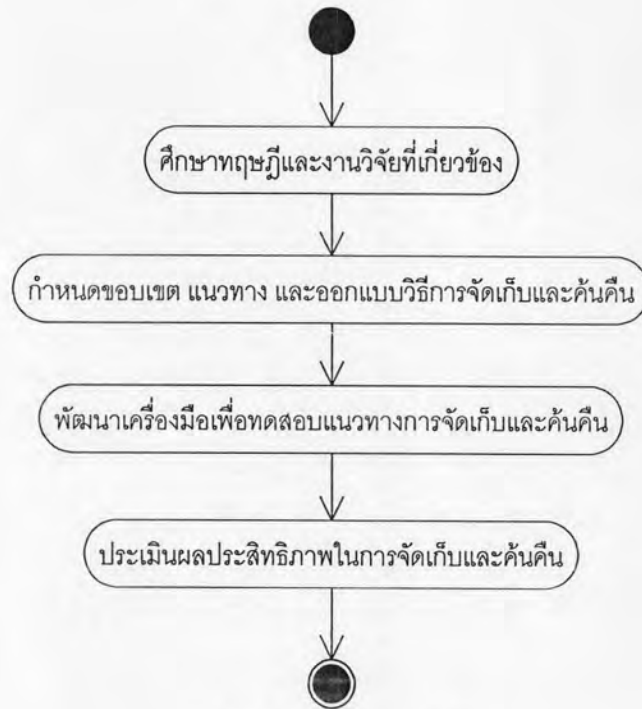




บทที่ 3

แนวคิดและวิธีการวิจัย

แนวทางการวิจัยแบ่งออกเป็น 4 ส่วนตามลำดับ ดังแสดงได้ตามรูปที่ 3.1



รูปที่ 3.1 แนวทางการวิจัย

ดังนั้นในส่วนของบทที่ 3 นี้จะกล่าวถึงแนวทางการวิจัย 4 ส่วนตามรูปที่ 3.1 ซึ่งมีรายละเอียดดังต่อไปนี้

