1
กำหนดการหักเงินพนักงานที่มาสายได้ทั้งแบบเป็นจำนวนเงิน และ คิดเป็นเท่าของค่าแรง	0	0	0	1	0	0	1	0	0

ตารางที่ 3.1 ตัวอย่างผลลัพธ์การนับคำ แสดงเฉพาะ 9 คำ (ต่อ)

ความต้องการเชิงหน้าที่ \ คำ	ฟิลด์	รองรับ	รหัส	พนักงาน	ตัวอักษร	ตัวเลข	จำนวน	บัตร	รูด
กำหนดจำนวนสวัสดิการของพนักงานแต่ละคนแต่ละประเภทแตกต่างกันได้	0	0	0	1	0	0	1	0	0
เก็บข้อมูลการ มาสาย ขาดงาน ไม่รูดบัตรเข้า ไม่รูดบัตรออก ไม่รูดบัตร ออกก่อนเวลางาน ได้	0	0	0	0	0	0	0	3	3
เก็บข้อมูลประวัติการฝึกอบรมของพนักงาน รองรับการฝึกอบรมภายในและภายนอก	0	1	0	1	0	0	0	0	0
เก็บข้อมูลแยกเป็นประเภทหรือชุดได้ไม่จำกัดจำนวน โดยกำหนดรหัสบริษัทได้	0	0	1	0	0	0	1	0	0
ข้อมูลจะถูกจัดเก็บไว้ในระบบโดยเรียกดูย้อนหลังได้เสมอ แม้จะมีการเปลี่ยนรหัสพนักงานไปแล้ว	0	0	1	1	0	0	0	0	0
บันทึกยอดรวมในแต่ละงวด, บันทึกเวลา จากเครื่องรูดบัตร	0	0	0	0	0	0	0	1	1
พนักงาน 1 คน สามารถมีได้ 5 รหัสพนักงาน	0	0	1	2	0	0	0	0	0
มีฟิลด์รองรับการคิดคำนวณภาษีตามกรมสรรพากร	1	1	0	0	0	0	0	0	0
รองรับการกำหนดกลุ่มพนักงาน เพื่อช่วยในการจัดโครงสร้างหน่วยงานของบริษัท	0	1	0	1	0	0	0	0	0
รองรับการกำหนดวิธีการจ่ายเงินพนักงาน เงินสด เช็ค โอนเข้าบัญชี	0	1	0	1	0	0	0	0	0

3.1.2.2 การคำนวณหาค่าน้ำหนักของคำ

การคำนวณหาค่าน้ำหนักของคำเป็นการนำค่าที่ได้จากการนับคำมาคำนวณ เพื่อให้ทราบน้ำหนักของแต่ละคำที่ตัดมาจากแต่ละความต้องการเชิงหน้าที่มีความสำคัญมากน้อยเพียงใด ผู้วิจัยนำทฤษฎีทีเอฟไอดีเอฟ (TF-IDF) มาช่วยหาค่าน้ำหนักของคำ เนื่องจากทฤษฎีทีเอฟไอดีเอฟจะช่วยตัดคำที่มีสำคัญน้อยออกไปได้ โดยการพิจารณาคำที่ปรากฏอยู่หลายความต้องการเชิงหน้าที่ ซึ่งคำเหล่านั้นมีโอกาสเป็นคำที่มีอยู่ทั่วไป และไม่ใช่คำสำคัญที่แท้จริง โดยเริ่มจากการนำจำนวนความถี่ของแต่ละคำที่ปรากฏในแต่ละความต้องการเชิงหน้าที่มา และจำนวนความต้องการเชิงหน้าที่ที่มีแต่ละคำปรากฏอยู่มาคำนวณตามสูตรสมการ (2.3) ในบทที่ 2

ตัวอย่างเช่น ความต้องการเชิงหน้าที่มีประโยคว่า “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” จะได้ค่าจากการนับคำและตัดคำหยุดออกไปสามารถแทนค่าได้ ดังนี้

{ฟิลด์, รองรับ, รหัส, พนักงาน, ตัวอักษร, ตัวเลข, จำนวน} = {2, 1, 1, 1, 1, 1, 1}

TF = $\{1+\log_2(2), 1+\log_2(1), 1+\log_2(1), 1+\log_2(1), 1+\log_2(1), 1+\log_2(1), 1+\log_2(1)\}$

= {2, 1, 1, 1, 1, 1, 1}

IDF = {7.3259, 4.5651, 5.5417, 1.9034, 8.1739, 8.1739, 4.5186}

TF-IDF = {2x7.3259, 1x4.5651, 1x5.5417, 1x1.9034, 1x8.1739, 1x8.1739, 1x4.5186}

= {14.6519, 4.5651, 5.5417, 1.9034, 8.1739, 8.1739, 4.5186}

นอกจากนั้นจะได้ค่าตามตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของคำ แสดงเฉพาะ 9 คำ ตามตารางที่ 3.2

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 3.2 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของคำ แสดงเฉพาะ 9 คำ

ความต้องการเชิงหน้าที่ \ คำ	ฟิลด์	รองรับ	รหัส	พนักงาน	ตัวอักษร	ตัวเลข	จำนวน	บัตร	รูด
มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว	14.6519	4.5651	5.5417	1.9034	8.1739	8.1739	4.5186	0	0
สามารถกำหนดรหัสพนักงาน และ หมายเลขบัตรพนักงาน ได้ทั้งตัวเลขและตัวอักษร เหมือนหรือแตกต่างกันได้	0	0	5.5417	3.8068	8.1739	8.1739	0	5.3259	0
สามารถกำหนดรหัสพนักงานและรหัสบัตรรูดหรือแสกนนิ้ว ได้เหมือนหรือต่างกันก็ได้	0	0	11.0833	1.9034	0	0	0	5.3259	5.9723
บันทึกพนักงานยกเว้นการรูดบัตร และบันทึกการออกบัตรแทน และสามารถบันทึกเหตุผลการไม่รูดบัตร	0	0	0	1.9034	0	0	0	13.7673	11.9446
รองรับการรูดบัตรเข้าออก 2 คู่เวลา หรือ 4 ครั้งต่อวัน (เวลาปกติ 1 คู่และช่วงล่วงเวลา 1 คู่)	0	4.5651	0	0	0	0	0	5.3259	5.9723
สามารถบันทึกพนักงานยกเว้นการรูดบัตร และบันทึกการออกบัตรแทนได้	0	0	0	1.9034	0	0	0	10.6519	5.9723
สามารถรองรับกับเครื่องรูดบัตรได้ทุกรุ่น ทุกยี่ห้อ	0	4.5651	0	0	0	0	0	5.3259	5.9723
กรณีไม่มีข้อมูลรหัสพนักงานในทะเบียนประวัติ สามารถกำหนดให้สร้างรหัสพนักงานขึ้น	0	0	11.0833	3.8068	0	0	0	0	0
การตั้งสิทธิในการลาพักร้อน การลาอื่นๆ โดยสามารถเข้ากับเครื่องรูดบัตรได้ทุกรุ่น	0	0	0	0	0	0	0	5.3259	5.9723
กำหนดการหักเงินพนักงานที่มาสายได้ทั้งแบบเป็นจำนวนเงิน และ คิดเป็นเท่าของค่าแรง	0	0	0	1.9034	0	0	4.5186	0	0

ตารางที่ 3.2 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของคำ แสดงเฉพาะ 9 คำ (ต่อ)

ความต้องการเชิงหน้าที่ \ คำ	ฟิลด์	รองรับ	รหัส	พนักงาน	ตัวอักษร	ตัวเลข	จำนวน	บัตร	รูด
กำหนดจำนวนสวัสดิการของพนักงานแต่ละคนแต่ละประเภทแตกต่างกันได้	0	0	0	1.9034	0	0	4.5186	0	0
เก็บข้อมูลการ มาสาย ขาดงาน ไม่รูดบัตรเข้า ไม่รูดบัตรออก ไม่รูดบัตร ออกก่อนเวลางาน ได้	0	0	0	0	0	0	0	13.7673	15.4382
เก็บข้อมูลประวัติการฝึกอบรมของพนักงาน รองรับการฝึกอบรมภายในและภายนอก	0	4.5651	0	1.9034	0	0	0	0	0
เก็บข้อมูลแยกเป็นประเภทหรือชุดได้ไม่จำกัดจำนวน โดยกำหนดรหัสบริษัทได้	0	0	5.5417	0	0	0	4.5186	0	0
ข้อมูลจะถูกจัดเก็บไว้ในระบบโดยเรียกดูย้อนหลังได้เสมอ แม้จะมีการเปลี่ยนรหัสพนักงานไปแล้ว	0	0	5.5417	1.9034	0	0	0	0	0
บันทึกยอดรวมในแต่ละงวด, บันทึกเวลา จากเครื่องรูดบัตร	0	0	0	0	0	0	0	5.3259	5.9723
พนักงาน 1 คน สามารถมีได้ 5 รหัสพนักงาน	0	0	5.5417	3.8068	0	0	0	0	0
มีฟิลด์รองรับการคิดคำนวณภาษีตามกรมสรรพากร	7.3259	4.5651	0	0	0	0	0	0	0
รองรับการกำหนดกลุ่มพนักงาน เพื่อช่วยในการจัดโครงสร้างหน่วยงานของบริษัท	0	4.5651	0	1.9034	0	0	0	0	0
รองรับการกำหนดวิธีการจ่ายเงินพนักงาน เงินสด เช็ค โอนเข้าบัญชี	0	4.5651	0	1.9034	0	0	0	0	0

3.1.2.3 การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่

การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่เป็นการนำผลรวมของคำในความต้องการเชิงหน้าที่ยกกำลังสองแล้วนำมาถอดรากที่สอง โดยใช้ความต้องการเชิงหน้าที่เป็นตัวแทนของเว็กเตอร์ของระยะถ่วงน้ำหนัก เพื่อให้ได้ค่าน้ำหนักของความต้องการเชิงหน้าที่คำนวณตามสูตรสมการ (2.4) ในบทที่ 2

ตัวอย่างเช่น ความต้องการเชิงหน้าที่มีประโยคว่า “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” จะมีค่าน้ำหนักของแต่ละคำ ดังนี้

น้ำหนักของคำ = {ฟิลด์, รองรับ, รหัส, พนักงาน, ตัวอักษร, ตัวเลข, จำนวน}

$$= \{14.6519,\ 4.5651,\ 5.5417,\ 1.9034,\ 8.1739,\ 8.1739,\ 4.5186\}$$

$$= \sqrt{(14.6519^2 + 4.5651^2 + 5.5417^2 + 1.9034^2 + 8.1739^2 + 8.1739^2 + 4.5186^2)}$$

$$= 20.5887$$

นอกจากนั้นจะได้ค่าตามตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่บางส่วน ตามตารางที่ 3.3

ตารางที่ 3.3 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่

ลำดับที่	ความต้องการเชิงหน้าที่	เซตคำ	น้ำหนัก
1	มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว	ฟิลด์,รหัส,จำนวน,รองรับ, ตัวอักษร,พนักงาน,ตัวเลข	20.5887
2	สามารถกำหนดรหัสพนักงาน และหมายเลขบัตรพนักงานได้ทั้งตัวเลขและตัวอักษร เหมือนหรือแตกต่างกันได้	ตัวอักษร,หมายเลข,เหมือน ,บัตร,ตัวเลข,พนักงาน,รหัส, แตกต่าง	19.9477
3	สามารถกำหนดรหัสพนักงานและรหัสบัตรรูดหรือแสกนนิ้ว ได้เหมือนหรือต่างกันก็ได้	รูด,นิ้ว,ต่างกัน,บัตร,พนักงาน ,รหัส,แสก,เหมือน	22.2185
4	บันทึกพนักงานยกเว้นการรูดบัตร และบันทึกการออกบัตรแทน และสามารถบันทึกเหตุผลการไม่รูดบัตร	ยกเว้น,บัตร,รูด,บันทึก,เหตุผล, ออก,แทน,พนักงาน	24.7557
5	รองรับการรูดบัตรเข้าออก 2 คู่เวลา หรือ 4 ครั้งต่อวัน (เวลาปกติ 1 คู่และช่วงล่วงเวลา 1 คู่)	บัตร,รองรับ,รูด	31.6762

ตารางที่ 3.3 ตัวอย่างผลลัพธ์การคำนวณหาค่าน้ำหนักของความต้องการเชิงหน้าที่ (ต่อ)

ลำดับที่	ความต้องการเชิงหน้าที่	เซตคำ	น้ำหนัก
6	สามารถบันทึกพนักงานยกเว้นการรูดบัตรและบันทึกการออกบัตรแทนได้	บันทึก,พนักงาน,ออก,แทน,บัตร,รูด,ยกเว้น	18.4728
7	สามารถรองรับกับเครื่องรูดบัตรได้ทุกรุ่นทุกยี่ห้อ	ยี่ห้อ,บัตร,รุ่น,รูด,เครื่อง,รองรับ	17.2827
8	กรณีไม่มีข้อมูลรหัสพนักงานในทะเบียนประวัติสามารถกำหนดให้สร้างรหัสพนักงานขึ้น	กรณี,สร้าง,ประวัติ,ทะเบียน,ข้อมูล,ไม่มี,กำหนดให้,รหัส,พนักงาน	19.4156
9	การตั้งสิทธิในการลาพักร้อน การลาอื่นๆ โดยสามารถเข้ากับเครื่องรูดบัตรได้ทุกรุ่น	พักร้อน,รุ่น,เครื่อง,การลา,สิทธิ,บัตร,เข้ากับ,รูด	23.8136
10	กำหนดการหักเงินพนักงานที่มาสายได้ทั้งแบบเป็นจำนวนเงินและ คิดเป็นเท่าของค่าแรง	จำนวน,หัก,ที่มา,สาย,พนักงาน,กำหนดการ,เงิน,ค่าแรง	16.6237
11	กำหนดจำนวนสวัสดิการของพนักงานแต่ละคนแต่ละประเภทแตกต่างกันได้	พนักงาน,แตกต่าง,สวัสดิการ,ประเภท,จำนวน,แต่ละคน	10.2251
12	เก็บข้อมูลการ มาสาย ขาดงาน ไม่รูดบัตรเข้า ไม่รูดบัตรออก ไม่รูดบัตร ออกก่อนเวลางาน ได้	มาสาย,บัตร,ก่อนเวลา,เก็บข้อมูล,งาน,รูด,ขาดงาน,ออก	26.0156
13	เก็บข้อมูลประวัติการฝึกอบรมของพนักงาน รองรับการฝึกอบรมภายในและภายนอก	ประวัติการ,พนักงาน,เก็บข้อมูล,รองรับ,ฝึกอบรม,ภายใน	15.2502
14	เก็บข้อมูลแยกเป็นประเภทหรือชุดได้ไม่จำกัดจำนวน โดยกำหนดรหัสบริษัทได้	ประเภท,รหัส,จำนวน,แยก,ไม่จำกัด,เก็บข้อมูล,บริษัท,ชุด	14.7924
15	ข้อมูลจะถูกจัดเก็บไว้ในระบบโดยเรียกดูย้อนหลังได้เสมอ แม้จะมีการเปลี่ยนรหัสพนักงานไปแล้ว	พนักงาน,ไว้,จัดเก็บ,เสมอ,ย้อนหลัง,ข้อมูล,ระบบ,รหัส	17.4852
16	บันทึกยอดรวมในแต่ละงวด, บันทึกเวลาจากเครื่องรูดบัตร	เวลา,บัตร,ยอดรวม,บันทึก,งวด,รูด,เครื่อง	15.4102
17	พนักงาน 1 คน สามารถมีได้ 5 รหัสพนักงาน	รหัส,พนักงาน,คน	9.0805
18	มีฟิลด์รองรับการคิดคำนวณภาษีตามกรมสรรพากร	รองรับ,ภาษี,ฟิลด์,คำนวณ,กรมสรรพากร	12.7575
19	รองรับการกำหนดกลุ่มพนักงานเพื่อช่วยในการจัดโครงสร้างหน่วยงานของบริษัท	หน่วยงาน,โครงสร้าง,รองรับ,บริษัท,พนักงาน	11.153

3.1.3 การสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต

การสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีตแบ่งเป็น 3 ขั้นตอนดังนี้

3.1.3.1 การหาความต้องการเชิงหน้าที่ที่คล้ายคลึงกัน (Similarity)

การหาความต้องการเชิงหน้าที่ที่คล้ายคลึงกันเป็นการหาความต้องการเชิงหน้าที่ที่คาดว่าจะเหมือนกันหรือซ้ำกัน เพื่อป้องกันผลลัพธ์ที่ได้รับจากการสืบค้นซ้ำกัน ซึ่งจะทำการแสดงผลลัพธ์ที่ได้รับที่ใกล้เคียงกันแค่อันเดียวเท่านั้น โดยจะนำมาความต้องการเชิงหน้าที่ที่อยู่ในระบบย่อยเดียวกันมาแทนตามสูตรสมการ (2.5) .ในบทที่ 2

