



# วิทยานิพนธ์

อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้าง  
แผนที่ความรู้

**Classification Algorithm for Knowledge Mapping of  
Organizational Knowledge Workers**

นางสาวบุษกร สุคนธวงศาโรจน์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

พ.ศ. 2551





## ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

ปริญญา

วิทยาการคอมพิวเตอร์

สาขา

วิทยาการคอมพิวเตอร์

ภาควิชา

เรื่อง อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้

Classification Algorithm for Knowledge Mapping of Organizational Knowledge Workers

نامผู้วิจัย นางสาวบุษกร สุคนธวงศาโรจน์

ได้พิจารณาเห็นชอบโดย

ประธานกรรมการ

( รองศาสตราจารย์อนงค์นาฏ ศรีวิหค, Ph.D. )

กรรมการ

( รองศาสตราจารย์ประสงค์ ปราณิตพลกรัง, Ph.D. )

กรรมการ

( ผู้ช่วยศาสตราจารย์สมชาย นำประเสริฐชัย, Ph.D. )

หัวหน้าภาควิชา

( ผู้ช่วยศาสตราจารย์ศิริกร จันทน์นวล, M.S. )

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

( รองศาสตราจารย์กัญจนา ชีระกุล, D.Agr. )

คณบดีบัณฑิตวิทยาลัย

วันที่ ..... เดือน ..... พ.ศ. ....

วิทยานิพนธ์

เรื่อง

อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้

Classification Algorithm for Knowledge Mapping of Organizational Knowledge Workers

โดย

นางสาวบุษกร สุคนธวงศาโรจน์

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

พ.ศ. 2551

บุษกร สุคนธวงศาโรจน์ 2551: อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้ ปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์) สาขาวิทยาการคอมพิวเตอร์ ภาควิชา วิทยาการคอมพิวเตอร์ ปรชานกรรมการที่ปรึกษา: รองศาสตราจารย์อนงค์นัญ ศรีวิหค, Ph.D.  
106 หน้า

การจัดการความรู้เป็นกระบวนการในการนำความรู้ที่มีอยู่ หรือความรู้ที่ได้จากการเรียนรู้ มาใช้ในการ แก้ไขปัญหา การกำหนดกลยุทธ์ และการตัดสินใจ ทำให้เกิดประโยชน์สูงสุดต่อองค์กร และสร้างรายได้ เปรียบในเชิงการแข่งขันกับหน่วยงานภายนอก โดยมีแผนที่ความรู้เป็นเครื่องมือที่ใช้ในองค์กรเพื่อสนับสนุน การจัดการความรู้ ช่วยให้เห็นว่าองค์กรมีความรู้อะไรบ้าง ความรู้ต่าง ๆ มีอยู่ที่ไหนและจะสามารถเข้าถึง ความรู้นั้นด้วยวิธีการใด แต่ก่อนที่จะมาถึงจุดนี้ได้จำเป็นต้องมีการจำแนกความเชี่ยวชาญในงานด้านต่าง ๆ โดย งานวิจัยนี้ได้นำหลักการท่าเหมืองข้อมูลเข้ามาใช้ในการจัดแบ่งประเภทความเชี่ยวชาญของผู้ปฏิบัติงาน และ ค้นหารูปแบบของข้อมูลที่น่าสนใจ เพื่อใช้ในการหาความสัมพันธ์ หรือลักษณะของข้อมูลที่เกิดขึ้นร่วมกัน งานวิจัยฉบับนี้มีจุดมุ่งหมายในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานตามคลาสเบลของความเชี่ยวชาญ และหาความสัมพันธ์ของคุณลักษณะที่เกี่ยวข้องกับความเชี่ยวชาญในสาขาอาชีพที่ทำการศึกษา ข้อมูลที่ใช้ใน การทดลองได้จากหน่วยงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สาขางานรองผู้ว่าการบัญชีและการเงิน สาขางาน รองผู้ว่าการผลิตไฟฟ้า และสาขางานรองผู้ว่าการเชื้อเพลิง โดยทำการทดลองกับกลุ่มตัวอย่าง 28 หน่วยงาน การ ทดลองขั้นตอนแรกใช้วิธีการจำแนกประเภทความเชี่ยวชาญเชิงเปรียบเทียบด้วยอัลกอริทึม 3 แบบ คือ (1) อัลกอริทึม Decision Tree (C4.5) (2) อัลกอริทึม OneR (3) อัลกอริทึม Naïve Bayes การทดลองเพื่อหา อัลกอริทึมที่มีความแม่นยำที่สุดในการทำนายความเชี่ยวชาญ โดยใช้ตัววัดประสิทธิภาพ 4 ตัว (1) Accuracy (2) Precision/Recall (3) F-measure และ (4) Root Mean-Squared Error (RMSE) จากการทดลองพบว่าอัลกอริทึมที่ ให้ค่าความถูกต้องในการทำนายคลาสของความเชี่ยวชาญสูงสุด คือ อัลกอริทึม C4.5 มี Accuracy 95.61% ค่า Precision เท่ากับ 0.957 ค่า Recall เท่ากับ 0.999 ค่า F-measure เท่ากับ 0.978 และค่า RMSE เท่ากับ 0.093

การทดลองขั้นที่สอง คือ การหาความสัมพันธ์ของคุณลักษณะที่มีผลต่อความเชี่ยวชาญของ ผู้ปฏิบัติงาน โดยใช้ Association Rule Mining (อัลกอริทึม Apriori) คุณลักษณะที่ศึกษาได้แก่ข้อมูลของคุณวุฒิ การศึกษา คณะที่จบการศึกษา ตำแหน่งงาน หน่วยงานที่ผู้ปฏิบัติงานสังกัด และอายุความเชี่ยวชาญ ผลจากการ วิเคราะห์นี้คือ (1) ตำแหน่งงานช่าง จะมีคุณวุฒิต่ำกว่าระดับปริญญาตรี (2) หน่วยงานและอายุความเชี่ยวชาญใน งานมีความสัมพันธ์กับความเชี่ยวชาญของผู้ปฏิบัติงาน (3) การปฏิบัติงานที่มีส่วนเกี่ยวข้องกันจะช่วยให้ ผู้ปฏิบัติงานมีความเชี่ยวชาญในงานเพิ่มขึ้น โดยความรู้ที่ได้จากการศึกษารั้งนี้สามารถนำไปใช้ในการสร้าง แผนที่ความรู้ขององค์กร ช่วยให้เกิดความเข้าใจความสัมพันธ์ระหว่างคุณลักษณะของแหล่งความรู้ต่าง ๆ ใน องค์กร ทำให้เกิดการเชื่อมโยงของความรู้เข้าด้วยกัน และช่วยเพิ่มประสิทธิภาพในการจัดการความรู้ขององค์กร

Budsakorn Sukonthawongsaroth 2008: Classification Algorithm for Knowledge Mapping of Organizational Knowledge Workers. Master of Science (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Anongnart Srivihok, Ph.D. 106 pages.