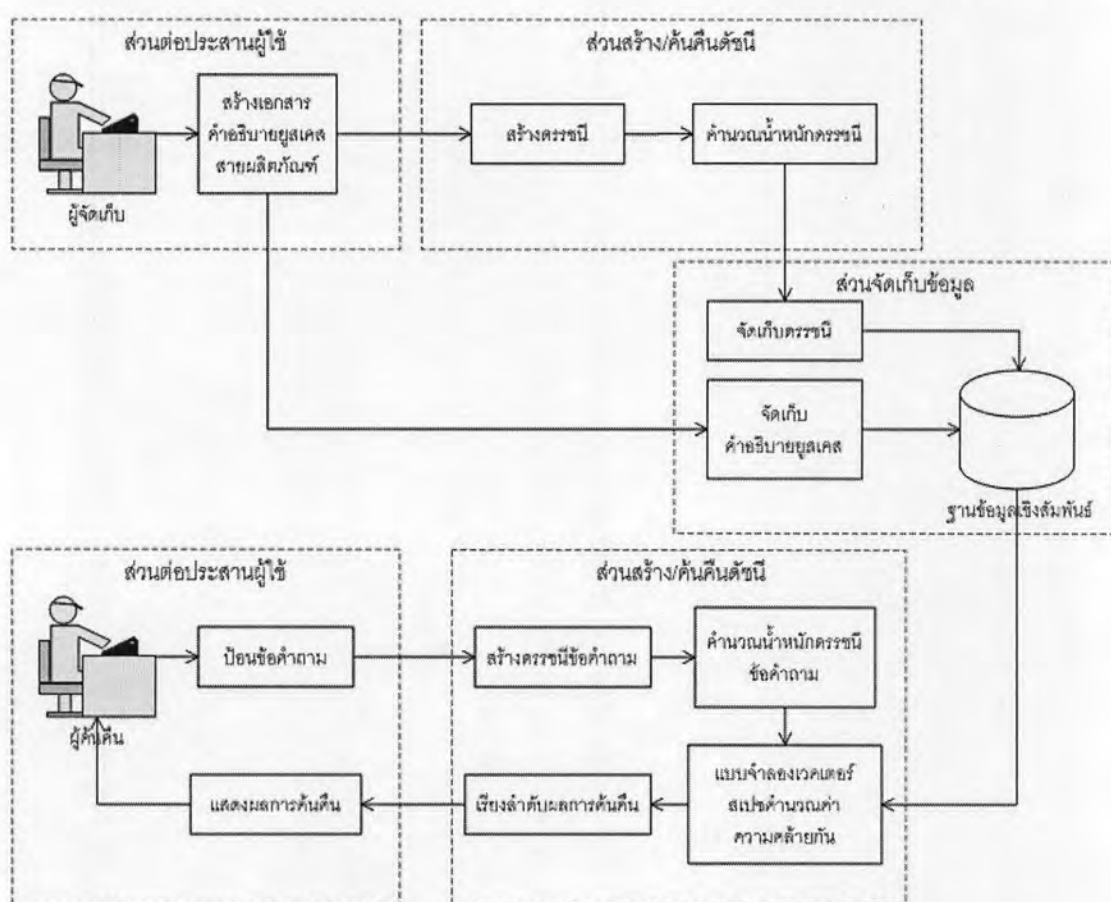
3.1 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทำการศึกษาทฤษฎีที่เกี่ยวข้องและงานวิจัยที่เกี่ยวข้องซึ่งประกอบไปด้วยวิศวกรรมสายผลิตภัณฑ์ซอฟต์แวร์บนพื้นฐานยูเอเอ็มแอลหรือพลัส และการค้นคืนสารสนเทศ ในส่วนของพลัสนั้นมีขั้นตอนที่กำหนดไว้ 5 ขั้นตอนด้วยกัน โดยเริ่มจากแบบจำลองความต้องการสายผลิตภัณฑ์ซอฟต์แวร์ แบบจำลองการวิเคราะห์สายผลิตภัณฑ์ซอฟต์แวร์ แบบจำลองการออกแบบสายผลิตภัณฑ์ซอฟต์แวร์ การเขียนโปรแกรมและการทดสอบคอมโพเนนต์ส่วนเพิ่ม จนถึง การทดสอบสายผลิตภัณฑ์ โดยในแต่ละขั้นตอนนี้จะเกิดสินทรัพย์ขึ้น แล้วนำไปจัดเก็บในคลังสายผลิตภัณฑ์ซอฟต์แวร์ สำหรับงานวิทยานิพนธ์นี้จะอาศัยระเบียบวิธีการในขั้นตอนแบบจำลองการวิเคราะห์สายผลิตภัณฑ์ซอฟต์แวร์เท่านั้น การจัดการกับสินทรัพย์ที่ได้ในขั้นตอนนี้จะอาศัย

หลักการค้นคืนสารสนเทศเข้าช่วย สินทรัพย์ที่กล่าวถึงนี้คือ เอกสารคำอธิบายยูสเคสสายผลิตภัณฑ์ เมื่อทำการจัดเก็บเอกสารคำอธิบายยูสเคสสายผลิตภัณฑ์แล้ว จะทำการสร้างตรรกะนี้ของเอกสาร ซึ่งจะใช้คำนวณค่าความคล้ายกันระหว่างตรรกะนี้ของเอกสารกับตรรกะนี้ของข้อความถามที่ใช้ในการค้นคืน แล้วนำผลที่ได้จากการค้นคืนมาแสดงแก่ผู้ใช้ต่อไป