ในงานวิจัยนี้ต้องการหาประโยคที่มีความคล้ายคลึงกันค่อนข้างมาก สำหรับยืนยันว่าเป็นประโยคเดียวกันหรือไม่ เพื่อไม่ให้ไปแสดงประโยคที่ไม่ซ้ำซ้อนกันจริง และไม่ควรตัดประโยคที่เป็นคนละประโยคออกไป จากการนำข้อมูลทดสอบมาคำนวณหาค่าความคล้ายคลึง พบว่าความต้องการเชิงหน้าที่ที่มีค่าความคล้ายคลึงมากกว่าหรือเท่ากับ 0.9 ส่วนใหญ่มีความหมายของประโยคที่คล้ายคลึงกัน แต่ถ้าค่าความคล้ายคลึงที่น้อยกว่า 0.9 ส่วนใหญ่มีความหมายของประโยคที่ไม่คล้ายคลึงกัน

ตัวอย่างเช่น ความต้องการเชิงหน้าที่มีประโยคว่า “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” และ “สามารถกำหนดรหัสพนักงาน และ หมายเลขบัตรพนักงานได้ทั้งตัวเลขและตัวอักษร เหมือนหรือแตกต่างกันได้”

จะได้ค่าน้ำหนักหลังจากตัดคำหยุดออกไปแล้ว จะได้ดังนี้

ประโยคที่ 1 = {ฟิลด์, รองรับ, รหัส, พนักงาน, ตัวอักษร, ตัวเลข, จำนวน}
= {14.6519, 4.5651, 5.5417, 1.9034, 8.1739, 8.1739, 4.5186}
ประโยคที่ 2 = {รหัส, พนักงาน, หมายเลข, บัตร, ตัวเลข, ตัวอักษร, เหมือน, แตกต่าง}
= {5.5417, 3.8068, 9.4959, 5.3259, 8.1739, 8.1739, 8.4959, 5.3259}

ขนาดของประโยคที่ 1
$= 14.6519^2+4.5651^2+5.5417^2+1.9034^2+8.1739^2+8.1739^2+4.5186^2$
= 423.8966

ขนาดของประโยคที่ 2
$= 5.5417^2+3.8068^2+9.4959^2+5.3259^2+8.1739^2+8.1739^2+8.4959^2+5.3259^2$
= 397.9103

ความคล้ายคลึงกัน

$$= \frac{(5.5417 \times 5.5417) + (1.9034 \times 3.8068) + (8.1739 \times 8.1739) + (8.1739 \times 8.1739)}{\sqrt{423.8966} \times \sqrt{397.9103}}$$

$$= \frac{171.5821}{410.6980}$$

$$= 0.4178$$

จากประโยคว่า “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” และ “สามารถกำหนดรหัสพนักงาน และ หมายเลขบัตรพนักงานได้ทั้งตัวเลขและตัวอักษรเหมือนหรือแตกต่างกันได้” ได้ผลลัพธ์ที่ได้มีค่าเท่ากับ 0.4178 ซึ่งความคล้ายคลึงกันน้อยกว่า 0.9 ดังนั้น “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” และ “สามารถกำหนดรหัสพนักงานและรหัสบัตรรูดหรือสแกนนิ้ว ได้เหมือนหรือต่างกันก็ได้” จึงไม่มีความคล้ายคลึงกัน

3.1.3.2 การคำนวณหาค่าสนับสนุนเชิงน้ำหนัก (W-Support)

การคำนวณหาค่าสนับสนุนเชิงน้ำหนักตามทฤษฎีของกฎความสัมพันธ์เชิงน้ำหนัก เป็นความน่าจะเป็นที่จะพบคำเหล่านี้เกิดร่วมกันหรือเกี่ยวข้องกัน ซึ่งนำค่าผลรวมของน้ำหนักของความต้องการเชิงหน้าที่ที่มีคำที่อยู่ในเซตคำนั้นที่เกิดร่วมกันทั้งหมดมาหารกับค่าผลรวมทั้งหมดของน้ำหนักของความต้องการเชิงหน้าที่ เพื่อหาโอกาสที่คำเหล่านั้นจะมีความสัมพันธ์หรือเกี่ยวข้องกัน คำนวณตามสูตรสมการ (2.5) ในบทที่ 2 โดยค่าสนับสนุนเชิงน้ำหนักยิ่งมากก็จะแสดงว่ามีโอกาสที่คำเหล่านั้นจะเกิดร่วมกัน ตั้งแต่ 2 ครั้งขึ้นไป จะได้ค่าตามตัวอย่างผลลัพธ์การคำนวณหาค่าสนับสนุนเชิงน้ำหนัก ตามตารางที่ 3.4 ซึ่งนำเอาคำหรือกลุ่มคำที่มีการเกิดขึ้นในความต้องการเชิงหน้าที่ ตั้งแต่ 2 ครั้งขึ้นไป มาแสดงเป็นเซตคำ แล้วนำค่าน้ำหนักของความต้องการเชิงหน้าที่ที่มีคำนั้นปรากฏอยู่มาคำนวณหาค่าสนับสนุนเชิงน้ำหนัก

ตัวอย่างเช่น คำว่า “ฟิลด์” ปรากฏอยู่บนความต้องการเชิงหน้าที่ “มีฟิลด์รองรับการกำหนดรหัสพนักงานได้ 2 ฟิลด์ ได้ทั้งตัวอักษรและตัวเลข จำนวน 30 ตัว” และ “สามารถเก็บฟิลด์ตัวเลือกได้ 20 ฟิลด์” ซึ่งมีค่าน้ำหนักของความต้องการเชิงหน้าที่เท่ากับ 20.5887 และ 17.1488 ตามลำดับ โดยมีค่าน้ำหนักของความต้องการเชิงหน้าที่ทั้งหมดในระบบเป็น 2716.6267 มาแทนสมการดังนี้

$$W_{support} = \frac{20.5887 + 17.1488}{2716.6267}$$

$$= 0.0139$$

ตารางที่ 3.4 ตัวอย่างผลลัพธ์การคำนวณหาค่าสนับสนุนเชิงน้ำหนัก

ลำดับที่	เซตคำ	ค่าสนับสนุน
1	ฟิลด์	0.0139
2	รองรับ	0.0248
3	รหัส	0.051
4	พนักงาน	0.4426
5	ตัวอักษร	0.0146
6	จำนวน	0.0112
7	บัตร	0.0431
8	รูด	0.0139

3.1.3.3 การคำนวณหาค่าความเชื่อมั่นเชิงน้ำหนัก (W-Confidence)

การคำนวณหาค่าเชื่อมั่นเชิงน้ำหนักตามทฤษฎีของกฎความสัมพันธ์เชิงน้ำหนัก เป็นความน่าจะเป็นที่จะพบคำ A แล้วจะเกิดคำ B ซึ่งนำค่าสนับสนุนเชิงน้ำหนักของเซตคำมาหารค่าสนับสนุนเชิงน้ำหนักของคำที่เกิดก่อน โดยใช้เซตคำที่มีอย่างน้อย 2 คำและเกิดร่วมกันมากกว่า 2 ครั้งนำมาสร้างความสัมพันธ์ที่มีโอกาสเป็นไปได้ว่าคำที่เกิดก่อน (Antecedent) แล้วเกิดเป็นคำที่เป็นผลลัพธ์ (Consequent) เพื่อหาโอกาสจะเกิดความสัมพันธ์หรือการเชื่อมต่อระหว่างคำคำนวณตามสูตรสมการ (2.6) ในบทที่ 2 โดยในงานวิจัยนี้กำหนดค่าสนับสนุนเชิงน้ำหนักขั้นต่ำเป็น 0.1 เพื่อให้ได้คำที่มีโอกาสเกิดร่วมกัน และมีค่าน้ำหนักไม่น้อยเกินไป ซึ่งจะทำให้ไม่ได้ค่าสนับสนุนความเชื่อมั่นเชิงน้ำหนักที่เหมาะสมได้ สำหรับใช้ในการคำนวณค่าสนับสนุนความเชื่อมั่นเชิงน้ำหนัก

ตัวอย่างผลลัพธ์นี้ได้แสดงตารางที่ 3.5 เช่น คำว่า “พนักงาน” มีความน่าจะเป็นที่เกิดคำว่า “ประวัติ” มาแทนสมการเป็น ค่าสนับสนุนเชิงน้ำหนักของเซตคำว่า “พนักงาน, ประวัติ” มีค่าเท่ากับ 0.1146 หารกับค่าสนับสนุนเชิงน้ำหนักของคำว่า “พนักงาน” มีค่าเท่ากับ 0.4426 ซึ่งจะได้สมการ ดังนี้

$$W_{Confidence} = \frac{0.1146}{0.4426}$$
$$= 0.2589$$

ตารางที่ 3.5 ตัวอย่างผลลัพธ์การคำนวณหาค่าความเชื่อมั่นเชิงน้ำหนัก

ลำดับที่	คำที่เกิดก่อน	คำที่เป็นผลลัพธ์	ค่าเชื่อมั่น
1	พนักงาน	ประวัติ	0.2589
2	พนักงาน	บันทึก	0.2417
3	ประวัติ	พนักงาน	0.6987
4	บันทึก	พนักงาน	0.5080
5	พนักงาน	สิทธิ์	1
6	พนักงาน, รูด	บันทึก	0.7604
7	พนักงาน, ทะเบียน	เงินเดือน	1
8	พนักงาน, อัตรา, หน่วยงาน	วิเคราะห์	0.7553
9	พนักงาน, การทำงาน, รูด	เวลา	1
10	พนักงาน, วิเคราะห์, อัตรา, หน่วยงาน	รายงาน	1

จากขั้นตอนนี้ จะได้โครงสร้างความสัมพันธ์ระหว่างคำ เพื่อสามารถหาคำที่มีความสัมพันธ์เกี่ยวข้องกัน ซึ่งจะทำให้ได้จากคำกลายเป็นกลุ่มคำหรือวลีได้ สำหรับขยายเป็นข้อคำถามในการค้นหาความต้องการเชิงหน้าที่ได้มากขึ้นและตรงตามต้องการมากขึ้น

3.2 การสืบค้นด้วยโครงสร้างความสัมพันธ์

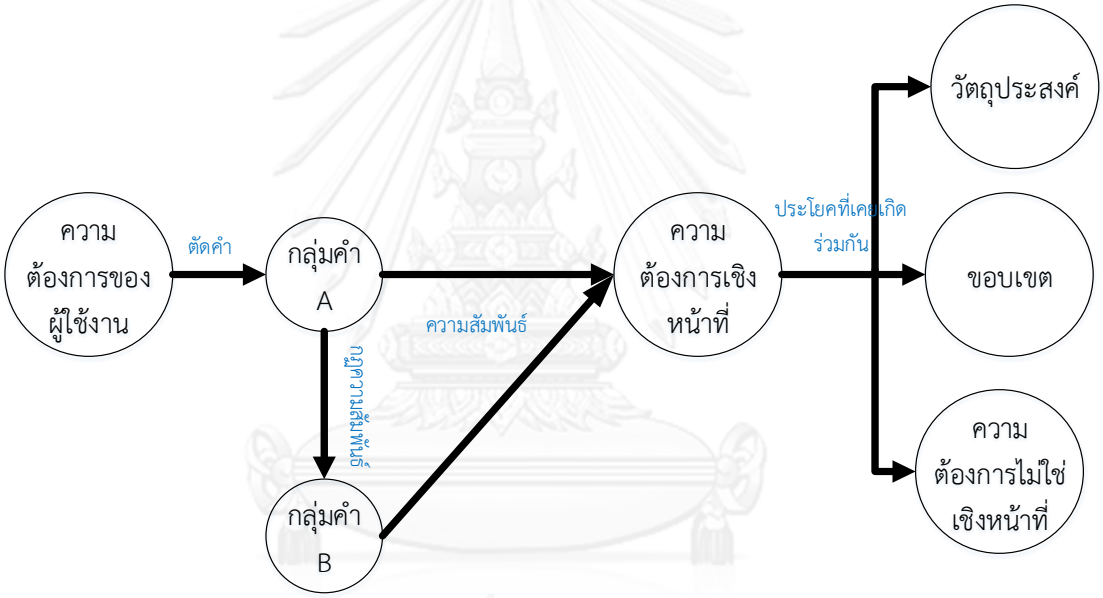
การสืบค้นด้วยโครงสร้างความสัมพันธ์เป็นการเทียบความสัมพันธ์ของคำจากความสัมพันธ์ของความต้องการเชิงหน้าที่ โดยใช้ทฤษฎีของกฎความสัมพันธ์เชิงน้ำหนัก เป็นเหมือนการขยายข้อคำถาม เพื่อให้ได้ผลในการสืบค้นมากขึ้น ซึ่งโครงสร้างความสัมพันธ์เป็นการเทียบข้อมูลที่มีในฐานข้อมูลด้วยความสัมพันธ์ด้วยอาศัยค่าความเชื่อมั่นที่เมื่อพบ A แล้วจะเกิด B ต่อมา ในการสืบค้นคำที่สัมพันธ์เกี่ยวเนื่องกัน โดยมีขั้นตอนดังนี้

1. วิศวกรความต้องการป้อนความต้องการของผู้ใช้งาน เพื่อสืบค้น
2. นำความต้องการของผู้ใช้งานที่ป้อนเข้าสำหรับการสืบค้นมาตัดออกเป็นคำ
3. นำคำสำคัญมาหาโครงสร้างความสัมพันธ์ของคำ ด้วยการหาค่าความเชื่อมั่นที่อย่างน้อย 0.9 เพื่อให้ได้คำที่เกี่ยวข้องกัน สำหรับค้นหาความต้องการเชิงหน้าที่ที่เกี่ยวข้องกันได้มากขึ้น

4. นำคำสำคัญที่ถูกค้นพบทั้งหมด เทียบหาความสัมพันธ์ระหว่างคำสำคัญเหล่านั้นกับความต้องการเชิงหน้าที่ เพื่อหาความต้องการเชิงหน้าที่ที่เกี่ยวข้องกัน สำหรับแสดงในเอกสารความต้องการซอฟต์แวร์เบื้องต้น

5. นำความต้องการเชิงหน้าที่ที่ได้รับ มาหาความต้องการไม่ใช่เชิงหน้าที่ ขอบเขต และวัตถุประสงค์ ด้วยการหาการที่ประโยคเหล่านั้นเคยปรากฏในเอกสารของโครงการร่วมกัน เพื่อนำมาแสดงในเอกสารความต้องการซอฟต์แวร์เบื้องต้น

6. นำเอาวัตถุประสงค์ ขอบเขต ความต้องการเชิงหน้าที่ และความต้องการไม่ใช่เชิงหน้าที่ออกมาในรูปแบบของเอกสารความต้องการซอฟต์แวร์เบื้องต้น และสามารถนำออกมาเป็นรูปแบบไฟล์ PDF



ภาพที่ 3.2 โครงสร้างความสัมพันธ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.3 ผลป้อนกลับของผู้ใช้งาน

ผลป้อนกลับของผู้ใช้งานเป็นการให้คะแนนของผู้ใช้งาน เพื่อเป็นการเพิ่มประสิทธิภาพให้กับผลลัพธ์ที่ได้ให้มีความใกล้เคียงกับความคิดของผู้ใช้งานมากขึ้น หลังจากผู้ใช้งานมีการสืบค้นและได้รับผลลัพธ์เรียบร้อยแล้ว ผู้วิจัยจะให้ผู้ใช้งานให้คะแนนกับผลลัพธ์ที่ได้แต่ละอัน โดยให้คะแนนเป็น 1 – 5 ตามตาราง 3.6

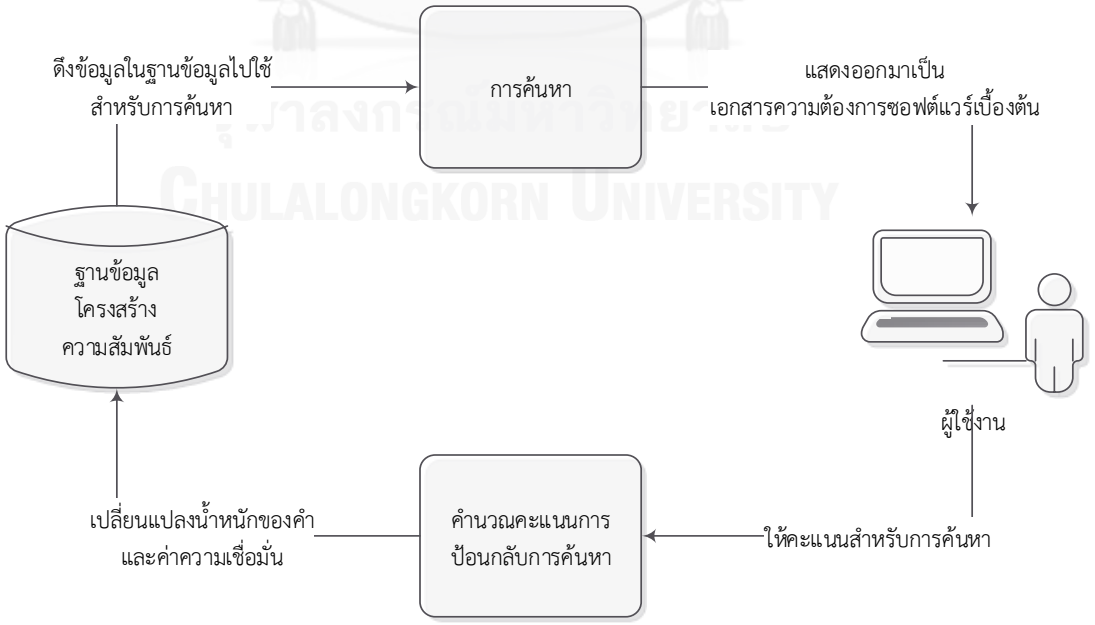
ตารางที่ 3.6 ความหมายของการให้คะแนน สำหรับผลป้อนกลับผู้ใช้งาน