Knowledge Management is a process that helps organizations to identify, select, organize, disseminate, and transfer important information and expertise that are the part of the organization's memory. The structuring of knowledge enables effective and efficient problem solving, strategic planning and decision making. A knowledge map is a navigation aid to explicit and tacit knowledge, supporting Knowledge Management and describing knowledge flows throughout an organization. However, the study of knowledge mapping in Thai organization is not widely conducted. Algorithms are used to classify attributes of knowledge workers to expert classes in order to generate a knowledge map. This paper proposes an approach for knowledge mapping of experts in organization by using data mining techniques. The expert classes in this paper include Power Generation Business, Account and Finance Deputy Governor and Fuel Deputy Governor, Electricity Generating Authority of Thailand. Data includes 28 sample groups and 805 instances. we classified domain knowledge by using Decision Trees (C4.5), OneR and Naïve Bayes. The prediction performances of three classifiers are measured by four indices used for evaluating the efficiency of classification. The indices include Precision, Recall, F-measure and Root mean-squared error (RMSE). There are two steps in this study. Step 1, three classification algorithms including Decision Tree (C4.5), OneR and Naïve Bayes were compared on their classification performances. Results show that Decision Tree (C4.5) algorithm is the best algorithm with 95.61% of accurate prediction, Precision (0.957), Recall (0.999), F-measure (0.978) and RMSE (0.093).

In Step 2, Association Rules Mining by Apriori algorithm was employed to explore positive association relationships with experts class such as Education level, Faculty graduated, Position, Division, Department, Section and Expert level. Results are as follows. First, if staff position is technician, the education level is lower than bachelor's degree. Second, the attributes include Division, Department, Section and Expert level can be used to identify classes of knowledge workers. Third, working on various related tasks will increase skills of knowledge workers. Finally, results from this study will be useful for developing a knowledge map. It involves locating important knowledge in the organization, understands relationships of knowledge components and increases the performance of knowledge management in organization.

---

Student's signature

---

Thesis Advisor's signature

\_\_\_\_ / \_\_\_\_ / \_\_\_\_

## กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. อนงค์นาฏ ศรีวิหค ประธานกรรมการที่ปรึกษา รองศาสตราจารย์ ดร. ประสงค์ ปราณิตพลกรัง กรรมการวิชาเอก และผู้ช่วยศาสตราจารย์ ดร. สมชาย นำประเสริฐชัย กรรมการวิชาการ ที่ทุกท่านให้คำแนะนำและให้ความช่วยเหลือในการทำวิทยานิพนธ์ จนสามารถทำได้เป็นผลสำเร็จ ขอกราบขอบพระคุณคณาจารย์ทุกท่าน ที่คอยประสิทธิ์ประสาทวิชาความรู้ แนะนำ สั่งสอน ให้ความใส่ใจ ห่วงใย แก่ข้าพเจ้าตลอดระยะเวลาการศึกษา

ขอขอบพระคุณ คุณพ่อ คุณแม่ และพี่ ๆ ของข้าพเจ้า ที่ให้กำลังใจอันเป็นแรงผลักดันให้มิย่อท้อกับปัญหา และอุปสรรค ที่ผ่านพามาในช่วงของการศึกษา จนสามารถประสบความสำเร็จในการศึกษาระดับปริญญาโท

ขอขอบคุณเพื่อน ๆ พี่ ๆ ภาควิชาวิทยาการคอมพิวเตอร์ ทั้งรุ่นเดียวกัน รุ่นพี่ ที่ช่วยเหลือและให้คำแนะนำในการทำวิทยานิพนธ์ตลอดมา ขอขอบพระคุณ การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย ที่ให้ข้อมูลในการทำวิจัย

สุดท้ายขอขอบคุณโครงการปริญญาโท ภาควิชาวิทยาการคอมพิวเตอร์ ที่ให้ทุนสนับสนุนในการเดินทางไปเสนอผลงานทางวิชาการและให้โอกาสข้าพเจ้าได้เรียนรู้จนสำเร็จการศึกษา

งานวิทยานิพนธ์นี้ ข้าพเจ้าหวังเป็นอย่างยิ่งว่าจะเป็นประโยชน์ต่อผู้ศึกษาค้นคว้าและผู้สนใจ หากมีข้อผิดพลาดประการใด ข้าพเจ้าน้อมรับไว้เพื่อนำไปใช้ในการปรับปรุงการทำงานต่อไป สำหรับความดีที่ได้รับจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้ามอบให้แก่ผู้มีพระคุณทุกท่าน

บุษกร สุคนธวงศาโรจน์

พฤษภาคม 2551

## สารบัญ

### หน้า

|                             |     |
|-----------------------------|-----|
| สารบัญ                      | (1) |
| สารบัญตาราง                 | (2) |
| สารบัญภาพ                   | (4) |
| คำอธิบายสัญลักษณ์และคำย่อ   | (5) |
| คำนำ                        | 1   |
| วัตถุประสงค์                | 3   |
| การตรวจเอกสาร               | 5   |
| อุปกรณ์และวิธีการ           | 43  |
| อุปกรณ์                     | 43  |
| วิธีการ                     | 43  |
| ผลและวิจารณ์                | 63  |
| ผล                          | 63  |
| วิจารณ์                     | 98  |
| สรุปและข้อเสนอแนะ           | 100 |
| สรุป                        | 100 |
| ข้อเสนอแนะ                  | 102 |
| เอกสารและสิ่งอ้างอิง        | 103 |
| ประวัติการศึกษา และการทำงาน | 106 |