3.2 สถาปัตยกรรมภาพรวมในการจัดเก็บและค้นคืน

สถาปัตยกรรมภาพรวมในการจัดเก็บและค้นคืนแสดงได้ดังรูปที่ 3.2



รูปที่ 3.2 โครงสร้างสถาปัตยกรรมการจัดเก็บและค้นคืน

โครงสร้างสถาปัตยกรรมการจัดเก็บและค้นคืน สามารถอธิบายส่วนการทำงานต่างๆ ได้ดังนี้

- 1) ส่วนต่อประสานผู้ใช้ แบ่งออกเป็น 2 คือส่วนของผู้จัดเก็บและส่วนของผู้ใช้ที่จะทำการค้นคืน

- (1) ส่วนของผู้จัดเก็บ ทำการสร้างเอกสารคำอธิบายยูสเคสสายผลิตภัณฑ์ แล้วป้อนข้อมูลผ่านส่วนต่อประสานตามโครงสร้างยูสเคส เพื่อนำมาจัดเก็บในฐานข้อมูลเชิงสัมพันธ์
- (2) ส่วนของผู้ใช้ เป็นส่วนที่ผู้ใช้จะทำการป้อนข้อความผ่านส่วนต่อประสาน เพื่อทำการค้นคืน แล้วนำผลจากการค้นคืนที่ได้กลับมาแสดงผลต่อผู้ใช้

2) ส่วนสร้าง/ค้นคืนตรรกะนี้ แบ่งเป็นส่วนของการจัดเก็บและส่วนของการค้นคืน ในส่วนของการจัดเก็บเป็นส่วนที่นำเอาข้อมูลที่ได้ในตอนต้นมาทำการสร้างตรรกะนี้และคำนวณน้ำหนักตรรกะนี้ ในส่วนของการค้นคืนจะเริ่มจากสร้างตรรกะนี้จากข้อความที่ผู้ใช้ป้อนเข้ามา แล้วนำมาคำนวณน้ำหนักตรรกะนี้ หลังจากนั้นทำการคำนวณค่าความคล้ายกันของข้อความที่ผู้ใช้ป้อน และเอกสารคำอธิบายยูสเคสโดยใช้แบบจำลองเวกเตอร์สเปซ เมื่อได้ผลลัพธ์จะทำการจัดเรียงผลการค้นคืนที่ได้ เพื่อนำมาแสดงผลให้ผู้ใช้เป็นลำดับต่อไป

3) ส่วนจัดเก็บข้อมูล เป็นส่วนที่นำเอาเอกสารคำอธิบายยูสเคส และตรรกะนี้ที่ได้ นำมาจัดเก็บ

3.3 กำหนดขอบเขต แนวทาง และออกแบบวิธีการจัดเก็บและค้นคืน

ขั้นตอนการกำหนดขอบเขต แนวทาง และการออกแบบวิธีการจัดเก็บและค้นคืนมี ลำดับดังรูปที่ 3.3 ซึ่งเริ่มจากการกำหนดขอบเขตและรูปแบบของข้อมูลที่ใช้ในการจัดเก็บ เมื่อ กำหนดขอบเขตข้อมูลแล้ว จะทำการออกแบบวิธีการจัดเก็บข้อมูล และทำการออกแบบวิธีการใน ส่วนของการค้นคืนข้อมูล สุดท้ายเป็นขั้นตอนในการกำหนดวิธีการในการประเมินประสิทธิภาพใน การค้นคืนของระบบ