คะแนน	ความหมาย
1	ผลลัพธ์ที่ได้ไม่มีความเกี่ยวข้องกับความต้องการเชิงหน้าที่
2	ผลลัพธ์ที่ได้ไม่เกี่ยวข้องกับความต้องการเชิงหน้าที่ แต่สามารถนำมาประยุกต์เพื่อให้ได้ความต้องการเชิงหน้าที่ใหม่ได้
3	ผลลัพธ์ที่ได้มีความเกี่ยวข้องกับความต้องการเชิงหน้าที่น้อย ต้องปรับปรุงข้อมูลมาก่อนนำมาใช้งาน
4	ผลลัพธ์ที่ได้มีความเกี่ยวข้องกับความต้องการเชิงหน้าที่มาก แต่ต้องปรับปรุงข้อมูลเล็กน้อยก่อนนำมาใช้งาน
5	ผลลัพธ์ที่ได้ตรงกับความต้องการเชิงหน้าที่ สามารถนำมาใช้งานได้

ผู้ใช้งานจะให้คะแนนตามความรู้สึกว่าความต้องการเชิงหน้าที่ที่ได้รับตรงตามความต้องการจากมากไปน้อย หลังจากนั้นจะนำคะแนนที่ได้รับมาเพิ่มหรือลดลำดับความสัมพันธ์ของส่วนนั้น โดยแยกเป็น 2 แบบดังนี้

1. กรณีให้คะแนนสำหรับผลลัพธ์ที่ได้มาจากคำที่รับค่าเข้ามา จะนำไปเพิ่มหรือลดต่อค่าน้ำหนักของคำที่อยู่ในความต้องการเชิงหน้าที่นั้น โดยจะส่งผลกับการเรียงลำดับของข้อมูลในส่วนของคำที่รับค่าเข้ามา

2. กรณีให้คะแนนสำหรับผลลัพธ์ที่ได้มาจากคำที่สัมพันธ์กัน จะนำไปเพิ่มหรือลดต่อค่าความเชื่อมั่นเชิงน้ำหนักของคำนั้น



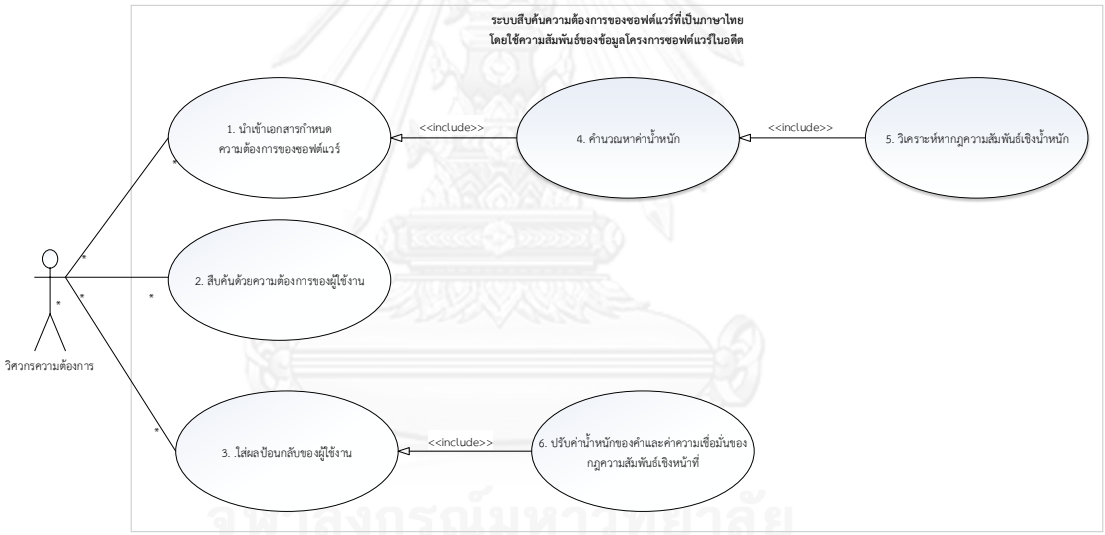
ภาพที่ 3.3 การป้อนกลับการสืบค้น

บทที่ 4
การพัฒนาเครื่องมือ

จากวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีตที่นำเสนอ ผู้วิจัยได้พัฒนาเครื่องมือเพื่อสนับสนุนวิธีการที่นำเสนอ โดยมีรายละเอียดของการพัฒนาเครื่องมือดังนี้

4.1 การออกแบบหน้าที่การทำงานของระบบ

ผู้วิจัยใช้แผนภาพยูสเคสแสดงฟังก์ชันการทำงานของเครื่องมือ และผู้ใช้งานที่มีปฏิสัมพันธ์กับเครื่องมือ เพื่อใช้อธิบายหน้าที่ของระบบ ดังภาพที่ 4.1



ภาพที่ 4.1 แผนภาพยูสเคสของระบบสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

จากภาพที่ 4.1 สามารถเขียนเป็นตารางอธิบายแผนภาพยูสเคสได้ดังนี้

ตารางที่ 4.1 คำอธิบายแผนภาพยูสเคสส่วนเพิ่มเอกสารกำหนดความต้องการของซอฟต์แวร์เข้าฐานข้อมูล

ชื่อยูสเคส: นำเข้าเอกสารกำหนดความต้องการของซอฟต์แวร์	รหัส: UC-001
ผู้กระทำหลัก: วิศวกรความต้องการ	ระดับความสำคัญ: มาก
ความสัมพันธ์: -	
คำอธิบาย: วิศวกรความต้องการนำเข้าเอกสารความต้องการของโครงการในอดีตลงในระบบ	
สิ่งกระตุ้น: คลิกปุ่มอัพโหลดไฟล์ เพื่อนำเข้าเอกสารความต้องการ ประเภทของสิ่งกระตุ้น: ภายนอก	
ขั้นตอนการทำงานปกติ: 1. นำเข้าเอกสารความต้องการของโครงการในอดีต โดยเลือกที่อยู่ของไฟล์ แล้วกดปุ่มอัพโหลดไฟล์ 2. ระบบดึงเอกสารความต้องการจากพาธที่กำหนด 3. ระบบตัดคำจากเอกสารความต้องการ 4. ระบบกำจัดคำหยุดจากเอกสารความต้องการ 5. ระบบแปลงคำพ้องความหมายเป็นรูปคำเดียวกัน 6. ระบบนับคำ แล้วเก็บลงฐานข้อมูล	

ตารางที่ 4.2 คำอธิบายแผนภาพยูสเคสส่วนการสืบค้นด้วยความต้องการของผู้ใช้งาน

ชื่อยูสเคส: สืบค้นด้วยความต้องการของผู้ใช้งาน	รหัส: UC-002
ผู้กระทำหลัก: วิศวกรความต้องการ	ระดับความสำคัญ: มาก
ความสัมพันธ์: -	
คำอธิบาย: วิศวกรความต้องการสืบค้น โดยใช้ความต้องการของผู้ใช้งานที่รวบรวม	
สิ่งกระตุ้น: ใส่คำค้น แล้วคลิกปุ่มสืบค้น ประเภทของสิ่งกระตุ้น: ภายนอก	
ขั้นตอนการทำงานปกติ: 1. วิศวกรความต้องการใส่คำหรือประโยคที่ต้องการสืบค้นลงในระบบ 2. คลิกปุ่ม สืบค้น 3. ระบบจะแสดงเอกสารความต้องการซอฟต์แวร์เบื้องต้น	

ตารางที่ 4.3 คำอธิบายแผนภาพยูสเคสส่วนผลป้อนกลับของผู้ใช้งาน

ชื่อยูสเคส: ใส่ผลป้อนกลับของผู้ใช้งาน	รหัส: UC-003
ผู้กระทำหลัก: วิศวกรความต้องการ	ระดับความสำคัญ: ปานกลาง
ความสัมพันธ์: -	
คำอธิบาย: วิศวกรความต้องการผลป้อนกลับของผู้ใช้งาน	
สิ่งกระตุ้น: เลือกค่าคะแนนความสัมพันธ์ที่ต้องการให้ แล้วคลิกปุ่มส่งผลป้อนกลับ ประเภทของสิ่งกระตุ้น: ภายนอก	
ขั้นตอนการทำงานปกติ: 1. วิศวกรความต้องการเลือกค่าคะแนน ตามความสัมพันธ์ของผลลัพธ์ที่ได้ 2. ระบบจะเก็บค่าคะแนนที่ได้เข้าสู่ฐานข้อมูล	

ตารางที่ 4.4 คำอธิบายแผนภาพยูสเคสส่วนคำนวณหาค่าน้ำหนัก

ชื่อยูสเคส: คำนวณหาค่าน้ำหนัก	รหัส: UC-004
ผู้กระทำหลัก: -	ระดับความสำคัญ: มาก
ความสัมพันธ์: Include: นำเข้าเอกสารกำหนดความต้องการของซอฟต์แวร์	
คำอธิบาย: คำนวณหาค่าน้ำหนัก	
สิ่งกระตุ้น: คลิกปุ่มประมวลผล ประเภทของสิ่งกระตุ้น: ภายนอก	
ขั้นตอนการทำงานปกติ: 1. ระบบนำคำที่มีในระบบมาคำนวณหาค่าน้ำหนักของคำ และความต้องการเชิงหน้าที่ 2. ระบบประมวลผลหาความต้องการเชิงหน้าที่ที่คล้ายคลึงกัน	

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

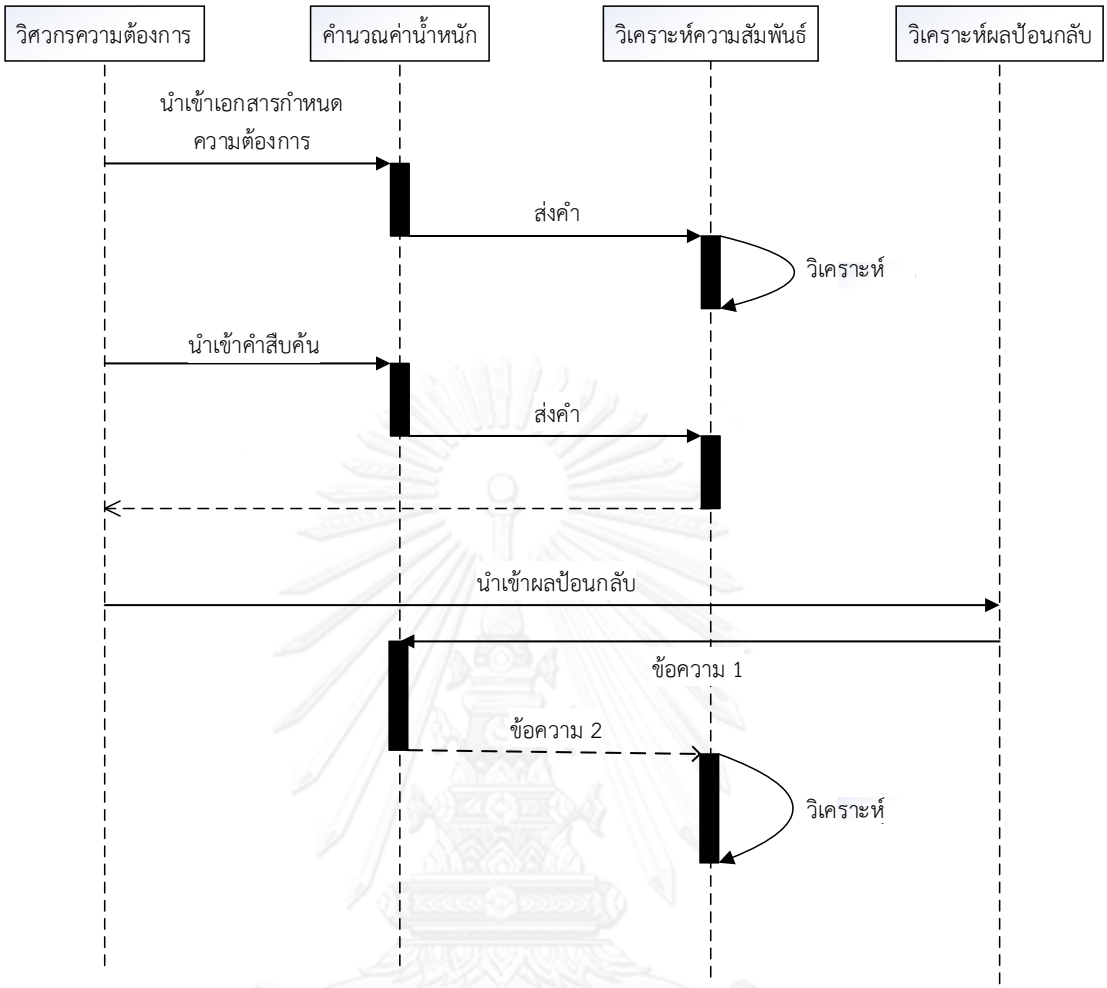
ตารางที่ 4.5 คำอธิบายแผนภาพยูสเคสส่วนวิเคราะห์โครงสร้างความสัมพันธ์

ชื่อยูสเคส: วิเคราะห์หากฎความสัมพันธ์เชิงน้ำหนัก	รหัส: UC-005
ผู้กระทำหลัก: -	ระดับความสำคัญ: มาก
ความสัมพันธ์: Include: สืบค้นด้วยความต้องการของผู้ใช้งาน คำนวณหาค่าน้ำหนัก และวิเคราะห์หาความเหมาะสมจากผลป้อนกับของผู้ใช้งาน	
คำอธิบาย: วิเคราะห์คำในระบบเพื่อสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต	
สิ่งกระตุ้น: คำนวณหาค่าน้ำหนักเสร็จสิ้น ประเภทของสิ่งกระตุ้น: ภายใน	
ขั้นตอนการทำงานปกติ: 1. ระบบนำคำที่มีในฐานข้อมูลมาคำนวณหาโครงสร้างความสัมพันธ์ 2. ระบบประมวลผลคำนวณหาค่าสนับสนุน และค่าความเชื่อมั่น	

ตารางที่ 4.6 คำอธิบายแผนภาพยูสเคสส่วนวิเคราะห์ผลการป้อนกลับ

ชื่อยูสเคส: วิเคราะห์หาความเหมาะสมจากผลป้อนกลับของผู้ใช้งาน	รหัส: UC-006
ผู้กระทำหลัก: -	ระดับความสำคัญ: ปานกลาง
ความสัมพันธ์: Include: ใส่ผลป้อนกลับของผู้ใช้งาน	
คำอธิบาย: วิเคราะห์การผลป้อนกลับของผู้ใช้งาน	
สิ่งกระตุ้น: คลิกปุ่มส่งผลป้อนกลับ เพื่อส่งค่าคะแนนการผลป้อนกลับของผู้ใช้งาน ประเภทของสิ่งกระตุ้น: ภายใน	
ขั้นตอนการทำงานปกติ: 1. ระบบเก็บค่าคะแนนการผลป้อนกลับของผู้ใช้งาน 2. ระบบคำนวณหาค่าน้ำหนักใหม่	

จากแผนภาพยูสเคสสามารถเขียนเป็นแผนภาพกิจกรรมเพื่อแสดงลำดับการทำงานและปฏิสัมพันธ์ระหว่างวัตถุ ตามลำดับของเหตุการณ์ที่เกิดขึ้นของระบบได้ดังภาพที่ 4.2



ภาพที่ 4.2 แผนภาพกิจกรรมของระบบสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

จากภาพที่ 4.2 สามารถอธิบายได้ดังขั้นตอนได้ดังต่อไปนี้

1. วิศวกรความต้องการนำเข้าเอกสารความต้องการของโครงการต่าง ๆ เข้าสู่ระบบ

2. ระบบเตรียมความพร้อมข้อมูลในเอกสารความต้องการที่วิศวกรความต้องการนำเข้าได้ผลลัพธ์เป็นคำสำคัญ

3. นำคำที่ได้จากการเตรียมความพร้อมข้อมูลในเอกสารความต้องการไปวิเคราะห์เพื่อสร้างโครงสร้างความสัมพันธ์ โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

4. วิศวกรความต้องการสืบค้นความต้องการจากการใส่ความต้องการของผู้ใช้งาน ซึ่งจะเข้าสู่การตัดคำและกำจัดคำหยุด แล้วจะตัดจากประโยค เป็นกลุ่มคำ

5. นำกลุ่มคำที่ได้ไปเปรียบเทียบกับความต้องการเชิงหน้าที่

6. นำกลุ่มคำที่ได้ไปเปรียบเทียบกับโครงสร้างความสัมพันธ์ เพื่อให้ได้กลุ่มคำที่เพิ่มขึ้น

