

ระบบฐานความรู้เพื่อการบำรุงรักษา พีซี
Knowledge Base for PC Maintenance



นายภมรศักดิ์ ปิยะวานิชย์สกุล

นายสุรเชษฐ์ เต็งสุจริตกุล

เลขหมู่.....
เลขทะเบียน 61340
วัน,เดือน,ปี 17 ก.ค. 2549

b.....
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2547

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบฐานความรู้เพื่อการบำรุงรักษา พีซี
Knowledge Base for PC Maintenance

โดย

นายภมรศักดิ์ ปิยะวานิชย์สกุล

นายสุรเชษฐ์ เต็งสุจริตกุล

อาจารย์ที่ปรึกษา

ดร. วรวัฒน์ ลิ้มโกศา

ปริญญาวิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2547

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาโท ปีการศึกษา 2547

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบฐานความรู้เพื่อการบำรุงรักษา พีซี

Knowledge Base for PC Maintenance

ผู้จัดทำ

1. นาย ภมรศักดิ์ ปิยะวานิชย์สกุล รหัสประจำตัว 44010357

2. นาย สุรเชษฐ์ เต็งสุจริตกุล รหัสประจำตัว 45015392



อาจารย์ที่ปรึกษา

(ดร. วรวัฒน์ ลิ้มโกศา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบฐานความรู้เพื่อการบำรุงรักษา พีซี

นายภมรศักดิ์ ปิยะวานิชย์สกุล 44010357

นายสุรเชษฐ์ เต็งสุจริตกุล 45015392

ดร. วรวัฒน์ ลิ้ม โภคา อาจารย์ที่ปรึกษา

ปีการศึกษา 2547

บทคัดย่อ

ในปัจจุบันอุปกรณ์คอมพิวเตอร์ได้เข้ามามีบทบาทในการดำรงชีวิตมากขึ้นอันได้แก่ ไม่ว่าจะ โรงงานหรือบริษัทใหญ่ค่างก็นำเอาคอมพิวเตอร์เข้ามาช่วยในการดำเนินงานเพื่อทำให้เกิดความรวดเร็ว และถูกต้องในการดำเนินงาน แน่นอนว่าอุปกรณ์เหล่านี้ในบางครั้งย่อมต้องเกิดปัญหาขึ้น ให้แก้ไข ซ่อมแซมบ้างดังนั้นจึงต้องมีแผนกไอทีเพื่อทำหน้าที่ในส่วนนี้ ซึ่งเราก็น่าจะมีโปรแกรมที่ช่วยจัดสรร ปัญหาต่างๆที่เข้ามาและทำการเก็บข้อมูลดังกล่าวเพื่อการวิเคราะห์ต่อไป

นอกจากนี้โครงการนี้ยังได้ทำเกี่ยวกับการค้นหาเอกสารเพื่อให้ได้ในเรื่องที่ผู้ใช้ต้องการ โดยนำ หลักการดึงข้อมูลมาใช้ ซึ่งการทำด้วยวิธีนี้จะทำให้ระบบมีความยืดหยุ่นในการค้นหามากขึ้น ส่วน ที่เป็นหัวใจของการทำค้นหาในฐานความรู้คือเราจะต้องทำการหาดัชนีที่เป็นใจความของเอกสารนั้นๆ ให้ได้และคำนวณค่าความตรงประเด็นของแต่ละเอกสารที่ตรงตามที่ใช้ต้องการให้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Knowledge Base for PC Maintenance

Pamonsak Piyawanitsakul

Surachet Tengsujarikul

Dr. Voravat Limpoka Advisor

Academic Year 2004

ABSTRACT

Recently, Hardware Computer have more important in your life such as In factory or company use computer hardware in working for quickly and correctly. But sometime Hardware Computer must any problem. We should program management about Keep and analyze Information

In addition, This project have Knowledge base for Search Relevant Document. It use Information retrieval system theory. Main problem in my project is to find index from String Which user put to system and calculate similarity of each document.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้คงไม่อาจเสร็จได้ด้วยดี หากไม่ได้รับความช่วยเหลือ และร่วมมือจากหลายๆ ฝ่ายด้วยกัน บุคคลแรกที่ต้องกล่าวถึงเพราะเป็นส่วนสำคัญที่ทำให้วิทยานิพนธ์นี้เสร็จลงได้ก็คืออาจารย์ วรวัฒน์ ลิ้ม โทคา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้ความเอาใจใส่ แนะนำ และช่วยเหลือเสมอมา ซึ่งต้องขอขอบพระคุณเป็นอย่างมาก

และต้องขอขอบพระคุณบุคคลสำคัญที่สุดที่ทำให้ข้าพเจ้ามีวันนี้ ก็คือ บิดา มารดา อันเป็นที่เคารพรักยิ่ง ซึ่งได้เลี้ยงดูผู้เขียนมาเป็นอย่างดี พร้อมทั้งให้โอกาสในการศึกษาอย่างเต็มที่ และยังให้กำลังใจเอาใจใส่เสมอมา ในทุกๆ ด้านอันหาที่เปรียบมิได้ ข้าพเจ้าขอระลึกในพระคุณอันสุดประมาทและขอกราบขอบพระคุณมา ณ ที่นี้

นายภมรศักดิ์ ปิยะวาณิชย์สกุล

นายสุรเชษฐ์ เต็งสุจริตกุล



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้าที่
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญภาพ	VII
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มา	1
1.2 วัตถุประสงค์ของงานวิจัย	1
1.3 วิธีการดำเนินงาน	2
บทที่ 2 ความรู้เบื้องต้นเกี่ยวกับการดึงข้อมูล	3
2.1 ความรู้เบื้องต้น	3
2.2 โครงสร้างของระบบดึงข้อมูล	4
2.3 แบบจำลองต่างๆ	5
2.3.1 แบบจำลองทางตรรกศาสตร์	6
2.3.2 แบบจำลองแบบมีค่าน้ำหนัก	7
2.4 หลักการคำนวณความตรงประเด็น	8
2.4.1 ค่าความเหมือน (Cosine Similarity Measure)	8
2.4.2 กระบวนการในการกำหนดค่าน้ำหนักให้แก่แต่ละดัชนี	10
บทที่ 3 การทำดัชนี	11
3.1 ความรู้เบื้องต้นในการหาดัชนี	11
3.2 การแยกคำ	11
3.3 การตัดคำ	11
3.4 การพิจารณารากศัพท์	12
3.4.1 พิจารณาจากความถี่ที่เกิดขึ้น	12
3.4.2 พิจารณาจำนวนคำที่แตกต่างของตัวอักษรที่ตามมา	12
3.4.3 สูตรวัดค่าความเหมือน	13
3.4.4 สูตรวัดความแตกต่าง	14
3.4.5 ใช้หลักการลบคำอุปสรรคออก	14
3.5 รูปแบบดัชนี	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้าที่
บทที่ 4 การประยุกต์ใช้งานในระบบฐานความรู้	15
4.1 ขั้นตอนในการหาดัชนีในข้อความ	15
4.1.1 การแยกคำ	15
4.1.2 การตัดคำ	15
4.2 การพิจารณาให้ค่าน้ำหนัก	18
4.3 การคำนวณหาค่าความเหมือน	20
บทที่ 5 การวิเคราะห์และออกแบบระบบ	22
5.1 ลักษณะโดยรวมของระบบ	22
5.2 ระบบแจ้งปัญหา	22
5.2.1 การแจ้งปัญหาของผู้ใช้	22
5.2.2 การตอบรับปัญหาของช่าง	23
5.3 ระบบประวัติการซ่อมแซม	24
5.4 ระบบฐานความรู้	24
5.4.1 การสร้างเอกสารวิธีการแก้ปัญหา	24
5.4.2 การใช้งานระบบฐานความรู้ของผู้ใช้	25
บทที่ 6 การจัดเก็บฐานข้อมูล	26
6.1 ตารางต่างๆในฐานข้อมูล	26
6.2 ความสัมพันธ์ของตารางทั้งหมดในฐานข้อมูล	30
บทที่ 7 การใช้งานโปรแกรม	31
7.1 การใช้งานของผู้ใช้ทั่วไป	31
7.2 การใช้งานของ MIS และ Admin	39
บทที่ 8 สรุปและแนวทางในการพัฒนาต่อ	47
8.1 สรุป	47
8.2 แนวทางในการพัฒนาต่อ	47
ภาคผนวก ก เครื่องมือที่ใช้ในการพัฒนาเว็บโปรแกรมในโครงการ	48
ภาคผนวก ข ภาษาที่ใช้ระบุความต้องการ	53
บรรณานุกรม	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

	หน้าที่
ตารางที่ 2-1 แสดงตารางเปรียบเทียบระหว่างระบบฐานข้อมูลและระบบการดึงข้อมูล	3
ตารางที่ 2-2 แสดงตัวอย่างการจัดเก็บข้อมูลดัชนีในแบบจำลองตรรกศาสตร์	6
ตารางที่ 3-1 แสดงการทำพจนานุกรมเก็บรากศัพท์	12
ตารางที่ 6-1 แสดงโครงสร้างตาราง T_Index	26
ตารางที่ 6-2 แสดง โครงสร้างตาราง T_Solve	26
ตารางที่ 6-3 แสดง โครงสร้างตาราง T_Location	26
ตารางที่ 6-4 แสดง โครงสร้างตาราง T_Department	26
ตารางที่ 6-5 แสดง โครงสร้างตาราง T_Machine	27
ตารางที่ 6-6 แสดง โครงสร้างตาราง T_Project	27
ตารางที่ 6-7 แสดง โครงสร้างตาราง T_ProblemType	27
ตารางที่ 6-8 แสดง โครงสร้างตาราง T_Section	27
ตารางที่ 6-9 แสดง โครงสร้างตาราง T_Type	27
ตารางที่ 6-10 แสดง โครงสร้างตาราง T_History	28
ตารางที่ 6-11 แสดง โครงสร้างตาราง T_Request	28
ตารางที่ 6-12 แสดง โครงสร้างตาราง T_Reply	29
ตารางที่ 6-13 แสดง โครงสร้างตาราง T_User	29

สารบัญภาพ

	หน้าที่
รูปที่ 2-1 แสดงระบบการดึงข้อมูลในภาพรวม	3
รูปที่ 2-2 แสดงโครงสร้างการดึงข้อมูล	4
รูปที่ 2-3 แสดงแบบจำลองต่างๆในระบบการดึงข้อมูล	6
รูปที่ 2-4 แสดงความสัมพันธ์ของกลุ่มดัชนีในรูปแบบเซต	7
รูปที่ 2-5 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก1	7
รูปที่ 2-6 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก2	8
รูปที่ 2-7 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก3	8
รูปที่ 2-8 แสดงแบบจำลองแบบมีค่าน้ำหนักในรูปแบบกราฟ	9
รูปที่ 2-9 แสดงตัวอย่างโงทย์แบบจำลองแบบมีค่าน้ำหนัก	9
รูปที่ 3-1 แสดงขั้นตอนการหาดัชนีจากข้อความ	11
รูปที่ 5-1 แสดงความสัมพันธ์ของระบบทั้งหมด	22
รูปที่ 5-2 แสดงขั้นตอนการแจ้งปัญหาของผู้ใช้	23
รูปที่ 5-3 แสดงขั้นตอนการตอบรับปัญหาของช่าง	23
รูปที่ 5-4 แสดงขั้นตอนทำงานของระบบประวัติการซ่อมแซม	24
รูปที่ 5-5 แสดงขั้นตอนทำงานของที่มาเอกสารวิธีการแก้ปัญหา	24
รูปที่ 5-6 แสดงขั้นตอนการใช้งานระบบฐานความรู้ของผู้ใช้	25
รูปที่ 6-1 แสดงความสัมพันธ์ของตารางทั้งหมดในฐานข้อมูล	30
รูปที่ 7-1 แสดงหน้าเว็บเข้าสู่ระบบ	31
รูปที่ 7-2 แสดงหน้าเว็บของการลงทะเบียน	32
รูปที่ 7-3 แสดงหน้าเว็บผลลัพธ์เมื่อทะเบียนเรียบร้อยแล้ว	32
รูปที่ 7-4 แสดงหน้าเว็บหลังทำการเข้าสู่ระบบแล้วของผู้ใช้	33
รูปที่ 7-5 แสดงหน้าเว็บของการแจ้งปัญหา	34
รูปที่ 7-6 แสดงหน้าเว็บของผลลัพธ์เมื่อทำการแจ้งปัญหาแล้ว	34
รูปที่ 7-7 แสดงหน้าเว็บของปัญหาที่มีสถานะเป็นการรอรับแจ้งปัญหา	35
รูปที่ 7-8 แสดงหน้าเว็บของส่วนฟังก์ชันการค้นหา	36
รูปที่ 7-9 แสดงหน้าเว็บผลลัพธ์ของฟังก์ชันการค้นหา	36
รูปที่ 7-10 แสดงหน้าเว็บของผลลัพธ์ของฟังก์ชันฐานความรู้	37
รูปที่ 7-11 แสดงหน้าเว็บในส่วนการแก้ไขรายละเอียดส่วนตัว	37
รูปที่ 7-12 แสดงหน้าเว็บของฟังก์ชันเมื่อลิ้มรสผ่าน	38
รูปที่ 7-13 แสดงหน้าเว็บของผลลัพธ์ของฟังก์ชันเมื่อลิ้มรสผ่าน	38
รูปที่ 7-14 แสดงหน้าเว็บแรกของการเข้าระบบในส่วนช่างและผู้บริหาร	39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ(ต่อ)

	หน้าที่
รูปที่ 7-15 แสดงหน้าเว็บของการดูรายละเอียดปัญหาของช่าง	40
รูปที่ 7-16 แสดงหน้าเว็บของการตอบกลับของช่าง	40
รูปที่ 7-17 แสดงหน้าเว็บของผลลัพธ์ของการใช้งานระบบฐานความรู้	41
รูปที่ 7-18แสดงหน้าเว็บของการค้นหาปัญหา	41
รูปที่ 7-19 แสดงหน้าเว็บของการค้นหาค่าใช้จ่าย	42
รูปที่ 7-20 แสดงหน้าเว็บของผลลัพธ์ที่ได้จากการค้นหา	42
รูปที่ 7-21 แสดงหน้าเว็บของการเพิ่มประวัติบุคคล	43
รูปที่ 7-22 แสดงหน้าเว็บของการค้นหาบุคคลเพื่อแก้ไข	43
รูปที่ 7-23 แสดงหน้าเว็บแสดงผลลัพธ์ของการค้นหา	44
รูปที่ 7-24 แสดงหน้าเว็บของการแก้ไขประวัติบุคคล	44
รูปที่ 7-25 แสดงหน้าเว็บของการเพิ่มผู้ใช้	45
รูปที่ 7-26 แสดงหน้าเว็บของการแก้ไขหรือลบผู้ใช้หรือช่าง	45
รูปที่ 7-27 แสดงหน้าเว็บของการเพิ่มข้อมูลเกี่ยวกับโรงงาน	46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มา

ในปัจจุบันอุปกรณ์คอมพิวเตอร์ได้เข้ามามีบทบาทในชีวิตประจำวันมากขึ้น ไม่ว่าจะเป็นโรงงานหรือว่าบริษัทต่างก็มีจำนวนการใช้อุปกรณ์คอมพิวเตอร์มาก ดังนั้นเกือบทุกบริษัทมักที่จะมีแผนกไอทีซึ่งมีหน้าที่ในการซ่อมแซมและแก้ไขปัญหาของอุปกรณ์ไอทีให้แก่พนักงานในบริษัท ด้วยสาเหตุที่กล่าวมาเราจึงคิดที่จะทำโปรแกรมที่คอยช่วยเหลือจัดการในการส่งปัญหา เป็นประวัติการซ่อมบำรุง เป็นตัวให้ความรู้ในเรื่องการซ่อมแซมอุปกรณ์ต่าง

ส่วนที่สำคัญเป็นหลักของโครงการนี้คือการพัฒนาระบบฐานความรู้ขึ้นเพื่อจุดประสงค์ให้ผู้ใช้สามารถค้นหาวิธีการแก้ปัญหาที่ต้องการได้อย่างถูกต้อง ซึ่งถ้าเป็นระบบฐานข้อมูลธรรมดาในระบบจะแค่ค้นหาที่ผู้ใช้กรอกไปค้นหาในฐานข้อมูลว่าตรงกับเอกสารใดโดยถ้าหากผู้ใช้กรอกข้อมูลที่ไม่ตรงกับเอกสารที่ต้องการแล้วนั้นก็ไม่สามารถค้นหาเจอได้ หรือไม่มีผลลัพธ์ที่ออกมามากเกินไปแล้วไม่ค่อยเกี่ยวข้องกับสิ่งที่เราต้องการค้นหา

เราจึงต้องมีการประยุกต์นำเอาทฤษฎีต่างๆมาใช้ในการพิจารณาค่าคล้าย การทำดัชนี การคำนวณหาค่าความตรงประเด็นต่างๆ ซึ่งถ้าเราทำที่กล่าวมาได้ทั้งหมดแล้วนั้นระบบของเราก็สามารถคำนวณผลลัพธ์ที่ออกมาเป็นเปอร์เซ็นต์ความถูกต้องให้ผู้ใช้ได้ด้วย

1.2 วัตถุประสงค์ของงานวิจัย

- ระบบแจ้งปัญหา (MIS Job Order System) :

หน้าที่หลักๆ ของส่วนนี้คือเป็นตัวกลางในการรับ-ส่งปัญหา(Job Order) ระหว่างผู้ใช้ (User) กับทางแผนกไอที (MIS) เพื่อให้ทางช่างทราบและเข้าไปแก้ปัญหาให้แก่ผู้ใช้ได้ นอกจากนี้ทางฝ่ายผู้ใช้งานก็ยังสามารถรับทราบถึงความก้าวหน้าของปัญหาที่ได้จากส่วนนี้

- ระบบประวัติการซ่อมแซม (Computer History System) :

ทำหน้าที่จัดเก็บรายละเอียดของการแจ้งปัญหาต่างๆเพื่อนำไปรวมกับประวัติของผู้ใช้แต่ละคน และ ข้อมูลของอุปกรณ์คอมพิวเตอร์ต่างๆเพื่อทำเป็นประวัติการซ่อมแซมเราสามารถนำข้อมูลในส่วนนี้ไปวิเคราะห์เพื่อพิจารณาในเรื่องต่างๆได้ไม่ว่าจะเป็น ค่าใช้จ่ายในการซ่อมบำรุงอุปกรณ์ของแผนกต่างๆ หรือการวิเคราะห์ความถี่ในการเกิดปัญหาแต่ละประเภท

- ระบบฐานความรู้ (Knowledge Base System) :

ระบบที่ทำการจัดเก็บวิธีการแก้ไขปัญหาต่างๆเพื่อทำเป็นฐานความรู้ให้แก่ผู้ใช้ โดยระบบจะพยายามค้นหาความรู้ที่ตรงตามผู้ใช้ต้องการให้ได้มากที่สุดเท่าที่มีในฐานข้อมูล โดยจะทำการวิเคราะห์ความตรงประเด็นของตัวเอกสารกับข้อมูลที่ผู้ใช้กรอกผ่านเข้ามา

1.3 วิธีการดำเนินงาน

เราทำการแบ่งการทำงานเป็น 2 ส่วนคือการทำในส่วนของ

- โปรแกรมซึ่งก็คือส่วนของระบบแจ้งปัญหา และ ระบบประวัติการซ่อมแซม ซึ่งจะต้องมีการออกแบบส่วนของฐานข้อมูลและและจัดเรียงลำดับเหตุการณ์ในแต่ละหน้าเว็บตามแต่ที่ผู้ใช้จะกระทำ
- ส่วนระบบฐานข้อมูลจะใช้หลักการทำดัชนีเพื่อหาดัชนีให้กับเอกสารในฐานข้อมูล และความต้องการของผู้ใช้ที่ป้อนเข้ามาเพื่อทำการเปรียบเทียบหาเอกสารที่มีความตรงประเด็นกับความต้องการของผู้ใช้ให้ ได้มากที่สุด นอกจากนี้ยังสามารถนำค่าความคล้ายของแต่ละเอกสารมาจัดลำดับ เพื่อแสดงให้ผู้ใช้เลือกได้อีกด้วย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ความรู้เบื้องต้นเกี่ยวกับการดึงข้อมูล

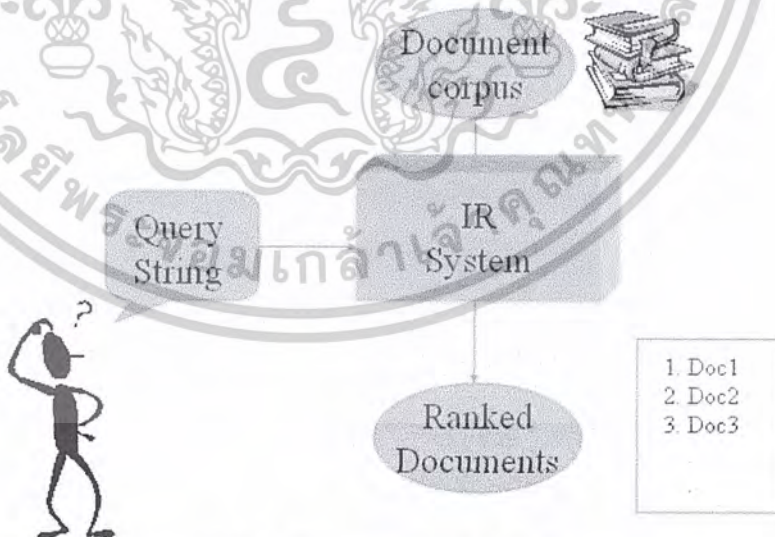
2.1 ความรู้เบื้องต้น

ในการจะเปรียบเทียบว่าเอกสารใดตรงประเด็นหรือคล้ายกับสิ่งที่ผู้ใช้ต้องการค้นหา นั้นนับว่าเป็นสิ่งที่ยาก จำเป็นต้องนำหลักการของระบบการดึงข้อมูล (Information retrieval System) เข้ามาช่วยให้ซึ่งมันจะทำให้ระบบของเรามีความแตกต่างจากระบบฐานข้อมูล (Database) ขรรรมคาค้างตาราง 2-1