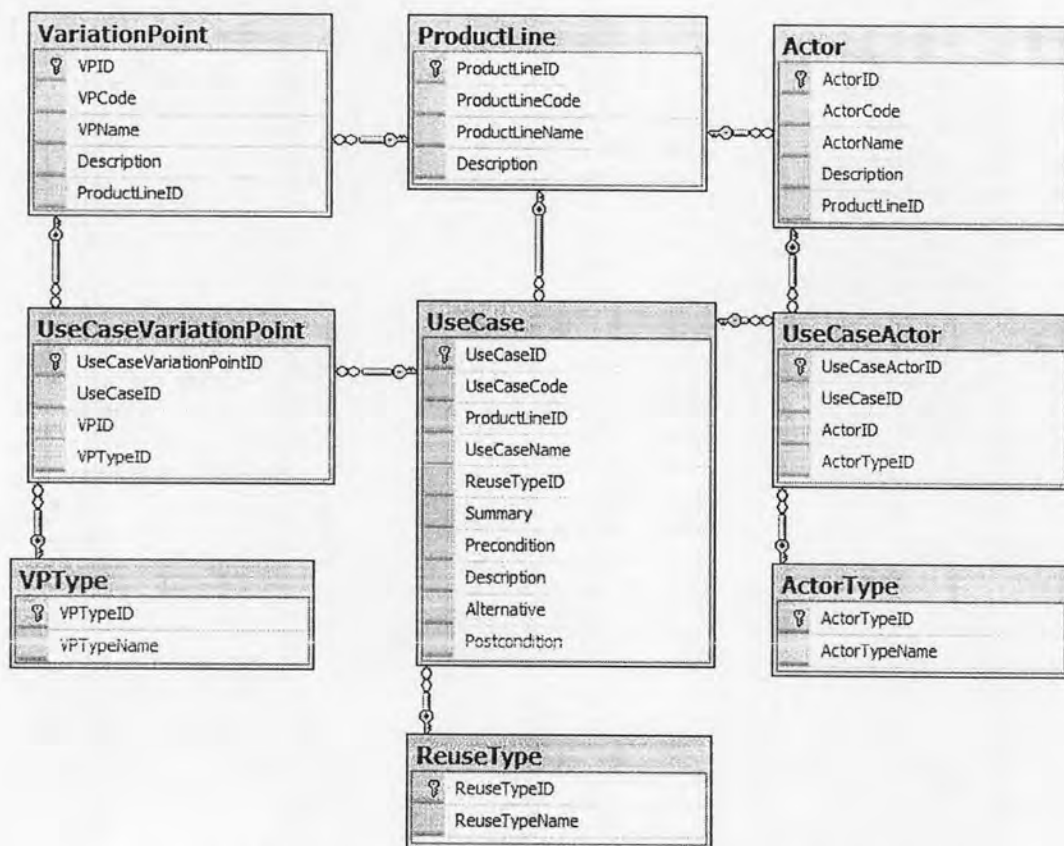
รูปที่ 3.3 ขั้นตอนการกำหนดขอบเขต แนวทางและการออกแบบวิธีการจัดเก็บและค้นคืน

3.3.1 กำหนดขอบเขตของข้อมูลสายผลิตภัณฑ์ซอฟต์แวร์ที่ต้องการจัดเก็บ

ข้อมูลในการจัดเก็บประกอบด้วย เอกสารคำอธิบายยูสเคสสายผลิตภัณฑ์ 10 ระบบ จำนวน 103 ยูสเคส เป็นภาษาอังกฤษ โดยอาศัยรูปแบบโครงสร้างเอกสารคำอธิบายยูสเคสสายผลิตภัณฑ์ตามกระบวนการพลัส

3.3.2 ออกแบบวิธีการจัดเก็บข้อมูล

การจัดเก็บข้อมูลสายผลิตภัณฑ์ซอฟต์แวร์จะจัดเก็บอยู่ในรูปแบบฐานข้อมูลเชิงสัมพันธ์แสดงได้ดังรูปที่ 3.4



รูปที่ 3.4 ความสัมพันธ์ของตารางที่ใช้จัดเก็บคำอธิบายยูสเคส

รายละเอียดของตารางที่ใช้จัดเก็บข้อมูลเอกสารคำอธิบายยูสเคสทั้ง 9 ตาราง แสดงไว้ในตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดของตารางที่ใช้จัดเก็บข้อมูลเอกสารคำอธิบายยูสเคส