## สารบัญตาราง

| ตารางที่                                                                                                                                              | หน้า |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 1 ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับแผนที่ความรู้                                                                                                    | 37   |
| 2 ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มข้อมูล (Clustering)                                                                                   | 39   |
| 3 ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับการจำแนกประเภทข้อมูล (Classification)                                                                            | 40   |
| 4 คุณลักษณะของข้อมูลที่ใช้ในการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงาน                                                                               | 51   |
| 5 คุณลักษณะของข้อมูลอายุงาน (Working year)                                                                                                            | 51   |
| 6 คุณลักษณะของข้อมูลตำแหน่งงาน (Position)                                                                                                             | 52   |
| 7 คุณลักษณะของข้อมูลคุณวุฒิการศึกษา (Education Level)                                                                                                 | 52   |
| 8 คุณลักษณะของข้อมูลคณะที่จบการศึกษา (Faculty graduated)                                                                                              | 53   |
| 9 คุณลักษณะของข้อมูลสาขาที่จบการศึกษา (Major)                                                                                                         | 53   |
| 10 คุณลักษณะของข้อมูลระดับฝ่าย (Division)                                                                                                             | 55   |
| 11 คุณลักษณะของข้อมูลระดับกอง (Department)                                                                                                            | 56   |
| 12 คุณลักษณะของข้อมูลระดับแผนก (Section)                                                                                                              | 58   |
| 13 คุณลักษณะของอายุความเชี่ยวชาญในงานที่ปฏิบัติ (Expert Level)                                                                                        | 62   |
| 14 คุณลักษณะของข้อมูลประเภทความเชี่ยวชาญ (Expert Class)                                                                                               | 62   |
| 15 การวัดประสิทธิภาพของอัลกอริทึม C4.5 อัลกอริทึม OneR อัลกอริทึม Naïve Bayes แบบเฉลี่ยโดยการแบ่งข้อมูลแบบ 10-Fold Cross-Validation ของข้อมูลชุดที่ 1 | 63   |
| 16 การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม C4.5                                                                          | 64   |
| 17 การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม OneR                                                                          | 65   |
| 18 การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม Naïve Bayes                                                                   | 66   |
| 19 การวัดประสิทธิภาพของอัลกอริทึม C4.5 อัลกอริทึม OneR อัลกอริทึม Naïve Bayes แบบเฉลี่ยโดยการแบ่งข้อมูลแบบ 10-Fold Cross-Validation                   | 67   |

### สารบัญตาราง (ต่อ)

| ตารางที่ |                                                                                                                                                                                       | หน้า |
|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 20       | เปรียบเทียบจำนวนข้อมูลที่ใช้ในการทดลองด้วยตัววัดประสิทธิภาพของ<br>อัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes แบบเฉลี่ยโดย<br>การแบ่งข้อมูลแบบ 10-Fold Cross-Validation | 69   |
| 21       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Boiler (กำหนดค่าสนับสนุนขั้นต่ำ<br>เท่ากับ 0.8)                                                                                                    | 71   |
| 22       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Civil (กำหนดค่าสนับสนุนขั้นต่ำ<br>เท่ากับ 0.9)                                                                                                     | 72   |
| 23       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Coal Ash (กำหนดค่าสนับสนุนขั้นต่ำ<br>เท่ากับ 0.95)                                                                                                 | 73   |
| 24       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Control Instrument (กำหนดค่า<br>สนับสนุน ขั้นต่ำ เท่ากับ 0.75)                                                                                     | 75   |
| 25       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Electrical (กำหนดค่าสนับสนุน ขั้นต่ำ<br>เท่ากับ 0.75)                                                                                              | 76   |
| 26       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Fuel Gas Desulfurization (FGD)<br>(กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.8 )                                                                            | 78   |
| 27       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน ICT (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ<br>0.65 )                                                                                                      | 80   |
| 28       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Management (กำหนดค่าสนับสนุนขั้น<br>ต่ำเท่ากับ 0.5)                                                                                                | 80   |
| 29       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Turbine (กำหนดค่าสนับสนุน ขั้นต่ำ<br>เท่ากับ 0.8 )                                                                                                 | 82   |
| 30       | กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Water Treatment (กำหนดค่า<br>สนับสนุน ขั้นต่ำเท่ากับ 0.9)                                                                                          | 83   |
| 31       | กฎความสัมพันธ์ของความเชี่ยวชาญทุกประเภท (กำหนดค่าสนับสนุน ขั้นต่ำ<br>เท่ากับ 0.15)                                                                                                    | 85   |

## สารบัญภาพ

| ภาพที่                                                                                   | หน้า |
|------------------------------------------------------------------------------------------|------|
| 1    แสดงกระบวนการค้นหาความรู้ในการทำเหมืองข้อมูล                                        | 9    |
| 2    การทำเหมืองข้อมูล                                                                   | 10   |
| 3    Classification Model ด้วยวิธี Decision Rules                                        | 11   |
| 4    ขั้นตอนการตรวจสอบความถูกต้องของโมเดล                                                | 12   |
| 5    การจำแนกประเภทผู้เชี่ยวชาญด้วยต้นไม้การตัดสินใจ                                     | 16   |
| 6    อัลกอริทึมต้นไม้การตัดสินใจ                                                         | 21   |
| 7    รูปแบบของกฎการตัดสินใจ                                                              | 21   |
| 8    ขั้นตอนการทำงานของอัลกอริทึม OneR                                                   | 22   |
| 9    ขั้นตอนการทำงานของอัลกอริทึม Naive Bayes                                            | 27   |
| 10    โครงสร้างการบังคับบัญชาของ กฟผ. เฉพาะหน่วยงานตัวอย่างที่นำข้อมูลมาใช้ในการทดลอง    | 45   |
| 11    โครงสร้างของงานวิจัยสำหรับทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน                       | 47   |
| 12    ต้นไม้การตัดสินใจสำหรับความเชี่ยวชาญด้าน Electrical และ ICT ด้วยอัลกอริทึม C4.5    | 68   |
| 13    ภาพรวมแผนที่ความรู้ขององค์กร กฟผ.                                                  | 88   |
| 14    แผนที่ความรู้ขององค์กร กฟผ. ระดับกอง                                               | 89   |
| 15    แผนที่ความรู้ขององค์กร กฟผ. แผนกประสิทธิภาพและแผนกบำรุงรักษาไฟฟ้า                  | 90   |
| 16    แผนที่ความรู้ขององค์กร กฟผ. แผนกเดินเครื่องกะ 1                                    | 91   |
| 17    แผนที่ความรู้ขององค์กร กฟผ. ฝ่ายบำรุงรักษาโรงไฟฟ้าแม่เมาะ                          | 92   |
| 18    แผนที่ความรู้ขององค์กร กฟผ. ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ                             | 94   |
| 19    แผนที่ความรู้ขององค์กร กฟผ. โรงไฟฟ้าน้ำพอง                                         | 95   |
| 20    แผนที่ความรู้ขององค์กร กฟผ. โรงไฟฟ้าภาคใต้ และโรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ | 96   |