	Databases	IR
Data	Structured	Unstructured
Fields	Clear semantics (SSN, age)	No fields (other than text)
Queries	Defined (relational algebra, SQL)	Free text ("natural language"), Boolean
Recoverability	Critical (concurrency control, recovery, atomic operations)	Downplayed, though still an issue
Matching	Exact (results are always correct)	Imprecise (need to measure effectiveness)

ตารางที่ 2-1 แสดงตารางเปรียบเทียบระหว่างระบบฐานข้อมูลและระบบการดึงข้อมูล

ระบบการดึงข้อมูลเป็นระบบที่ช่วยทำการวิเคราะห์และค้นหาเพื่อดึงเอาเอกสารจากในส่วนของฐานข้อมูลที่คาดว่าตรงกับที่ผู้ใช้ต้องการออกมาแสดงให้แก่ผู้ใช้โดยอาจมีการจัดทำลำดับ (Ranking) ด้วย ดังรูปที่ 2-1



รูปที่ 2-1 แสดงระบบการดึงข้อมูลในภาพรวม

หลักการในการพิจารณาความตรงประเด็น

- 1) ตรงตามหัวข้อ (Being on the proper subject)
- 2) เป็นข้อมูลที่เป็นปัจจุบัน (Being timely)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 3) เป็นข้อมูลที่มาจากแหล่งที่น่าเชื่อถือ (Being authoritative from a trusted source)
- 4) เป็นข้อมูลที่ทำให้ผู้ใช้พึงพอใจ (Satisfying the goals of the user)

ปัญหาในการใช้ดัชนี(Keyword)

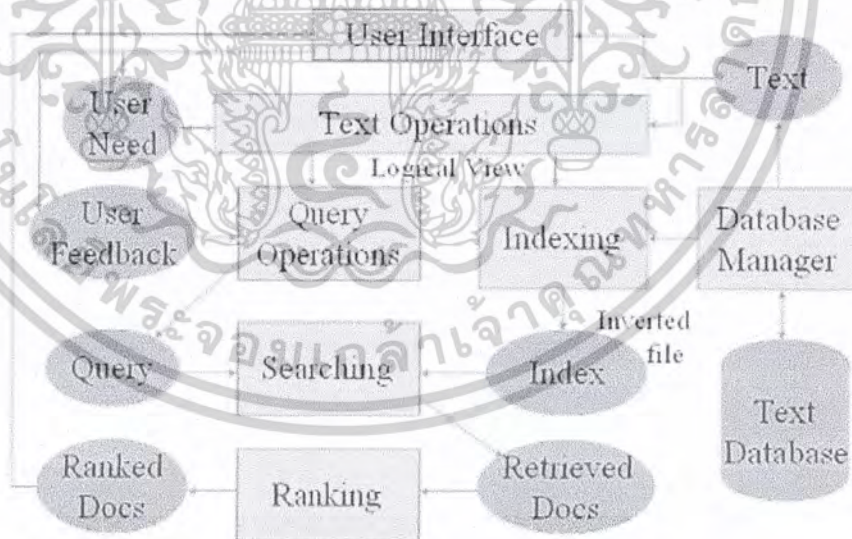
- 1) อาจมีความคลุมเครือของคำที่มีความหมายเหมือนกันหรือคำพ้อง(Synonymous terms)
ตัวอย่างเช่น “restaurant” กับ “cafe”
“PRC” vs. “China”
- 2) คำบางคำอาจมีความหมายที่ไม่ชัดเจนหรือกำกวม(Ambiguous terms)
ตัวอย่างเช่น “bat” หมายถึงเบสบอลหรืออาจหมายถึงค้างคาว
“Apple” หมายถึงชื่อบริษัทหรืออาจหมายถึงชื่อผลไม้
“bit” หมายถึงเป็นหน่วยของข้อมูลหรืออาจหมายถึงการแสดงการเผาผลาญ

ปัญหาทั่วไปในการดึงข้อมูล

1. เราจะแปลงความต้องการให้อยู่ในรูปของตรรกศาสตร์(logic)ได้อย่างไร
2. เราจะนำเอาดัชนี(Index)ไปใช้เพื่อหาคำตอบได้อย่างไร
3. เราจะนำผลลัพธ์ที่ได้ไปจัดลำดับตามความตรงประเด็นได้อย่างไร

2.2 โครงสร้างของระบบดึงข้อมูล

ในระบบทุกระบบจำเป็นต้องมีขั้นตอนในการดำเนินงานเพื่อเป็นแบบแผนของการดึงข้อมูลประกอบด้วยขั้นตอนต่างๆดังรูป2-2 ซึ่งในระบบ



รูปที่ 2 -2 แสดงโครงสร้างการดึงข้อมูล

อธิบายโครงสร้าง

Text Operations : กระบวนการวิเคราะห์ข้อความเพื่อหาดัชนีประกอบไปด้วย

- Stopword removal : กระบวนการตัดคำ
- Stemming : กระบวนการวิเคราะห์คำคล้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Indexing : กระบวนการทำจัดเก็บดัชนีให้มีโครงสร้าง(Inverted file)

Searching : กระบวนการค้นหาตามความต้องการ(query) จากโครงสร้างดัชนี

Ranking : กระบวนการคำนวณค่าความตรงประเด็นให้กับทุกเอกสาร

User Interface : ส่วนจัดการติดต่อกับผู้ใช้ทั้ง

- Query input และ Document output : เป็นส่วนรับความต้องการและแสดงเอกสาร
- Relevance feedback : ป้อนข้อมูลย้อนกลับของผู้ใช้
- Visualization of results : แสดงผลลัพธ์

Query Operations : กระบวนการเปลี่ยนแปลงความต้องการให้ดีขึ้น

- การใช้พจนานุกรม
- การใช้การป้อนข้อมูลย้อนกลับจากผู้ใช้

ความรู้อื่นๆที่เกี่ยวข้อง

1. การจัดการฐานข้อมูล(Database Management)
2. ศาสตร์ของข้อมูล(Library and Information Science)
3. จักรกลอัจฉริยะ(Artificial Intelligence)
4. กระบวนการของภาษารธรรมชาติ(Natural Language Processing)
5. การเรียนรู้(Machine Learning)

2.3 แบบจำลองต่างๆ(Modeling)

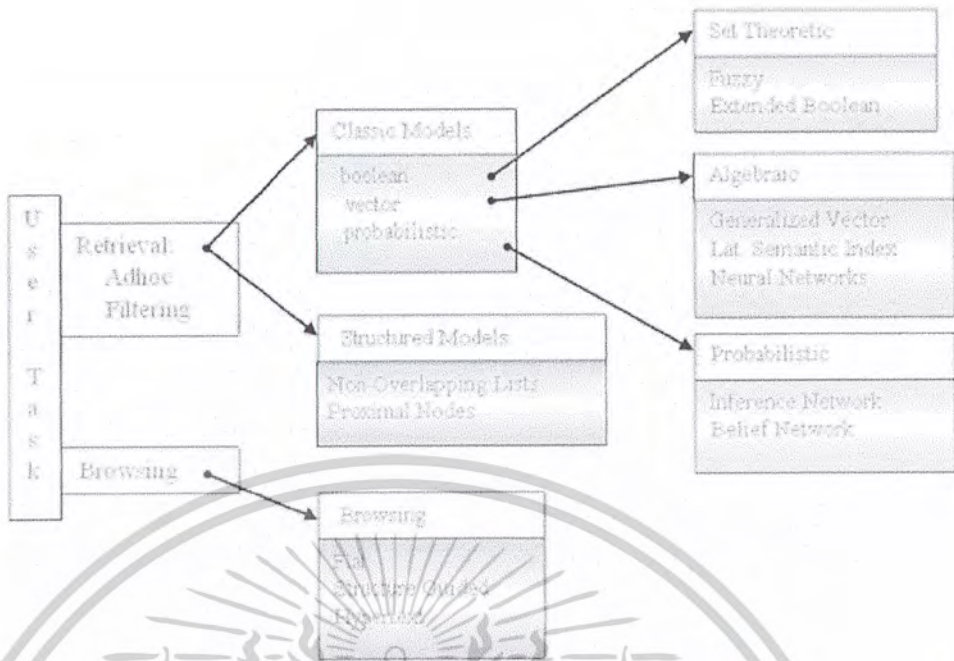
การทำเรียงลำดับ(Ranking) นั้นคือการนำเอกสารที่ผู้ใช้ต้องการขึ้นมาแสดง โดยทำการเรียงลำดับตามลำดับความน่าจะเป็นที่จะตรงประเด็นกับที่ผู้ใช้ต้องการ ซึ่งมีหลักการดังนี้

- ความตรงกับดัชนีของเรา
- การแชร์น้ำหนักเช่น คำว่า“Network”ในเอกสารด้านการเกษตรย่อมต้องมีน้ำหนักหรือความตรงประเด็นที่น้อยกว่าในเอกสารด้านคอมพิวเตอร์
- ความเป็นไปได้

แนวคิดของแบบจำลองการดึงข้อมูล

- แต่ละเอกสารจะถูกแทนด้วยกลุ่มของดัชนี
- ดัชนีคือคำศัพท์ที่เราใช้แทนใจความสำคัญในแต่ละเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-3 แสดงแบบจำลองต่างๆในระบบการดึงข้อมูล

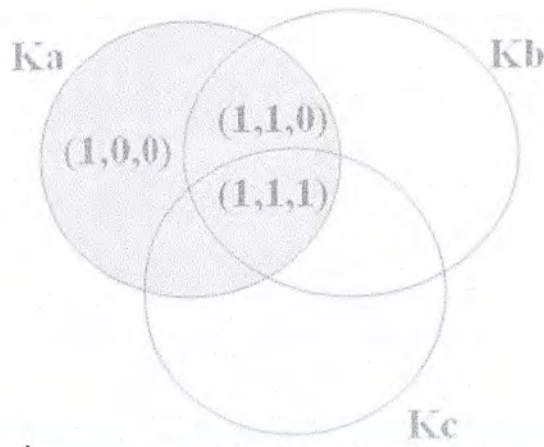
2.3.1 แบบจำลองทางตรรกศาสตร์ (Boolean Model):

เป็นรูปแบบพื้นฐานที่นำหลักการของตรรกศาสตร์ (Boolean) มาใช้ในการให้คำแนะนำนักกล่าวคือ เราจะนำเอกสารต่างๆมาจัดเก็บและหาดัชนีในแต่ละเอกสารแล้วแทนค่าในดัชนีนั้นด้วย 0 หรือ 1 เมื่อมีการป้อนความต้องการเข้ามาเราก็จะแปลงข้อความนั้นให้เป็นความต้องการในรูปแบบตรรกศาสตร์ (Boolean query) แล้วนำมาเทียบกับสิ่งที่ข้อมูลดัชนีที่เราจัดเก็บเพื่อหาคำตอบให้แก่ผู้ใช้งานดังตัวอย่างในตารางที่ 2-2 และรูปที่ 2-4

ดัชนี \ เอกสาร	dog (Ka)	cat (Kb)	tiger (Kc)
Doc 1	1	1	0
Doc 2	0	1	1
Doc 3	1	0	1
Doc 4	1	1	1

ตารางที่ 2-2 แสดงตัวอย่างการจัดเก็บข้อมูลดัชนีในแบบจำลองตรรกศาสตร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-4 แสดงความสัมพันธ์ของกลุ่มดัชนีในรูปแบบเซต

ตัวอย่างการค้นหา

ถ้าผู้ใช้ต้องการหาเอกสารที่เกี่ยวกับ Dog และ Tiger โดยที่เอกสารนั้นจะต้องไม่เกี่ยวกับ Cat เราสามารถแปลงเป็น Boolean Query ได้เป็น

$$Q = (ka \wedge kb) \vee \neg kc \rightarrow \text{ได้ผลลัพธ์คือ Doc 3}$$

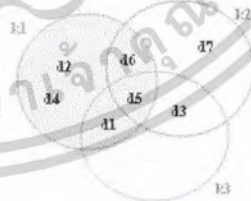
ข้อเสียของจำลองทางตรรกศาสตร์

1. ไม่สามารถจัดลำดับได้ เนื่องจากมีการจัดเก็บค่าน้ำหนักแค่ 0 หรือ 1
2. การแปลงความต้องการให้อยู่ในรูปความหมายทางตรรกศาสตร์นั้นทำได้เฉพาะบางความต้องการเท่านั้น

2.3.2 แบบจำลองแบบมีค่าน้ำหนัก (Vector Model)

มีลักษณะที่คล้ายกับแบบจำลองทางตรรกศาสตร์แต่แทนที่เราจะมีค่าน้ำหนักแค่ 0 กับ 1 นั้นเราสามารถใส่ค่าน้ำหนักตามน้ำหนักความตรงประเด็นของดัชนีนั้นๆ ซึ่งหากทำตามรูปแบบนี้แล้วเราสามารถนำน้ำหนักของแต่ละดัชนี ไปคำนวณหาค่าความตรงประเด็น (Degree of similarity) เพื่อนำมาทำลำดับของความตรงประเด็น ตัวอย่างของรูปแบบจำลองแบบมีค่าน้ำหนักแสดงดังรูปที่ 2-5, 2-6 และ 2-7

The Vector Model: Example 1

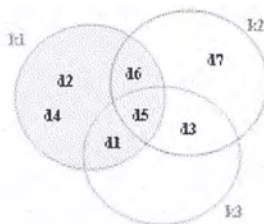


	k1	k2	k3	q * dj
d1	1	0	1	2
d2	1	0	0	1
d3	0	1	1	2
d4	1	0	0	1
d5	1	1	1	3
d6	1	1	0	2
d7	0	1	0	1
q	1	1	1	

รูปที่ 2-5 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

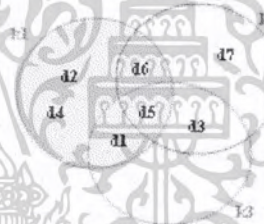
The Vector Model:
Example II



	k1	k2	k3	$q \cdot d_j$
d1	1	0	1	4
d2	1	0	0	1
d3	0	1	1	5
d4	1	0	0	1
d5	1	1	1	6
d6	1	1	0	3
d7	0	1	0	2
q	1	2	3	

รูปที่ 2-6 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก2

The Vector Model:
Example III



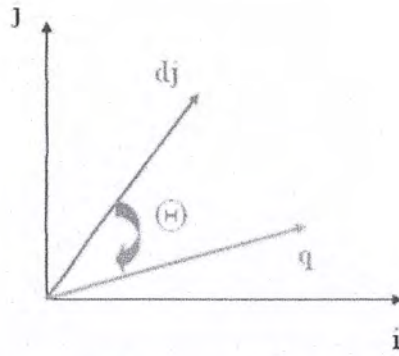
	k1	k2	k3	$q \cdot d_j$
d1	2	0	1	5
d2	1	0	0	1
d3	0	1	3	11
d4	2	0	0	2
d5	1	2	4	17
d6	1	2	0	5
d7	0	5	0	10
q	1	2	3	

รูปที่ 2-7 แสดงตัวอย่างแบบจำลองแบบมีค่าน้ำหนัก3

2.4 หลักการคำนวณความตรงประเด็น

2.4.1 ค่าความเหมือน (Cosine Similarity Measure)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-8 แสดงแบบจำลองแบบมีค่าน้ำหนักในรูปแบบกราฟ

จากรูปที่ 2-8 แกน j และ i คือดัชนี

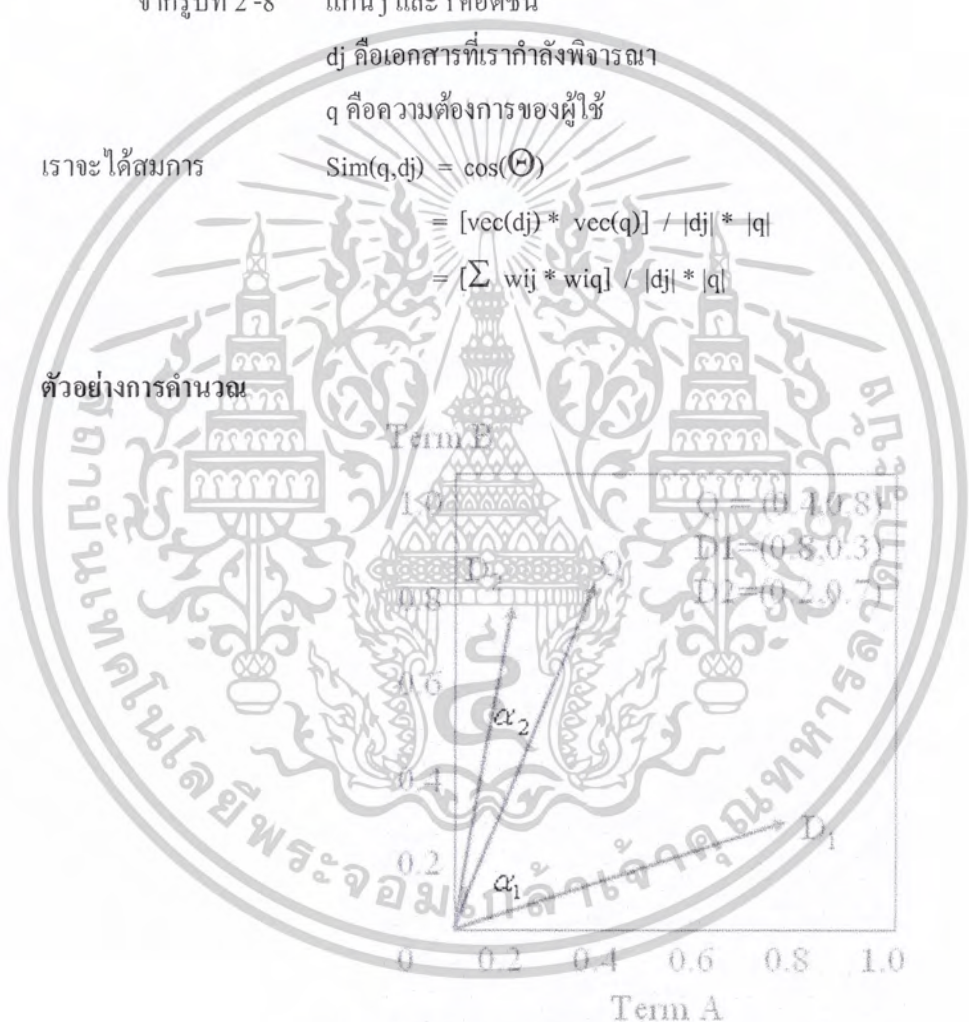
d_j คือเอกสารที่เรากำลังพิจารณา

q คือความต้องการของผู้ใช้

เราจะได้สมการ

$$\begin{aligned} \text{Sim}(q,d_j) &= \cos(\Theta) \\ &= [\text{vec}(d_j) \cdot \text{vec}(q)] / |d_j| \cdot |q| \\ &= [\sum w_{ij} \cdot w_{iq}] / |d_j| \cdot |q| \end{aligned}$$

ตัวอย่างการคำนวณ



รูปที่ 2-9 แสดงตัวอย่างโจทย์แบบจำลองแบบมีค่าน้ำหนัก

$$D_i = (d_{i1}, w_{d1}; d_{i2}, w_{d2}; \dots; d_{it}, w_{dit})$$

$$Q = (q_{i1}, w_{qi1}; q_{i2}, w_{qi2}; \dots; q_{it}, w_{qit})$$

จากรูปที่ 2-9 เราจะได้สมการค่า $\text{CosSim}(d_j, q) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^l (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^l w_{ij}^2} \cdot \sqrt{\sum_{i=1}^l w_{iq}^2}}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned} \text{sim}(Q, D_2) &= \frac{(0.4 \cdot 0.2) + (0.8 \cdot 0.7)}{\sqrt{[(0.4)^2 + (0.8)^2] \cdot [(0.2)^2 + (0.7)^2]}} \\ &= \frac{0.64}{\sqrt{0.42}} = 0.98 \\ \text{sim}(Q, D_1) &= \frac{0.56}{\sqrt{0.58}} = 0.74 \end{aligned}$$

สรุป เอกสาร D2 มีความตรงประเด็นมากกว่าเอกสาร D1

2.4.2 กระบวนการในการกำหนดค่าน้ำหนักให้แก่แต่ละดัชนี

เราจะใช้หลักการของ TF-IDF ดังนี้ $w_{ij} = \text{tf}(i,j) * \text{idf}(i)$

โดยที่ $\text{tf}(i,j)$ เป็นตัวแสดงความถี่ของดัชนีในแต่ละเอกสาร

$\text{idf}(i)$ เป็นตัวแสดงส่วนกลับของความถี่ของเอกสาร

ตัวอย่าง สมมติว่า เรามี Doc1 โดยเราต้องการหาว่า Doc1 นี้ควรให้ค่าน้ำหนักแก่ดัชนี Printer, PC, Monitor อย่างละเท่าไร

- ขั้นแรกเราพิจารณาว่า Doc1 นั้นประกอบคำอะไรบ้างอย่างละเท่าไร

→ Printer(3 คำ), PC(2 คำ), Monitor(1 คำ)

- ขั้นที่สองสมมติว่าเรามี Document อยู่ทั้งหมด 10,000 เอกสารและแต่ละเอกสารมีความถี่ของ

แต่ละดัชนีดังนี้

→ Printer(50 document), PC(1300 document), Monitor(250 document)

สรุป:

Printer : $\text{tf} = 3/3$; $\text{idf} = \ln(10000/50) = 5.3$; $\text{tf-idf} = 5.3$

PC : $\text{tf} = 2/3$; $\text{idf} = \ln(10000/1300) = 2.0$; $\text{tf-idf} = 1.3$

Monitor : $\text{tf} = 1/3$; $\text{idf} = \ln(10000/250) = 3.7$; $\text{tf-idf} = 1.2$