ชื่อตาราง	คำอธิบาย
UseCase	เก็บข้อมูลหลักของยูสเคส
Actor	เก็บข้อมูลผู้กระทำ
UseCaseActor	แสดงความสัมพันธ์ระหว่างยูสเคสกับผู้กระทำที่เกี่ยวข้องกับยูสเคส
ActorType	เก็บประเภทผู้กระทำ
ProductLine	เก็บข้อมูลสายผลิตภัณฑ์
ReuseType	เก็บประเภทการนำกลับมาใช้
VariationPoint	แสดงส่วนแปรผันทั้งหมดที่เกิดขึ้นในแต่ละยูสเคส
UseCaseVariationPoint	แสดงความสัมพันธ์ระหว่างยูสเคสกับส่วนแปรผัน
VPTYPE	เก็บประเภทของส่วนแปรผัน

รายละเอียดโครงสร้างข้อมูลของแต่ละตารางแสดงได้ดังต่อไปนี้

ตารางที่ 3.2 โครงสร้างข้อมูลของตาราง UseCase

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>UseCaseID</u>	คีย์หลักของตารางยูสเคส	int	PK
UseCaseCode	รหัสยูสเคส	nvarchar(50)	
ProductLineID	คีย์หลักของตารางสายผลิตภัณฑ์	int	
UseCaseName	ชื่อยูสเคส	nvarchar(1000)	
ReuseTypeID	รหัสประเภทการนำกลับมาใช้	tinyint	
Summary	รายละเอียดสรุป	nvarchar(MAX)	
Precondition	เงื่อนไขก่อน	nvarchar(MAX)	
Description	คำอธิบาย	nvarchar(MAX)	
Alternative	ทางเลือก	nvarchar(MAX)	
Postcondition	เงื่อนไขสิ้นสุด	nvarchar(MAX)	

ตารางที่ 3.3 โครงสร้างข้อมูลของตาราง Actor

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>ActorID</u>	คีย์หลักของตารางผู้กระทำ	int	PK
ActorCode	รหัสแผนภาพยูสเคส	nvarchar(50)	
ActorName	ชื่อผู้กระทำ	nvarchar(1000)	
Description	คำอธิบาย	nvarchar(MAX)	
ProductLineID	รหัสประเภทการนำกลับมาใช้	int	

ตารางที่ 3.4 โครงสร้างข้อมูลของตาราง UseCaseActor

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>UseCaseActorID</u>	คีย์หลักของตารางผู้กระทำต่อยูสเคส	int	PK
UseCaseID	คีย์หลักของตารางยูสเคส	int	
ActorID	คีย์หลักของตารางผู้กระทำ	int	
ActorTypeID	รหัสประเภทผู้กระทำ	tinyint	

ตารางที่ 3.5 โครงสร้างข้อมูลของตาราง ActorType

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>ActorTypeID</u>	รหัสประเภทผู้กระทำ	tinyint	PK
ActorTypeName	ชื่อประเภทผู้กระทำ	nvarchar(1000)	

ตารางที่ 3.6 โครงสร้างข้อมูลของตาราง ProductLine

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>ProductLineID</u>	คีย์หลักของตารางสายผลิตภัณฑ์	int	PK
ProductLineCode	รหัสแผนภาพสายผลิตภัณฑ์	nvarchar(50)	
ProductLineName	ชื่อสายผลิตภัณฑ์	nvarchar(1000)	
Description	คำอธิบาย	nvarchar(MAX)	

ตารางที่ 3.7 โครงสร้างข้อมูลของตาราง ReuseType

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>ReuseTypeID</u>	รหัสประเภทการนำกลับมาใช้	tinyint	PK
ReuseTypeName	ชื่อประเภทการนำกลับมาใช้	nvarchar(1000)	