7. นำกลุ่มคำที่ได้เพิ่มใหม่ไปเปรียบเทียบกับความต้องการเชิงหน้าที่

8. นำความต้องการเชิงหน้าที่ไปหาความสัมพันธ์กับส่วนอื่น ๆ ของเอกสารความต้องการซอฟต์แวร์

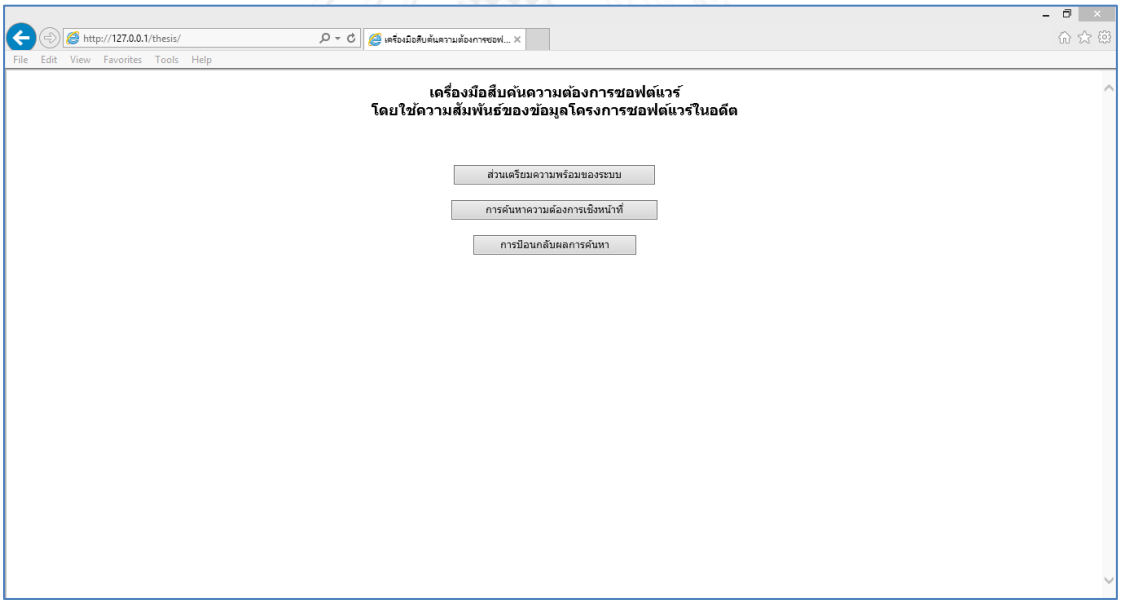
9. ระบบแสดงผลลัพธ์ของการสืบค้นในรูปแบบความต้องการเชิงฟังก์ชัน จากการเปรียบเทียบความสัมพันธ์ กลับไปยังวิศวกรความต้องการ ในรูปแบบเอกสารความต้องการซอฟต์แวร์

10. วิศวกรความต้องการให้ค่าคะแนนจากผลลัพธ์การสืบค้นลงในระบบวิเคราะห์การป้อนกลับ

11. ระบบวิเคราะห์การป้อนกลับนำค่าคะแนนที่ได้ไปปรับปรุงค่าความสัมพันธ์ของโครงสร้างความสัมพันธ์

4.2 การออกแบบส่วนต่อประสานของเครื่องมือ

ส่วนต่อประสานของเครื่องมือ สนับสนุนวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ข้อมูลของโครงการซอฟต์แวร์ในอดีต ประกอบด้วย 3 ส่วนหลัก ดังภาพที่ 4.3



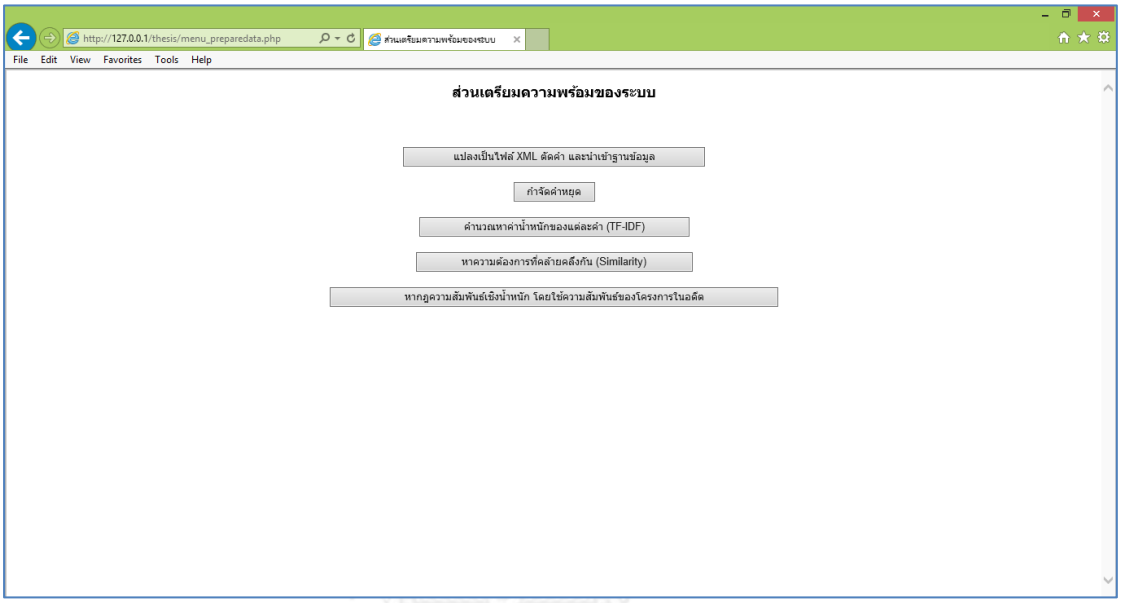
ภาพที่ 4.3 หน้าจอหลักของเครื่องมือ

4.2.1 ส่วนการเตรียมความพร้อมของระบบ

ส่วนเตรียมความพร้อมของระบบเป็นส่วนที่จะนำข้อมูลของโครงการในอดีตมาประมวลผลให้ระบบเตรียมความพร้อมก่อนจะทำการสืบค้นได้ ซึ่งประกอบด้วย 5 ส่วนดังนี้

1. แปลงเป็นไฟล์ XML ตัดคำ และนำเข้าฐานข้อมูล

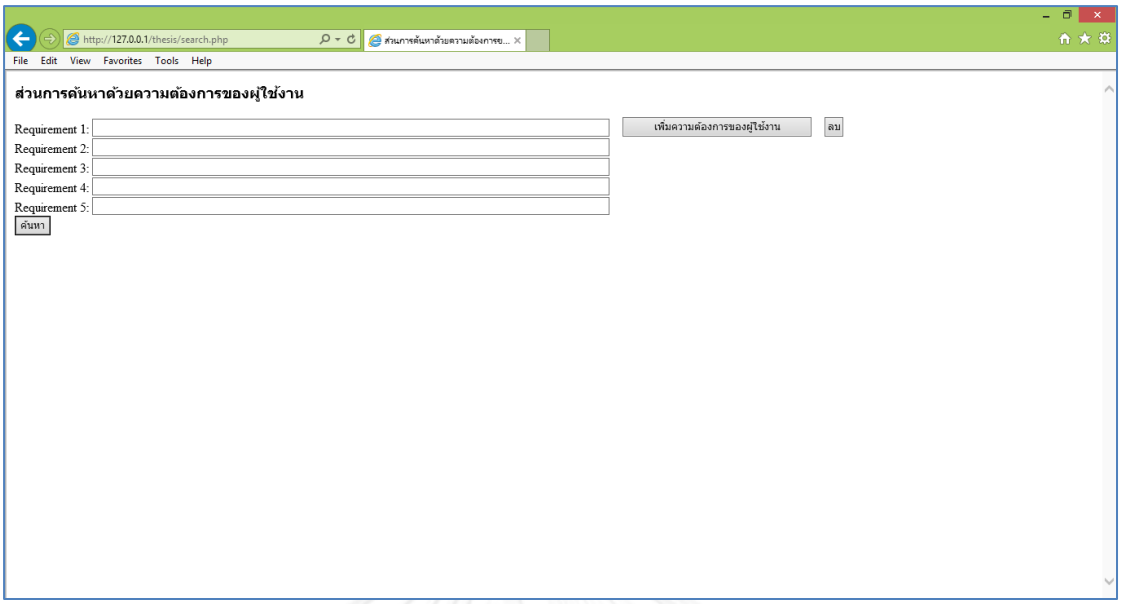
2. กำจัดคำหยุด
3. คำนวณหาค่าน้ำหนักของแต่ละคำ (TF-IDF)
4. หาค่าความต้องการที่คล้ายคลึงกัน (Similarity)
5. หากฎความสัมพันธ์เชิงน้ำหนัก โดยใช้ความสัมพันธ์ของโครงการในอดีต ดังภาพที่ 4.4



ภาพที่ 4.4 หน้าจอของส่วนการเตรียมความพร้อมของระบบ

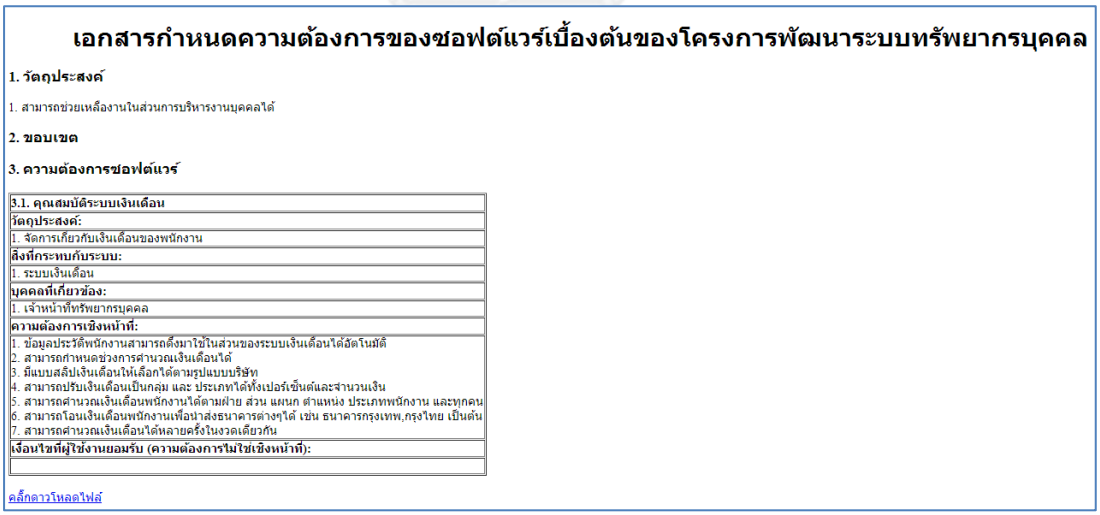
4.2.2 ส่วนการสืบค้น

ส่วนสืบค้นด้วยความต้องการของผู้ใช้งานเป็นส่วนที่สามารถให้วิศวกรความต้องการนำความต้องการของผู้ใช้งานมาป้อนเข้าสู่ระบบ ซึ่งระบบจะนำไปตัดคำ และกำจัดคำหยุด เพื่อหาความต้องการเชิงหน้าที่ แล้วเอากลุ่มคำที่ได้ไปหาคำที่มีความสัมพันธ์ เพื่อนำไปขยายหาความต้องการเชิงหน้าที่ที่คาดว่าจะเกี่ยวข้องเพิ่มได้อีก หลังจากนั้นจึงนำความต้องการเชิงหน้าที่ที่ได้รับไปหาส่วนต่าง ๆ ของเอกสารความต้องการซอฟต์แวร์ เพื่อประกอบเป็นรูปแบบไฟล์ PDF เอกสารความต้องการซอฟต์แวร์เบื้องต้น เพื่อให้วิศวกรความต้องการนำไปตรวจทานกับผู้ใช้งาน เพื่อหาความต้องการซอฟต์แวร์ที่แท้จริงต่อไปดังภาพที่ 4.5



ภาพที่ 4.5 หน้าจอของส่วนการสืบค้นด้วยความต้องการของผู้ใช้งาน

จากภาพที่ 4.5 เป็นส่วนการสืบค้นด้วยความต้องการของผู้ใช้งาน เมื่อป้อนความต้องการของผู้ใช้งานที่ต้องการสืบค้นแล้ว ระบบจะนำไปประมวลผล เพื่อให้ได้เอกสารความต้องการซอฟต์แวร์เบื้องต้นออกมา ดังภาพที่ 4.6



เอกสารกำหนดความต้องการของซอฟต์แวร์เบื้องต้นของโครงการพัฒนาระบบทรัพยากรบุคคล

1. วัตถุประสงค์

1. สามารถช่วยเหลืองานในส่วนการบริหารงานบุคคลได้

2. ขอบเขต

3. ความต้องการซอฟต์แวร์

3.1. คุณสมบัติระบบเงินเดือน
วัตถุประสงค์:
1. จัดการเกี่ยวกับเงินเดือนของพนักงาน
สิ่งที่กระทบกับระบบ:
1. ระบบเงินเดือน
บุคคลที่เกี่ยวข้อง:
1. เจ้าหน้าที่ทรัพยากรบุคคล
ความต้องการเชิงหน้าที่:
1. ข้อมูลประวัติพนักงานสามารถดึงมาใช้ในส่วนของระบบเงินเดือนได้อัตโนมัติ 2. สามารถกำหนดช่วงการคำนวณเงินเดือนได้ 3. มีแบบสลิปเงินเดือนให้เลือกได้ตามรูปแบบบริษัท 4. สามารถปรับเงินเดือนเป็นกลุ่ม และ ประเภทได้ทั้งเปอร์เซ็นต์และจำนวนเงิน 5. สามารถคำนวณเงินเดือนพนักงานได้ตามฝ่าย ส่วน แผนก ตำแหน่ง ประเภทพนักงาน และทุกคน 6. สามารถโอนเงินเดือนพนักงานเพื่อนำส่งธนาคารต่างๆได้ เช่น ธนาคารกรุงเทพ,กรุงไทย เป็นต้น 7. สามารถคำนวณเงินเดือนได้หลายครั้งในงวดเดียวกัน
เงื่อนไขที่ผู้ใช้งานยอมรับ (ความต้องการไม่ใช่เชิงหน้าที่):

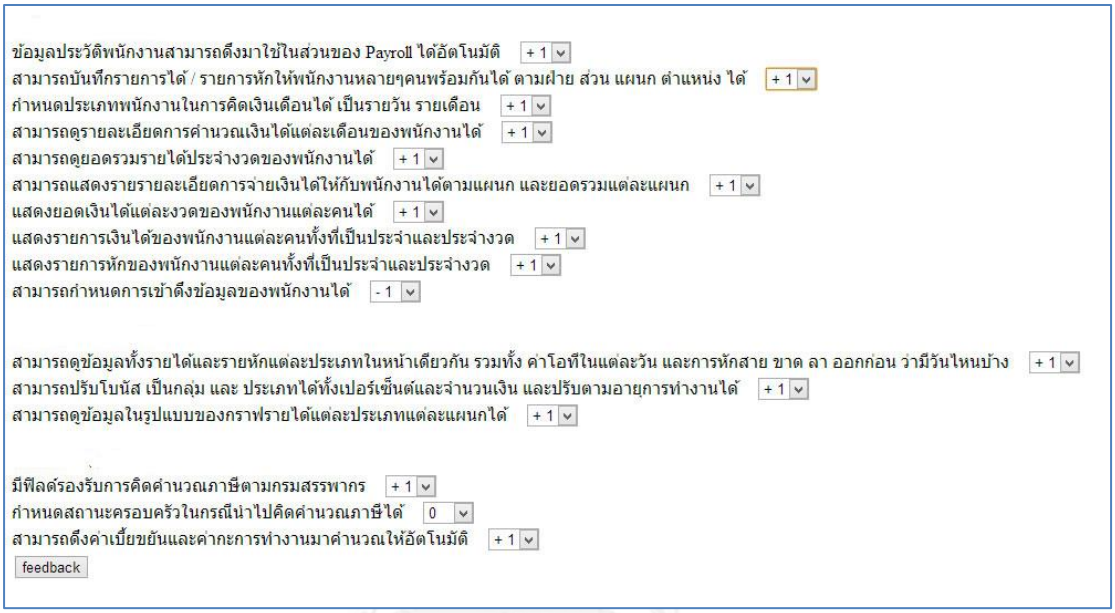
คลิ๊กดาวโหลดไฟล์

ภาพที่ 4.6 หน้าจอของส่วนการแสดงผลของการสืบค้น

จากภาพที่ 4.6 เป็นการแสดงผลออกมาเป็นเอกสารความต้องการซอฟต์แวร์เบื้องต้น ซึ่งสามารถนำออกมาเป็นไฟล์ PDF ได้

4.2.3 ส่วนผลป้อนกลับของผู้ใช้งาน

ส่วนผลป้อนกลับของผู้ใช้งานเป็นส่วนการให้คะแนนจากผู้ใช้งาน เพื่อปรับปรุงประสิทธิภาพของโครงสร้างความสัมพันธ์ให้มีการสืบค้นความต้องการเชิงหน้าที่ให้มีความแม่นยำมากขึ้น ดังภาพที่ 4.7