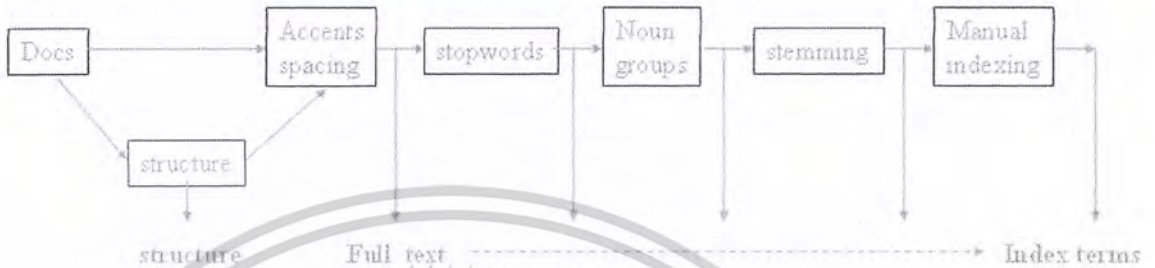
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การทำดัชนี

3.1 ความรู้เบื้องต้นในการหาดัชนี

ในส่วนนี้จะพูดถึงกระบวนการนำข้อความมาเข้ากระบวนการเพื่อให้ได้ดัชนีใจความสำคัญของข้อความนั้นๆออกมาโดยจะประกอบไปด้วยขั้นตอนต่างๆดังรูป 3-1



รูป 3-1 แสดงขั้นตอนการหาดัชนีจากข้อความ

1. การแยกคำ(Lexical Analysis of the text)
2. การตัดคำ(Elimination of stopwords)
3. การพิจารณารากศัพท์(Stemming)
4. รูปแบบดัชนี(Selection of index terms)

3.2 การแยกคำ

เราต้องทำแยกข้อความที่เข้ามาออกให้เป็นคำๆ ให้ได้โดยมีหลักการดังนี้พิจารณาจากสัญลักษณ์ต่างๆดังนี้

1. ช่องว่าง(Space) : เป็นวิธีใช้ในการแบ่งคำที่นิยมใช้กัน โดยเฉพาะในภาษาอังกฤษเนื่องจากภาษาอังกฤษมีการแบ่งคำโดยใช้ช่องว่างอยู่แล้ว
2. ตัวเลข(Digits) : โดยส่วนใหญ่ตัวเลขจะเป็นคำดัชนีที่ไม่ดีนักแต่ก็มีบางคำที่จำเป็นต้องประกอบและลงท้ายด้วยตัวเลขเช่น “B6”, “B12” เป็นศัพท์เกี่ยวกับวิตามิน
3. ขีดสั้น(Hyphens) : เป็นการยากที่ใช้ในการแบ่งเพราะว่า
 - อาจเป็นจุดแบ่งคำ เช่นคำว่า state-of-the-art
 - อาจเป็นคำๆเดียวกัน เช่นคำว่า F-16, Jean-Claude
4. เครื่องหมายวรรคตอน(punctuation marks) : การใช้สัญลักษณ์(,) ในการแบ่งคำ
5. การใช้ขนาดตัวอักษร(the case of the letters) : พิมพ์ใหญ่หรือว่าพิมพ์เล็กเช่น
CatDog → Cat,Dog

3.3 การตัดคำ

เราจะกำจัดคำที่อยู่ในจำพวกได้แก่

- เป็นคำที่มักจะปรากฏอยู่ในเกือบทุกเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- คำนำหน้านาม(articles), บุพบท(prepositions), สันธาน (conjunctions)
การกำจัดคำพวกนี้จะช่วยให้เราลดขนาดในการพิจารณาหาคำที่เป็นดัชนีลงได้

3.4 การพิจารณารากศัพท์

เราจะพยายามที่จะเก็บเฉพาะคำที่เป็นรากศัพท์เท่านั้น เพื่อประโยชน์ในการพิจารณาความเหมือนของคำตัวอย่างของคำที่มีความหมายเหมือนกันแต่เขียนไม่เหมือนกัน

connect, connected, connecting, connection, connections
effectiveness --> effective --> effect
picnicking --> picnic

วิธีในการพิจารณา

- พิจารณาคด้วยคน (Manual)
- ทำเป็นกฎเอาไว้(Automatic)

ข้อดีคือมีประสิทธิภาพและช่วยลดขนาดของการจัดเก็บ index

ข้อเสียคือ มันอาจทำให้คำศัพท์ที่เป็นวลีคือประกอบด้วยหลายคำมาต่อกันแล้วได้ความหมายใหม่ ถูกทำลายไป

- การเก็บทำเป็นพจนานุกรม (Evaluation) คือสร้างเป็นตารางจัดเก็บ ดังตารางที่ 3-1

Term	Stem
engineering	engineer
engineered	engineer
engineer	engineer

ตารางที่ 3-1 แสดงการทำพจนานุกรมเก็บรากศัพท์

จะมีความถูกต้องสูงโดยที่ประสิทธิภาพขึ้นอยู่กับข้อมูลที่ใส่เข้าไปแต่มันจะมีขนาดและใช้ทรัพยากรสูงตามไปด้วย

กฎในการพิจารณารากศัพท์

3.4.1) พิจารณาจากความถี่ที่เกิดขึ้น

ตัวอย่างการทำเช่น

คำศัพท์	ความถี่	
1. user	15	
2. users	1	
3. used	3	โดยที่ผู้ใช้จะเลือกคำที่จะมาเป็นรากศัพท์จากจำนวน
4. using	2	คำที่ปรากฏบ่อยที่สุด ในที่นี้คือ user

3.4.2) พิจารณาจำนวนคำที่แตกต่างของตัวอักษรที่ตามมา(Successor Variety)

ตัวอย่างเช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำศัพท์ในข้อความ: able, axle, accident, ape, about

จำนวนคำที่แตกต่างของตัวอักษรที่ตามมาของคำว่า apple

ตำแหน่งแรก: มี 4 (b, x, c, p)

ตำแหน่งสอง: มี 1 (e)

แนวคิด :

จำนวนคำที่แตกต่างของตัวอักษรที่ตามมาของส่วนของคำใดที่ลดลงอย่างรวดเร็วและเพิ่มขึ้นอย่างรวดเร็วจนถึงขีดสุดคำ(sharply increase)นั้นแหละเป็นรากศัพท์

ตัวอย่างการใช้งาน

คำศัพท์ที่จะทดสอบหารากศัพท์: READABLE

คำศัพท์ทั้งหมด : ABLE, BEATABLE, FIXABLE, READS, READABLE, READING, RED, ROPE, RIPE

คำอุปสรรค	จำนวนคำที่แตกต่างของตัวอักษรที่ตามมา	ตัวอักษร
R	3	E, O, I
RE	2	A, D
REA	1	D
READ	3	A, I, S
READA	1	B
READAB	1	L
READABL	1	E
READABLE	1	blank

3.4.3) สูตรวัดค่าความเหมือน(n-gram stemmers)

ใช้วัดว่าคำศัพท์ 2 คำมีความเหมือนกันเท่าไรตามสูตร

$$S = \frac{2C}{A+B}$$

โดยที่

A : จำนวนของส่วนพจน์ที่ไม่ซ้ำกันในคำศัพท์คำแรก

B : จำนวนของส่วนพจน์ที่ไม่ซ้ำกันในคำศัพท์คำสอง

C : จำนวนของส่วนพจน์ที่ไม่ซ้ำกันที่มีในทั้งส่วน A และ B

ตัวอย่าง

statistics => st ta at ti is st ti ic cs

พจน์ที่ไม่ซ้ำกัน => at cs ic is st ta ti

statistical => st ta at ti is st ti ic ca al

พจน์ที่ไม่ซ้ำกัน => al at ca ic is st ta ti

$$S = \frac{2C}{A+B} = \frac{2*6}{7+8} = 0.80$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถนำไปประยุกต์ใช้โดยทำเป็นตารางวัดความเหมือน(similarity matrix) เพื่อใช้ในการตัดสินใจว่าคำศัพท์คำใดเหมือนกับคำใด

	word1	word2	word3	...	wordn-1
word1					
wrod2	S21				
word3	S31	S32			
.					
.					
wordn	Sn1	Sn2	Sn3	...	Sn(n-1)

3.4.4) สูตรวัดความแตกต่าง(Q-gram stemmers)

คล้ายกับ n-gram แต่ต้องกำหนดค่า Q และผลที่ได้จะแสดงค่าความต่างของคำศัพท์ทั้งสองตัวอย่างเช่น

กำหนด $Q=3$ ($L+Q-1$)

statistics $\Rightarrow 10+3-1=12$

{ #s,#st,sta,tat,ati,tis,ist,sti,tic,ics,cs#,s## }

statistical $\Rightarrow 11+3-1=13$

{#s,#st,sta,tat,ati,tis,ist,sti,tic,ica,cal,al#,l#}

Q-gram Distance = 4

3.4.5) ใช้หลักการลบคำอุปสรรคออก(Affix Removal Stemmers)

การลบ suffixes หรือ prefixes จากคำศัพท์เพื่อเปลี่ยนให้อยู่ในรูปของรากศัพท์เช่น ถ้าคำศัพท์นั้นจบด้วย "ies" แต่ไม่ใช่ "eies" หรือ "aies" แล้ว "ies" \rightarrow "y" ถ้าคำศัพท์นั้นจบด้วย "es" แต่ไม่ใช่ "acs", "ees", หรือ "oes" แล้ว "es" \rightarrow "e" ถ้าคำศัพท์นั้นจบด้วย "s" แต่ไม่ใช่ "us" หรือ "ss" แล้ว "s" \rightarrow NULL

3.5 รูปแบบดัชนี

คำเดี่ยว(Single words)

คำศัพท์เดี่ยวบางคำก็มีลักษณะที่กำวมและไม่เฉพาะเจาะจงซึ่งอาจทำให้เกิดข้อผิดพลาดได้

วลี(Phrases)

คำศัพท์บางประเภทควรที่จะเป็นวลีเนืองหากแยกออกเป็นคำๆแล้วจะทำให้ความหมายเปลี่ยนไป

มีทั้งคำเดี่ยวและวลี(Single words and Phrases)

ใช้ทั้ง2แบบร่วมกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การประยุกต์ใช้งานในระบบฐานความรู้

เราได้นำหลักการของแบบจำลองแบบมีค่าน้ำหนัก มาประยุกต์ใช้ใน โครงงานของเราเพื่อช่วย ค้นหาเอกสารที่ใกล้เคียงกับความต้องการของ ผู้ใช้ และยังสามารถจัดลำดับความตรงประเด็นให้กับ ผู้ใช้งานอีกด้วย

ขั้นตอนในการพิจารณาผลลัพธ์มีดังนี้

1. เราทำการพิจารณาค่าของกลุ่มของดัชนีให้ในแต่ละเอกสารในฐานความรู้ ข้อมูลและข้อมูลความต้องการที่ผู้ใช้ได้กรอกเข้ามา โดยที่จะต้องทำการคำนวณหาค่าน้ำหนักในแต่ละดัชนีด้วย
2. เราจะทำการคำนวณหาความเหมือนระหว่างความต้องการของผู้ใช้กับเอกสารทั้งหมด
3. เรานำค่าความเหมือนที่ได้ในข้อ 3 มาจัดเรียงลำดับตามค่าความเหมือนเพื่อแสดงเป็นผลลัพธ์แก่ ผู้ใช้

4.1 ขั้นตอนในการหาดัชนีในข้อความตามรูปที่ 3-1 มีดังนี้

4.1.1 การแยกคำ

เราจะต้องทำการแยกข้อความที่เข้ามาออกเป็นคำๆก่อน โดยใช้หลักที่ว่า

- เราจะใช้ช่องว่าง เครื่องหมายวรรคตอน และ เครื่องหมายทับในการแบ่งเป็นคำเช่น Hello! I can't use pc in ECC/Olala. → "Hello!", "I", "can't", "use", "pc", "in", "ECC", "Olala."
- เมื่อเราได้คำมาแล้วเรายังจะต้องนำคำดังกล่าวมาพิจารณาว่าคำนั้นลงท้ายหรือขึ้นต้นด้วยสัญลักษณ์ที่ไม่มีความหมาย("?", ":", "!",) หรือเปล่าซึ่งเราจะตัดมันทิ้งไปด้วยเช่น "Hello!", "I", "can't", "use", "pc", "in", "ECC", "Olala." → "Hello", "I", "can't", "use", "pc", "in", "ECC", "Olala"

4.1.2 การตัดคำ

เราก็จะนำคำที่ได้จากขั้นแรกมากำจัดคำที่ไม่มีมีความหมายหรือเป็นคำที่เกิดขึ้นบ่อยๆในทุกข้อความ โดยเรามีชุดของคำดังกล่าวไว้อยู่แล้ว ตามทฤษฎีของ THOMAS ซึ่งมีประมาณ 425 คำเรียงตามตัวอักษรดังนี้

A

a, about, above, according, across, actually, adj, after, afterwards, again, against, albeit, all, almost, alone, along, already, also, although, always, among, amongst, an, and, another, any, anyhow, anyone, anything, anywhere, are, aren't, around, as, at.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

B

b, be, became, because, become, becomes, becoming, been, before, beforehand, begin, beginning, behind, being, below, beside, besides, between, beyond, billion, both, but, by.

C

c, can, can't, cannot, caption, co, co., could, couldn't.

D

d, did, didn't, do, does, doesn't, don't, down, during.

E

e, each, eg, eight, eighty, either, else, elsewhere, end, ending, enough, etc, even, ever, every, everyone, everything, everywhere, except.

F

f, few, fifty, first, five, for, former, formerly, forty, found, , four, from, further.

G-H

g, h, had, has, hasn't, have, haven't, he, he'd, he'll, he's, hence, her, here, here's, hereafter, hereby, herein, hereupon, hers, herself, him, himself, his, how, however, hundred.

I

i, i'd, i'll, i'm, i've, ie, if, in, inc., indeed, instead, into, is, isn't, it, it's, its, itself.

J-K-L

j, k, l, last, later, latter, latterly, least, less, let, let's, like, likely, ltd.

M

m, made, make, makes, many, maybe, me, meantime, meanwhile, might, million, miss, more, moreover, most, mostly, mr, mrs, much, must, my, myself.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

N

n, namely, neither, never, nevertheless, next, nine, ninety, no, nobody, none, nonetheless, noone, nor, not, nothing, now, nowhere.

O

o, of, off, often, on, once, one, one's, only, onto, or, other, others, otherwise, our, ours, ourselves, out, over, overall, own.

P-Q

p, per, perhaps, q.

R

r, rather, recent, recently.

S

s, same, seem, seemed, seeming, seems, seven, seventy, several, she, she'd, she'll, she's, should, shouldn't, since, six, sixty, so, some, somehow, someone, something, sometime, sometimes, somewhere, still, stop, such.

T

t, taking, ten, than, that, that'll, that's, that've, the, their, them, themselves, then, thence, there, there'd, there'll, there're, there's, there've, thereafter, thereby, therefore, therein, thereupon, these, they, they'd, they'll, they're, they've, thirty, this, those, though, thousand, three, through, throughout, thru, thus, to, together, too, toward, towards, trillion, twenty, two.

U

u, under, unless, unlike, unlikely, until, up, upon, us, used, using.

V

v, very, via.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

W

w, was, wasn't, we, we'd, we'll, we're, we've, well, were, weren't, what, what'll, what's, what've, whatever, when, whence, whenever, where, where's, whereafter, whereas, whereby, wherein, whereupon, wherever, whether, which, while, whither, who, who'd, who'll, who's, whoever, whole, whom, whomever, whose, why, will, with, within, without, won't, would, wouldn't.

X-Y-Z

x, y, yes, yet, you, you'd, you'll, you're, you've, your, yours, yourself, yourselves, z.

อื่นๆ

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

หลังจากทำตามขั้นตอนในหัวข้อ 4.1.2 แล้วเราจะถือว่าคำที่เหลือคือดัชนีซึ่งเป็นตัวแสดงใจความสำคัญของข้อความนั้นๆ เช่น

Hello! I can't use pc in ECC/Olala. → "pc", "ecc", "olala"

4.2 การพิจารณาให้คำนำหน้า

การที่เราจะพิจารณาคำนำหน้าให้แต่ละดัชนีตามรูปแบบของ Tf-idf ในหัวข้อ 2.4.2 เราต้องมีการจัดเก็บความถี่ของแต่ละดัชนีซึ่งกระบวนการนั้นๆ เราต้องมีการพิจารณาคำคล้ายเพื่อช่วยในการเก็บค่าความถี่ของดัชนีไม่ให้มีการซ้ำซ้อน โดยเราใช้หลัก 2 อย่างคือ

- เราพิจารณาตามพจนานุกรมที่ทำขึ้น โดยจะจัดเก็บคำที่มีความหมายเดียวกันไว้ด้วยกันซึ่งคำนั้นไม่สามารถถูกตรวจสอบได้โดยใช้หลักกฎคำอุปสรรคเช่น

PC, Computer, Desktop → พวกนี้ถือว่าเป็นคำๆ เดียวกัน

notebook, Desknote → พวกนี้ถือว่าเป็นคำๆ เดียวกัน

- เราทำการใส่กฎของคำอุปสรรคเหล่านี้เข้าไปดังนี้

กฎจะเป็นคำทางด้านซ้ายมือเช่น sses>ss มีความหมายว่าจะลงท้ายด้วย"sses" หรือ "ss" ก็มีความหมายเหมือนกันแต่ถ้าเป็น s>* มีความหมายว่าจะลงท้ายด้วย"s" หรือไม่ ก็มีความหมายเหมือนกัน ส่วนกลุ่มคำด้านขวาคือตัวอย่างคำตามกฎ

SSes > SS caresses > caress

IES > I ponies > poni

ties > ti

SS > SS caress > caress

S > * cats > cat

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

EED > EE	agreed > agree
ED > *	plastered > plaster
	bled > bled
ING > *	motoring > motor
	sing > sing
AT > ATE	conflat(ed) > conflate
BL > BLE	troubl(ed) > trouble
IZ > IZE	siz(ed) > size
Y > I	happy > happi
ATIONAL > ATE	relational > relate
TIONAL > TION	conditional > condition
	rational > rational
ENCI > ENCE	valenci > valence
ANCI > ANCE	hesitanci > hesitance
IZER > IZE	digitizer > digitize
ABLI > ABLE	conformabli > conformable
ALLI > AL	radicalli > radical
ENTLI > ENT	differentli > different
ELI > E	vileli > vile
OUSLI > OUS	analogousli > analogous
IZATION > IZE	vietnamization > vietnamize
ATION > ATE	predication > predicate
ATOR > ATE	operator > operate
ALISM > AL	feudalism > feudal
IVENESS > IVE	decisiveness > decisive
FULNESS > FUL	hopefulness > hopeful
OUSNESS > OUS	callousness > callous
ALITI > AL	formaliti > formal
IVITI > IVE	sensitiviti > sensitive
BILITI > BLE	sensibiliti > sensible
ICATE > IC	triplicate > triplic
ATIVE > *	formative > form
ALIZE > AL	formalize > formal
ICITI > IC	electriciti > electric

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ICAL > IC	electrical > electric
FUL > *	hopeful > hope
NESS > *	goodness > good
AL > *	revival > reviv
ANCE > *	allowance > allow
ENCE > *	inference > infer
ER > *	airliner > airlin
IC > *	gyroscopic > gyroscop
ABLE > *	adjustable > adjust
IBLE > *	defensible > defens
ANT > *	irritant > irrit
EMENT > *	replacement > replac
MENT > *	adjustment > adjust
ENT > *	dependent > depend
(*S or *T) ION > *	adoption > adopt
OU > *	homologou > homolog
ISM > *	communism > commun
ATE > *	activate > activ
ITI > *	angulariti > angular
OUS > *	homologous > homolog
IVE > *	effective > effect
IZE > *	bowdlerize > bowdler

4.3 การคำนวณหาค่าความเหมือน

เป็นขั้นตอนการพิจารณาว่าเอกสารใดมีความตรงประเด็นกับสิ่งที่ผู้ใช้ต้องการ โดยเราต้องทำการคำนวณหาจากสูตรค่าความเหมือน ตามหัวข้อที่ 2.4 โดยค่าที่เราต้องการใช้คือค่าน้ำหนักในแต่ละดัชนีในหัวข้อที่ 4.2 ของทั้งส่วนความต้องการของผู้ใช้และส่วนเอกสารยกตัวอย่างเช่น

- ขั้นแรกเราพิจารณาว่าในแต่ละเอกสารนั้นประกอบค่าอะไรบ้างอย่างละเท่าไร

เอกสาร 1 → Printer(2 คำ), PC(1 คำ), Monitor(1 คำ), Scanner(1 คำ)

เอกสาร 2 → Printer(3 คำ), Monitor(2 คำ)

ความต้องการผู้ใช้ → Monitor(1 คำ), Scanner(1 คำ), Printer(1 คำ)

- ขั้นที่สองสมมติว่าเรามีเอกสารอยู่ทั้งหมด 10,000 เอกสารและแต่ละเอกสารมีความถี่ของแต่ละดัชนีดังนี้

→ Printer(50 เอกสาร), PC(1300 เอกสาร), Monitor(250 เอกสาร), Scanner(200 เอกสาร)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่บนสื่อออนไลน์

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปของความต้องการ: Printer : tf= 1/1; idf= ln(10000/50) = 5.3 ; tf-idf = 5.3
 Monitor: tf= 1/1; idf= ln(10000/250) = 3.68 ; tf-idf = 3.68
 Scanner: tf= 1/1; idf= ln(10000/200) = 3.91 ; tf-idf = 3.91

สรุปของเอกสาร1: Printer : tf= 2/2; idf= ln(10000/50) = 5.3 ; tf-idf = 5.3
 PC : tf= 1/2; idf= ln(10000/1300) = 2.04; tf-idf = 1.02
 Monitor: tf= 1/2; idf= ln(10000/250) = 3.68 ; tf-idf = 1.844
 Scanner: tf= 1/2; idf= ln(10000/200) = 3.91 ; tf-idf = 1.95