ตารางที่ 3.8 โครงสร้างข้อมูลของตาราง VariationPoint

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>VPID</u>	คีย์หลักของตารางส่วนแปรผัน	int	PK
VPCode	รหัสแผนภาพส่วนแปรผัน	nvarchar(50)	
VPName	ชื่อส่วนแปรผัน	nvarchar(1000)	
Description	คำอธิบาย	nvarchar(MAX)	
ProductLineID	รหัสสายผลิตภัณฑ์	int	

ตารางที่ 3.9 โครงสร้างข้อมูลของตาราง UseCaseVariationPoint

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>UseCaseVariationPointID</u>	รหัสยูสเคสส่วนแปรผัน	int	PK
UseCaseID	คีย์หลักของตารางยูสเคส	int	
VPID	คีย์หลักของตารางส่วนแปรผัน	int	
VPTypID	รหัสประเภทส่วนแปรผัน	tinyint	

ตารางที่ 3.10 โครงสร้างข้อมูลของตาราง VPTyp

คุณลักษณะ	คำอธิบาย	ชนิด	คีย์
<u>VPTypID</u>	รหัสประเภทส่วนแปรผัน	tinyint	PK
VPTypName	ชื่อประเภทส่วนแปรผัน	nvarchar(1000)	

3.2.3 ออกแบบวิธีการค้นคืนข้อมูล

การค้นคืนข้อมูลจะใช้แบบจำลองเวกเตอร์สเปซ ซึ่งเป็นการคำนวณค่าความคล้ายกันระหว่างเวกเตอร์ของตัวเอกสารและเวกเตอร์ของข้อความ โดยวิธีในการค้นคืนจะมี 2 แบบคือ

- 1) การค้นคืนโดยใช้คำสำคัญ คำนวณค่าความคล้ายกันตามสมการที่ 3
- 2) การค้นคืนโดยพิจารณาจากโครงสร้างเอกสารคำอธิบายยูสเคส มีการคำนวณค่าความคล้ายกันระหว่างข้อความ และเอกสารตามสมการดังนี้

$$\text{SIM}(UC_i, \text{Query}_j) = \frac{\sum_{m=1}^e [\text{SIM}(E_{mi}, E_{mj}) * WE_m]}{\text{TotalWeight}} \quad (7)$$

เมื่อ	SIM(UC _i , Query _j)	แทน ค่าคะแนนความคล้ายระหว่างข้อความกับยูสเคส
e		แทน จำนวนส่วนประกอบของเอกสารคำอธิบายยูสเคสที่ i และข้อความที่ j
WE _m		แทน ค่าน้ำหนักของส่วนประกอบที่ m ในโครงสร้างเอกสารคำอธิบายยูสเคส ซึ่งระบุโดยผู้ใช้
TotalWeight		แทน ค่าผลรวมของน้ำหนักทั้งหมดที่ระบุโดยผู้ใช้
SIM(E _{mi} , E _{mj})		แทน ค่าความคล้ายกันระหว่างแต่ละส่วนประกอบของยูสเคสที่ i กับข้อความที่ j โดยคำนวณได้จากสมการที่ 2

ในการค้นคืนทั้ง 2 แบบนั้น จะค้นคืนทั้งในส่วนของยูสเคสทั่วไป และยูสเคสสายผลิตภัณฑ์ และนำผลการค้นคืนที่ได้มาเปรียบเทียบระหว่างยูสเคสทั่วไป และยูสเคสสายผลิตภัณฑ์ว่าให้ผลการค้นคืนแตกต่างกันหรือไม่ อย่างไร โดยจะกำหนดโจทย์คำถามขึ้น 30 โจทย์ โดยแต่ละโจทย์จะใช้ข้อความ 3 ข้อคำถามในการค้นคืน

3.2.4 กำหนดวิธีการประเมินประสิทธิภาพในการจัดเก็บและค้นคืน การประเมินผลจะใช้วิธีการคำนวณค่าระลึกค่าความแม่นยำและค่าเฉลี่ยฮาร์โมนิก ซึ่งมีสูตรการคำนวณดังสมการต่อไปนี้