### คำอธิบายสัญลักษณ์และคำย่อ

|      |   |                                                       |
|------|---|-------------------------------------------------------|
| กฟผ. | = | การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย                         |
| สพฟ. | = | ส่วนบริหารคุณภาพและเพิ่มผลผลิตสายงานผลิตไฟฟ้า         |
| ปวช. | = | ประกาศนียบัตรวิชาชีพ                                  |
| ปวส. | = | ประกาศนียบัตรวิชาชีพชั้นสูง                           |
| WEKA | = | Waikato Environment for Knowledge Analysis            |
| RMSE | = | Root-Mean-Square-Error                                |
| FGD  | = | Fuel Gas Desulfurization                              |
| ICT  | = | Information Communication Technology                  |
| C4.5 | = | อัลกอริทึมของการจำแนกประเภทข้อมูลของต้นไม้การตัดสินใจ |

## อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้

### Classification Algorithm for Knowledge Mapping of Organizational Knowledge Workers

#### คำนำ

ปัจจุบัน หน่วยงานต่าง ๆ ได้เห็นความสำคัญในการจัดการความรู้ (Knowledge Management) เพิ่มขึ้น โดยจัดกิจกรรมให้พนักงานเข้ามามีส่วนร่วมในการแลกเปลี่ยน เรียนรู้ จากประสบการณ์ทางตรง และจากประสบการณ์ที่ได้จากผู้ร่วมงาน จนสามารถนำมารวบรวม จัดทำเป็นวิธีปฏิบัติที่ยอดเยี่ยม (Best Practice) เมื่อดำเนินการมาถึงระดับที่สามารถสร้างจิตสำนึก ให้กับพนักงานที่จะยอมแลกเปลี่ยนเรียนรู้กันอย่างสันทัดใจแล้ว หน่วยงานนั้นจะต้องมีการจับ ความรู้ (Capture Knowledge) และทำการเปลี่ยนแปลงความรู้ที่ได้มา ให้อยู่ในรูปแบบที่สามารถ นำไปใช้งานและสามารถเข้าถึงข้อมูลนั้นได้ง่าย ซึ่งพฤติกรรมเหล่านี้จะช่วยหล่อหลอมผู้คนใน องค์กรให้มีวัฒนธรรมที่เอื้อเพื่อแผ่ต่อกัน ยอมรับฟังความคิดเห็นของกันและกัน พร้อมทั้งจะ แลกเปลี่ยนเรียนรู้ร่วมกัน และเมื่อนั้นองค์กรจะสามารถก้าวไปสู่การเป็นองค์กรแห่งการเรียนรู้ (Learning Organization)

ข้อมูลที่ได้จากการแลกเปลี่ยนเรียนรู้ จะมีการจัดเก็บลงในฐานข้อมูล และเราจะนำข้อมูลที่ได้ จัดเก็บไว้มาใช้แก้ปัญหาในงานที่ปฏิบัติ ช่วยเพิ่มประสิทธิภาพในการทำงาน รวมถึงนำมาสร้าง เป็นองค์ความรู้ขององค์กร เพื่อช่วยให้องค์กรมีแหล่งความรู้ ที่ผู้ปฏิบัติงานจะสามารถดึงเอาความรู้ ที่มีอยู่นั้น มาช่วยขับเคลื่อนให้องค์กรทำงานอย่างมีประสิทธิภาพและประสิทธิผลที่ดียิ่งขึ้น โดย ต้องเริ่มจากการค้นหาความรู้จากฐานข้อมูล แล้วทำการวิเคราะห์ความรู้ที่ค้นหามาได้ด้วยเทคนิคที่ นิยมนำมาใช้ในการวิเคราะห์ข้อมูล ได้แก่ เทคนิคการทำเหมืองข้อมูล

การทำเหมืองข้อมูล สามารถจัดกลุ่มกว้างๆได้เป็น 2 กลุ่ม คือ กลุ่มที่ 1 ใช้เหมืองข้อมูลเพื่อ การทำนาย เป็นการนำความรู้ที่เรียนรู้จากข้อมูลที่มีอยู่มาใช้ประโยชน์ในการทำนายข้อมูลใหม่ที่จะ เกิดขึ้นในอนาคต และกลุ่มที่ 2 ใช้เหมืองข้อมูล เพื่อการอธิบายเป็นการค้นหารูปแบบของข้อมูลที่ น่าสนใจ เพื่อใช้ในการหาความสัมพันธ์ หรือลักษณะที่เชื่อมโยงกันของข้อมูล ซึ่งงานวิจัยฉบับนี้ ได้ใช้หลักการจำแนกข้อมูลด้วยเทคนิคการทำเหมืองข้อมูล โดยแบ่งออกเป็น 2 กลุ่ม เช่นเดียวกับ

การแบ่งกลุ่มของการทำเหมืองข้อมูลทั่วไป ทั้งนี้ เพื่อนำความรู้ที่ได้ไปใช้ในการสร้างแผนที่ความรู้ขององค์กร

ความรู้ขององค์กรมีจำนวนมากมาย วิธีหนึ่งที่เราสามารถนำความรู้ที่มีอยู่มาใช้ให้เกิดประโยชน์สูงสุด คือ การทำแผนที่ความรู้ (Knowledge Maps) แผนที่ความรู้ เป็นเครื่องมือที่ช่วยให้พนักงานในองค์กรทราบว่า ในองค์กรมีความรู้อะไรบ้าง จะหาความรู้ที่ต้องการได้ที่ไหน ใครเป็นผู้รู้ หรือเป็นผู้สร้างความรู้ และเมื่อพบปัญหา อุปสรรคในการปฏิบัติงาน แผนที่ความรู้จะเป็นสิ่งที่ช่วยผู้ปฏิบัติงานให้พบผู้รู้ / ผู้เชี่ยวชาญ ได้อย่างถูกต้อง ถูกทาง สามารถช่วยเหลือ หรือ ให้คำแนะนำในการแก้ไขปัญหาต่าง ๆ ได้ทันกาล หรือเกิดความเสียหายน้อยสุด เนื่องจากแผนที่ความรู้ได้มีการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงานไว้อย่างชัดเจน