สรุปของเอกสาร2: Printer : tf= 3/3; idf= ln(10000/50) = 5.3 ; tf-idf = 5.3
 Monitor: tf= 2/3; idf= ln(10000/250) = 3.68 ; tf-idf = 2.45

แล้วเราก็ทำการหาค่าความเหมือนจากสูตร

$$\text{CosSim}(d_j, q) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^r (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^r w_{ij}^2} \cdot \sqrt{\sum_{i=1}^r w_{iq}^2}}$$

ระหว่าง ความต้องการกับเอกสารที่1

$$\begin{aligned} \text{sim}(Q, D1) &= \frac{(5.3 \cdot 5.3) + (3.68 \cdot 1.844) + (3.91 \cdot 1.95) + (0 \cdot 1.02)}{\sqrt{[(5.3)^2 + (3.68)^2 + (3.91)^2]} \cdot \sqrt{[(5.3)^2 + (1.02)^2 + (1.844)^2 + (1.95)^2]}} \\ &= \frac{42.5}{\sqrt{2095.59}} = 0.928 \end{aligned}$$

ระหว่าง ความต้องการกับเอกสารที่2

$$\begin{aligned} \text{sim}(Q, D2) &= \frac{(5.3 \cdot 5.3) + (3.68 \cdot 2.45) + (3.91 \cdot 0) + (0 \cdot 0)}{\sqrt{[(5.3)^2 + (3.68)^2 + (3.91)^2]} \cdot \sqrt{[(5.3)^2 + (2.45)^2]}} \\ &= \frac{37.106}{\sqrt{1863.76}} = 0.86 \end{aligned}$$

ดังนั้นเอกสาร1มีความตรงประเด็นมากกว่าเอกสารชุด2ในปฏิบัติจริงๆแล้วเราจำเป็นต้องหาค่าความเหมือนของเอกสารทั้งหมดเลยแล้วจึงบอกได้ว่าเอกสารใดที่ตรงประเด็นที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

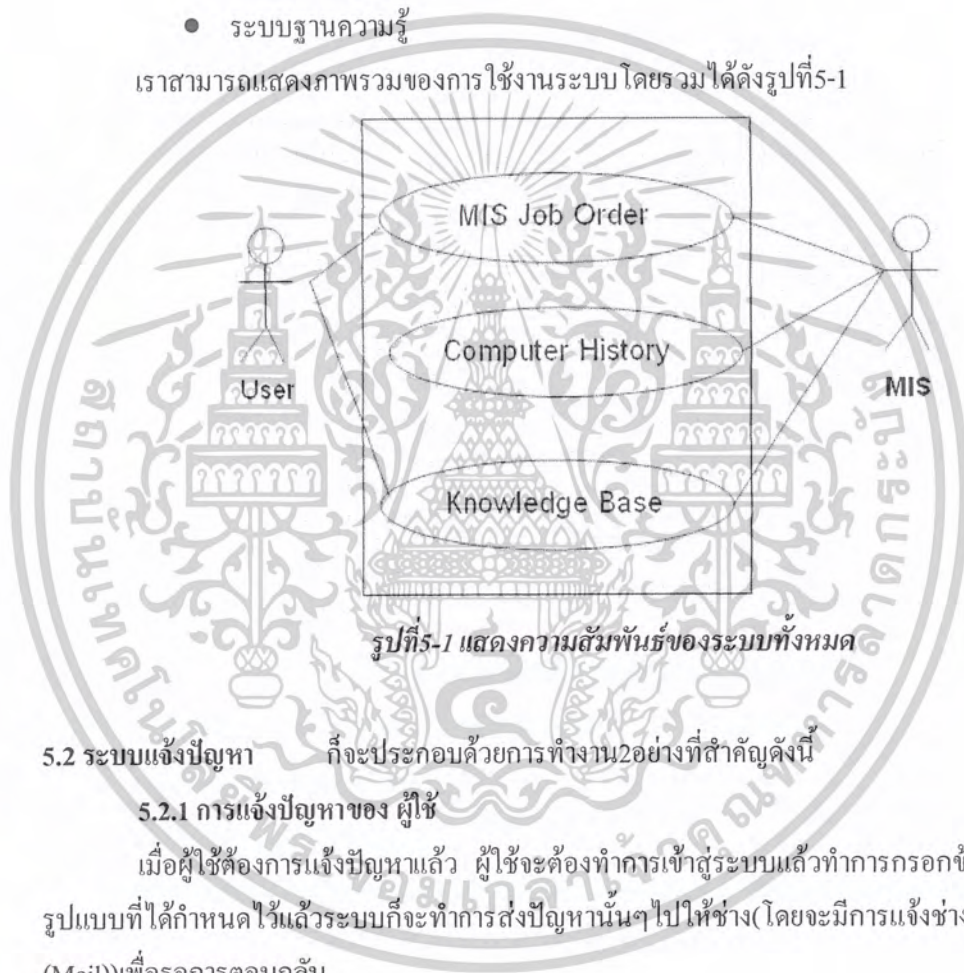
บทที่ 5

การวิเคราะห์และออกแบบระบบ

5.1 ลักษณะโดยรวมของระบบ

1. เป็นระบบแบบโปรแกรมเว็บ(Web Application)
2. ใช้โปรแกรม Microsoft SQL Server 2000 เป็นตัวจัดเก็บฐานข้อมูล
3. แบ่งระบบออกเป็น 3 ระบบหลักๆคือ
 - ระบบแจ้งปัญหา
 - ระบบประวัติการซ่อมแซม
 - ระบบฐานความรู้

เราสามารถแสดงภาพรวมของการใช้งานระบบโดยรวมได้ดังรูปที่5-1

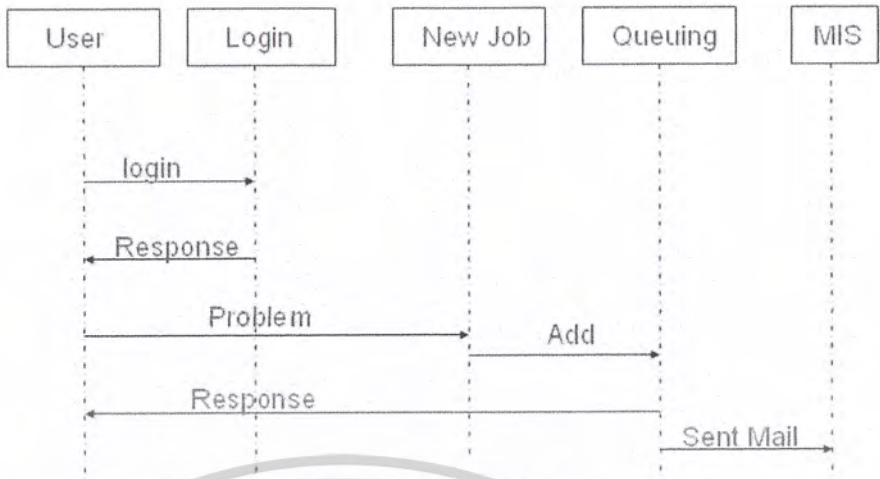


รูปที่5-1 แสดงความสัมพันธ์ของระบบทั้งหมด

5.2 ระบบแจ้งปัญหา ก็จะประกอบด้วยการทำงาน2อย่างที่สำคัญดังนี้

5.2.1 การแจ้งปัญหาของผู้ใช้

เมื่อผู้ใช้ต้องการแจ้งปัญหาแล้ว ผู้ใช้จะต้องทำการเข้าสู่ระบบแล้วทำการกรอกข้อมูลปัญหาตามรูปแบบที่ได้กำหนดไว้แล้วระบบก็จะทำการส่งปัญหานั้นๆไปให้ช่าง(โดยจะมีการแจ้งช่างทาง เมล์ (Mail))เพื่อรอการตอบกลับ



รูปที่ 5-2 แสดงขั้นตอนการแจ้งปัญหาของผู้ใช้

5.2.2 การตอบรับปัญหาของช่าง

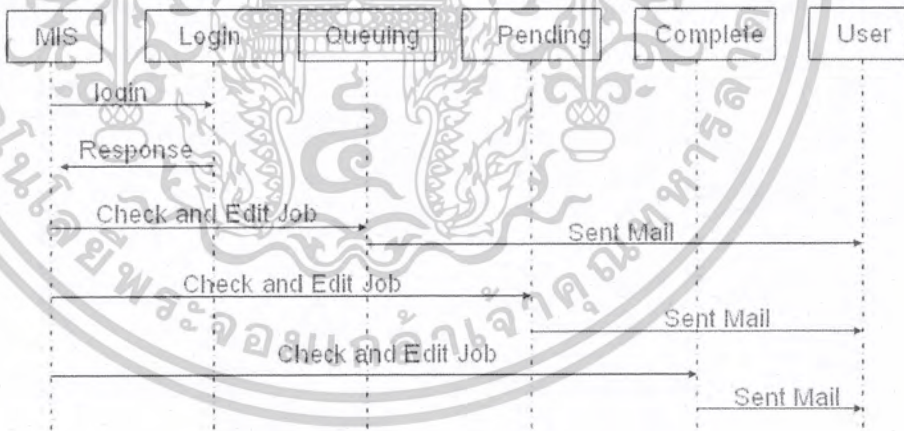
ช่างจะทำการตรวจสอบปัญหาต่างๆที่เข้ามาและทำการตอบกลับ ไปยังผู้ใช้เพื่อให้ทราบถึง ความก้าวหน้าโดยระบบ ได้มีการแบ่งสถานะของปัญหาไว้ 3 ขั้นตอนคือ

Queuing : ปัญหานั้นๆอยู่ระหว่างการรอช่างตรวจสอบ

Pending : ปัญหานั้นๆอยู่ระหว่างการซ่อมแซม

Complete : ปัญหานั้นๆถูกแก้ไขสมบูรณ์แล้ว

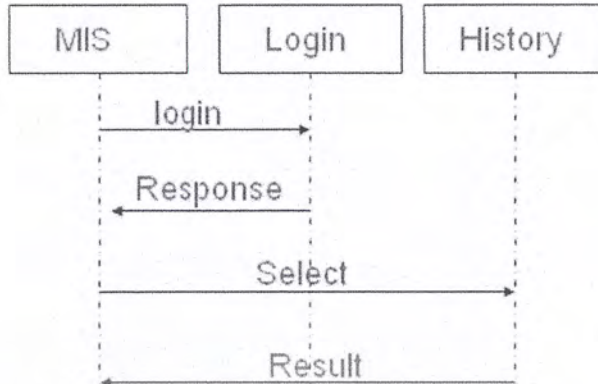
ซึ่งในแต่ละขั้นก็จะมี การส่งแจ้งเตือนความก้าวหน้าไปยังเมลล์ของผู้ใช้ด้วย



รูปที่ 5-3 แสดงขั้นตอนการตอบรับปัญหาของช่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ระบบประวัติการซ่อมแซม

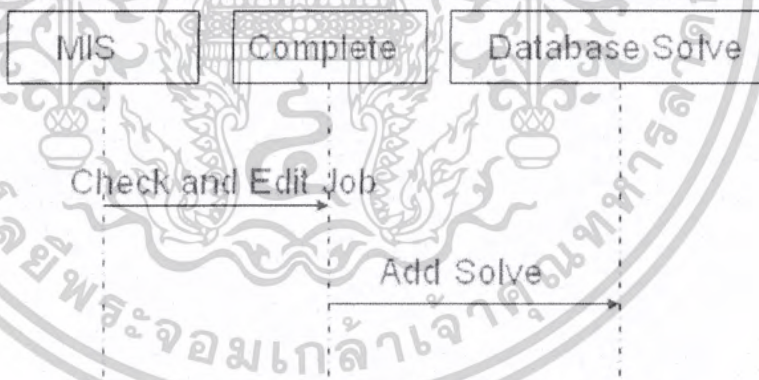


รูปที่ 5-4 แสดงขั้นตอนการทำงานของระบบประวัติการซ่อมแซม

ช่างจะทำการเลือกว่าจะดูข้อมูลใดบ้าง แล้วระบบจะทำการค้นหาแล้วแสดงผลตามที่ช่างต้องการ

5.4 ระบบฐานความรู้ ประกอบด้วยการทำงาน 2 อย่างที่สำคัญดังนี้

5.4.1 การสร้างเอกสารวิธีการแก้ปัญหา

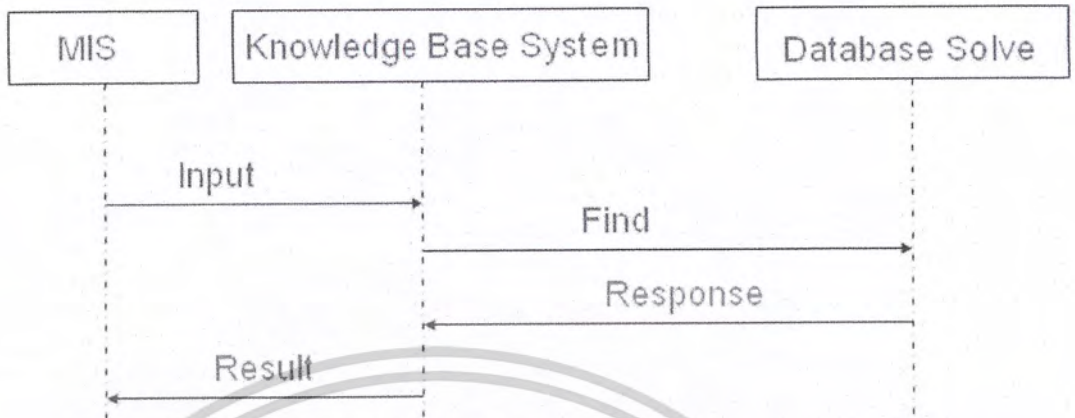


รูปที่ 5-5 แสดงขั้นตอนการทำงานของที่มาเอกสารวิธีการแก้ปัญหา

เมื่อช่างได้ทำการแก้ปัญหาเรียบร้อยแล้ว ระบบจะนำวิธีการแก้ปัญหาที่ช่างระบุไว้ในแต่ละปัญหานั้นมาทำการเก็บเพื่อทำเป็นเอกสารวิธีการแก้ปัญหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.4.2 การใช้งานระบบฐานความรู้ของผู้ใช้



รูปที่ 5-6 แสดงขั้นตอนการใช้งานระบบฐานความรู้ของผู้ใช้

ผู้ใช้งานทำการป้อนข้อมูลในสิ่งที่ต้องรู้แล้วระบบจะทำการค้นในฐานข้อมูลเพื่อให้ได้ผลลัพธ์ตามที่ผู้ใช้งานต้องการออกมาแสดง

นอกจากนี้ ระบบหลักข้างต้นแล้ว ยังมีฟังก์ชันย่อยๆ มากมาย เช่น การลงทะเบียน , การแก้ไขข้อมูลส่วนตัว , การแก้ไขพิกัดมาสเตอร์ไฟล์ (department, location, section) , การค้นหา ปัญหา เป็นต้น

บทที่ 6

การจัดเก็บฐานข้อมูล

6.1 ตารางต่างๆในฐานข้อมูล

- ตาราง T_Index ทำหน้าที่จัดเก็บดัชนีและเป็นพจนานุกรมดัชนีด้วยมีโครงสร้างดังตารางที่ 6-1

T_Index
Index_ID(int) (PK)
Index(ntext)

ตารางที่ 6-1 แสดงโครงสร้างตาราง T_Index

- ตาราง T_Solve ทำหน้าที่จัดเก็บฐานข้อมูลความรู้ที่ผู้ใช้ต้องการทราบและยังทำการจัดเก็บกลุ่มดัชนีของเอกสารนั้นๆด้วยมีโครงสร้างดังตารางที่ 6-2

T_Solve
Solve_ID(int) (PK)
Solve(ntext)
KeyWord(ntext)

ตารางที่ 6-2 แสดงโครงสร้างตาราง T_Solve

- ตาราง T_Location ทำหน้าที่จัดเก็บว่าใน โรงงานมี Location อะไรบ้างมีโครงสร้างดังตารางที่ 6-3

T_Location
Location_ID(int) (PK)
Location (nvarchar)

ตารางที่ 6-3 แสดงโครงสร้างตาราง T_Location

- ตาราง T_Department ทำหน้าที่จัดเก็บว่าใน โรงงานมี Department อะไรบ้างมีโครงสร้างดังตารางที่ 6-4

T_Department
Department_ID(int) (PK)
Department (nvarchar)

ตารางที่ 6-4 แสดงโครงสร้างตาราง T_Department

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ตาราง T_Machine ทำหน้าที่จัดเก็บว่าในโรงงานมี Machine อะไรบ้างมีโครงสร้างดังตารางที่ 6-5

T_Machine
Machine_ID(int) (PK)
Machine (nvarchar)

ตารางที่ 6-5 แสดงโครงสร้างตาราง T_Machine

- ตาราง T_Project ทำหน้าที่จัดเก็บว่าในโรงงานมี Project อะไรบ้างมีโครงสร้างดังตารางที่ 6-6

T_Project
Project_ID(int) (PK)
Project (nvarchar)

ตารางที่ 6-6 แสดงโครงสร้างตาราง T_Project

- ตาราง T_ProblemType ทำหน้าที่จัดเก็บว่าในโรงงานมี ProblemType อะไรบ้างมีโครงสร้างดังตารางที่ 6-7

T_ProblemType
ProblemType_ID(int) (PK)
ProblemType (nvarchar)

ตารางที่ 6-7 แสดงโครงสร้างตาราง T_ProblemType

- ตาราง T_Section ทำหน้าที่จัดเก็บว่าในโรงงานมี Section อะไรบ้างมีโครงสร้างดังตารางที่ 6-8

T_Section
Section_ID(int) (PK)
Section (nvarchar)

ตารางที่ 6-8 แสดงโครงสร้างตาราง T_Section

- ตาราง T_Type ทำหน้าที่จัดเก็บว่าในโรงงานมี Type อะไรบ้างมีโครงสร้างดังตารางที่ 6-9

T_Type
Type_ID(int) (PK)
Type (nvarchar)

ตารางที่ 6-9 แสดงโครงสร้างตาราง T_Type

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาดูเท่านั้น ไม่อนุญาติให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ตาราง T_History ทำหน้าที่จัดเก็บว่าประวัติข้อมูลของอุปกรณ์คอมพิวเตอร์นั้นๆมีใครสร้าง
ดังตารางที่6-10

T_History	
NO (int)	(PK)
ComputerName	(ntext)
Type	(nvarchar)
Dept_ID(int)	(FK)
Section_ID(int)	(FK)
Location_ID(int)	(FK)
Project_ID(int)	(FK)
OS	(nvarchar)
Model	(nvarchar)
CPU	(nvarchar)
RAM	(nvarchar)
HDD	(nvarchar)
Service	(nvarchar)
Remark	(ntext)

ตารางที่6-10 แสดงโครงสร้างตาราง T_History

- ตาราง T_Request ทำหน้าที่จัดเก็บว่าข้อมูลรายละเอียดของแต่ละการแจ้งปัญหาการ
ซ่อมแซมที่ผู้ใช้ส่งเข้ามา มีโครงสร้างดังตารางที่6-11

T_Request	
Request_ID(int)	(PK)
User_ID(int)	(FK)
Priority	(nvarchar)
ProblemType_ID(int)	(FK)
Machine_ID	(FK)
ComputerNumber	
Title	(nvarchar)
Detail	(ntext)
Dates	(int)
Times	(nvarchar)
Status	(nvarchar)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Type(nvarchar)
Location_Req(nvarchar)

ตารางที่6-11 แสดงโครงสร้างตาราง T_Request

- ตารางT_Reply ทำหน้าที่จัดเก็บว่าข้อมูลรายละเอียดของการตอบกลับของช่างไปยังผู้ใช้ในแต่ละครั้งมีโครงสร้างดังตารางที่6-12

T_Reply
Reply_ID(int) (PK)
Request_ID(int) (FK)
User_ID(int) (FK)
ProblemType_ID(int) (FK)
Status(nvarchar)
Cost(nvarchar)
Detail(ntext)
Dates(int)
Times(nvarchar)

ตารางที่6-12 แสดงโครงสร้างตาราง T_Reply

- ตารางT_User ทำหน้าที่จัดเก็บว่าข้อมูลรายละเอียดของประวัติของผู้ใช้ที่ทำการลงทะเบียนแต่ละคนมีโครงสร้างดังตารางที่6-13

T_User
User_ID(int) (PK)
Login(nvarchar)
Password(nvarchar)
Name(nvarchar)
Department_ID(int)(FK)
Section_ID(int)(FK)
Location_ID(int)(FK)
Tel(nvarchar)
Email(nvarchar)
Status(nvarchar)

ตารางที่6-13 แสดงโครงสร้างตาราง T_User

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7

การใช้งานโปรแกรม

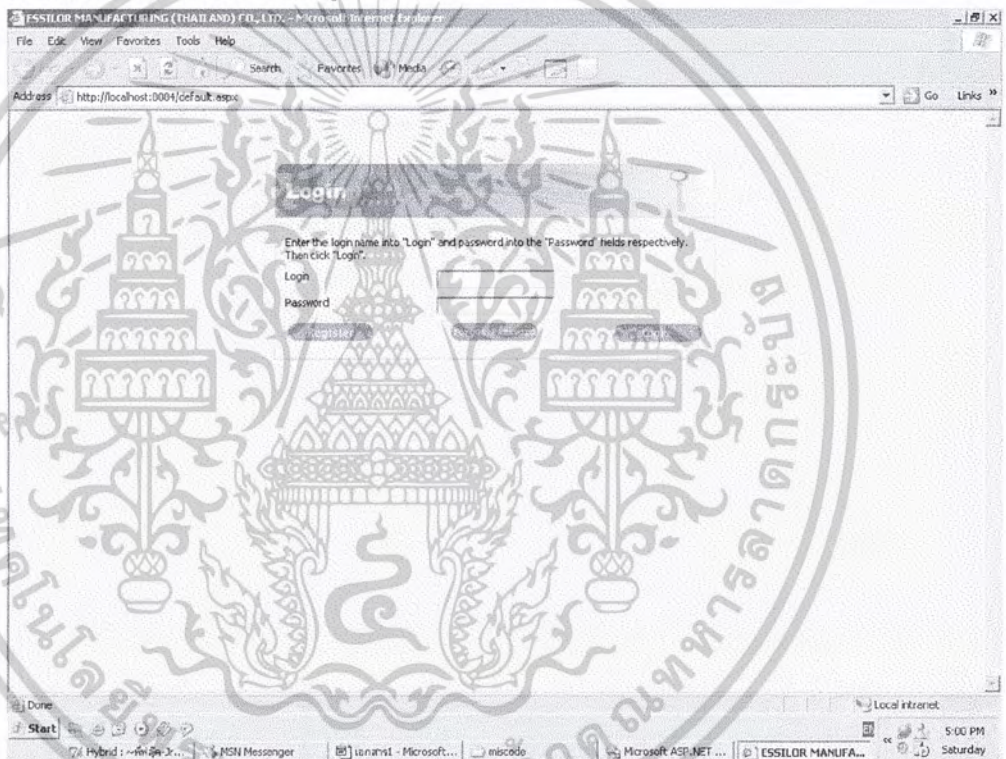
การใช้งานโปรแกรม

โปรแกรมจะแบ่งผู้ใช้ออกเป็น 3 ประเภท

- 1 ทางด้านของ ผู้ใช้ทั่วไป(User)
- 2 ทางด้านของ ช่าง(MIS)
- 3 ทางด้านของ ผู้บริหาร(Admin)