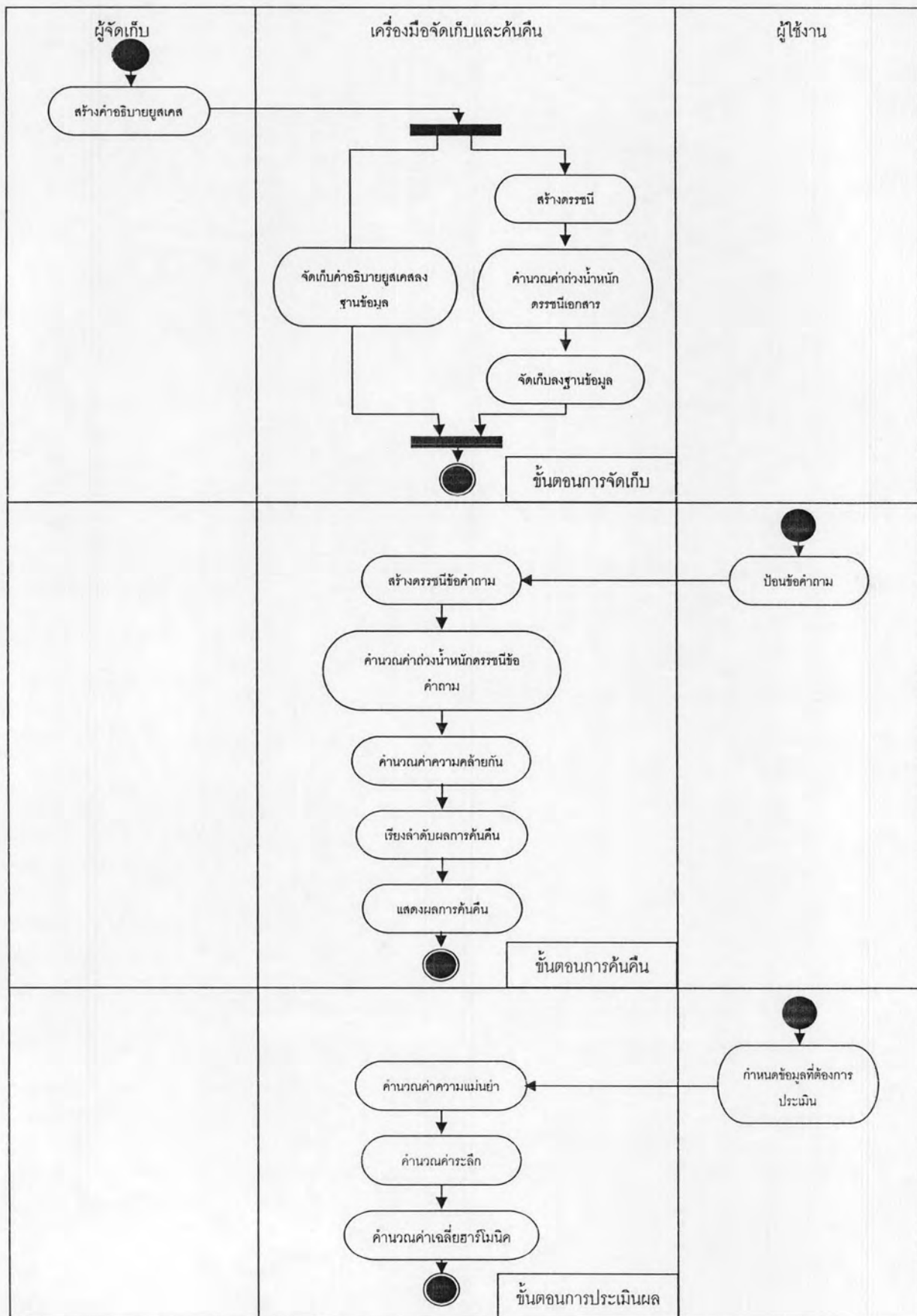
$$\text{ค่าความแม่นยำ (precision)} = \frac{|Ra|}{|A|}$$