ภาพที่ 4.7 หน้าจอของส่วนผลป้อนกลับของผู้ใช้งาน

จากภาพที่ 4.7 เป็นส่วนผลป้อนกลับของผู้ใช้งาน ซึ่งเป็นการให้คะแนนจากผู้ใช้งาน หลังจากสืบค้น เป็นการสอบถามว่าผู้ใช้งานได้ความต้องการเชิงหน้าที่ที่ถูกต้องหรือไม่ เพื่อเป็นการเรียนรู้พัฒนาประสิทธิภาพของโครงสร้างความสัมพันธ์

4.3 สภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือ

สภาพแวดล้อมที่ใช้ในการพัฒนาเครื่องมือเพื่อการพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต ผู้วิจัยแนะนำฮาร์ดแวร์ และซอฟต์แวร์ ซึ่งมีรายละเอียดขั้นต่ำดังนี้

4.3.1 ฮาร์ดแวร์

ฮาร์ดแวร์สำหรับพัฒนาเครื่องมือเพื่อการพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต มีรายละเอียดขั้นต่ำดังนี้

1. เครื่องคอมพิวเตอร์ หน่วยประมวลผล อินเทล เพนเทียม 2.80 กิกะเฮิร์ทซ์ (Intel® Pentium® Processor 2.80 GHz)

2. หน่วยความจำ (Memory) 2.0 กิกะไบต์

3. จานบันทึกแบบแข็ง (Hard disk) ความจุ 160 กิกะไบต์

4.3.2 ซอฟต์แวร์

ซอฟต์แวร์สำหรับพัฒนาเครื่องมือเพื่อการพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต มีรายละเอียดดังนี้

1. ระบบปฏิบัติการไมโครซอฟท์วินโดวส์ 7 (Microsoft Windows 7)

2. อาปาเช่ เว็บเซิร์ฟเวอร์ (Apache Web Server) รุ่น 2.2.4 เพื่อให้เครื่องคอมพิวเตอร์มีคุณสมบัติเป็นเว็บเซิร์ฟเวอร์ (Web Server)

3. ภาษาพีเอชพี (PHP Language) รุ่น 5.2.6 สำหรับพัฒนาเว็บเพจแบบไดนามิกด้วยภาษาพีเอชพี (PHP)

4. มายเอสคิวแอล (MySQL) รุ่น 5.0.51b สำหรับจัดเก็บข้อมูลของเครื่องมือ

5. เจดีเค (JDK – Java Development Kit) รุ่น 7u2 สำหรับวินโดว์ เพื่อพัฒนาโปรแกรมด้วยภาษาจาวา (Java)

6. เกล (Grails) รุ่น 1.3.7 เป็นเฟรมเวิร์คเปิดต้นรหัสสำหรับพัฒนาเว็บแอปพลิเคชั่นที่เขียนด้วยภาษากรูวี ซึ่งอยู่บนแพลตฟอร์มจาวา (Java Platform)

7. โทดฟอร์มายเอสคิวแอล (Toad for MySQL) รุ่น 6.7.0.1524 เพื่อบริหารจัดการฐานข้อมูลมายเอสคิวแอล ผ่านโปรแกรมจัดการ

8. โปรแกรมอีดิพลัส (EditPlus) รุ่น 3.50 เพื่อเขียนรหัสต้นฉบับ (Source Code) ภาษาพีเอชพี

9. โปรแกรมเล็กซ์โต (LexTo) ของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ เพื่อตัดคำภาษาไทย ถูกพัฒนาด้วยภาษาจาวา แต่ผู้วิจัยนำมารันเป็นเว็บเซอร์วิช (Web Service) ด้วยภาษากรูวี่ เพื่อเชื่อมต่อกับเว็บแอปพลิเคชั่นที่ถูกพัฒนาด้วยภาษาพีเอชพีได้

4.3.3 ฐานข้อมูล

ฐานข้อมูลสำหรับพัฒนาเครื่องมือเพื่อการพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต มีรายละเอียดดังตาราง ภาคผนวก ข

4.4 การทดสอบการทำงานของเครื่องมือ

การทดสอบเครื่องมือที่พัฒนาขึ้น ได้ทำการทดสอบแบบกล่องดำ (Black-Box Testing) ซึ่งสนใจเฉพาะข้อมูลนำออกจากการตอบสนองของระบบ เพื่อทดสอบว่าเครื่องมือสามารถใช้ในการทดสอบการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

การสนับสนุนการดำเนินงานตามขั้นตอนของกรอบงานสนับสนุนการวิเคราะห์ช่องว่างที่นำเสนอ โดยมีขั้นตอนการทดสอบเครื่องมือดังนี้

1) กำหนดวัตถุประสงค์การทดสอบ

2) ออกแบบกรณีทดสอบ แสดงตัวอย่างได้ดังตารางที่ 4.7

3) เตรียมข้อมูลสำหรับการทดสอบ ใช้กรณีศึกษาเกี่ยวกับการทำซอฟต์แวร์ดีพลอยเมนต์

4) ทำการทดสอบ ตามขั้นตอนและกรณีทดสอบ

5) สรุปผลการทดสอบ

ตารางที่ 4.7 ตัวอย่างกรณีทดสอบการสืบค้นความต้องการเชิงหน้าที่ด้วยคำสำคัญ

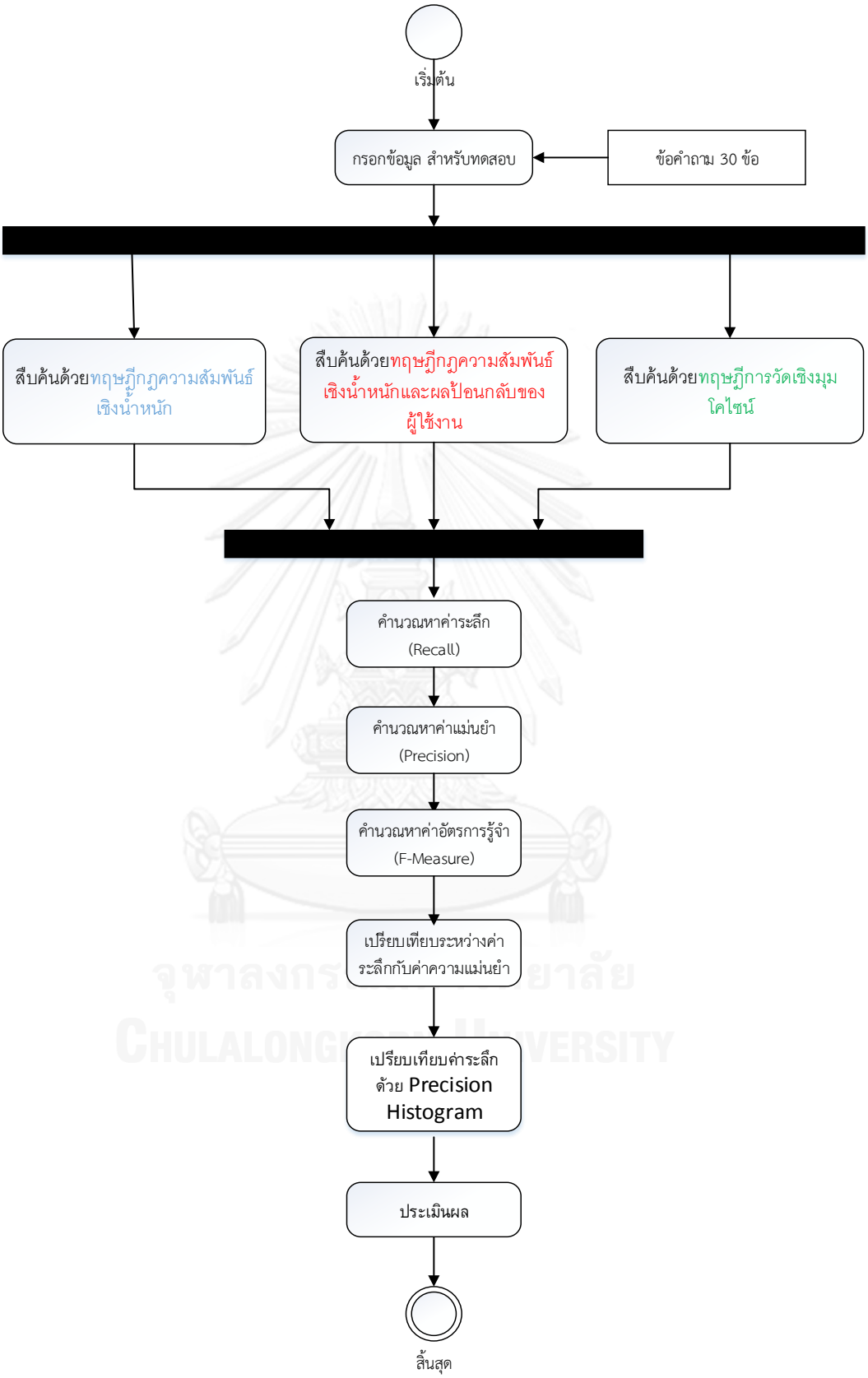
หัวข้อ	รายละเอียด
ชื่อระบบ	ระบบการสืบค้นความต้องการเชิงหน้าที่ด้วยคำสำคัญ โดยใช้โครงสร้างความสัมพันธ์
หน้าที่หลัก	การสืบค้นความต้องการเชิงหน้าที่ด้วยคำสำคัญ โดยใช้โครงสร้างความสัมพันธ์
กรณีทดสอบ	ใช้โครงสร้างความสัมพันธ์
วัตถุประสงค์	เพื่อทดสอบการสืบค้นความต้องการเชิงหน้าที่ด้วยคำสำคัญ โดยใช้โครงสร้างความสัมพันธ์
ผู้ใช้งาน	วิศวกรความต้องการ
ข้อมูลนำเข้า	คำสำคัญจากความต้องการของผู้ใช้งาน
ผลลัพธ์ที่คาดหวัง	- สามารถแสดงความต้องการเชิงหน้าที่มีคำสำคัญที่ใช้สำหรับสืบค้นปรากฏอยู่ได้ - สามารถแสดงความต้องการเชิงหน้าที่ที่มีความสัมพันธ์กับคำสำคัญที่ถูกนำมาใช้สำหรับสืบค้น จากโครงสร้างความสัมพันธ์ได้
ผลการทดสอบ	เครื่องมือสามารถแสดงความต้องการเชิงหน้าที่ทั้งที่มีคำสำคัญที่ใช้สำหรับสืบค้นและที่มีความสัมพันธ์กันในโครงสร้างความสัมพันธ์
สรุปผลการทดสอบ	☑ ผ่าน ☐ ไม่ผ่าน

บทที่ 5
การประเมินผลลัพธ์ที่ได้จากวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต

จากแนวคิดวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต โดยนำวิธีการสืบค้นในงานวิจัยนี้มาเปรียบเทียบกับการสืบค้น เพื่อให้ทราบวิธีการสืบค้นที่เหมาะสม สำหรับใช้ในการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีตหรือไม่

ซึ่งจะมีขั้นตอนการประเมินวิธีการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต แสดงได้ดังภาพที่ 5.1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 5.1 ภาพรวมการประเมินผลลัพธ์

5.1 ข้อมูลของการทำทดสอบ

ข้อมูลของการทดสอบจะทำการแยกออกเป็น 6 หมวด โดยแบ่งตามคุณสมบัติของระบบในโครงการของซอฟต์แวร์ด้านบริหารทรัพยากรบุคคล แต่ละหมวดจะใช้ 5 ข้อคำถาม มีทั้งหมด 30 ข้อคำถาม ดังตามภาคผนวก ค

5.2 วิธีการสืบค้นความต้องการเชิงหน้าที่

จากการนำทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมาใช้ในการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลในอดีต ผู้วิจัยได้ทำการวัดประสิทธิผลในการสืบค้นความต้องการเชิงหน้าที่ 3 วิธี ดังนี้

1. การสืบค้นความต้องการเชิงหน้าที่ โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต ด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก

2. การสืบค้นความต้องการเชิงหน้าที่ โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต ด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน

3. การสืบค้นความต้องการเชิงหน้าที่ โดยใช้ความคล้ายคลึงกัน ด้วยทฤษฎีการวัดเชิงมุมโคไซน์

ซึ่งนำผลลัพธ์ความต้องการจากทั้งสามวิธีนำมาคำนวณค่าระลึก (Recall) ค่าความแม่นยำ (Precision) ค่าอัตราการรู้จำ (F-measure) และ Precision Histograms

5.3 วิธีประเมินผลการทดลอง

5.3.1 ค่าระลึก

การประเมินผลการทดลองด้วยการค่าระลึก เพื่อวัดประสิทธิภาพของแต่ละวิธีการในการสืบค้นผลลัพธ์ที่เกี่ยวข้องออกมาได้ โดยการนำจำนวนผลลัพธ์ที่ค้นพบแล้วตรงตามต้องการทั้งหมดหารกับจำนวนผลลัพธ์ที่ตรงตามต้องการทั้งหมด ค่าระลึกที่ได้อยู่ระหว่าง 0 กับ 1 ซึ่งยิ่งได้ค่าระลึกมากก็ยิ่งดี ดังสมการที่ (5.1)

$$\text{Re}\,call = \frac{|R \cap A|}{|R|} \qquad (5.1)$$

จากสมการที่ (5.1) ข้อคำถาม 1 “เก็บข้อมูลประวัติพนักงาน” ใช้วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักพบจำนวนผลลัพธ์ที่ค้นพบแล้วตรงตามต้องการทั้งหมด ($|R \cap A|$) เท่ากับ 16 และจำนวนผลลัพธ์ที่ตรงตามต้องการทั้งหมด ($|R|$) เท่ากับ 45 มาแทนสมการเป็น 16 / 45 เท่ากับ 0.3556

5.3.2 ค่าความแม่นยำ

การประเมินผลการทดลองด้วยการค่าความแม่นยำ เพื่อวัดประสิทธิภาพของแต่ละวิธีการในความแม่นยำในการสืบค้นผลลัพธ์ โดยการนำจำนวนผลลัพธ์ที่ค้นพบแล้วตรงตามต้องการทั้งหมดหารกับจำนวนผลลัพธ์ที่ถูกค้นพบทั้งหมด ค่าความแม่นยำที่ได้อยู่ระหว่าง 0 กับ 1 ซึ่งยิ่งได้ค่าความแม่นยำมากก็ยิ่งดี ดังสมการที่ (5.2)

$$\Pr ecision = \frac{|R \cap A|}{|A|} \qquad (5.2)$$

จากสมการที่ (5.2) ข้อคำถาม 1 “เก็บข้อมูลประวัติพนักงาน” ใช้วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักพบจำนวนผลลัพธ์ที่ค้นพบแล้วตรงตามต้องการทั้งหมด ($|R \cap A|$) เท่ากับ 16 และจำนวนผลลัพธ์ที่ถูกค้นพบทั้งหมด ($|A|$) เท่ากับ 29 มาแทนสมการเป็น 16 / 29 เท่ากับ 0.5517

5.3.3 ค่าอัตราการรู้จำ

การประเมินผลการทดลองด้วยการค่าอัตราการรู้จำ เพื่อหาค่าเฉลี่ยของค่าระลึกและค่าแม่นยำ โดยการนำสองคูณผลคูณของค่าระลึกและค่าแม่นยำหารกับผลบวกของค่าระลึกและค่าแม่นยำ ค่าอัตราการรู้จำที่ได้อยู่ระหว่าง 0 กับ 1 ซึ่งยิ่งได้ค่าอัตราการรู้จำมากก็ยิ่งดี ดังสมการที่ (5.3)

$$F - Measure = 2 \times \frac{\Pr ecision \times \mathrm{Re}\, call}{\Pr ecision + \mathrm{Re}\, call} \qquad (5.3)$$

จากสมการที่ (5.3) ข้อคำถาม 1 “เก็บข้อมูลประวัติพนักงาน” ใช้วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักพบค่าระลึกเท่ากับ 0.3556 และค่าความแม่นยำเท่ากับ 0.5517 มาแทนสมการเป็น (2 x 0.5517 x 0.3556) / (0.5517 + 0.3556) เท่ากับ 0.4324

5.3.4 Precision Histogram

การประเมินผลการทดลองด้วย Precision Histogram เพื่อเป็นการเปรียบเทียบขั้นตอนวิธี (Algorithm) 2 วิธีการ โดยนำค่าระลึกของแต่ละข้อคำถามของทั้ง 2 วิธีการมาเปรียบเทียบกัน ดังสมการที่ (5.4)

$$RP_{A/B}(i) = RP_A(i) - RP_B(i) \qquad (5.4)$$

จากสมการที่ (5.4) ข้อคำถาม 1 “เก็บข้อมูลประวัติพนักงาน” เมื่อใช้วิธีการที่ 1 วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก ได้ค่าระลึกเท่ากับ 0.3556 และวิธีการที่ 2 วิธีการสืบค้นด้วยทฤษฎีการวัดเชิงมุมโคไซน์ ได้ค่าระลึกเท่ากับ 0.0889 มาแทนสมการเป็น 0.3556 – 0.0889 จะได้ค่าเท่ากับ 0.2667