7.1 การใช้งานของผู้ใช้ทั่วไป

เมื่อเริ่มใช้โปรแกรมจะมีการโชว์หน้าของการเข้าสู่ระบบ(login) โดยจะมีอยู่ 3 ปุ่ม ให้เลือก โดยจะมีปุ่ม Register, Login, Forgot Password ดังรูป 7-1



รูปที่ 7-1 แสดงหน้าเว็บเข้าสู่ระบบ

โดยผู้เริ่มแรกจะต้องทำการลงทะเบียน ก่อนถึงจะสามารถมีชื่อและรหัสผ่านได้ โดยรูปแบบของการลงทะเบียนนั้น จะเป็นดังรูป 7-2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

REGISTER ESSILOR MANUFACTURING (THAI) AND CO., LTD. - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://localhost:8084/register.aspx

Register

FirstName : ***

LastName : ***

Login : ***

Password : ***

Re Password : ***

Department : A01-Asia Operation ***

Section : 0001-Administration ***

Location : EMTC1 ***

Telephone : ***

Email : ***

Submit Reset

Done Local Intranet 5:03 PM Saturday

รูปที่ 7-2 แสดงหน้าเว็บของการลงทะเบียน

ซึ่งจะเห็นได้ว่าจะมีเครื่องหมาย *** (สีแดง) นั้น คือเราจะต้องทำการกรอกข้อมูลใส่ไว้ในช่องนั้น และในการใส่ชื่อนั้น ชื่อสามารถซ้ำกันได้ แต่ถ้าชื่อซ้ำนั้นนามสกุลห้ามซ้ำกันด้วย ถ้ากรอกข้อมูลชื่อนั้นจะทำให้ไม่สามารถลงทะเบียนได้ และ ชื่อผู้ใช้นั้น จะไม่สามารถซ้ำกับชื่อคนที่เคยลงทะเบียนไปแล้วนั้นได้ เมื่อเราทำการกด submit แล้วนั้น จะได้ดังรูป 7-3

http://localhost:8084/welcome.aspx - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://localhost:8084/welcome.aspx

ขณะนี้ท่านได้ทำการ Login เรียบร้อยแล้ว
โปรดตรวจสอบ Login และ Password ของท่านด้วย
ขอบคุณค่ะ

How system already register new account for you
Pls check your login and password
thank you

Login : ff
Password : hh

คลิกข้างล่างเพื่อกลับสู่หน้า Login
Click below button for login page

REGISTER SUCCESS

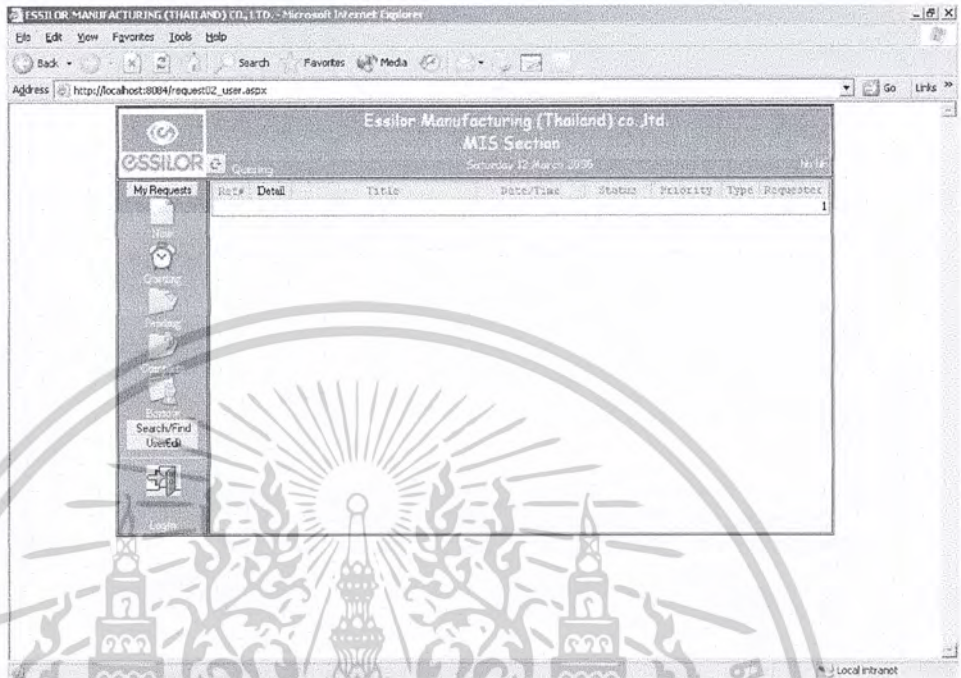
Click here to Login

Done Local Intranet

รูปที่ 7-3 แสดงหน้าเว็บผลลัพธ์เมื่อทะเบียนเรียบร้อยแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แล้วจะมีปุ่มให้เราคลิกไปยังหน้าของการลงทะเบียนอีก ซึ่งตอนนี้เราจะสามารถ ใส่ชื่อผู้ใช้และรหัสผ่าน ได้ โดยเอาชื่อและรหัสที่เราได้ทำการลงทะเบียนไปแล้วนั้น กรอกลงไปแล้วทำการเข้าสู่ระบบได้ทันที เมื่อทำการ เข้าสู่ระบบแล้วจะได้ดังรูป7-4



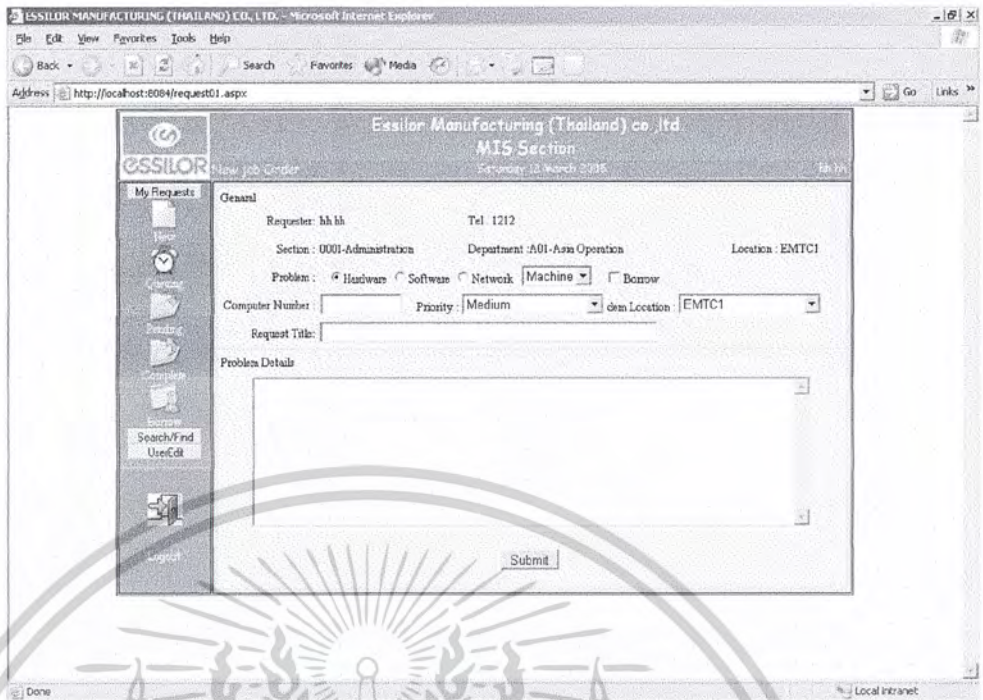
รูปที่7-4 แสดงหน้าเว็บหลังทำการเข้าสู่ระบบแล้วของผู้ใช้

ซึ่งจะมีปุ่มอยู่ด้วยกันอยู่ 3 ส่วน จะมีดังนี้

1. My Requests ซึ่งจะมี New, Queuing, Pending, Complete, Borrow
2. Search/Find ซึ่งจะมี Best Solutions และ Search
3. User Edit

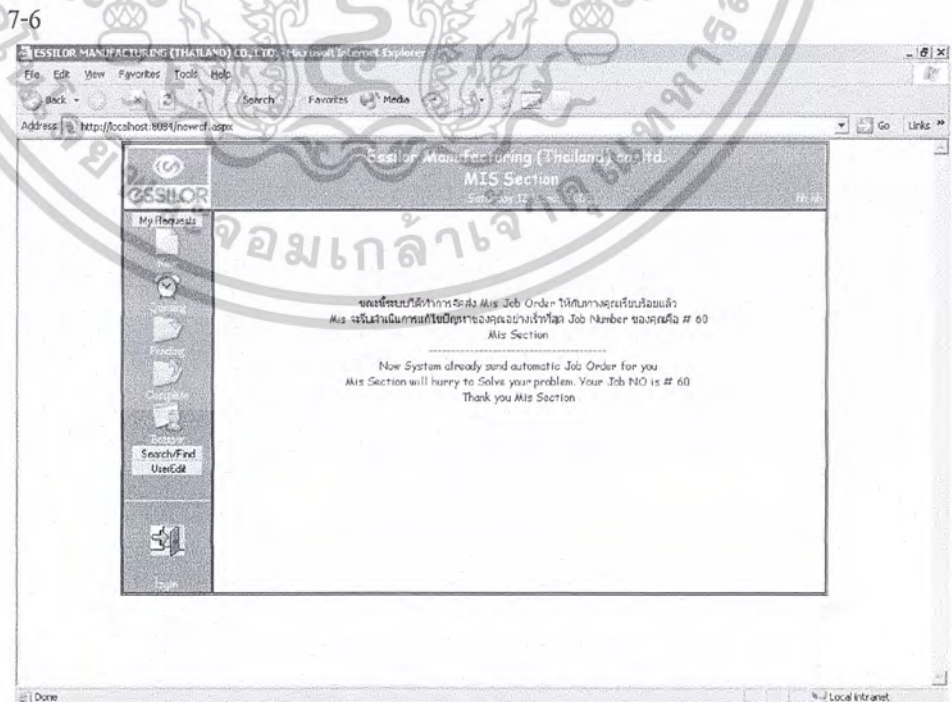
1. My Requests จะมีส่วนประกอบดังนี้

สร้างเอกสารแจ้งปัญหา(New) นั้นจะเป็นการส่งแจ้งปัญหาไปให้กับฝ่ายของช่างและผู้บริหารซึ่งมีรูปดังรูปที่7-5



รูปที่ 7-5 แสดงหน้าเว็บของการแจ้งปัญหา

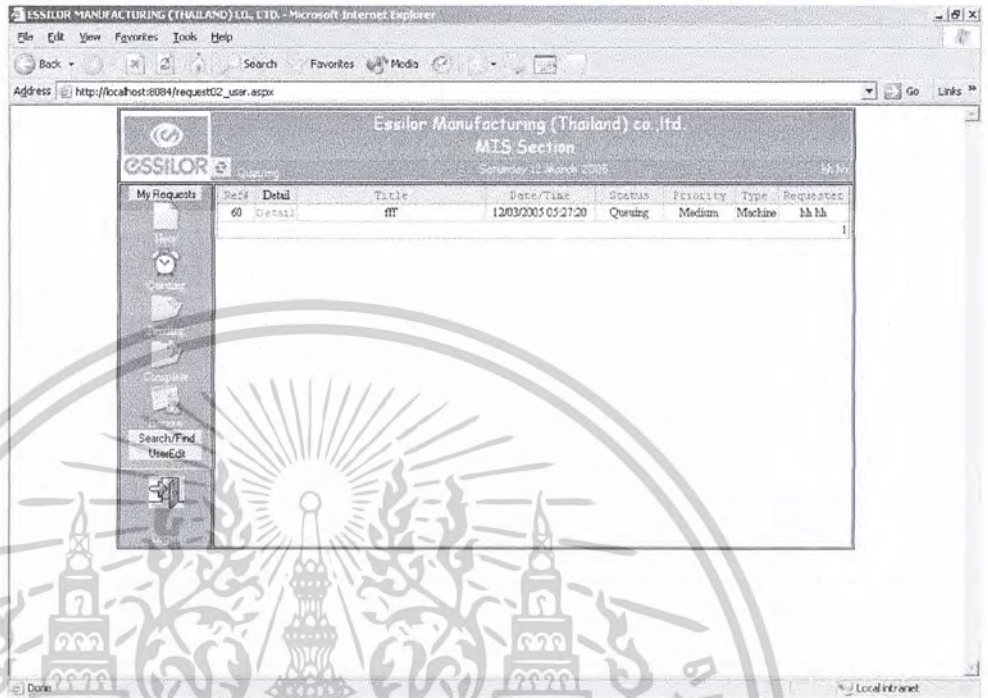
เมื่อเห็นดังรูป 7-5 แล้วนั้น จะมีส่วนที่เราได้กรอกไปแล้ว คือที่กรอกไปตอนลงทะเบียน ซึ่งเราไม่ต้องมากรอกใหม่อีกรอบ และส่วนที่เราจะต้องทำการกรอกข้อมูลลงไปนั้น จะมีให้เดิกว่าแจ้งปัญหาว่าเป็นชนิดอะไร และว่าเรา จะทำการยืมไหม ถ้าเรากรอกว่ายืมส่วนของ Computer Number ก็ไม่ต้องกรอกข้อมูลนั้นลงไป และ Priority คือ ว่าสิ่งที่เราส่ง ไปให้ฝ่ายช่างและผู้บริหารว่านั่นเป็นงานที่เร่งด่วนหรือไม่เร่งด่วน และบอกตำแหน่งว่าจะให้ไป ช่อมที่ไหน และหัวข้อที่จะทำการซ่อม และ รายละเอียดของปัญหาด้วย เมื่อกด Submit แล้ว จะได้ดังรูป



รูปที่ 7-6 แสดงหน้าเว็บของผลลัพธ์เมื่อทำการแจ้งปัญหาแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำออกเผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อได้ดังรูป 7-6 แล้ว โปรแกรมจะทำการ เชื่อมกลับไป ที่ หน้าของรอรับแจ้งปัญหา (Queuing) ให้โดยอัตโนมัติและจะทำการส่งไป 2ทาง โดยเมลล์และ โปรแกรมนี้ด้วยไปบอกฝ่าย ของช่างและผู้บริหารด้วยเพื่อที่จะ ได้รู้ว่า สถานะตอนนี้เราเป็นอะไร ดังรูปที่ 7-7



รูปที่ 7-7 แสดงหน้าเว็บของปัญหาที่มีสถานะเป็นการรอรับแจ้งปัญหา

เมื่อฝ่ายช่างตอบกลับมาแล้วนั้น จะมีสถานะอยู่อีก 2 สถานะ ซึ่งจะมี รอการซ่อม (Pending) และแก้ปัญหาเสร็จแล้ว(Complete) ซึ่งเมื่อกี้เราได้ทำการส่งปัญหาไปแล้วนั้นถ้าหากว่าช่าง ได้ทำการตอบกลับมาแล้วนั้นปัญหาที่แต่ก่อนเคยอยู่ใน รอรับแจ้งปัญหาจะ ไปอยู่ส่วน ของรอการซ่อม หรือ แก้ปัญหาเสร็จแล้วแทน ซึ่งเราสามารถดูรายละเอียดได้ว่า ฝ่ายช่าง ได้ตอบ อะไรกลับมาหรือ ใ้เห็นได้ ใจอะไรให้เรานั้นเอง

โดยที่เราสามารถรู้ได้จากทางเมลล์และ โปรแกรมว่าทางช่าง ได้ตอบกลับมาแล้วหรือยัง แต่ถ้าหากว่าเราส่งปัญหาไปแบบขี้มมันจะไม่ไปอยู่ในสถานะของรอรับแจ้งปัญหา แต่จะไปอยู่ในสถานะของการขี้มแทน และเมื่อช่างตอบกลับมาจะไม่ไปอยู่ในส่วนของการซ่อม และ แก้ปัญหาเสร็จแล้ว ซึ่งจะอยู่แค่สถานะของขี้มเท่านั้น แต่จะมีบอกวาสถานะตอนนี้เป็น รอ การซ่อม หรือ แก้ปัญหาเสร็จแล้ว

2. Search/Find โดยในส่วนของค้นหาจะมีดังรูปที่ 7-8

รูปที่ 7-8 แสดงหน้าเว็บของส่วนฟังก์ชันการค้นหา

ซึ่งในการค้นหานี้จะ ค้นหาออกมาเป็นการแจ้งปัญหาและจะไม่สามารถไปดู การแจ้งปัญหา ของคนอื่น ได้ ซึ่งจะมีให้เลือกค้นหาว่าจะเอาแบบ ไหนแล้วเมื่อกด Search แล้วจะได้ดังรูปที่ 7-9

Ref#	Detail	Title	Que. Time	Status	Priority	Requester
60	Detail	HF	12/03/05 05:27:20	Queung	Medium	hh hh

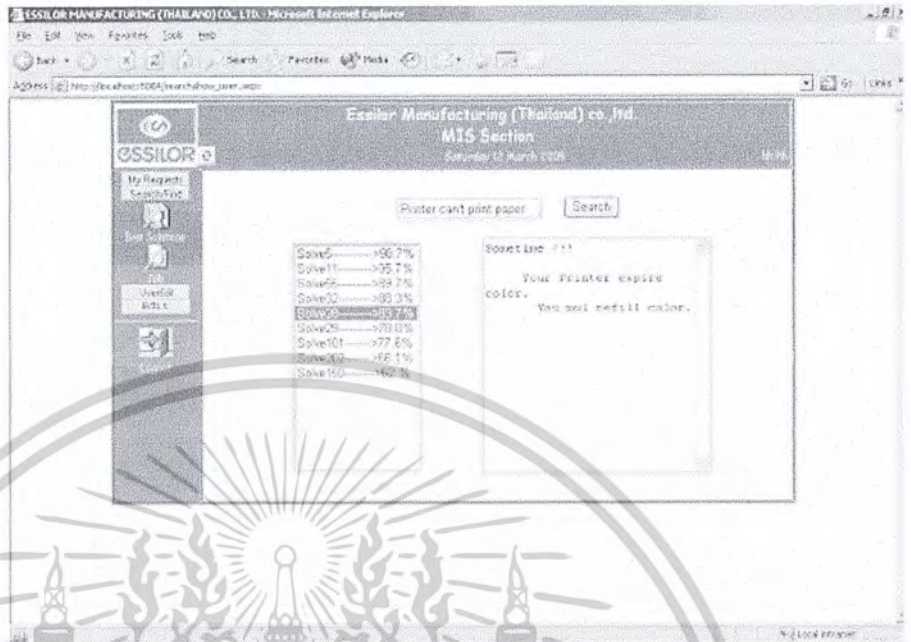
รูปที่ 7-9 แสดงหน้าเว็บผลลัพธ์ของฟังก์ชันการค้นหา

โดยรูปที่ออกมาจะออกมา แค่เป็น Job Order เท่านั้น

ฐานความรู้ (Best Solution) : เป็น โปรแกรมให้ความรู้แก่ผู้ใช้ โดยให้ผู้ใช้กรอกสิ่งที่ต้องการแล้วพกดคลิก Search

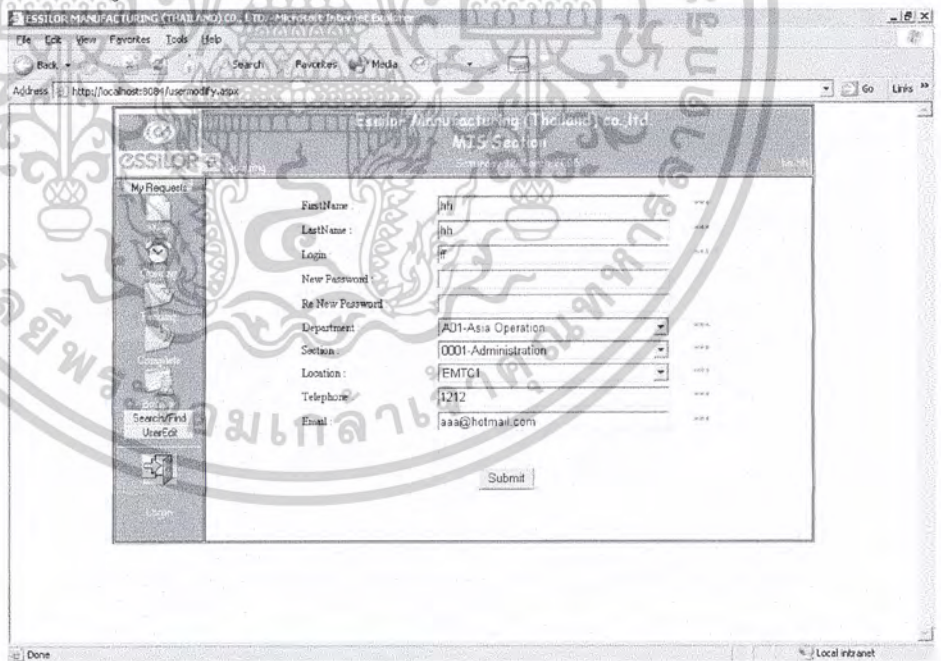
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบจะทำการค้นหาเอกสารวิธีการแก้ปัญหาที่คิดว่าตรงกับที่ผู้ใช้ต้องการออกมาให้โดยเรียงตามลำดับความตรงประเด็นดังรูปที่ 7-10