$$\text{ค่าระลึก (recall)} = \frac{|Ra|}{|R|}$$

$$\text{ค่าเฉลี่ยฮาร์โมนิก } F(j) = \frac{2}{\frac{1}{r(j)} + \frac{1}{P(j)}}$$

3.4 พัฒนาเครื่องมือเพื่อทดสอบแนวทางการจัดเก็บและค้นคืน รวมถึงประเมินผล ประสิทธิภาพในการจัดเก็บและค้นคืน

กิจกรรมในการพัฒนาเครื่องมือเพื่อทดสอบแนวทางการจัดเก็บและค้นคืนแสดง เป็นแผนภาพได้ดังรูป 3.5



รูปที่ 3.5 แผนภาพกิจกรรมเครื่องมือที่ใช้ในการจัดเก็บและค้นคืน

ในส่วนนี้จะขอกล่าวรวมถึงกิจกรรมการประเมินผลประสิทธิภาพในการจัดเก็บและคั่นคืนด้วย เนื่องจากแผนภาพได้รวมขั้นตอนในส่วนของการประเมินผลไว้แล้ว สามารถแบ่งได้เป็น 3 ขั้นตอนหลักคือ

1) ขั้นตอนการจัดเก็บ เริ่มจากผู้จัดเก็บทำการสร้างเอกสารคำอธิบายยูนิตเอสซีขึ้นใหม่ หรือจัดเก็บเอกสารคำอธิบายยูนิตเอสซีที่มีอยู่แล้ว ในการจัดเก็บจะผ่านการสร้างดรรชนีของเอกสาร เพื่อใช้เป็นตัวแทนของเอกสารแต่ละตัว แล้วทำการคำนวณค่าถ่วงน้ำหนักของดรรชนีเอกสาร ซึ่งขั้นตอนดังกล่าวจะกระทำโดยใช้เครื่องมือในการสร้างดรรชนี และคำนวณค่าถ่วงน้ำหนักดรรชนีเอกสาร

2) ขั้นตอนการคั่นคืน เริ่มจากผู้ใช้งานจะทำการป้อนข้อความผ่านเครื่องมือในการจัดเก็บและคั่นคืน แล้วเครื่องมือจะทำการสร้างดรรชนีของข้อความที่ผู้ใช้งานทำการป้อนเข้ามา แล้วจึงทำการคำนวณค่าถ่วงน้ำหนักดรรชนีคำถาม จากนั้นใช้แบบจำลองเวกเตอร์สเปซในการเปรียบเทียบความคล้ายกันระหว่างดรรชนีของเอกสารและข้อความที่ผู้ใช้ป้อนเข้ามานั้น เพื่อให้ได้ผลลัพธ์ที่มีความคล้ายกัน แล้วนำผลลัพธ์ที่ได้มาเรียงลำดับตามความคล้ายกัน ซึ่งทำการเรียงจากผลลัพธ์ที่มีความคล้ายกันมากที่สุดก่อนตามลำดับ เมื่อทำการจัดเรียงแล้วนำผลลัพธ์ที่ได้นั้นมาแสดงต่อผู้ใช้งาน

3) ขั้นตอนการประเมินผล เป็นขั้นตอนสุดท้าย เพื่อทำการประเมินว่าเครื่องมือที่ใช้นั้นสามารถหาผลลัพธ์ได้ตรงตามความต้องการของผู้ใช้งานได้เพียงใด โดยใช้การคำนวณค่าระลึค่าความแม่นยำและค่าเฉลี่ยฮาร์โมนิคตามสมการที่กล่าวมาแล้วในหัวข้อที่ 3.2.4