5.3.5 ผลลัพธ์ประเมินผลจากค่าระลึก ค่าความแม่นยำ และค่าอัตรารู้จำ

การประเมินผลจากค่าระลึก ค่าความแม่นยำ และค่าอัตรารู้จำ ทั้งสามวิธีการสืบค้น ได้ค่าตามตารางที่ 5.1

ตารางที่ 5.1 ผลลัพธ์การประเมินจากค่าระลึก ค่าความแม่นยำ และค่าอัตราการรู้จำ

ข้อคำถาม	วิธีการสืบค้นความต้องการเชิงหน้าที่								
	กฎความสัมพันธ์เชิงน้ำหนัก			กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน			วัดความคล้ายคลึงเชิงมุมโคไซน์		
	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ
1	0.5671	0.2514	0.3483	0.5671	1	0.7237	0.4329	0.2423	0.3107
2	0.5183	0.2374	0.3257	0.5183	1	0.6827	0.4939	0.2813	0.3584
3	0.3537	0.2086	0.2624	0.3537	1	0.5225	0.2561	0.2069	0.2289
4	0.628	0.1813	0.2814	0.628	1	0.7715	0.5183	0.2033	0.2921
5	0.4451	0.1978	0.2739	0.4451	1	0.616	0.3659	0.2027	0.2609
6	0.6912	0.0856	0.1524	0.6912	1	0.8174	0.5588	0.0898	0.1548
7	0.3088	0.0814	0.1288	0.3088	1	0.4719	0.2059	0.069	0.1033
8	0.5	0.0835	0.1432	0.5	1	0.6667	0.3235	0.1068	0.1606
9	0.2647	0.072	0.1132	0.2647	1	0.4186	0.2059	0.0819	0.1172
10	0.3971	0.1849	0.2523	0.3971	1	0.5684	0.2059	0.1333	0.1618
11	0.2816	0.451	0.3467	0.2816	1	0.4395	0.2245	0.4331	0.2957

ตารางที่ 5.1 ผลลัพธ์การประเมินจากค่าระลึก ค่าความแม่นยำ และค่าอัตราการรู้จำ (ต่อ)

ข้อคำถาม	วิธีการสืบค้นความต้องการเชิงหน้าที่								
	กฎความสัมพันธ์เชิงน้ำหนัก			กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน			วัดความคล้ายคลึงเชิงมุมโคไซน์		
	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ	ค่าระลึก	ค่าความแม่นยำ	ค่าอัตราการรู้จำ
12	0.6041	0.2561	0.3597	0.6041	1	0.7532	0.1714	0.4516	0.2485
13	0.2327	0.5481	0.3266	0.2327	1	0.3775	0.1306	0.4571	0.2032
14	0.5592	0.5074	0.532	0.5592	1	0.7173	0.2653	0.5603	0.3601
15	0.1796	0.4231	0.2521	0.1796	1	0.3045	0.1714	0.4516	0.2485
16	0.3913	0.1579	0.225	0.3913	1	0.5625	0.2087	0.12	0.1524
17	0.5739	0.0946	0.1624	0.5739	1	0.7293	0.3739	0.0872	0.1414
18	0.2957	0.1	0.1495	0.2957	1	0.4564	0.1913	0.0767	0.1095
19	0.487	0.2051	0.2887	0.487	1	0.655	0.2	0.1271	0.1554
20	0.6087	0.1296	0.2137	0.6087	1	0.7568	0.2696	0.2793	0.2743
21	0.7213	0.234	0.3534	0.7213	1	0.8381	0.3279	0.4938	0.3941
22	0.7213	0.2115	0.3271	0.7213	1	0.8381	0.3443	0.1981	0.2515
23	0.7459	0.2407	0.364	0.7459	1	0.8545	0.3443	0.525	0.4158
24	0.8115	0.1514	0.2552	0.8115	1	0.8959	0.3033	0.1233	0.1754
25	0.7951	0.1745	0.2861	0.7951	1	0.8858	0.3689	0.1923	0.2528
26	0.4	0.093	0.1509	0.4	1	0.5714	0.15	0.0435	0.0674
27	0.45	0.0191	0.0367	0.45	1	0.6207	0.1	0.0064	0.012
28	0.25	0.0294	0.0526	0.25	1	0.4	0	0	0
29	0.15	0.1429	0.1463	0.15	1	0.2609	0	0	0
30	0.1	0.0093	0.017	0.1	1	0.1818	0	0	0

จากตารางที่ 5.1 พบว่าค่าระลึกทั้ง 30 ข้อคำถามได้ผลลัพธ์ในทิศทางเดียวกัน วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน และวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีค่าระลึกที่เท่ากัน และมากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ ทั้ง 30 ข้อคำถาม โดยมีค่าระลึกเฉลี่ย คือ 0.4678 และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ มีค่าระลึกเฉลี่ย คือ 0.2571

ค่าความแม่นยำในวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน ทั้ง 30 ข้อคำถาม มีค่าความแม่นยำมากที่สุดใน 3 วิธีการ โดยมีค่าความ

แม่นยำเฉลี่ย คือ 1 ส่วนวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ มีค่าความแม่นยำที่ใกล้เคียงกัน จากการประเมินด้วย 30 ข้อคำถาม พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีค่าความแม่นยำที่มากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์จำนวน 17 ข้อคำถาม แต่ค่าความแม่นยำเฉลี่ยทั้ง 30 ข้อคำถาม พบว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์มีค่าความแม่นยำเฉลี่ยมากกว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ โดยมีค่าความแม่นยำเฉลี่ย คือ 0.2081 และวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีค่าความแม่นยำเฉลี่ย คือ 0.1921

ซึ่งเมื่อนำค่าระลึกและค่าแม่นยำ มาเทียบเป็นค่าอัตราการรู้จำ พบว่าในวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน ทั้ง 30 ข้อคำถาม มีค่าอัตราการรู้จำมากที่สุดใน 3 วิธีการ ซึ่งค่าอัตราการรู้จำเฉลี่ย คือ 0.612 วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีค่าอัตราการรู้จำเฉลี่ยรองลงมา คือ 0.2376 และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ มีค่าอัตราการรู้จำเฉลี่ยน้อยที่สุด คือ 0.1969 โดย 22 ข้อคำถามจากทั้งหมด 30 ข้อคำถาม พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์มีค่าอัตราการรู้จำมากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์

จะเห็นได้ว่าค่าระลึก ค่าแม่นยำ และค่าอัตราการรู้จำของวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งานมีค่ามากที่สุดจึงเป็นวิธีที่ดีที่สุดจากทั้งสามวิธีการ ที่ให้ได้ความต้องการของซอฟต์แวร์ที่ตรงตามต้องการและครบถ้วนมากที่สุดจากทั้งสามวิธีการ

5.3.6 ผลลัพธ์ประเมินผลการทดลองด้วยการเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำ

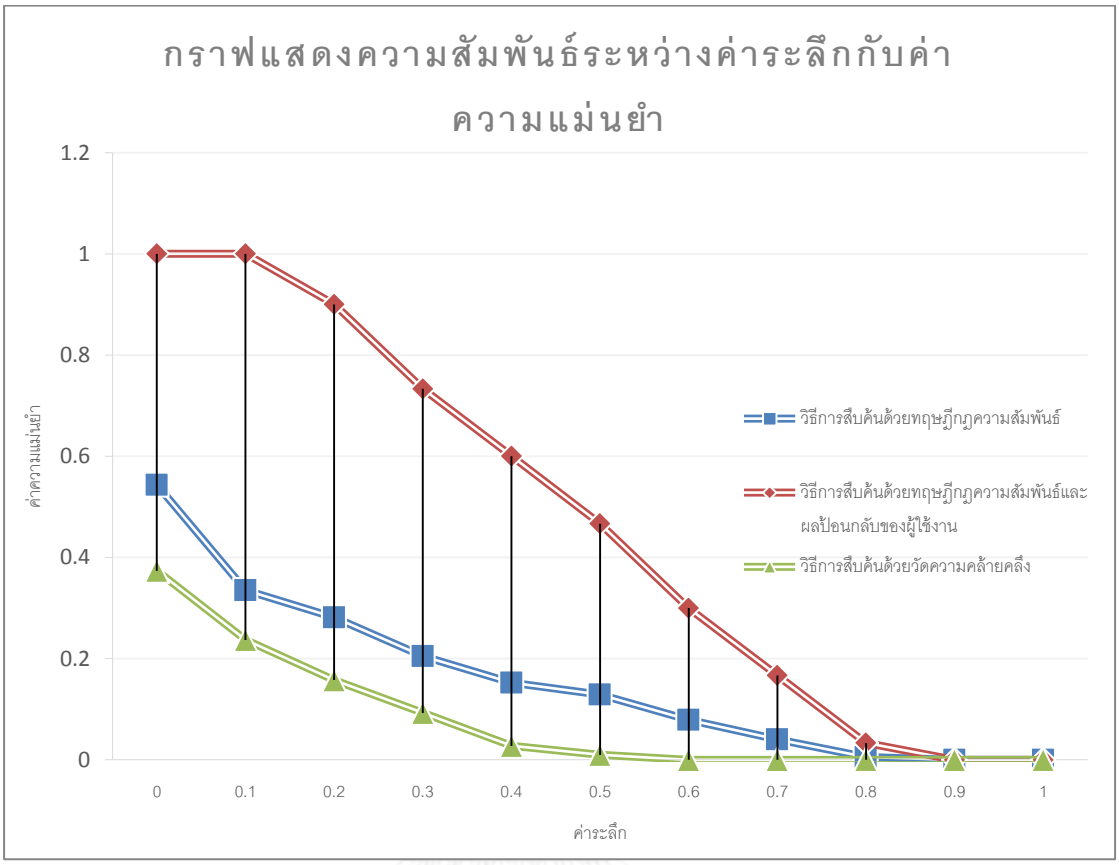
การประเมินผลการทดลองด้วยการเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำ เพื่อให้ทราบว่าเมื่อค่าระลึกเท่าไร แล้วจะได้รับค่าความแม่นยำสูงสุดเท่าไร โดยนำเอาผลลัพธ์ของวิธีการนั้นแต่ละข้อคำถามมาเฉลี่ยกัน เพื่อให้ทราบถึงค่าความแม่นยำโดยประมาณ แต่ละช่วงค่าระลึก เช่น ความต้องการเชิงหน้าที่ที่ได้จากการสืบค้น ในช่วงที่ได้ค่าระลึกอยู่ระหว่าง 0 กับ 0.0999 ถือว่าเป็นค่าระลึกที่ 0 ซึ่งนำค่าความแม่นยำที่ได้ทั้งหมดในช่วงนั้นเฉลี่ยกันเป็น 0.5431 เป็นต้น ตามตาราง ภาคผนวก ง จำนวน 30 ข้อคำถาม ซึ่งนำตารางทั้งหมดมาเฉลี่ยได้ตามตารางที่ 5.3 ซึ่งสามารถเขียนในรูปกราฟความสัมพันธ์ได้ดังภาพที่ 5.2

ตารางที่ 5.2 ผลลัพธ์ค่าเฉลี่ยการเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำ

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึงเชิง มุมโคไซน์
0	0.5431	1	0.3737
10	0.3353	1	0.2369
20	0.2815	0.9	0.1576
30	0.2047	0.7333	0.0927
40	0.1524	0.6	0.0275
50	0.1289	0.4667	0.0099
60	0.0788	0.3	0
70	0.0404	0.1667	0
80	0.0052	0.0333	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จากตารางที่ 5.2 สามารถนำมาแปลงเป็นรูปกราฟ ดังภาพที่ 5.2



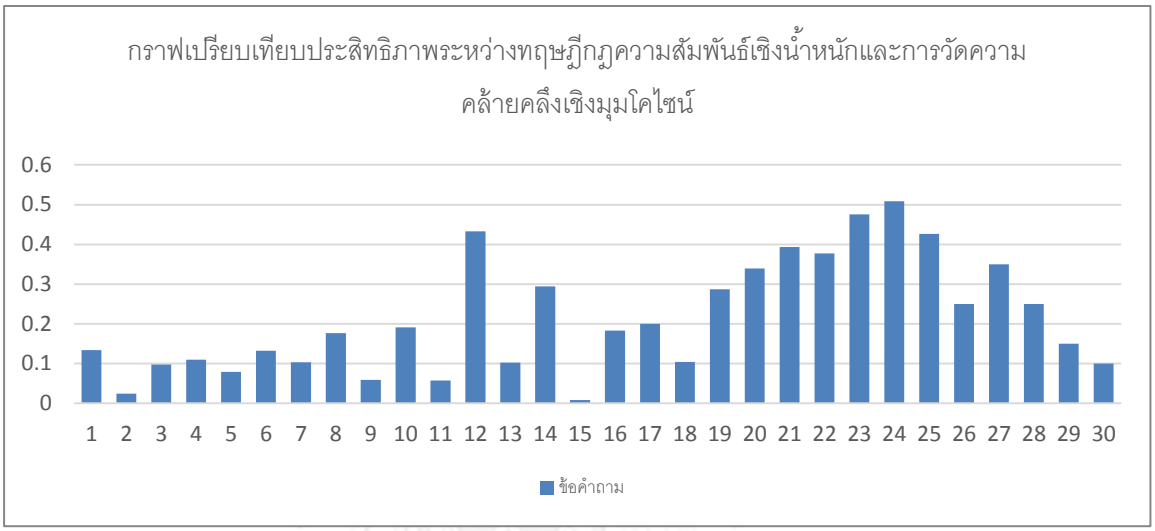
ภาพที่ 5.2 กราฟแสดงความสัมพันธ์ระหว่างค่าระลึกกับค่าความแม่นยำ

จากภาพที่ 5.2 พบว่าในแต่ละช่วงของค่าระลึกเดียวกันนั้น วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน มีความแม่นยำมากที่สุดจาก 3 วิธีการ รองลงมาเป็นวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ และวิธีการสืบค้นด้วยวัดความคล้ายคลึงตามลำดับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5.3.7 ผลลัพธ์ประเมินผลการทดลองด้วย Precision Histogram

การประเมินผลเปรียบเทียบทั้งสองวิธีการสืบค้นด้วย Precision Histogram แต่ละข้อคำถาม ซึ่งจะทำการเปรียบเทียบระหว่างทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและการวัดความคล้ายคลึงเชิงมุมโคไซน์ สามารถนำมาสร้างกราฟเปรียบเทียบได้ ตามภาพที่ 5.3



ภาพที่ 5.3 กราฟเปรียบเทียบประสิทธิภาพระหว่างทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและการวัดความคล้ายคลึงเชิงมุมโคไซน์

จากภาพที่ 5.3 กราฟเปรียบเทียบประสิทธิภาพระหว่างทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและการวัดความคล้ายคลึงเชิงมุมโคไซน์ พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมีค่าระลึกที่มากกว่ามากกว่าการวัดความคล้ายคลึงเชิงมุมโคไซน์ ในทุกข้อคำถาม ดังนั้นวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมีโอกาสที่จะพบความต้องการของซอฟต์แวร์ที่ต้องการมากกว่าการวัดความคล้ายคลึงเชิงมุมโคไซน์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 6
สรุปผลการวิจัย

จากการที่ผู้วิจัยนำทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมาใช้ในการพัฒนาการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต สามารถสรุปผลการวิจัย และอภิปรายผลการวิจัยได้ดังรายละเอียดต่อไปนี้

6.1 สรุปผลการวิจัย

จากการประเมินผลการสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต ทั้ง 3 วิธีการ โดยใช้ข้อมูลทดสอบจากเอกสารกำหนดความต้องการซอฟต์แวร์ของระบบบริหารทรัพยากรบุคคล พบว่าวิธีการการสืบค้นความต้องการของซอฟต์แวร์ ด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน สามารถได้รับความต้องการของซอฟต์แวร์ตรงตามต้องการมากที่สุด รองลงมาเป็นทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก และทฤษฎีการวัดเชิงมุมโคไซน์ ตามลำดับ

จากตารางที่ 5.1 ในบทที่ 5 ผลลัพธ์การประเมินจากค่าระลึก ค่าความแม่นยำ และค่าอัตราการรู้จำ สามารถสรุปผลการวิจัย ได้ดังนี้

จากผลลัพธ์การประเมินจากค่าระลึก ทั้ง 3 วิธีการ พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน และวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีค่าระลึกเฉลี่ยที่มากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ ดังนั้นวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักสามารถได้รับความต้องการซอฟต์แวร์ที่ตรงตามต้องการมากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์

จากผลลัพธ์การประเมินจากค่าความแม่นยำ ทั้ง 3 วิธีการ พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งานมีค่าความแม่นยำเฉลี่ยที่มากที่สุด วิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ และวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์มีค่าความแม่นยำเฉลี่ยรองลงมา เท่ากับ 0.2081 และ 0.1921 ตามลำดับ แต่จากข้อคำถาม 30 ข้อ พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ มีจำนวน 17 ใน 30 ข้อคำถาม พบว่ามีความแม่นยำมากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ ดังนั้นวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ และวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์มีค่าความแม่นยำที่ใกล้เคียงกัน