มูลเหตุจูงใจที่ทำให้เกิดงานวิจัยฉบับนี้ เกิดขึ้นจากความต้องการที่จะช่วยลดปัญหา และอุปสรรคในการทำงาน โดยได้เล็งเห็นว่าผู้ปฏิบัติงานในองค์กรมีความรู้ความสามารถเฉพาะด้านแต่ยังไม่ได้มีการรวบรวมแหล่งความรู้ และผู้รู้ที่จะสามารถพัฒนาต่อเป็นผู้เชี่ยวชาญ มาช่วยให้การทำงานเป็นไปอย่างมีประสิทธิภาพ เลือกใช้คนให้ถูกต้องเหมาะสมกับความสามารถ นำมาสู่ความเสี่ยงในการทำงาน สูญเสียโอกาสที่องค์กรจะแข่งขันกับหน่วยงานภายนอกได้ จึงได้ทำการศึกษางานวิจัยในเรื่องการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงาน เพื่อนำไปใช้ในการสร้างแผนที่ความรู้ขององค์กร จากการศึกษางานวิจัยที่เกี่ยวข้องทั้งในประเทศไทยและต่างประเทศ นั้น มีการวิจัยค่อนข้างน้อยมาก ซึ่งยังไม่พบงานวิจัยที่มีผู้ใช้อัลกอริทึมของต้นไม้การตัดสินใจ (Decision Tree) โดยอัลกอริทึม C4.5 กฎการตัดสินใจ (Decision Rules) โดยอัลกอริทึม OneR หรือใช้หลักการทางสถิติ โดยอัลกอริทึม Naïve Bayes มาเป็นตัวดำเนินการในงานวิจัยการสร้างแผนที่ความรู้ขององค์กร

## วัตถุประสงค์

1. เพื่อศึกษาและเปรียบเทียบประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes
2. เพื่อให้ได้อัลกอริทึมที่เหมาะสมกับการสร้างแผนที่ความรู้
3. หาความสัมพันธ์ของคุณลักษณะของข้อมูลในแต่ละคลาสของความเชี่ยวชาญ โดยใช้ Association Rule

## ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถเลือกอัลกอริทึมที่เหมาะสมนำไปใช้ในการสร้างแผนที่ความรู้ได้
2. สามารถค้นพบความเชี่ยวชาญหลักของหน่วยงานได้
3. สามารถรู้ได้ว่าหน่วยงานใดขาดแคลนผู้เชี่ยวชาญในเรื่องใด
4. สามารถระบุความเชี่ยวชาญในงานของหน่วยงานต่าง ๆ ได้
5. สามารถหาความสัมพันธ์ของคุณลักษณะ (attribute) ของแต่ละคลาส

### ขอบเขตของงานวิจัย

1. ข้อมูลจากหน่วยงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง ปี 2550 จำนวน 805 รายการ ประกอบด้วยข้อมูลบุคคลกับข้อมูลความเชี่ยวชาญของผู้ปฏิบัติงาน จำนวน 10 คุณลักษณะ
2. จำแนกความเชี่ยวชาญของผู้ปฏิบัติงานเป็น 10 ประเภท
3. ระดับความเชี่ยวชาญของผู้ปฏิบัติงาน แบ่งเป็น 4 ระดับ

## การตรวจเอกสาร

### ความรู้เบื้องต้นเกี่ยวกับงานวิจัย

#### การจัดการความรู้

การจัดการความรู้ คือ การรวบรวมองค์ความรู้ที่มีอยู่ในองค์กร ตัวบุคคล และเอกสาร มาพัฒนาให้เป็นระบบ เพื่อให้ทุกคนในองค์กรสามารถเข้าถึงความรู้ และพัฒนาตนเองให้เป็นผู้รู้ รวมทั้งปฏิบัติงานได้อย่างมีประสิทธิภาพ อันจะส่งผลให้องค์กรมีความสามารถในเชิงแข่งขันกับคู่แข่งกันได้สูงสุด โดยเราสามารถจำแนกความรู้ได้ 2 ประเภท (Nanoka and Takeuchi, 1995) คือ

1. ความรู้ที่ฝังอยู่ในคน (Tacit Knowledge) เป็นความรู้ที่ได้จากประสบการณ์ พรสวรรค์ หรือสัญชาตญาณของแต่ละบุคคลในการทำความเข้าใจในสิ่งต่าง ๆ เป็นความรู้ที่ไม่สามารถถ่ายทอดออกมาเป็นคำพูดหรือลายลักษณ์อักษรได้โดยง่าย เช่น ทักษะในการทำงาน งานฝีมือ หรือการคิดเชิงวิเคราะห์ บางครั้งจึงเรียกว่าเป็นความรู้แบบนามธรรม

2. ความรู้ที่ชัดแจ้ง (Explicit Knowledge) เป็นความรู้ที่สามารถรวบรวม ถ่ายทอดได้ โดยผ่านวิธีต่าง ๆ เช่น การบันทึกเป็นลายลักษณ์อักษร ทฤษฎี คู่มือต่าง ๆ และบางครั้งเรียกว่าเป็นความรู้แบบรูปธรรม

#### องค์ประกอบสำคัญของการจัดการความรู้ (Knowledge Process)

1. คน ถือว่าเป็นองค์ประกอบที่สำคัญที่สุดเพราะเป็นแหล่งความรู้ และเป็นผู้นำความรู้ไปใช้ให้เกิดประโยชน์

2. เทคโนโลยี เป็นเครื่องมือเพื่อให้คนสามารถค้นหา จัดเก็บ แลกเปลี่ยน รวมทั้งนำความรู้ไปใช้ได้ง่าย และรวดเร็วขึ้น

3. กระบวนการของความรู้ เป็นการบริหารจัดการ เพื่อนำความรู้จากแหล่งความรู้ไปให้ผู้รู้ทำให้เกิดการปรับปรุง ต่อยอดความรู้ และเกิดนวัตกรรม

### กระบวนการจัดการความรู้ (Knowledge Management Process)

สำนักคณะกรรมการพัฒนาระบบราชการ [ก.พ.ร.] (2548) อธิบายความคิดรวบยอดของการจัดการความรู้ว่า “การจัดการความรู้ เป็นการบริหารจัดการเพื่อให้คนที่ต้องการใช้ความรู้ได้รับความรู้ที่ต้องการ ในเวลาที่ต้องการ เพื่อให้บรรลุเป้าหมายการทำงาน”