รูปที่ 7-10 แสดงหน้าเว็บของผลลัพธ์ของฟังก์ชันฐานความรู้

3. แก้ไขประวัติ (User Edit) ซึ่งจะเป็นการแก้ไขส่วนที่เราได้ทำการลงทะเบียนไปแล้วนั้นได้ โดยจะมีรูปที่ 7-11



รูปที่ 7-11 แสดงหน้าเว็บในส่วนการแก้ไขรายละเอียดส่วนตัว

ซึ่งเมื่อแก้ไขเสร็จแล้วนั้น เมื่อเราทำการกด ยืนยัน ไปแล้วนั้น จะทำให้เราต้องไปทำการเข้าระบบอีกรอบเพื่อทำการปรับปรุงข้อมูลนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลืมรหัส(Forgot Password)

คือ เมื่อผู้ใช้ลืมรหัสของตัวเองนั้นผู้ใช้สามารถที่จะใส่ชื่อตัวเองที่เคยลงทะเบียนไว้ เพื่อที่จะให้ระบบส่งเมลล์มาออกกับผู้ใช้ได้ว่า รหัสและชื่อ อะไรได้ดังรูป7-12

ForGot Password

Firstname :

Lastname :

รูปที่ 7-12 แสดงหน้าเว็บของฟังก์ชันเมื่อลืมรหัสผ่าน โดยเมื่อกด send แล้วนั้น จะ ได้รูปที่7-13

ขณระบบได้ทำการส่งเมลล์เรียบร้อยแล้ว
โปรดตรวจสอบ Login และ Password ของท่านด้วย
ขอบคุณค่ะ

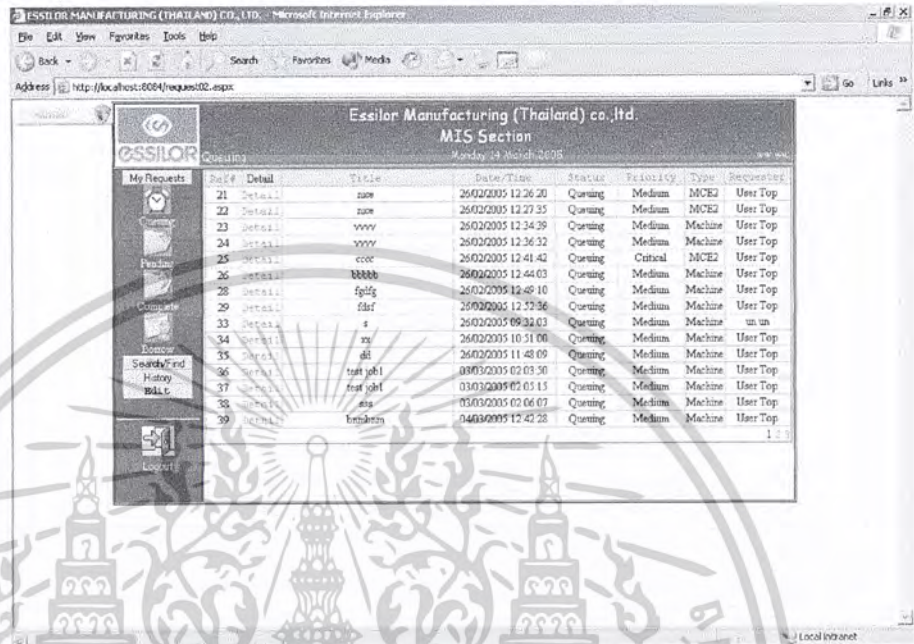
How system already send mail for you
Pls check your login and password
thank you

รูปที่ 7-13 แสดงหน้าเว็บของผลลัพธ์ของฟังก์ชันเมื่อลืมรหัสผ่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.2 การใช้งานของ MIS และ Admin

ซึ่งการใช้งานของช่างนั้น จะต้องให้ผู้บริหารนั้นทำการลงทะเบียนเท่านั้นถึงจะสามารถที่จะเข้าสู่ระบบเป็นช่างได้ เมื่อเข้าสู่ระบบได้แล้วจะได้ดังรูปที่ 7-14



รูปที่ 7-14 แสดงหน้าเว็บแรกของการเข้าระบบในส่วนช่างและผู้บริหาร

โดยจะมีหัวข้อหลักอยู่ 4 ประเภทคือ

1. My Requests ซึ่งจะมีส่วนประกอบอยู่ 4 อย่างคือ

1.1 Queuing คือเมื่อ user ส่ง แจ้งปัญหาเข้ามานั้นปัญหานั้นจะเข้ามาอยู่ใน รอคอบรับ นั้นเอง

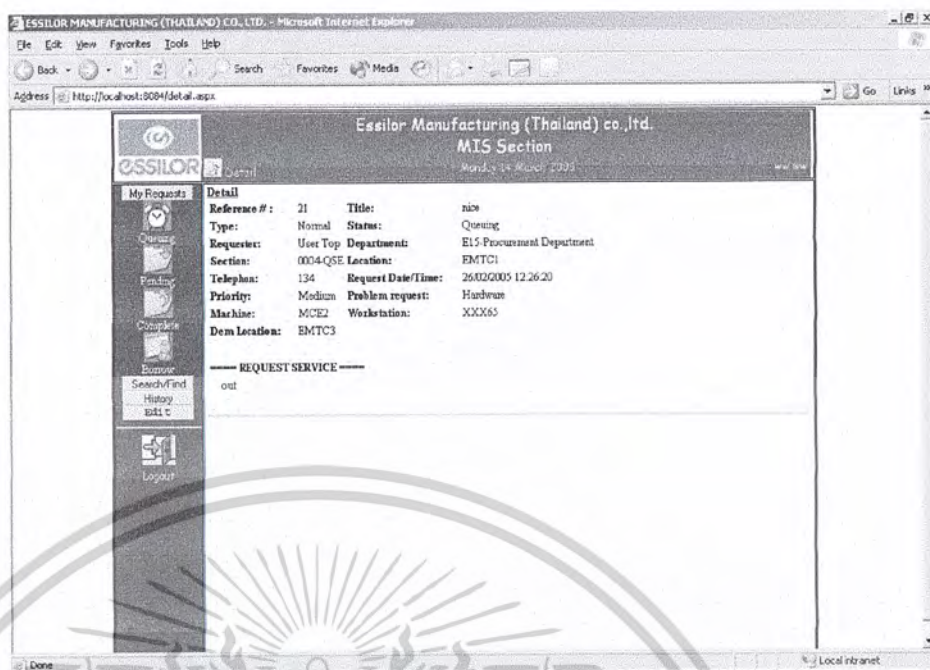
1.2 Pending คือ เมื่อ ช่างทำการตอบกลับผู้ใช้นั้นจะมี สถานะ ให้เลือกว่า จะเลือกเป็น รอคอบหรือซ่อมหรือซ่อมเสร็จถ้าเลือกเป็นรอกอบซ่อมนั้นปัญหาที่ทำการตอบกลับไปในนั้นจะไปอยู่ที่ รอคอบซ่อมทันที

1.3 Complete คือ เมื่อ ช่างทำการตอบกลับผู้ใช้นั้นจะมี สถานะ ให้เลือกว่า จะเลือกเป็น รอคอบหรือซ่อมหรือซ่อมเสร็จถ้าเลือกเป็นซ่อมเสร็จนั้นปัญหาที่ทำการตอบกลับไปในนั้นจะไปอยู่ที่ ซ่อมเสร็จทันที

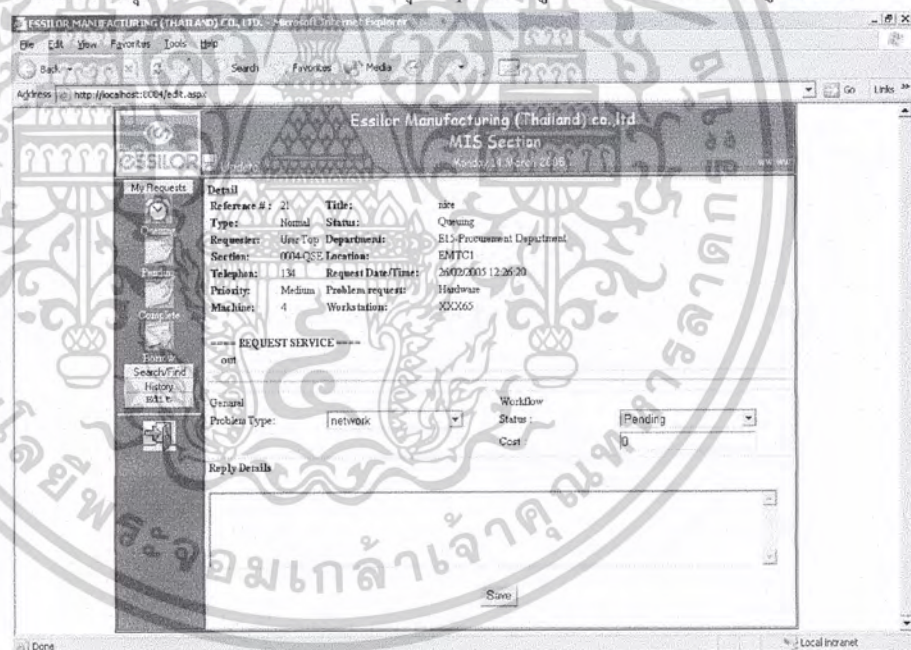
1.4 Borrow คือเมื่อผู้ใช้ส่งปัญหาเข้ามานั้น โดยจะส่งเป็นแบบยืมนั้น ปัญหานั้นจะเข้ามาอยู่ ใน การยืมนั้นเอง แต่ถ้าหากว่าช่างตอบกลับ ไปให้ ผู้ใช้นั้นปัญหานั้น จะไม่ไปอยู่รอกอบหรือซ่อมหรือซ่อมเสร็จแต่จะอยู่ที่การยืมเช่นเดิม ซึ่งถ้าหากว่าเราเปลี่ยนสถานะแล้วในหน้าของการ ยืมที่ ปัญหา นั้นๆ จะเปลี่ยนตาม สถานะที่เราเลือกนั้น

วิธีการตอบกลับของทาง ช่างนั้น ให้เราดูรายละเอียดก่อน แล้วจะได้รูปที่ 7-15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-15 แสดงหน้าเว็บของการดูรายละเอียดปัญหาของช่าง
แล้วสังเกตปุ่มข้างบนทางซ้าย ซึ่งจะมีปุ่ม Update อยู่ เมื่อเรากดแล้วจะได้รูปที่ 7-16



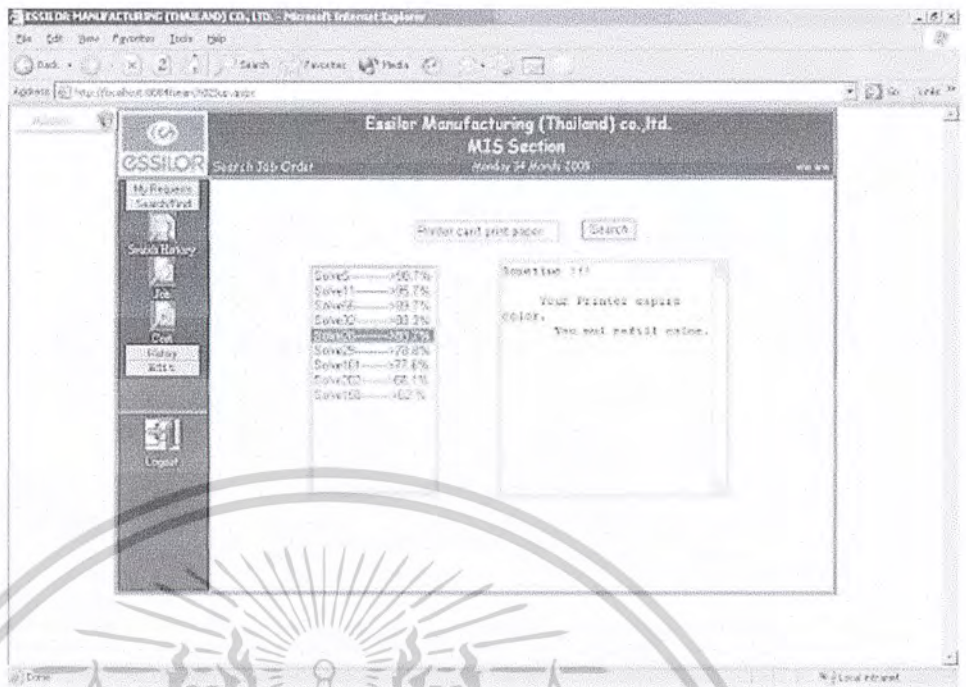
รูปที่ 7-16 แสดงหน้าเว็บของการตอบกลับของช่าง
ซึ่งหน้าที่จะประกอบไปด้วยชนิดของ problem, Status, Cost และ Reply ซึ่งเมื่อกรอก
ข้อมูลเสร็จแล้วนั้น ก็จะ เชื่อม ไปที่หน้าของ รอรการตอบรับเพื่อทำการตอบกลับปัญหาอื่นๆ
ต่อไป

2. ส่วนของการ Search/Find จะมีส่วนประกอบอยู่ 3 อย่างคือ

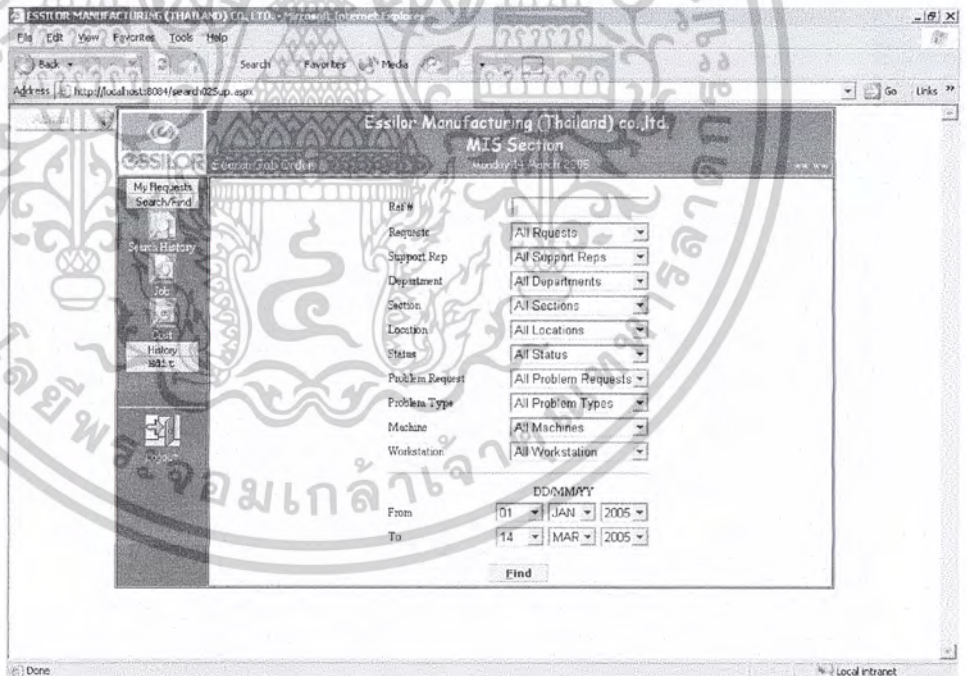
Best Solutions ซึ่งจะเป็นการค้นหา Sol เก่าๆที่เคย sol ไปแล้วนั้นขึ้นมาแสดงให้เราดูดัง

รูปที่ 7-17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



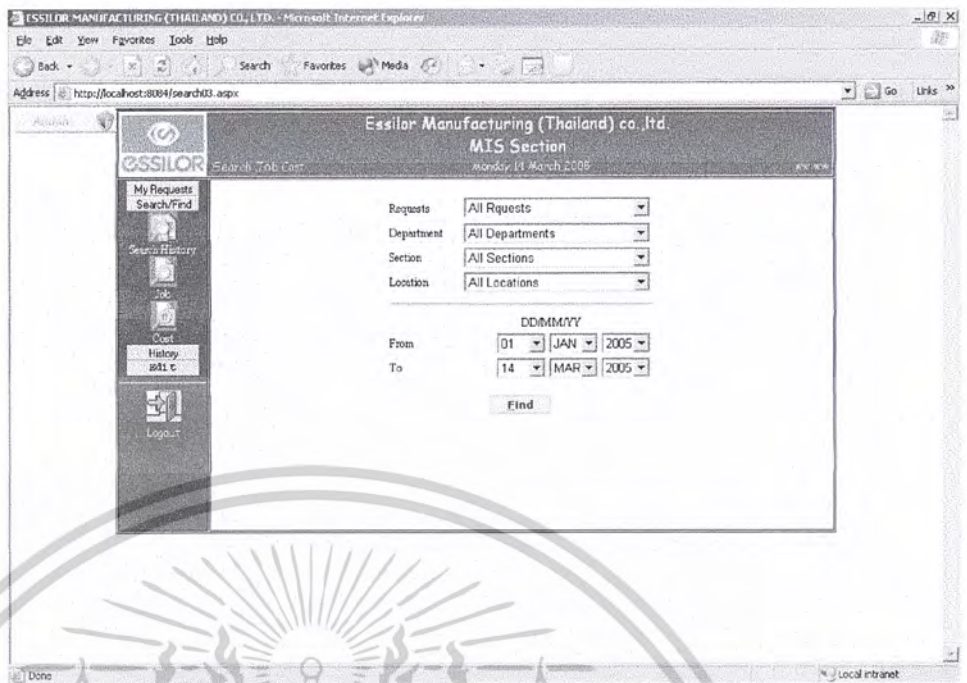
รูปที่ 7-17 แสดงหน้าเว็บของผลลัพธ์ของการใช้งานระบบฐานความรู้ Search Job order ซึ่งจะเป็นการ Search ปัญหาทุกอย่างตั้งแต่ส่ง ปัญหาหรือตอบกลับ ไป ตั้งแต่วันไหนถึงวันไหนได้ โดยหน้าของการ ค้นหาจะได้ดังรูปที่ 7-18



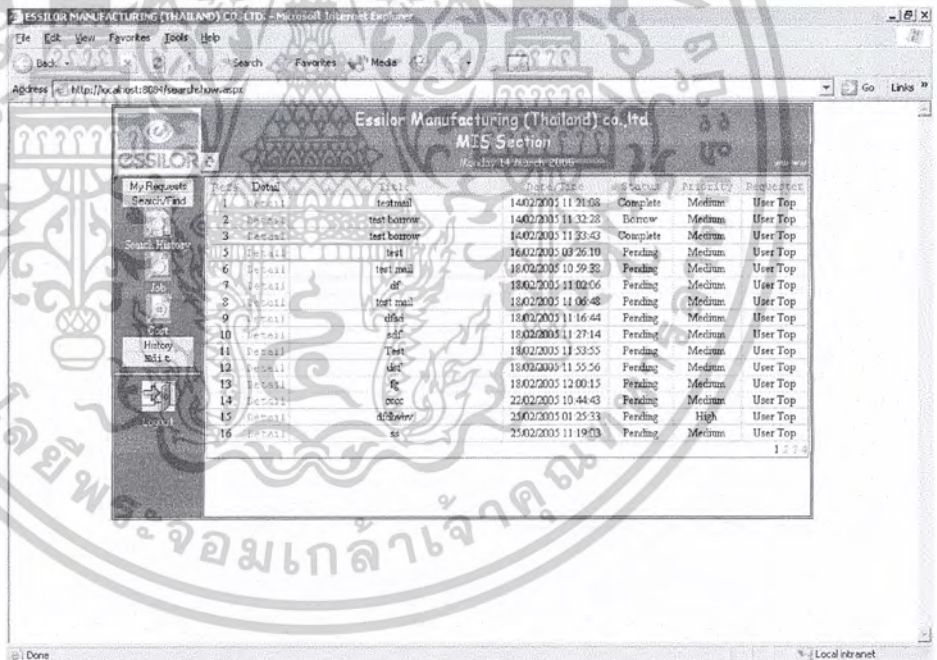
รูปที่ 7-18 แสดงหน้าเว็บของการค้นหาปัญหา

Search ค่า Cost โดยจะคล้ายกับ Search job order แต่หน้าตาคั่นหน้านั้นจะไม่เหมือนกัน โดยจะได้ดังรูปที่ 7-19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-19 แสดงหน้าเว็บของการค้นหาค่าใช้จ่าย
เมื่อกด Search แล้วนั้นจะได้ดังรูป 7-20



รูปที่ 7-20 แสดงหน้าเว็บของผลลัพธ์ที่ได้จากการค้นหา

3. History โดยจะมีส่วนประกอบอยู่ 3 ส่วนคือ

Add History จะเป็นการเก็บประวัติของเครื่อง computer นั้นไว้โดยจะมีรูปแบบการกรอกข้อมูลดังรูปที่ 7-21

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ESSILOR Manufacturing (Thailand) Co., Ltd.
MIS Section
Monday, 14 March 2005

Computer Name: Project:

User: Type:

Dept: A01-Asia Operation OS:

Section: 0001-Administration Model:

Location: EMTC1 CPU:

HDD: RAM:

Remark: Service:

Submit

รูปที่ 7-21 แสดงหน้าเว็บของการเพิ่มประวัติบุคคล

Edit History คือการแก้ไขข้อมูลที่เราได้ทำการเพิ่มไปแล้วว่าจะแก้ไขยังไง
ซึ่งเราจะต้องทำการ ค้นหาก่อนว่าจะทำการ แก้ไขอะไร โดยจะได้ดังรูปที่ 7-22