จากผลลัพธ์การประเมินจากค่าอัตราการรู้จำ ทั้ง 3 วิธีการ พบว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งานมีค่าอัตราการรู้จำเฉลี่ยที่มากที่สุดเท่ากับ 0.612 วิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์ และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ มีค่าอัตราการรู้จำรองลงมา เท่ากับ 0.2376 และ 0.1969 ตามลำดับ ดังนั้นทฤษฎีกฎความสัมพันธ์จะได้ความต้องการซอฟต์แวร์ที่ตรงตามความต้องการและแม่นยำมากว่าทฤษฎีวัดความคล้ายคลึงเชิงมุมโคไซน์

จากภาพที่ 5.2 ในบทที่ 5 เป็นการประเมินด้วยการเทียบค่าระลึกกับค่าความแม่นยำ ซึ่งแสดงให้เห็นภาพโดยรวมของทั้ง 3 วิธีการ สามารถสรุปวิธีการสืบค้นที่มีประสิทธิผล เรียงลำดับจากมากไปน้อย คือ ทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน ทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก และการวัดความคล้ายคลึงเชิงมุมโคไซน์ เนื่องจากโครงสร้างความสัมพันธ์จากทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักเป็นการอิงกับความสัมพันธ์ของข้อมูล โดยมีการเทียบค่าความเชื่อมั่นที่เป็นค่าที่จะเกิดโอกาสเมื่อพบคำนี้แล้วจะเกิดคำนั้นตามมา รวมถึงคำที่มีความสัมพันธ์มีความเป็นคำเฉพาะของระบบนั้น ยิ่งมีโอกาสคำที่ถูกขยายจะได้ผลลัพธ์ที่ถูกต้องตามไปด้วย ส่วนการผลป้อนกลับของผู้ใช้งาน เป็นส่วนที่ช่วยเสริมทำให้ค่าความแม่นยำมากขึ้น โดยให้ผู้ใช้งานมีส่วนร่วมในการปรับปรุงโครงสร้างความสัมพันธ์ให้ดีขึ้น จึงทำให้การเทียบค่าระลึกกับค่าความแม่นยำของโครงสร้างความสัมพันธ์จากทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งาน มีประสิทธิภาพดีที่สุดใน 3 วิธีการ

จากภาพที่ 5.3 ในบทที่ 5 เป็นผลการประเมินด้วยวิธี Precision Histogram ซึ่งเปรียบเทียบระหว่างวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ จากผลลัพธ์ที่ได้พบว่าค่าความระลึกทั้งข้อคำถาม 30 ข้อของวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักมีค่ามากกว่าวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ทั้งหมด ซึ่งสามารถสรุปได้ว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักสามารถช่วยให้ได้ความต้องการซอฟต์แวร์ที่ตรงตามต้องการได้มากกว่าทฤษฎีวัดความคล้ายคลึงเชิงมุมโคไซน์

ดังนั้นจากผลการประเมินทั้งหมดสามารถสรุปได้ว่าวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนักและผลป้อนกลับของผู้ใช้งานสามารถช่วยให้ได้ความต้องการซอฟต์แวร์ที่ตรงตามต้องการมากที่สุดใน 3 วิธีการ รองลงมาเป็นวิธีการสืบค้นด้วยทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก และวิธีการสืบค้นด้วยวัดความคล้ายคลึงเชิงมุมโคไซน์ ตามลำดับ

6.2 อภิปรายผลการวิจัย

การสืบค้นความต้องการของซอฟต์แวร์ที่เป็นภาษาไทย โดยใช้ความสัมพันธ์ของข้อมูลโครงการซอฟต์แวร์ในอดีต แบ่งออกเป็น 2 ส่วน ดังนี้

6.2.1 ทฤษฎีกฎความสัมพันธ์เชิงน้ำหนัก

ช่วยสร้างโครงสร้างความสัมพันธ์ โดยอาศัยเอกสารกำหนดความต้องการซอฟต์แวร์ของโครงการในอดีต ทำให้ได้ความต้องการซอฟต์แวร์ที่มีความสัมพันธ์กับข้อคำถามที่ใช้ในการสืบค้น เพื่อได้รับความต้องการซอฟต์แวร์ที่มากขึ้น ซึ่งขั้นตอนการรวบรวมความต้องการซอฟต์แวร์จำเป็นต้องรวบรวมและวิเคราะห์ความต้องการซอฟต์แวร์ที่คาดว่าเป็นส่วนหนึ่งของซอฟต์แวร์ เพื่อนำไปวิเคราะห์ร่วมกันผู้ใช้งานให้มากที่สุด สำหรับได้รับความต้องการที่ตรงตามความต้องการของผู้ใช้งานให้มากที่สุด

จากการที่ได้รับความต้องการซอฟต์แวร์ที่ตรงตามความต้องการของผู้ใช้งานมากขึ้น ทำให้โอกาสที่จะรวบรวมความต้องการซอฟต์แวร์ไม่ครบถ้วนลดน้อยลง หากไม่มีความต้องการซอฟต์แวร์เพิ่มภายหลัง จึงทำให้โอกาสที่โครงการล่าช้าลดน้อยลง

6.2.2 ทฤษฎีผลป้อนกลับของผู้ใช้งาน

ช่วยให้โครงสร้างความสัมพันธ์มีความแม่นยำมากขึ้น ซึ่งทำให้ความต้องการซอฟต์แวร์ที่ได้รับตรงตามความต้องการของผู้ใช้งานมากขึ้น หลังจากดำเนินโครงการเสร็จ เพื่อปรับปรุงโครงสร้างความสัมพันธ์มีความแม่นยำมากขึ้นในการสืบค้นความต้องการซอฟต์แวร์ครั้งต่อไป

6.3 ปัญหาและข้อจำกัดของงานวิจัย

1. การตัดคำภาษาไทยในงานวิจัยยังไม่สามารถตัดคำได้ถูกต้อง 100% ซึ่งทำให้ข้อมูลตั้งต้นขาดความสมบูรณ์ถูกต้องและส่วนที่เป็นใจความสำคัญอาจถูกบิดเบือนไป ซึ่งมีผลกับการสร้างโครงสร้างความสัมพันธ์ให้มีคุณภาพ

2. คำศัพท์ คำหยุดและคำที่มีความหมายเหมือนกันของภาษาไทยมีฐานข้อมูลที่ถูกรวบรวมไว้พร้อมใช้งานจำนวนน้อย รวมถึงภาษาไทยที่มีการเปลี่ยนแปลงของภาษาได้ตลอดเวลาเนื่องมาจากภาษาต่างประเทศ และภาษาวัยรุ่น เป็นต้น

3. โครงสร้างความสัมพันธ์เป็นการนำข้อมูลที่มีอยู่ในฐานข้อมูลมาสืบค้นความสัมพันธ์กัน หากฐานข้อมูลขนาดใหญ่ก็สามารถทำให้โครงสร้างความสัมพันธ์ที่ถูกสร้างขึ้นมีความแม่นยำและความน่าเชื่อถือได้มากขึ้น

6.4 แนวทางการวิจัยในอนาคต

1. นำข้อมูลการเปลี่ยนแปลงความต้องการของซอฟต์แวร์ ระยะเวลาในการดำเนินโครงการ ปัญหาต่างๆในโครงการ และข้อมูลอื่นๆในโครงการซอฟต์แวร์ เป็นต้น นำมาเก็บไว้เป็นฐานข้อมูล เพื่อนำมาแสดงเป็นส่วนเสนอแนะเพิ่มเติมของเอกสารกำหนดความต้องการซอฟต์แวร์เบื้องต้น ซึ่งสามารถทำให้ทราบถึงปัญหาหรือข้อควรระวังที่จะเกิดขึ้นในระหว่างดำเนินโครงการล่วงหน้า และช่วยให้การดำเนินโครงการมีประสิทธิภาพและรวดเร็วมากขึ้น

2. นำความสัมพันธ์ของข้อมูลมาวิเคราะห์ สำหรับจัดกลุ่มองค์ความรู้ให้เป็นหมวดหมู่ที่ชัดเจนมากขึ้น ซึ่งสามารถช่วยลดเวลาในการสืบค้นความต้องการซอฟต์แวร์ของโครงการขนาดใหญ่

3. รองรับข้อคำถามเชิงปฏิเสธ เช่น ไม่ เป็นต้น มาช่วยในการรวบรวมความต้องการเชิงหน้าที่ที่ผู้ใช้งานไม่ต้องการออกไป เพื่อให้ระบบสามารถนำสิ่งที่ผู้ใช้งานไม่ต้องการออกไปจากเอกสารกำหนดความต้องการของซอฟต์แวร์ และลดความต้องการของซอฟต์แวร์ที่ไม่ตรงตามผู้ใช้งาน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายการอ้างอิง

[1] Joachim Karlsson, Software Requirements Prioritizing, IEEE Int. Conf. on Requirements Engineering - RE , pp. 110-116, 1996.

[2] Hajar Mat Jani, Jalan Kajang-Puchong, Selangor Darul Ehsan, Applying Case-Based Reasoning to software requirements specifications quality analysis system, 2010.

[3] Wayne C. Lim, Effects of Reuse on Quality, Productivity, and Economics, IEEE Software - SOFTWARE, vol. 11, no. 5, pp. 23-30, 1994.

[4] IEEE Recommended Practice for Software Requirements Specifications. IEEE Standard 830-1998.

[5] Warawee Kesorn, Similarity Measurement of Thai Documents Using Natural Language Processing, Chiang Mai University, 2010.

[6] Leonard Bloomfield, Language, New York : Henry Holt and Co., 1933.

[7] ปราณี กุลละวณิชย์, กัลยา ติงศภัทิย์ สุดาพร ลักษณียนาวิน และอมรา ประสิทธิ์รัฐสินธุ์, ภาษาทัศนา, พิมพ์ครั้งที่ 2. กรุงเทพฯ : วัฒนชัยการพิมพ์, 1992.

[8] Wuu Yang, Chey-Woei Tsay, Jien-Tsai Chan, On the applicability of the longest-match rule in lexical analysis, Computer Languages, Systems & Structures - CL , vol. 28, no. 3, pp. 273-288, 2002.

[9] Choochart Haruechaiyasak, Sarawoot Kongyoung, Chaianun Damrongrat, LearnLexTo: A Machine-Learning Based Word Segmentation for Indexing Thai Texts, International Conference on Information and Knowledge Management - CIKM, pp. 85-88, 2008.

[10] Patcharanut Daowadung, Yaw-Huei Chen, Stop Word in Readability Assessment of Thai Text, 12th IEEE International Conference on Advanced Learning Technologies, 2012.

[11] Ho Chung Wu, Robert Wing Pong Luk, Kam-fai Wong, Kui-lam Kwok, Interpreting TF-IDF term weights as making relevance decisions, ACM Transactions on Information Systems - TOIS, vol. 26, no. 3, pp. 1-37, 2008.

[12] Baeza-Yates & Ribeiro-Neto, Modern Information Retrieval, 2nd Edition, 2010.

[13] Ramakrishnan Srikant, Rakesh Agrawal, Mining Generalized Association Rules, Very Large Data Bases - VLDB, pp. 407-419, 1995.

[14] Rakesh Agrawal, Tomasz Imieliński, Arun N. Swami, Mining association rules between sets of items in large databases, Sigmod Record, vol. 22, no. 2, pp. 207-216, 1993.

[15] C. H. Cai, Ada Wai-chee Fu, C. H. Cheng, W. W. Kwong, Mining Association Rules with Weighted Items, International Database Engineering and Application Symposium - IDEAS, pp. 68-77, 1998.

[16] V.Vidya, Mining Weighted Association Rule using FP-tree, International Journal on Computer Science and Engineering (IJCSE), 2013.

[17] G.D. Ramkumar, Sanjay Ranka, Shalom Tsur, Weighted Association Rules: Model and Algorithm, Knowledge Discovery and Data Mining - KDD, 1997.

[18] Simone Stumpf, Erin Sullivan, Erin Fitzhenry, Ian Oberst, Weng-Keen Wong, Margaret Burnett, Integrating Rich User Feedback into Intelligent User Interfaces, IUI '08: Proceedings of the 13th international conference on intelligent user interfaces, 2008.

[19] Anas Mahmoud, Nan Niu, An Experimental Investigation of Reusable Requirements Retrieval, Information Reuse and Integration - IRI, 2010.

[20] Tao Zhang, Byungjeong Lee, Complementary Classification Techniques based Personalized Software Requirements Retrieval with Semantic Ontology and User Feedback, Conference on Computer and Information Technology - IEEECIT, pp. 1358-1363, 2010

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ก

คำอภิธานศัพท์

คำศัพท์	คำอธิบาย
วิศวกรความต้องการ	ผู้สร้างและจัดการข้อกำหนดความต้องการด้านระบบหรือซอฟต์แวร์
TF-IDF	Term Frequency-Inverse Document Frequency เป็นการวัดค่าน้ำหนักที่เอาไว้ประเมินค่าความสำคัญของคำนั้น ๆ ในกลุ่มของเอกสาร
กฎความสัมพันธ์เชิงน้ำหนัก	รูปแบบความสัมพันธ์ของข้อมูล 2 ชุดหรือมากกว่า เพื่อให้ทราบว่าถ้ามีข้อมูล A แล้วจะเกิดข้อมูล B ตามมา เพื่อคาดการณ์สิ่งที่จะเกิดล่วงหน้าได้ง่ายขึ้น
คำหยุด	คำที่เป็นคำขยายให้แก่คำอื่น ๆ แต่ไม่มีความหมายในตัวเอง ซึ่งจะไม่นำไปรวมกับการสืบค้น
ค่าเชื่อมั่น	ความน่าจะเป็นที่เกิดเหตุการณ์ A แล้วจะเกิดเหตุการณ์ B
ค่าสนับสนุน	ความน่าจะเป็นที่เกิดร่วมกันของเหตุการณ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ข

รายละเอียดโครงสร้างตารางที่ทำการจัดเก็บในฐานข้อมูล

ตารางอธิบายโครงสร้างข้อมูลในงานวิจัยนี้มีทั้งสิ้น 21 ตาราง โดยมีรายละเอียดโครงสร้างของแต่ละตารางดังแสดงต่อไปนี้

ตารางที่ ข.1 โครงสร้างข้อมูลของตาราง functional

คุณลักษณะ	คำอธิบาย	ชนิด
func_id	รหัสความต้องการเชิงหน้าที่	int(11)
func_name	ความต้องการเชิงหน้าที่	varchar(333)
fweight	ค่าน้ำหนักของความต้องการเชิงหน้าที่	float
avgweight	ค่าเฉลี่ยน้ำหนักของความต้องการเชิงหน้าที่	float
proj_id	รหัสโครงการ	int(11)
req_id	รหัสระบบย่อย	int(11)

ตารางที่ ข.2 โครงสร้างข้อมูลของตาราง idf

คุณลักษณะ	คำอธิบาย	ชนิด
term_id	รหัสคำ	int(11)
part	ประเภทหัวข้อ	varchar(20)
idf	ค่าความถี่ส่วนกลับของเอกสาร	float

ตารางที่ ข.3 โครงสร้างข้อมูลของตาราง itemset

คุณลักษณะ	คำอธิบาย	ชนิด
itemset_id	รหัสกลุ่มคำ	int(11)
itemset	กลุ่มรหัสคำ	varchar(100)
level	ระดับของกลุ่มคำ	int(11)
funcset	กลุ่มรหัสความต้องการเชิงหน้าที่	text
req_group	กลุ่มระบบย่อย	varchar(100)
sup	จำนวนครั้งที่กลุ่มคำปรากฏ	int(11)
wsupport	ค่าสนับสนุนเชิงน้ำหนัก	float

ตารางที่ ข.4 โครงสร้างข้อมูลของตาราง non_functional

คุณลักษณะ	คำอธิบาย	ชนิด
nfunc_id	รหัสความต้องการไม่ใช่เชิงหน้าที่	int(11)
nfunc_name	ความต้องการไม่ใช่เชิงหน้าที่	varchar(20)
proj_id	รหัสโครงการ	int(11)
req_id	รหัสระบบย่อย	int(11)

ตารางที่ ข.5 โครงสร้างข้อมูลของตาราง proj_purpose

คุณลักษณะ	คำอธิบาย	ชนิด
proj_purpose_id	รหัสวัตถุประสงค์ของโครงการ	int(11)
proj_purpose	วัตถุประสงค์ของโครงการ	varchar(333)
proj_id	รหัสโครงการ	int(11)

ตารางที่ ข.6 โครงสร้างข้อมูลของตาราง proj_scope

คุณลักษณะ	คำอธิบาย	ชนิด
proj_scope_id	รหัสขอบเขตของโครงการ	int(11)
proj_scope	ขอบเขตของโครงการ	varchar(333)
proj_id	รหัสโครงการ	int(11)

ตารางที่ ข.7 โครงสร้างข้อมูลของตาราง project

คุณลักษณะ	คำอธิบาย	ชนิด
proj_id	รหัสโครงการ	int(11)
proj_name	ชื่อโครงการ	varchar(333)