เป้าหมายของกระบวนการจัดการความรู้ เป็นหัวใจสำคัญในกระบวนการจัดการความรู้ที่จะเป็นกรอบกำหนดให้การจัดการความรู้ไม่หลงประเด็น ไม่เดินทางผิด ไปสู่จุดหมายปลายทางได้รวดเร็ว ทันเหตุการณ์ ทันเวลา สมองความต้องการได้เฉียบพลัน ดังนั้นการกำหนดเป้าหมายถือเป็นหัวใจในการบอกความสำเร็จในการจัดการความรู้ได้อย่างชัดเจนมีหลักการในการกำหนดเป้าหมาย เมื่อองค์กรมีความรู้ที่หลากหลาย เราจะสามารถค้นหาว่าในองค์กรเรามีความรู้อะไรบ้าง ความรู้เหล่านั้นอยู่ในรูปแบบใด ใครเป็นผู้รู้ และความรู้อะไรบ้างที่องค์กรมีความจำเป็นต้องใช้ ทั้งนี้ เพื่อช่วยให้ทราบว่า องค์กรขาดความรู้อะไรบ้าง โดยอาศัยแผนที่ความรู้ (Knowledge Maps) เป็นเครื่องมือที่จะช่วยให้รู้ว่ามีความรู้ใดที่สำคัญต่อองค์กร และจัดลำดับความสำคัญของความรู้เหล่านั้น ทั้งนี้ เพื่อเป็นข้อมูลให้ผู้บริหารใช้ในการวางแผนและกำหนดนโยบายในองค์กรเรื่องการจัดการความรู้ และบริหารทรัพยากรต่าง ๆ ให้เป็นไปอย่างมีประสิทธิภาพและประสิทธิผล

### แผนที่ความรู้ (Knowledge Maps)

แผนที่ความรู้ จึงเป็นเครื่องมือที่ช่วยให้เห็นภาพรวมของคลังความรู้ขององค์กร ทำให้ทราบว่าในองค์กรมีความรู้อะไรบ้าง ความรู้ต่าง ๆ อยู่ที่ไหน และความรู้เหล่านั้น มีส่วนที่ซ้ำซ้อนกับหน่วยงานต่าง ๆ หรือไม่ นอกจากนี้ ยังใช้เป็นโครงสร้างพื้นฐานทางความรู้ที่องค์กรสามารถใช้เป็นฐานในการต่อยอดขยายความรู้ในเรื่องต่าง ๆ ได้อย่างเป็นระบบ

แผนที่ความรู้ จึงเป็นเครื่องมือที่ช่วยในการจัดการความรู้ ประเภทความรู้ชัดเจน ความรู้ที่ฝังอยู่ในตัวบุคคล และเป็นเครื่องมือที่ช่วยเหลือในการเรียนรู้และการตัดสินใจ (Cognitive Tools) จัดเก็บหรือสืบค้นความรู้ในเชิงตำแหน่งของความรู้ ซึ่งจะช่วยส่งเสริมกระบวนการจัดการความรู้ (Tergan, 2003)

แผนที่ความรู้ช่วยให้บุคคล กลุ่มบุคคล องค์กร ใช้ชุดความรู้เดิมในการสังเคราะห์ความรู้ขึ้นมาใหม่ โดยผ่านกระบวนการระดมสมอง และแทนความรู้ (Knowledge Representations) จัดเก็บความรู้ การเพิ่มเติม สร้างความรู้ จากเครือข่ายด้วยระบบการเชื่อมต่อแผนที่ความรู้ ผ่านทางช่องทางสื่อสารความรู้ (Knowledge Communication) ทำให้เกิดการสังสมความรู้จากภาคส่วนต่างๆ การแลกเปลี่ยนความรู้ ผ่านช่องทางการสื่อสาร ช่วยส่งเสริมประสิทธิภาพการใช้ความรู้ เช่น การสืบค้น การติดตามความรู้ เป็นเครื่องมือที่ใช้ในการเรียนรู้ต่อยอดความรู้ วางแผนการจัดการความรู้ที่มีบทบาทสำคัญต่อองค์กร

### ประเภทของแผนที่ความรู้

แผนที่ความรู้ (Knowledge maps) จำแนกได้ 4 ระดับ (Plumley, 2003) โดยพิจารณาจากลำดับชั้นของความรู้และแหล่งความรู้

ระดับที่ 1 แผนที่ความรู้เชิงแนวคิด (Concept knowledge maps) การจำแนกความรู้เชิงแนวความคิดและเนื้อหาตามลำดับชั้น ในลักษณะที่เป็นอนุกรมวิธานความรู้ เพื่อใช้จัดการเชิงความรู้ เชิงเนื้อหา (Content management) ของบุคคลและองค์กร ว่ามีบริบทเชิงเนื้อหาที่กลุ่ม (Category) ที่ด้าน ความหลากหลาย ความกว้าง ลึกของความรู้

ระดับที่ 2 แผนที่ความรู้เชิงกระบวนการ (Procedural knowledge maps) ความรู้ หรือแหล่งของความรู้ ที่ถูกจัดเก็บในระหว่างกระบวนการทำงานหรือกิจกรรมในองค์กร หรือบุคคล เช่น กระบวนการวิจัยและพัฒนา

ระดับที่ 3 แผนที่ความรู้เชิงแหล่งความรู้ (Competency knowledge maps) ตำแหน่ง แหล่งของภูมิปัญญาในระดับบุคคลและองค์กร ที่เป็ความรู้ประเภท Explicit Knowledge และความรู้ประเภท Tacit Knowledge ที่ถูกจัดเก็บในฐานข้อมูล

ระดับที่ 4 แผนที่สังคมเครือข่ายความรู้ (Social network mapping or Social network analysis) การเชื่อมต่อแลกเปลี่ยนความรู้กับบุคคล หรือองค์กรอื่น ๆ มีการแบ่งปันความรู้ผ่านเครือข่ายที่สนใจความรู้ในประเด็นเดียวกัน ผ่านทางช่องทางการสื่อสาร ซึ่งจะก่อเกิดเป็นแผนที่ความรู้ที่แตกหน่อ และสะสมความรู้ได้ตลอดเวลาไม่สิ้นสุด

## การทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) คือ กระบวนการค้นหาความรู้ที่มีอยู่ในฐานข้อมูลขนาดใหญ่ โดยที่องค์กรต่าง ๆ มักจะมองข้ามความสำคัญของข้อมูลที่เก็บไว้ มองเพียงด้านเดียว คือ จะต้องเก็บข้อมูลให้ได้จำนวนมากไว้ก่อน มิได้คิดว่าจะนำความรู้ที่มีอยู่ในฐานข้อมูล ออกมาใช้งานให้เกิดประโยชน์ต่อองค์กร หรือไม่ทราบว่ามีความรู้ใดบ้างที่แฝงอยู่กับข้อมูลที่ได้จัดเก็บไว้ ดังนั้น การทำเหมืองข้อมูล จึงเป็นกระบวนการของการค้นกรองข้อมูลที่มีอยู่ให้เป็นความรู้ เพื่อนำไปใช้ทำนายแนวโน้ม พฤติกรรม หรือลักษณะที่น่าสนใจของข้อมูลเหล่านี้ เช่น รูปแบบ ความสัมพันธ์ ลักษณะการเปลี่ยนแปลง หรือลักษณะที่ผิดปกติของข้อมูล โดยอาศัยข้อมูลในอดีตมาใช้ในการศึกษาความเป็นไปได้ที่จะเกิดขึ้นในปัจจุบัน ซึ่งวิธีการต่าง ๆ ที่นำมาใช้ในการทำเหมืองข้อมูล จะมีวัตถุประสงค์ที่แตกต่างกันไป ทั้งนี้ขึ้นอยู่กับผลลัพธ์ที่ต้องการ ดังนั้นจึงควรนำเสนอวิธีการที่หลากหลายสำหรับเป้าหมายที่ต่างกัน เพื่อให้ได้ผลลัพธ์ที่เหมาะสมตามความต้องการ (Larose, 2005)

การทำเหมืองข้อมูลมีความหมายใน 2 มุมมอง โดยมุมมองแรก เป็นมุมมองทางวิชาการ นักวิจัยจะมองเป็นกระบวนการที่เกี่ยวข้องกับการค้นหารูปแบบความสัมพันธ์ของข้อมูลเท่านั้น ส่วนมุมมองที่สอง เป็นมุมมองในเชิงธุรกิจ ซึ่งจะแทนความหมายของเหมืองข้อมูลว่าเป็นขั้นตอนโดยทั่วไปมีขั้นตอนต่าง ๆ ดังนี้

1. การเตรียมข้อมูล (Data Preprocessing) ประกอบด้วย 3 ขั้นตอน (Fayyad, 1996) ได้แก่

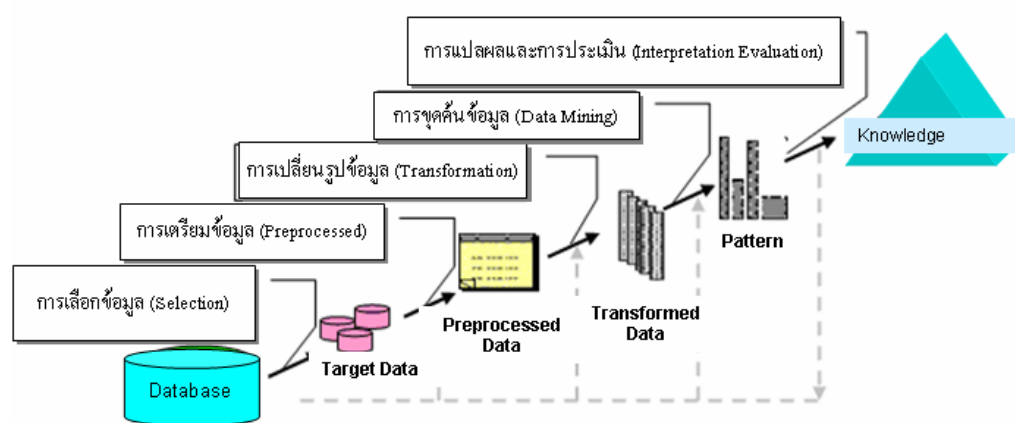
1.1 การคัดเลือกข้อมูล (Data Selection) เป็นการระบุแหล่งข้อมูลที่น่าไปใช้ทำเหมืองข้อมูล รวมถึงการคัดเลือกข้อมูลที่ต้องการจากฐานข้อมูล เพื่อนำไปพิจารณาในเบื้องต้น

1.2 การทำความสะอาดข้อมูล (Data Cleaning) เป็นกระบวนการที่ทำให้เกิดความมั่นใจในคุณภาพของข้อมูลที่จะนำไปวิเคราะห์ว่ามีความถูกต้อง โดยขั้นตอนนี้จะทำการลบข้อมูลที่ไม่เกี่ยวข้อง ข้อมูลซ้ำซ้อน ข้อมูลที่ไม่ถูกต้องออก

1.3 การแปลงข้อมูล (Data transformation) การแปลงข้อมูล และการจัดข้อมูลให้อยู่ในรูปแบบที่เหมาะสม สำหรับการนำไปใช้วิเคราะห์ตามอัลกอริทึมและแบบจำลองที่ใช้ในการทำเหมืองข้อมูลต่อไป

2. การทำเหมืองข้อมูล ขั้นตอนนี้จะนำเทคนิคต่าง ๆ มาใช้ในการดึงข้อมูลที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่

3. การแปลผลและการประเมินผล เป็นการแปลผลที่ได้ในขั้นตอนการทำเหมืองข้อมูลให้อยู่ในรูปของความรู้ ที่สามารถนำไปใช้ประโยชน์ในการวิเคราะห์หาความสัมพันธ์ของข้อมูล และเรื่อง ๆ ที่ต้องการได้ และมีการประเมินผลลัพธ์ที่ได้ว่ามีความเหมาะสม หรือตรงกับวัตถุประสงค์ที่ต้องการหรือไม่

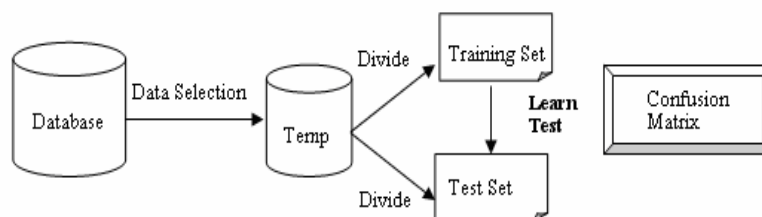


ภาพที่ 1 แสดงกระบวนการค้นหาความรู้ในการทำเหมืองข้อมูล

ที่มา: Fayyad *et al.* (1996)

### ขั้นตอนการทำเหมืองข้อมูล

การทำเหมืองข้อมูลมีหลายเทคนิค ผู้วิจัยได้เลือกใช้เทคนิคการจำแนกข้อมูล (Classification) เป็นกระบวนการสร้างโมเดลจัดการข้อมูล คือ การแยกประเภทข้อมูล หรือการทำนายประเภทของข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ โดยการสร้างกฎเพื่อช่วยในการตัดสินใจจากข้อมูลที่มีอยู่ เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของข้อมูลที่ยังไม่เคยเห็นมาก่อน (Unseen Data)