ESSILOR Manufacturing (Thailand) Co., Ltd.
MIS Section
Monday, 14 March 2005

Department: All Department

Section: All Section

Location: All Location

Project: All Project

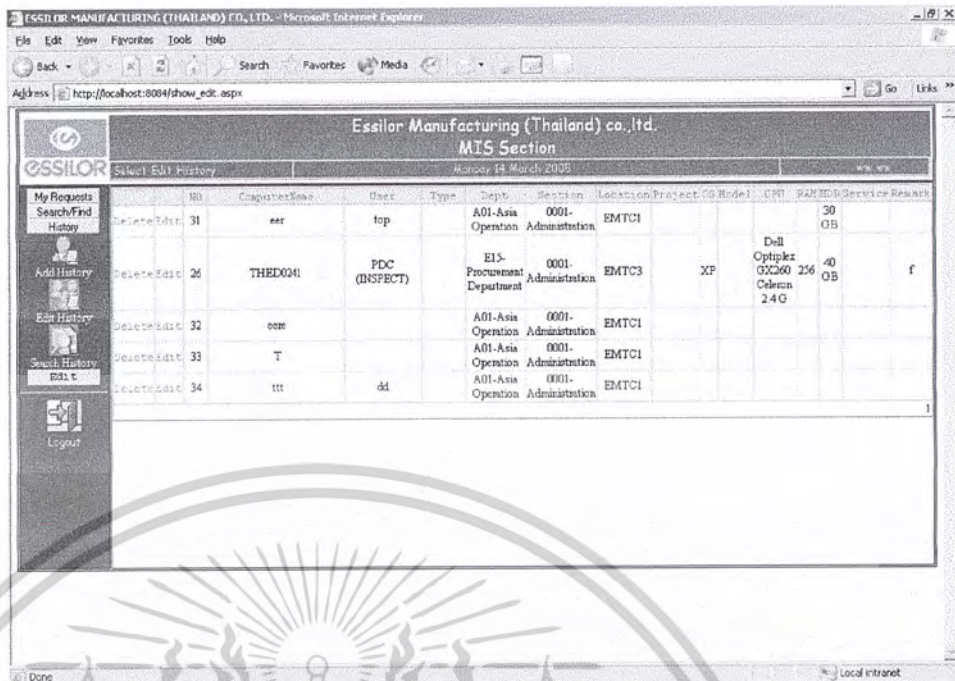
Type: All Type

Find

รูปที่ 7-22 แสดงหน้าเว็บของการค้นหาบุคคลเพื่อแก้ไข

เมื่อกดหาออกมาแล้วจะได้ดังรูปที่ 7-23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

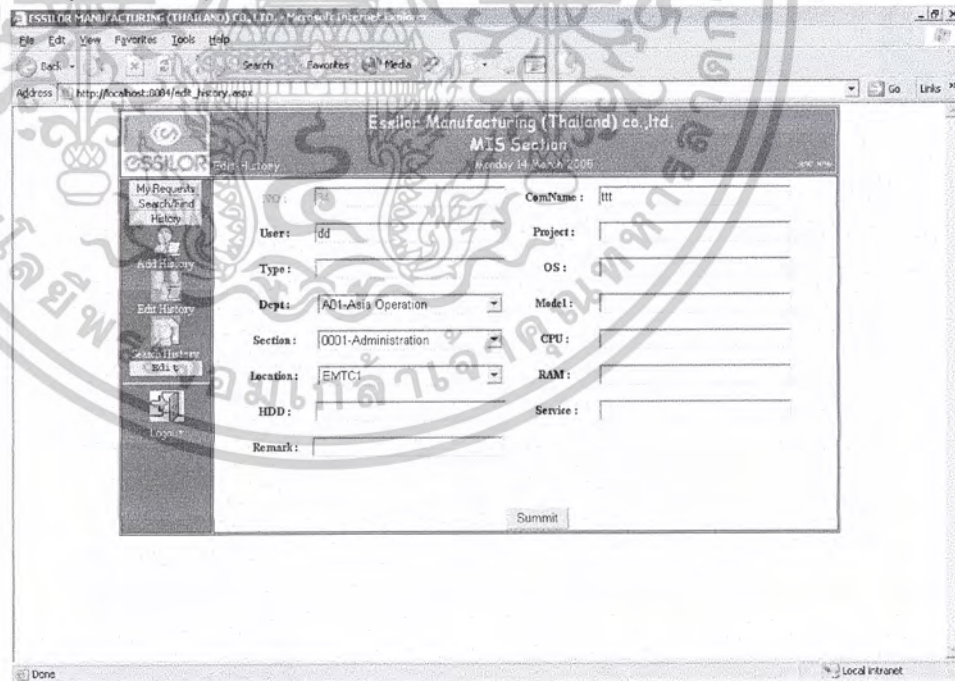


รูปที่ 7-23 แสดงหน้าเว็บแสดงผลพัสดุของการค้นหา

โดยรูปออกมาจะได้ประวัติของคอมพิวเตอร์ที่เราค้นหาซึ่งจะมีปุ่ม อยู่ 2 ปุ่ม คือปุ่ม Delete และปุ่ม Edit

ปุ่ม Delete นั้น คือจะเป็นลบข้อมูลของคอมพิวเตอร์นั้นทิ้งไป

ปุ่ม Edit คือ ปุ่มแก้ไขข้อมูลของคอมพิวเตอร์นั้น โดยเมื่อกดแล้วจะได้ดังรูปที่ 7-24



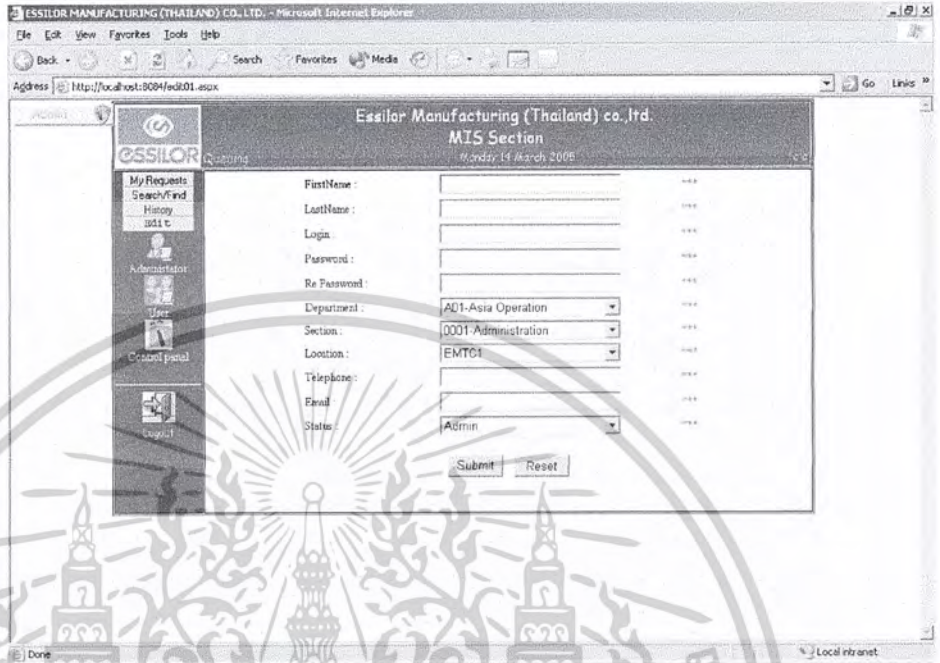
รูปที่ 7-24 แสดงหน้าเว็บของการแก้ไขประวัติบุคคล

Search History นั้นจะเหมือนกับ แก้ไขประวัติ แต่จะไม่มี ปุ่ม Delete และปุ่ม Edit

4. Edit จะมีอยู่ 3 อย่างคือ

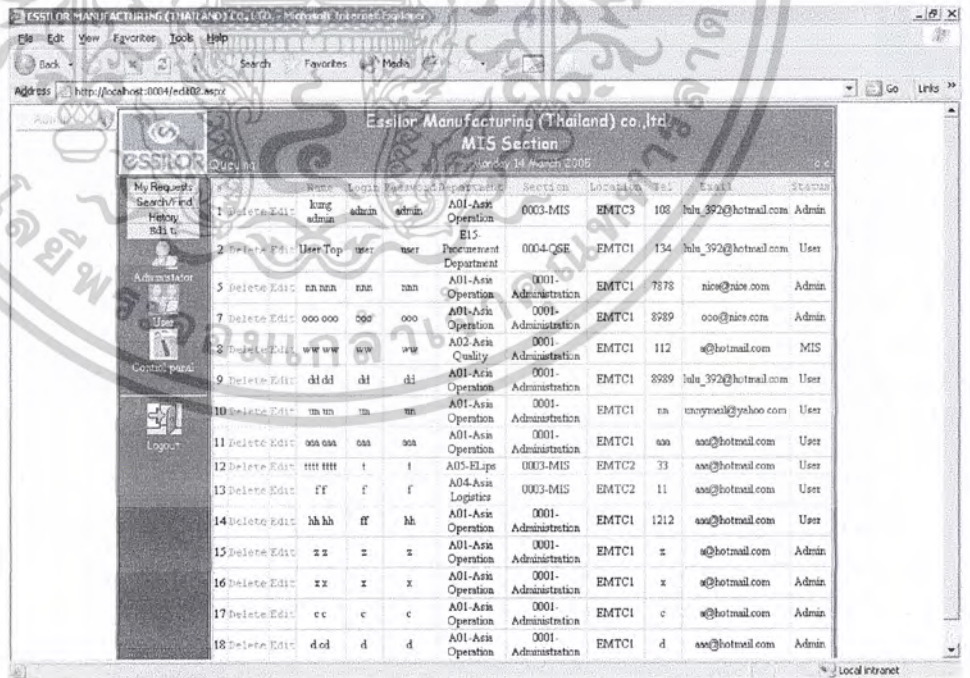
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Admin คือจะเป็น ปุ่มเอาไว้สำหรับในการ ลงทะเบียนนั้น ซึ่งจะสามารถ ลงทะเบียน ได้ทั้งช่าง ผู้ใช้ ผู้บริหาร โดยที่ปุ่มนี้จะมองเห็นเฉพาะ ผู้บริหารเท่านั้น โดย รูปแบบการ เพิ่มจะได้ดังรูปที่ 7-25



รูปที่ 7-25 แสดงหน้าเว็บของการเพิ่มผู้ใช้

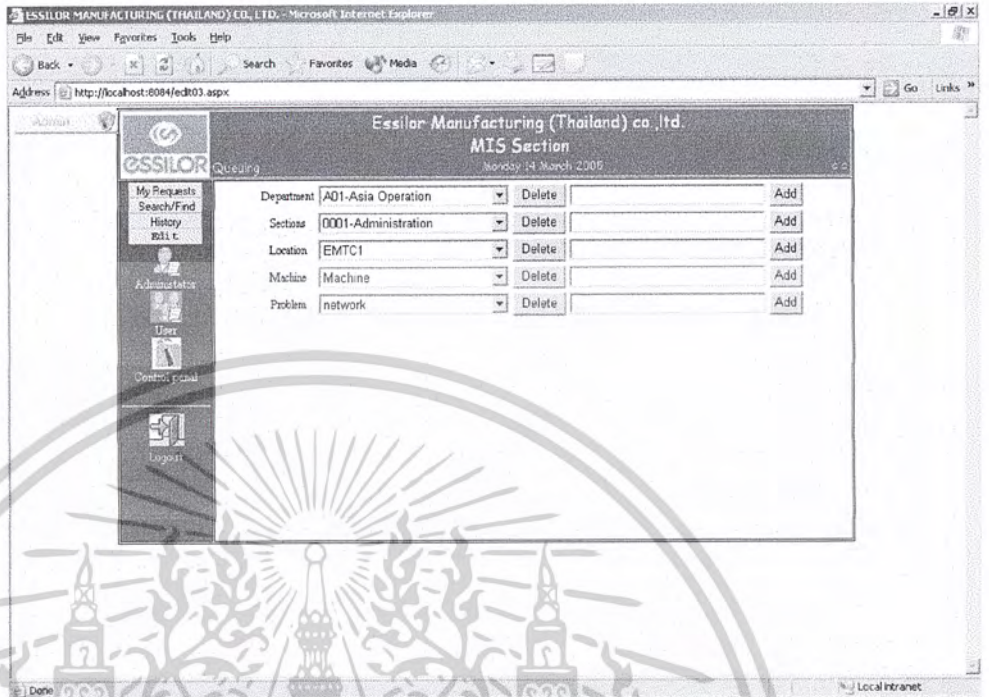
User คือจะเป็นการดู ลบ แก้ไข ข้อมูลต่างๆของผู้ใช้หรือช่างที่ได้ทำการ ลงทะเบียน ไปแล้ว โดยจะได้ดังรูปที่ 7-26



รูปที่ 7-26 แสดงหน้าเว็บของการแก้ไขหรือลบผู้ใช้หรือช่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Control Panel คือ ปุ่มที่เอาไว้ทำการ Add Delete ข้อมูล เช่น Section
Department location และอื่นๆ ซึ่งสามารถดูจากรูปที่ 7-27



รูปที่ 7-27 แสดงหน้าเว็บของการเพิ่มข้อมูลเกี่ยวกับโรงงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 8

สรุปและแนวทางในการพัฒนาต่อ

8.1 สรุป

โครงการที่ได้ทำมาสามารถสรุปเป็นส่วนดีและส่วนที่ควรปรับปรุงดังนี้

ข้อดี

- ระบบแจ้งปัญหาที่มีการรวบรวมข้อมูลทีละเย็บซึ่งอาจเป็นประโยชน์แก่ช่างในการหาวิธีแก้ไขปัญหาได้และนอกจากนี้ระบบมีส่วนช่วยในการแบ่งงานของช่างแต่ละบุคคลให้ง่ายและไม่สับสน
- ระบบประวัติการซ่อมแซมมีประโยชน์ในการนำข้อมูลที่ได้ไปให้ผู้บริหารวิเคราะห์หาสาเหตุหรือวิธีป้องกันให้ปัญหาลดลงได้
- ระบบฐานความรู้มีส่วนช่วยให้ผู้ใช้เข้าใจปัญหาของตนเองได้มากขึ้นซึ่งทำให้สามารถแก้ไขปัญหาเองได้หรือสามารถเข้าใจกระบวนการซ่อมแซมของช่างได้

ข้อเสีย

- ระบบฐานความรู้ต้องอาศัยเวลาและฐานข้อมูลที่มากพอที่จะทำให้ระบบมีความเสถียรในการวัดค่าความตรงต่อประเด็น
- ระบบฐานความรู้ต้องอาศัยทรัพยากรในการคำนวณค่อนข้างมากเพื่อให้ได้ค่าความเหมือนที่ถูกต้อง

8.2 แนวทางในการพัฒนาต่อ

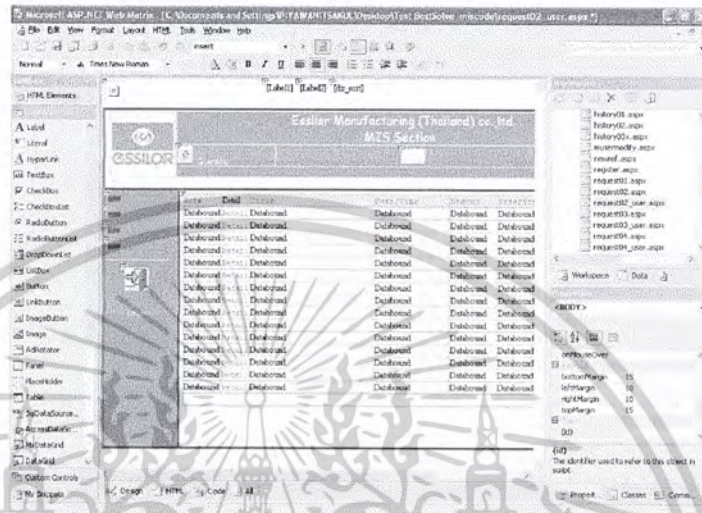
- เราอาจเพิ่มหลักการของการปรับปรุงความต้องการของผู้ใช้โดยอาศัยหลักการตอบกลับของผู้ใช้ รายละเอียดอยู่ในภาคผนวก ข
- เราเพิ่มหลักในการจัดเก็บดัชนีให้มีโครงสร้างเพื่อที่จะง่ายและรวดเร็วในการค้นหาค้นหานี้
นั้นๆ

ภาคผนวก ก

เครื่องมือที่ใช้ในการพัฒนาเว็บโปรแกรมในโครงการ

1. Microsoft Asp.net Web Matrix

ใช้เพื่อเป็น Tool ในพัฒนาเว็บ โปรแกรมด้วยภาษา ASP.NET (VB)



ตัวอย่างการเขียนโปรแกรม

1.1 การเพิ่มข้อมูลในฐานข้อมูล(Insert)

ตัวอย่างของฟังก์ชันการเพิ่มข้อมูลในตารางT_Index

```
Function MyInsertIndex(ByVal index As String) As Integer
    Dim connectionString As String = "server=(local); trusted_connection=true;
    database='misjob_order'"
    Dim dbConnection As System.Data.IDbConnection = New
    System.Data.SqlClient.SqlConnection(connectionString)
    Dim queryString As String = "INSERT INTO [T_Index] ([Index],[freq])
    VALUES (@Index,1)"
    Dim dbCommand As System.Data.IDbCommand = New
    System.Data.SqlClient.SqlCommand
    dbCommand.CommandText = queryString
    dbCommand.Connection = dbConnectio
    Dim dbParam_index As System.Data.IDataParameter = New
    System.Data.SqlClient.SqlParameter
    dbParam_index.ParameterName = "@Index"
    dbParam_index.Value = index
    dbParam_index.DbType = System.Data.DbType.String
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

dbCommand.Parameters.Add(dbParam_index)
Dim rowsAffected As Integer = 0
dbConnection.Open
Try
    rowsAffected = dbCommand.ExecuteNonQuery
Finally
    dbConnection.Close
End Try
Return rowsAffected
End Function

```

1.2 การลบข้อมูลฐานข้อมูล(delete)

ตัวอย่างของฟังก์ชันการลบข้อมูลในตารางT_Index ตามค่าID ที่ส่งเข้ามา

```

Function Delete_Index(ByVal index_ID As Integer) As Integer
    Dim connectionString As String = "server='piya\un'; trusted_connection=true;
    database='misjob_order'"
    Dim dbConnection As System.Data.IDbConnection = New
    System.Data.SqlClient.SqlConnection(connectionString) Dim queryString As
    String = "DELETE FROM [T_Index] WHERE ([T_Index].[Index_ID] = @Index_ID)"
    Dim dbCommand As System.Data.IDbCommand = New
    System.Data.SqlClient.SqlCommand
    dbCommand.CommandText = queryString
    dbCommand.Connection = dbConnection
    Dim dbParam_index_ID As System.Data.IDataParameter = New
    System.Data.SqlClient.SqlParameter
    dbParam_index_ID.ParameterName = "@Index_ID"
    dbParam_index_ID.Value = index_ID
    dbParam_index_ID.DbType = System.Data.DbType.Int32
    dbCommand.Parameters.Add(dbParam_index_ID)
    Dim rowsAffected As Integer = 0
    dbConnection.Open
    Try
        rowsAffected = dbCommand.ExecuteNonQuery
    Finally
        dbConnection.Close

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

End Try
Return rowsAffected
End Function

```

1.3 การปรับปรุงข้อมูลฐานข้อมูล(update)

ตัวอย่างของฟังก์ชันการปรับปรุงข้อมูลในตารางT_Index ที่IDนั้นที่ส่งเข้ามาโดยแก้ไขตามค่าของIndexที่ส่งเข้ามา

```

Function Update_Index(ByVal index_ID As Integer, ByVal index As String) As
Integer
Dim connectionString As String = "server='piya'un'; trusted_connection=true;
database='misjob_order'"
Dim dbConnection As System.Data.IDbConnection = New
System.Data.SqlClient.SqlConnection(connectionString)
Dim queryString As String = "UPDATE [T_Index] SET [Index]=@Index WHERE
([T_Index].[Index_ID] = @Index_ID)"
Dim dbCommand As System.Data.IDbCommand = New
System.Data.SqlClient.SqlCommand
dbCommand.CommandText = queryString
dbCommand.Connection = dbConnection
Dim dbParam_index_ID As System.Data.IDataParameter = New
System.Data.SqlClient.SqlParameter
dbParam_index_ID.ParameterName = "@Index_ID"
dbParam_index_ID.Value = index_ID
dbParam_index_ID.DbType = System.Data.DbType.Int32
dbCommand.Parameters.Add(dbParam_index_ID)
Dim dbParam_index As System.Data.IDataParameter = New
System.Data.SqlClient.SqlParameter
dbParam_index.ParameterName = "@Index"
dbParam_index.Value = index
dbParam_index.DbType = System.Data.DbType.String
dbCommand.Parameters.Add(dbParam_index)

```

```
Dim rowsAffected As Integer = 0
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

dbConnection.Open
Try
    rowsAffected = dbCommand.ExecuteNonQuery
Finally
    dbConnection.Close
End Try
Return rowsAffected
End Function

```

1.3 การดึงข้อมูลจากฐานข้อมูล(select)

ตัวอย่างของฟังก์ชันการดึงข้อมูลIndex ในตารางT_Index

```

Function ReturnIndexAll() As System.Data.DataSet
    Dim connectionString As String = "server='piya'un'; trusted_connection=true;
    database='misjob_order'"
    Dim dbConnection As System.Data.IDbConnection = New
    System.Data.SqlClient.SqlConnection(connectionString)
    Dim queryString As String = "SELECT [T_Index].[Index] FROM [T_Index]"
    Dim dbCommand As System.Data.IDbCommand = New
    System.Data.SqlClient.SqlCommand
    dbCommand.CommandText = queryString
    dbCommand.Connection = dbConnection
    Dim dataAdapter As System.Data.IDbDataAdapter = New
    System.Data.SqlClient.SqlDataAdapter
    dataAdapter.SelectCommand = dbCommand
    Dim dataSet As System.Data.DataSet = New System.Data.DataSet
    dataAdapter.Fill(dataSet)
    Return dataSet
End Function