ตารางที่ ข.8 โครงสร้างข้อมูลของตาราง req_actor

คุณลักษณะ	คำอธิบาย	ชนิด
req_actor_id	รหัสผู้เกี่ยวข้อง	int(11)
req_actor	ผู้เกี่ยวข้อง	varchar(333)
proj_id	รหัสโครงการ	int(11)
req_id	รหัสระบบย่อย	int(11)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ข.9 โครงสร้างข้อมูลของตาราง req_group

คุณลักษณะ	คำอธิบาย	ชนิด
req_group_id	รหัสกลุ่มระบบย่อย	int(11)
req_id	รหัสระบบย่อย	int(11)

ตารางที่ ข.10 โครงสร้างข้อมูลของตาราง req_purpose

คุณลักษณะ	คำอธิบาย	ชนิด
req_purpose_id	รหัสวัตถุประสงค์ของระบบย่อย	int(11)
req_purpose	วัตถุประสงค์ของระบบย่อย	varchar(333)
proj_id	รหัสโครงการ	int(11)
req_id	รหัสระบบย่อย	int(11)

ตารางที่ ข.11 โครงสร้างข้อมูลของตาราง req_system

คุณลักษณะ	คำอธิบาย	ชนิด
req_system_id	รหัสสิ่งที่กระทบระบบ	int(11)
req_system	สิ่งที่กระทบระบบ	varchar(333)
proj_id	รหัสโครงการ	int(11)
req_id	รหัสระบบย่อย	int(11)

ตารางที่ ข.12 โครงสร้างข้อมูลของตาราง req_user

คุณลักษณะ	คำอธิบาย	ชนิด
srs_id	รหัสเอกสารกำหนดความต้องการของซอฟต์แวร์	int(11)
req_user_id	รหัสความต้องการของผู้ใช้งาน	int(11)
seq_id	ลำดับ	int(11)
term_id	รหัสคำ	int(11)

ตารางที่ ข.13 โครงสร้างข้อมูลของตาราง req_user_functional

คุณลักษณะ	คำอธิบาย	ชนิด
srs_id	รหัสเอกสารกำหนดความต้องการของซอฟต์แวร์	int(11)
termset	กลุ่มคำ	text
sweight	ค่าน้ำหนักของคำ	float
func_id	รหัสความต้องการเชิงหน้าที่	int(11)
func_name	ความต้องการเชิงหน้าที่	varchar(333)
avgweight	ค่าเฉลี่ยน้ำหนักของความต้องการเชิงหน้าที่	float
feedback	ค่าผลป้อนกลับ	int(11)

ตารางที่ ข.14 โครงสร้างข้อมูลของตาราง req_user_relationship

คุณลักษณะ	คำอธิบาย	ชนิด
srs_id	รหัสเอกสารกำหนดความต้องการของซอฟต์แวร์	int(11)
wrule_id	รหัสกฎความสัมพันธ์	int(11)
feedback	ค่าผลป้อนกลับ	int(11)

ตารางที่ ข.15 โครงสร้างข้อมูลของตาราง requirement

คุณลักษณะ	คำอธิบาย	ชนิด
req_id	รหัสระบบย่อย	int(11)
req_name	ระบบย่อย	varchar(333)
proj_id	รหัสโครงการ	int(11)

ตารางที่ ข.16 โครงสร้างข้อมูลของตาราง similarity

คุณลักษณะ	คำอธิบาย	ชนิด
sim_id	รหัสความคล้ายคลึง	int(11)
part	หัวข้อ	varchar(20)
part_id_a	รหัสที่ 1	int(11)
part_id_b	รหัสที่ 2	int(11)
sim	ค่าความคล้ายคลึง	float

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ข.17 โครงสร้างข้อมูลของตาราง stopword

คุณลักษณะ	คำอธิบาย	ชนิด
stop_id	รหัสคำหยุด	int(11)
stop_name	คำหยุด	varchar(333)

ตารางที่ ข.18 โครงสร้างข้อมูลของตาราง synonym

คุณลักษณะ	คำอธิบาย	ชนิด
syn_id	รหัสคำเหมือน	int(11)
root_name	คำหลัก	varchar(330)
syn_name	คำเหมือน	varchar(330)

ตารางที่ ข.19 โครงสร้างข้อมูลของตาราง term

คุณลักษณะ	คำอธิบาย	ชนิด
term_id	รหัสคำ	int(11)
term_name	คำ	varchar(333)
stop	เป็นคำหยุดหรือไม่	char(1)

ตารางที่ ข.20 โครงสร้างข้อมูลของตาราง term_weight

คุณลักษณะ	คำอธิบาย	ชนิด
term_weight_id	รหัสน้ำหนักคำ	int(11)
part	หัวข้อ	varchar(20)
part_id	รหัสของหัวข้อนั้น	int(11)
term_id	รหัสคำ	int(11)
num	จำนวน	int(11)
tf	ค่าความถี่ของคำ	float
weight	ค่าน้ำหนักของคำ	float
feedback	ค่าเฉลี่ยผลป้อนกลับ	float

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ข.21 โครงสร้างข้อมูลของตาราง wrule

คุณลักษณะ	คำอธิบาย	ชนิด
wrule_id	รหัสกฎความสัมพันธ์	int(11)
req_group	รหัสกลุ่มระบบย่อย	varchar(100)
itemset_id	รหัสกลุ่มคำ	int(11)
itemset_ant_id	รหัสกลุ่มคำที่	int(11)
antecedent	คำตั้งต้น	varchar(100)
consequent	คำที่เกิดตามหลัง	varchar(100)
wconfidence	ค่าความเชื่อมั่นเชิงน้ำหนัก	float
feedback	ค่าผลป้อนกลับ	int(11)

ตารางที่ ข.22 โครงสร้างข้อมูลของตาราง wrule_antecedent

คุณลักษณะ	คำอธิบาย	ชนิด
wrule_id	รหัสกฎความสัมพันธ์	int(11)
antecedent	คำตั้งต้น	int(11)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ค

ข้อคำถามสำหรับทำการทดสอบ

1. คุณสมบัติประวัติพนักงาน คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 1: เก็บข้อมูลประวัติพนักงาน

ข้อคำถาม 2: เก็บข้อมูลประวัติการทำงานของพนักงานในแต่ละคน

ข้อคำถาม 3: บันทึกรายละเอียดเพิ่มเติมและรูปภาพ

ข้อคำถาม 4: ปรับปรุงแก้ไขข้อมูลพนักงาน

ข้อคำถาม 5: บักทึกวันลาออกของพนักงาน

2. คุณสมบัติระบบบันทึกเวลา คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 6: บันทึกตารางการทำงานของพนักงาน

ข้อคำถาม 7: บันทึกการลางาน

ข้อคำถาม 8: บันทึกการขออนุญาตทำงานล่วงเวลา

ข้อคำถาม 9: บันทึกการขออนุญาตไม่ลงเวลาทำงาน

ข้อคำถาม 10: คำนวณเวลาสาย

3. คุณสมบัติระบบเงินเดือน คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 11: คำนวณเงินเดือน

ข้อคำถาม 12: คำนวณค่าทำงานล่วงเวลา

ข้อคำถาม 13: คำนวณภาษี

ข้อคำถาม 14: หักเงินมาสาย

ข้อคำถาม 15: พิมพ์หนังสือรับรองเงินเดือน

4. คุณสมบัติระบบฝึกอบรม คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 16: บันทึกรายละเอียดฝึกอบรม

ข้อคำถาม 17: บันทึกข้อมูลพนักงานที่ผ่านการฝึกอบรมแล้ว

ข้อคำถาม 18: ประเมินพนักงาน

ข้อคำถาม 19: บันทึกค่าใช้จ่ายในการอบรม

ข้อคำถาม 20: พิมพ์รายงานผลการอบรม

5. คุณสมบัติระบบสวัสดิการ คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 21: เก็บข้อมูลสวัสดิการ

ข้อคำถาม 22: บันทึกการเบิกสวัสดิการ

ข้อคำถาม 23: กำหนดรายละเอียดสวัสดิการ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ข้อคำถาม 24: กำหนดโควตาหรือสิทธิการเบิกแต่ละประเภทสวัสดิการให้พนักงานแต่ละคน

ข้อคำถาม 25: วิเคราะห์ข้อมูลของสวัสดิการแต่ละประเภท

6. คุณสมบัติระบบประเมินผล คำที่ใช้สืบค้นมีดังนี้

ข้อคำถาม 26: ประเมินผลการทำงาน

ข้อคำถาม 27: วิเคราะห์การทำงานของพนักงาน

ข้อคำถาม 28: พิมพ์ผลการประเมิน

ข้อคำถาม 29: กำหนดหัวข้อการประเมิน

ข้อคำถาม 30: ประเมินประจำเดือนหรือปี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ง

ผลลัพธ์การเปรียบเทียบระหว่างค่าระลึกกับค่าความแม่นยำแต่ละข้อคำถาม

ตารางที่ ง.1 ผลลัพธ์ของข้อคำถาม 1 เก็บข้อมูลประวัติพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.5161	1	1
10	0.5833	1	0.198
20	0.4853	1	0.2076
30	0.3562	1	0.2338
40	0.3073	1	0.2432
50	0.2689	1	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.2 ผลลัพธ์ของข้อคำถาม 2 เก็บข้อมูลประวัติพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.5556	1	1
10	0.3947	1	0.2553
20	0.3737	1	0.2438
30	0.3012	1	0.2783
40	0.2588	1	0.2948
50	0.2594	1	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.3 ผลลัพธ์ของข้อคำถาม 3 บันทึกรายละเอียดเพิ่มเติมและรูปภาพ

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.25
10	0.3091	1	0.2148
20	0.2353	1	0.2316
30	0.2152	1	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.4 ผลลัพธ์ของข้อคำถาม 4 ปรับปรุงแก้ไขข้อมูลพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	1
10	0.3279	1	0.225
20	0.2222	1	0.1921
30	0.1762	1	0.1896
40	0.1829	1	0.199
50	0.1867	1	0.2043
60	0.1874	1	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.5 ผลลัพธ์ของข้อคำถาม 5 บักทึกวันลาออกของพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.5	1	1
10	0.3455	1	0.2336
20	0.2008	1	0.2
30	0.218	1	0.2086
40	0.2083	1	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.6 ผลลัพธ์ของข้อคำถาม 6 บันทึกตารางการทำงานของพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.3333
10	0.3171	1	0.2692
20	0.3556	1	0.0816
30	0.3243	1	0.0925
40	0.2718	1	0.0876
50	0.2121	1	0.0929
60	0.1108	1	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.7 ผลลัพธ์ของข้อคำถาม 7 บันทึกการลางาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.3333	1	0.1818
10	0.1905	1	0.0828
20	0.1414	1	0.0725
30	0.0875	1	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.8 ผลลัพธ์ของข้อคำถาม 8 บันทึกการขออนุญาตทำงานล่วงเวลา

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.4	1	0.1818
10	0.5	1	0.1887
20	0.5	1	0.1389
30	0.21	1	0.1214
40	0.0881	1	0
50	0.0842	1	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.9 ผลลัพธ์ของข้อคำถาม 9 บันทึกการขออนุญาตไม่ลงเวลาทำงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.5	1	0.2
10	0.0946	1	0.0839
20	0.0842	1	0.0838
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.10 ผลลัพธ์ของข้อคำถาม 10 คำนวณเวลาสาย

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.6
10	0.9	1	0.6364
20	0.6429	1	0.1414
30	0.5122	1	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.11 ผลลัพธ์ของข้อคำถาม 11 คำนวณเงินเดือน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.75	1	0.5
10	0.619	1	0.4257
20	0.5	1	0.4331
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ ง.12 ผลลัพธ์ของข้อคำถาม 12 คำนวณค่าทำงานล่วงเวลา

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.4364
10	0.4746	1	0.4565
20	0.4365	1	0
30	0.3237	1	0
40	0.3234	1	0
50	0.2991	1	0
60	0.2611	1	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.13 ผลลัพธ์ของข้อคำถาม 13 คำนวณภาษี

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.6429
10	0.6944	1	0.5556
20	0.6364	1	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.14 ผลลัพธ์ของข้อคำถาม 14 หักเงินมาสาย

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	1
10	0.8125	1	0.5952
20	0.6404	1	0.5789
30	0.6446	1	0
40	0.5556	1	0
50	0.5365	1	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.15 ผลลัพธ์ของข้อคำถาม 15 พิมพ์หนังสือรับรองเงินเดือน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.7692	1	1
10	0.4938	1	0.4516
20	0	0	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ ง.16 ผลลัพธ์ของข้อคำถาม 16 บันทึกรายละเอียดฝึกอบรม

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.2188	1	0.1429
10	0.2169	1	0.1485
20	0.232	1	0.1224
30	0.2023	1	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.17 ผลลัพธ์ของข้อคำถาม 17 บันทึกข้อมูลพนักงานที่ผ่านการฝึกอบรมแล้ว

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.2	1	0.101
10	0.0932	1	0.125
20	0.107	1	0.125
30	0.0904	1	0.1014
40	0.0998	1	0
50	0.0998	1	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.18 ผลลัพธ์ของข้อคำถาม 18 ประเมินพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.375	1	0.1134
10	0.1624	1	0.1104
20	0.1324	1	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.19 ผลลัพธ์ของข้อคำถาม 19 บันทึกค่าใช้จ่ายในการอบรม

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.3226	1	0.0894
10	0.3393	1	0.1222
20	0.3676	1	0.1271
30	0.3947	1	0
40	0.3309	1	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.20 ผลลัพธ์ของข้อคำถาม 20 พิมพ์รายงานผลการอบรม

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.4167	1	0.1719
10	0.1176	1	0.2391
20	0.1553	1	0.2818
30	0.1841	1	0
40	0.1888	1	0
50	0.1877	1	0
60	0.1372	1	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.21 ผลลัพธ์ของข้อคำถาม 21 เก็บข้อมูลสวัสดิการ

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.3429
10	0.3478	1	0.4146
20	0.3784	1	0.48
30	0.3704	1	0.5
40	0.3488	1	0
50	0.3862	1	0
60	0.3895	1	0
70	0.245	1	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.22 ผลลัพธ์ของข้อคำถาม 22 บันทึกการเบิกสวัสดิการ

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.5238	1	0.1026
10	0.3692	1	0.1384
20	0.4416	1	0.1773
30	0.4458	1	0.1991
40	0.4348	1	0
50	0.4671	1	0
60	0.3541	1	0
70	0.2157	1	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.23 ผลลัพธ์ของข้อคำถาม 23 กำหนดรายละเอียดสวัสดิการ

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	1	1	0.6667
10	0.4444	1	0.6667
20	0.4412	1	0.5185
30	0.4216	1	0.5395
40	0.35	1	0
50	0.3744	1	0
60	0.39	1	0
70	0.3399	1	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.24 ผลลัพธ์ของข้อคำถาม 24 กำหนดโควตาหรือสิทธิการเบิกแต่ละประเภทสวัสดิการให้พนักงานแต่ละคน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.1967	1	0.05
10	0.2097	1	0.0889
20	0.2297	1	0.1229
30	0.2375	1	0.1233
40	0.2727	1	0
50	0.2776	1	0
60	0.2862	1	0
70	0.1702	1	0
80	0.1556	1	0
90	0	0	0
100	0	0	0

ตารางที่ ง.25 ผลลัพธ์ของข้อคำถาม 25 วิเคราะห์ข้อมูลของสวัสดิการแต่ละประเภท

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.4	1	0.0805
10	0.2376	1	0.1257
20	0.2602	1	0.1667
30	0.2327	1	0.1931
40	0.2283	1	0
50	0.2281	1	0
60	0.2485	1	0
70	0.2404	1	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.26 ผลลัพธ์ของข้อคำถาม 26 ประเมินผลการทำงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.0588	1	0.0182
10	0.1111	1	0.0469
20	0.1613	1	0
30	0.1628	1	0
40	0.0941	1	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.27 ผลลัพธ์ของข้อคำถาม 27 วิเคราะห์การทำงานของพนักงาน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.0103	1	0.0049
10	0.0155	1	0.0088
20	0.0239	1	0
30	0.0287	1	0
40	0.0281	1	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.28 ผลลัพธ์ของข้อคำถาม 28 พิมพ์ผลการประเมิน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.0233	1	0
10	0.0667	1	0
20	0.061	1	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

ตารางที่ ง.29 ผลลัพธ์ของข้อคำถาม 29 กำหนดหัวข้อการประเมิน

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.2	1	0
10	0.25	1	0
20	0	0	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ง.30 ผลลัพธ์ของข้อคำถาม 30 ประเมินประจำเดือนหรือปี

ค่าระลึก	ค่าความแม่นยำ		
	กฎความสัมพันธ์เชิงน้ำหนัก	กฎความสัมพันธ์และผลป้อนกลับของผู้ใช้งาน	วัดความคล้ายคลึง
0	0.0238	1	0
10	0.0192	1	0
20	0	0	0
30	0	0	0
40	0	0	0
50	0	0	0
60	0	0	0
70	0	0	0
80	0	0	0
90	0	0	0
100	0	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียนวิทยานิพนธ์

นายปองลิขิต สิงห์ชัย เกิดวันที่ 23 เมษายน พ.ศ.2525 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาบัณฑิต หลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ จากมหาวิทยาลัยอัสสัมชัญ เมื่อ พ.ศ. 2548 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมซอฟต์แวร์ ณ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2554

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University