```

1.4 ตัวอย่างโค้ดการส่งเมล

```

mail.To = un_kmitl@kmitl.ac.th
mail.From = kung_essilor@essilor.co.th
mail.Subject = "MIS Job Order"
mail.Body = "Test Program"
SmtpMail.SmtpServer = "localhost"

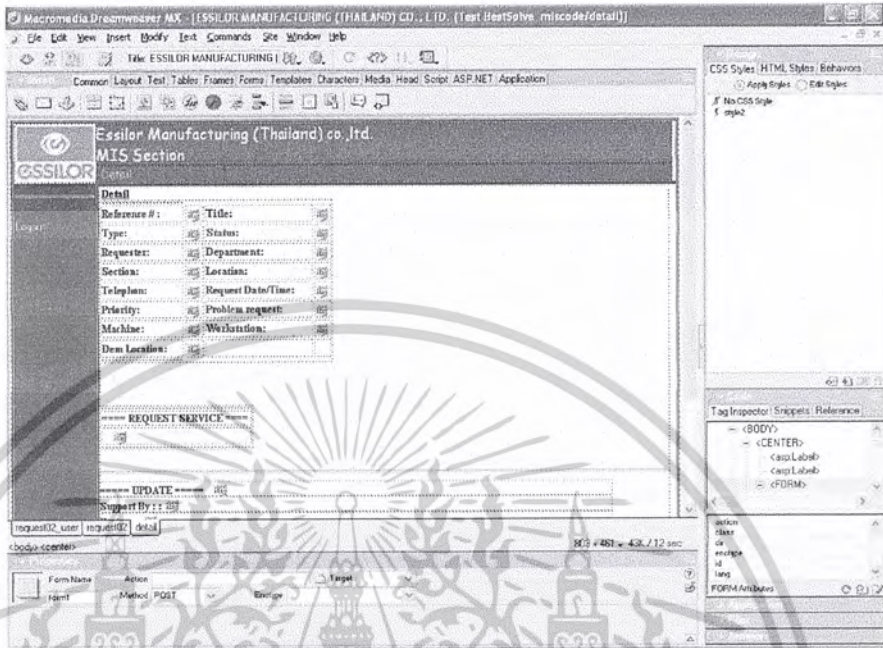
```

```
SmtpMail.Send(mail)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

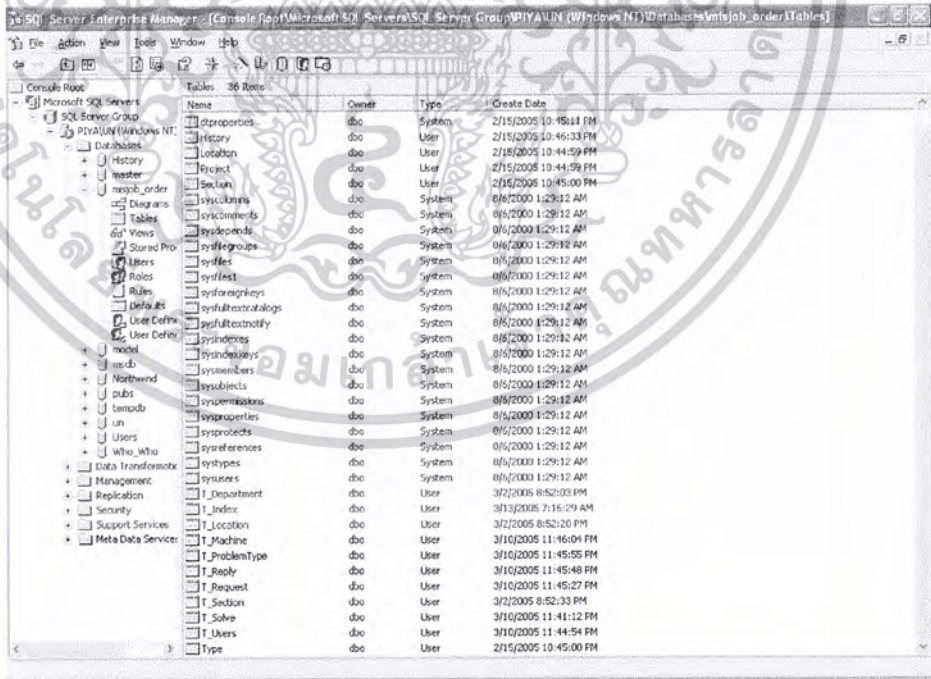
2. Macromedia Dreamweaver MX

ใช้เพื่อเป็นเครื่องมือในการออกแบบตกแต่ง เว็บไซต์ โปรแกรม



3. Microsoft SQL Server 2000

เป็นระบบฐานข้อมูลที่มีคนนิยมใช้ในองค์กรขนาดใหญ่



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

ภาษาที่ใช้ระบุความต้องการ(Query Languages)

รูปแบบพื้นฐานของ ดัชนี ที่นิยมใช้ใน ระบุความต้องการ

1) คำเดี่ยว (Single-word Queries)

ถือว่าเอกสารประกอบด้วย Word หลายๆคำมาต่อกัน โดยใช้สัญลักษณ์ต่างๆในการแบ่งคำ เช่น ช่องว่าง - / เป็นต้น

ข้อเสียคือยากที่จะวิเคราะห์ในการแบ่งคำเนื่องจากอาจมีประคำศัพท์ที่หากแยกกันแล้ว อาจทำให้ความหมายเปลี่ยนไป เช่น on-line

2) ข้อความ(Context Queries)

เป็นการค้นหาคำที่อยู่ใน context สามารถแบบได้เป็น 2 ประเภท

Phrase

เป็นการค้นหาแบบค้นคำที่ต่อกันและตัดคำที่ไม่มีความหมายออกไป

Proximity

เป็นการวิเคราะห์ความเหมือนของคำด้วย

3) ข้อความทางตรรกศาสตร์(Boolean Queries)

ใช้และ (and), หรือ (or), ไม่ (not) คือจะต้องทำการแปลงความต้องการ ที่ผู้ใช้กรอกให้อยู่ในรูปของตรรกศาสตร์ก่อนไปเปรียบเทียบกับฐานข้อมูลต่อไป

ข้อเสียคือผู้ใช้ทั่วไปมักไม่เข้าใจและไม่นิยมใช้ Boolean logic นอกจากนี้ยังนำผลที่ได้มาจัดลำดับ ไม่ได้ด้วย

4) Natural Language

เป็นการระบุความต้องการแบบ วลี คือถามแต่ผู้ใช้จะป้อนเข้ามา ระบบจะทำการประมวลผลโดยใช้รูปแบบของการให้น้ำหนักในแต่ละดัชนี ที่จะไม่เท่ากันในการแยกความตรงประเด็น

กระบวนการปรับปรุงความต้องการ

เป็นกระบวนการเปลี่ยนแปลงความต้องการที่ผู้ใช้ป้อนเข้ามา เพื่อให้ได้ผลลัพธ์ที่มีความตรงประเด็นมากขึ้น

ตัวอย่างของผลลัพธ์ที่ไม่ดี

- ข้อมูลให้ผลลัพธ์ออกมาน้อยเกินไปหรือไม่ก็มากเกินไป
- ผู้ใช้โดยทั่วไปมักจะเฝ้ามองผลลัพธ์บางอย่างเพียงด้านเดียว
- ผู้ใช้มัก ไม่ต้องการผลลัพธ์ที่ออกมาไป

ดังนั้นจึงต้องมีการเรียงลำดับตามความต้องประเด็นของเอกสารเพื่อให้ตรงตามที่ผู้ใช้ต้องการมาก

ที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัญหาของ ข้อมูลที่ผู้ใช้กรอก

- สิ่งที่ป้อนเข้ามาน้อยไป
- ส่วนมากแล้วสิ่งที่ผู้ใช้ป้อนเข้ามาให้ระบบนั้นจะน้อยไปประมาณ 1 ถึง 3 คำ ซึ่งอาจทำให้เราไม่เข้าใจความต้องการที่แท้จริงของผู้ใช้ได้
- สิ่งที่ป้อนเข้ามามีลักษณะที่กำกวม
- ในบางคำศัพท์ที่ผู้ใช้ป้อนเข้านั้นอาจมีได้หลากหลายความหมาย ซึ่งระบบอาจนำไปตีความผิดได้ เช่น คำว่า "Saturn" อาจหมายถึง ดาวเคราะห์หรือรถก็ได้

แนวทางในปรับปรุงข้อมูลที่ผู้ใช้กรอกนั้นทำได้ 2 แบบคือ

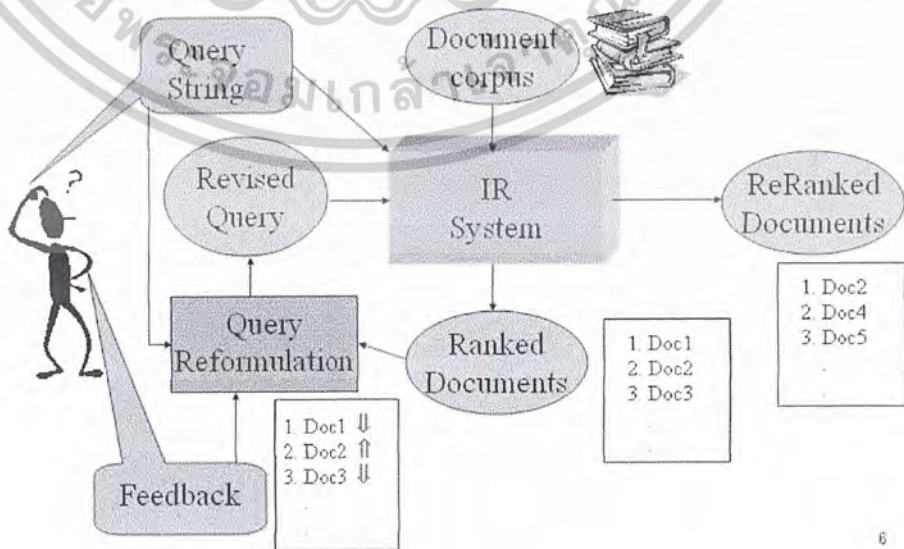
- สร้างข้อมูลที่ผู้ใช้กรอกใหม่ด้วยการป้อนข้อมูลกลับของผู้ใช้ (User feedback)
- สร้างข้อมูลที่ผู้ใช้กรอกใหม่ด้วย ใช้สูตร (Automatically)

โดยที่วิธีทั้ง 2 ข้างต้นนอกจะมีกระบวนการทำงาน 2 ขั้นตอนเหมือนกันคือ

- การคำนวณค่าน้ำหนักแต่ละดัชนี ในข้อมูลที่ผู้ใช้กรอกใหม่เพื่อให้ได้ข้อมูลที่ผู้ใช้กรอกใหม่ที่ตรงประเด็นมากขึ้น
- การเปลี่ยนแปลงค่าน้ำหนักดัชนีเดิมที่ผู้ใช้ป้อนเข้ามาตามที่เราได้ค่านวใหม่ หลักการป้อนข้อมูลกลับของผู้ใช้

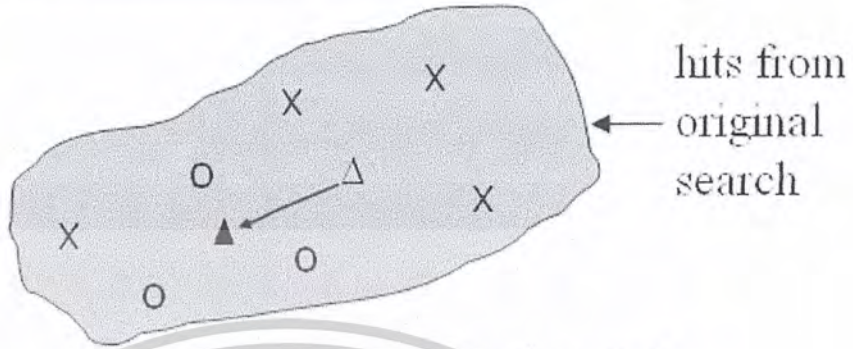
มีหลักการดังนี้ เมื่อผู้ใช้ทำการป้อนความต้องการเข้าไปในระบบระบบก็จะส่งผลลัพธ์ที่จัดทำารเรียงลำดับออกมาให้ ผู้ใช้ทำการเปลี่ยนแปลงผลลัพธ์ที่ออกมาโดยทำการเลือกและจัดเรียงลำดับเอกสารใหม่ตามที่ผู้ใช้ต้องการ (ในตอนนี้จะเรียกว่าข้อมูลย้อนกลับ จากผู้ใช้) แล้วระบบจะนำชุดผลลัพธ์ที่ได้ใหม่จากผู้ใช้ไปทำการประมวลผลเพื่อทำการเปลี่ยนแปลงความต้องการของผู้ใช้ที่ป้อนเข้ามาใหม่ เพื่อให้ได้ตามที่ต้องการ ระบบจะกระทำกระบวนการดังกล่าวซ้ำๆกัน ไปเรื่อยๆจนกว่าผู้ใช้จะพอใจ ดังรูปข้างล่าง

Relevance Feedback Architecture



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำการเปลี่ยนแปลงตัวความต้องการนั้นก็เปรียบเสมือนกับที่เราทำการย้ายเจ้าตัวความต้องการให้ไปอยู่ใกล้กับกลุ่มเอกสารที่ตรงประเด็นนั่นเอง



- X : เอกสารที่ไม่ตรงประเด็น
- O : เอกสารที่ตรงประเด็น
- △ : คำถามที่ผู้ใช้ป้อนเดิม
- ▲ : คำถามหลังจากการคำนวณใหม่

กระบวนการเปลี่ยนแปลงคำถาม

เป็นการเพิ่มคำใหม่เข้าไปในคำถามเดิม เช่น “วิทยาศาสตร์” อาจทำการเปลี่ยนไปเป็น “วันวิทยาศาสตร์” จะได้มีความตรงประเด็นมากขึ้น

การเปลี่ยนแปลงค่าน้ำหนักของแต่ละคำสำคัญในคำถามใหม่ โดยอาจทำการเพิ่มน้ำหนักให้กับคำสำคัญในคำถามถ้าคำนั้นมีมีความตรงประเด็น ในทางกลับกันก็จะทำการลดค่าน้ำหนักคำสำคัญในคำถามถ้าคำนั้นไม่ตรงประเด็น

เราสามารถคำนวณการเปลี่ยนแปลงคำถามได้โดยใช้หลักการของพีชคณิต(vector algebra) ซึ่งมีอัลกอริทึมหลายแบบดังนี้

Rocchio Method

$$Q_1 = \alpha Q_0 + \frac{\beta}{n_1} \sum_{d_j \in D_r} \vec{d}_j - \frac{\gamma}{n_2} \sum_{d_j \in D_n} \vec{d}_j$$

- โดยที่ Q_0 คือ Vectorของคำถามเดิม
- D_r คือ กลุ่มของเอกสารที่ตรงประเด็น
- D_n คือ กลุ่มของเอกสารที่ไม่ตรงประเด็น
- n_1 คือ จำนวนของเอกสารที่ผู้ใช้เลือกว่าตรงประเด็น
- n_2 คือ จำนวนของเอกสารที่ผู้ใช้เลือกว่าไม่ตรงประเด็น
- α, β และ γ คือ ค่าคงที่ของเอกสารที่ตรงประเด็นและไม่ตรงประเด็น(โดยปกติจะตั้งค่าไว้ที่ 1, 0.75, 0.25 ตามลำดับ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการคำนวณด้วย Rocchio

$$R_1 = (.030, 0.00, 0.00, .025, .025, .050, 0.00, 0.00, .120)$$

$$R_2 = (.020, .009, .020, .002, .050, .025, .100, .100, .120)$$

เอกสารที่ตรงประเด็น

$$S_1 = (.030, .010, .020, 0.00, .005, .025, 0.00, .020, 0.00)$$

$$Q = (0.00, 0.00, 0.00, 0.00, .500, 0.00, .450, 0.00, .950)$$

เอกสารที่ไม่ตรงประเด็น

คำถามเริ่มต้น

$$\alpha = 1$$

$$\beta = 0.75 \quad \text{ค่าคงที่}$$

$$\gamma = 0.25$$

การคำนวณด้วย Rocchio

$$Q_{new} = \alpha \times Q + \left(\frac{\beta}{2} \times (R_1 + R_2) \right) - \left(\frac{\gamma}{1} \times S_1 \right)$$

ผลลัพธ์ feedback query

$$Q_{new} = (0.011, 0.000875, 0.002, 0.01, 0.527, 0.022, 0.488, 0.033, 1.04)$$

Ide Regular Method

$$\bar{q}_1 = \alpha \bar{q}_0 + \beta \sum_{\forall d_j \in D_r} \bar{d}_j - \gamma \sum_{\forall d_j \in D_n} \bar{d}_j$$

α คือ ค่าน้ำหนักสำหรับคำถามเดิม

β คือ ค่าน้ำหนักสำหรับเอกสารที่ตรงประเด็น

γ คือ ค่าน้ำหนักสำหรับเอกสารที่ไม่ตรงประเด็น

ตัวอย่างการคำนวณ

Original Query : (5,0,3,0,1)

Document D1 Relevant : (2,1,2,0,0)

Document D2 Nonrelevant : (1,0,0,0,2)

$$\alpha = 0.50 \quad \beta = 0.25$$

$$Q' = Q + 0.5D1 - 0.25D2$$

$$= (5, 0, 3, 0, 1) + 0.5(2, 1, 2, 0, 0) - 0.25(1, 0, 0, 0, 2)$$

$$= (5.75, 0.50, 4.0, 0.0, 0.5)$$

จากรูปแสดงผลการคำนวณค่าคำสำคัญของคำถามใหม่



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความน่าจะเป็นการปรับปรุ้ค่าน้ำหนัก(Probabilistic Re-Weighting)

หลังจากการที่เราทำการคำนวณน้ำหนักในแต่ละคำสำคัญในคำถามใหม่แล้วนั้นเราจำเป็นต้องคำนวณหาค่าความเหมือน ใหม่อีกด้วยตามรูปแบบ ดังนี้

สมมติว่าแต่ละคำสำคัญเป็นอิสระจากกันและเป็นแบบเอกสารแบบ มีแค่จริงกับเท็จ(binary index)

$$sim(d_j, q) \approx \sum_{i=1}^l w_{i,q} \times w_{i,j} \times \left(\log \frac{P(k_i | R)}{1 - P(k_i | R)} + \log \frac{1 - P(k_i | \bar{R})}{P(k_i | \bar{R})} \right)$$

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

Dr คือกลุ่มของเอกสารที่ตรงประเด็น

Dr,i คือกลุ่มของเอกสารที่ตรงประเด็นที่มีคำ ki อยู่

ni คือจำนวนของเอกสารที่อยู่ในฐานข้อมูลที่มีคำ ki อยู่

ซึ่งมันก็ยังคงต้องอาศัยหลักการป้อนข้อมูลกลับอยู่

ตัวอย่างการคำนวณของ ความน่าจะเป็น(Probabilistic Re-Weighting)

Q : Gold Silver Truck

D1 : Shipment of gold damaged in fire. (Relevance)

D2 : Delivery of silver arrived in a silver truck.

D3 : Shipment of gold arrived in a truck. (Relevance)

D4 : My gold truck gone by silver ship.

Var	Gold	Silver	Truck
N	4	4	4
n	3	2	3
Dr	2	2	2
Dri	2	0	1

Doc	Gold	Silver	Truck
D1	1	0	0
D2	0	1	1
D3	1	0	1
D4	1	1	1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$P(k_{gold} | R) = \frac{2 + \frac{3}{4}}{2+1} = 0.917$$

$$P(k_{silver} | R) = \frac{0 + \frac{2}{4}}{2+1} = 0.167$$

$$P(k_{gold} | \bar{R}) = \frac{3 - 2 + \frac{3}{4}}{4 - 2 + 1} = 0.583$$

$$P(k_{silver} | \bar{R}) = \frac{2 - 0 + \frac{2}{4}}{4 - 2 + 1} = 0.833$$

$$P(k_{Truck} | R) = \frac{1 + \frac{3}{4}}{2+1} = 0.583$$

$$P(k_{Truck} | \bar{R}) = \frac{3 - 1 + \frac{3}{4}}{4 - 2 + 1} = 0.917$$

$$sim(d_1, q) = 1 \times 1 \times \left(\log \frac{0.917}{1 - 0.917} + \log \frac{1 - 0.583}{0.583} \right) + 0 + 0 = 0.89$$

$$Sim(d_2, q) = -2.29$$

$$Sim(d_3, q) = 0$$

$$Sim(d_4, q) = -1.39$$

ลำดับใหม่ = D1, D3, D4, D3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

Frakes, William.B. (1992): Stemming Algorithms. In: Information Retrieval. Data Structures & Algorithms. William B. Frakes & Ricardo BaezaYates(eds.). PTR PrenticeHall:Englewood Cliffs, New Jersey. S. 131-160.

Kowalski, Gerald J. & Maybury, Mark T. (eds.) (2000): Information Storage and Retrieval Systems. Theory and Implementation. 2nd ed. Kluwer Academic Publishers: Boston u.a.

PorterAlgorithmus and Implementierung in Snowball fur zahlreiche Sprachen

www.snowball.tartarus.org

PorterAlgorithmus and Implementierung in verschiedenen Programmiersprachen furs Englische

www.tartarus.org/~martin/PorterStemmer/

PorterAlgorithmus and Implementierung in verschiedenen Programmiersprachen furs Franzosische

snowball.tartarus.org/french/stemmer.html

Rossen Kowatschew, Johannes Lang (2002): Stemming. Seminarreferat

www.rzuser.uniheidelberg.de/~jlang/Stemming/

Kiduk Yang, Denqi Song, Wooseob Jeoung, Rong Tang: Nice Stemmer. Projekt

www.ils.unc.edu/yangk/nstem.htm