## ภาพที่ 2 การทำเหมืองข้อมูล

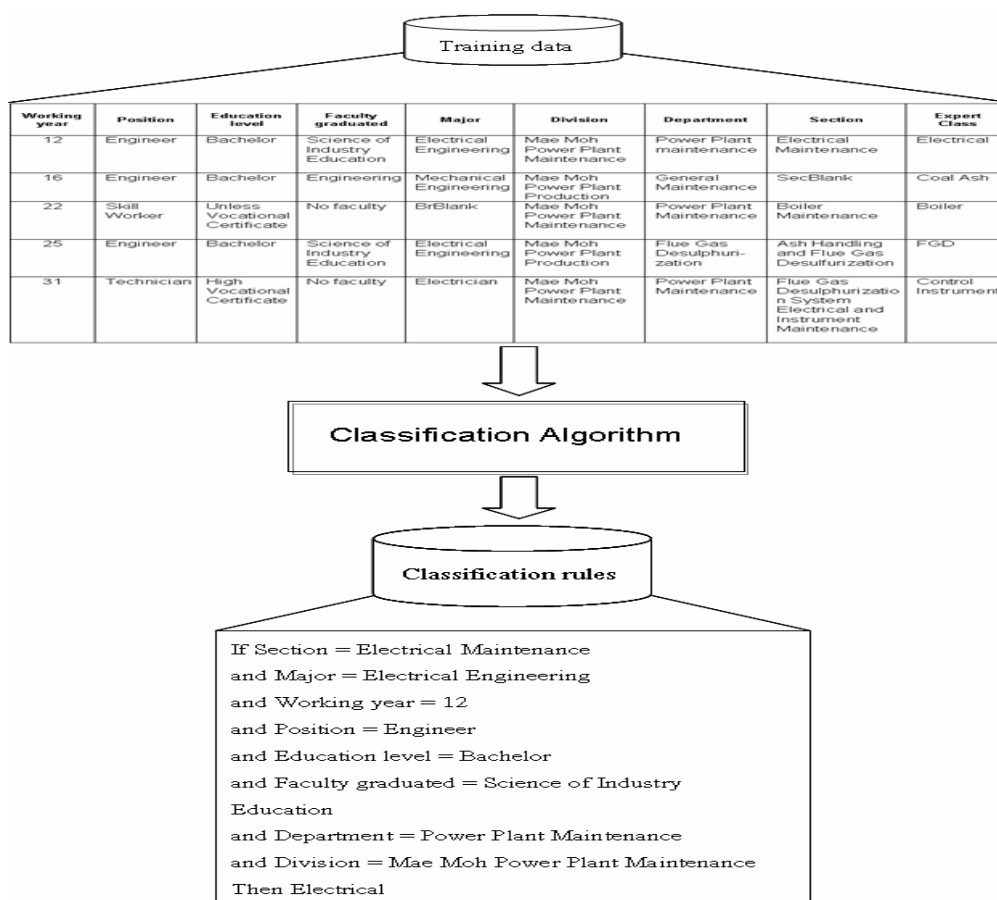
ที่มา: Kantardzic (2003)

ข้อมูลที่บรรจุในฐานข้อมูลได้รับการ Preprocess เพื่อเตรียมข้อมูลและดึงข้อมูลที่ต้องการใช้งานออกมา โดยแบ่งข้อมูลออกเป็น 2 ส่วน ได้แก่ ข้อมูลสอนและข้อมูลทดสอบ

ส่วนที่ 1 เป็นข้อมูลที่เรียกว่า ข้อมูลสอนเพื่อให้ระบบเกิดการเรียนรู้ว่ามีข้อมูลใดควรอยู่ในคลาสเดียวกันบ้างโดยสร้างเป็น โมเดล เมื่อได้โมเดลออกมาแล้ว ก็นำข้อมูลส่วนที่ 2 มาใช้สำหรับทดสอบ เพื่อนำมาตรวจสอบความถูกต้องของโมเดลว่าผลของการแยกประเภทออกมาถูกต้องหรือไม่ และจะปรับปรุงโมเดลจนกว่าจะได้รับความถูกต้องในระดับที่น่าพอใจ จากนั้นจะนำโมเดลที่ได้ไปพิจารณาข้อมูลที่ยังไม่ได้จำแนก เพื่อทำนายกลุ่มของข้อมูล

โมเดลที่ถูกปรับปรุงและตรวจสอบความถูกต้องแล้ว เรียกว่าโมเดลจำแนกประเภทข้อมูล (Classifier Model) การนำเสนอกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล ผู้วิจัยได้นำเสนอในรูปแบบของแผนภูมิต้นไม้ซึ่งเรียกว่า ต้นไม้การตัดสินใจ (decision tree) เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล ต้นไม้การตัดสินใจจะมีลักษณะคล้ายโครงสร้างต้นไม้ที่แต่ละโนดแสดงคุณลักษณะ (Attribute) แต่ละกิ่งแสดงเงื่อนไขในการทดสอบ และโนดปลาย (leaf node) แสดงกลุ่มที่กำหนดไว้ แนวทางการจำแนกประเภทข้อมูลโดยใช้วิธีการสร้างต้นไม้การตัดสินใจ เป็นวิธีการจำแนกประเภทข้อมูลที่ได้รับความนิยมอย่างแพร่หลาย ประสิทธิภาพของวิธีการจำแนกประเภทข้อมูลแบบนี้ได้รับการพัฒนาอย่างต่อเนื่องโดยเฉพาะอย่างยิ่งในเรื่องของความถูกต้องในการจำแนกประเภทข้อมูลและความเร็วในการจำแนกประเภทข้อมูล โดยได้แบ่งกระบวนการจำแนกออกเป็น 3 ขั้นตอน ซึ่งเป็นโมเดลมาตรฐานของการจำแนกข้อมูล ดังนี้ (Kantardzic, 2003)

1. Model construction (Learning) เป็นขั้นตอนการสร้างโมเดลโดยเรียนรู้จากข้อมูลที่ได้กำหนดคลาสไว้เรียบร้อยแล้ว (training data) ซึ่งโมเดลที่ได้อาจแสดงในรูปของต้นไม้การตัดสินใจที่ประกอบด้วยโนดราก กิ่งและโนดใบ แต่ละโนดจะแสดงค่าของคุณลักษณะข้อมูล (Attribute) และโนดใบ จะแสดงคลาสที่ทำนายได้

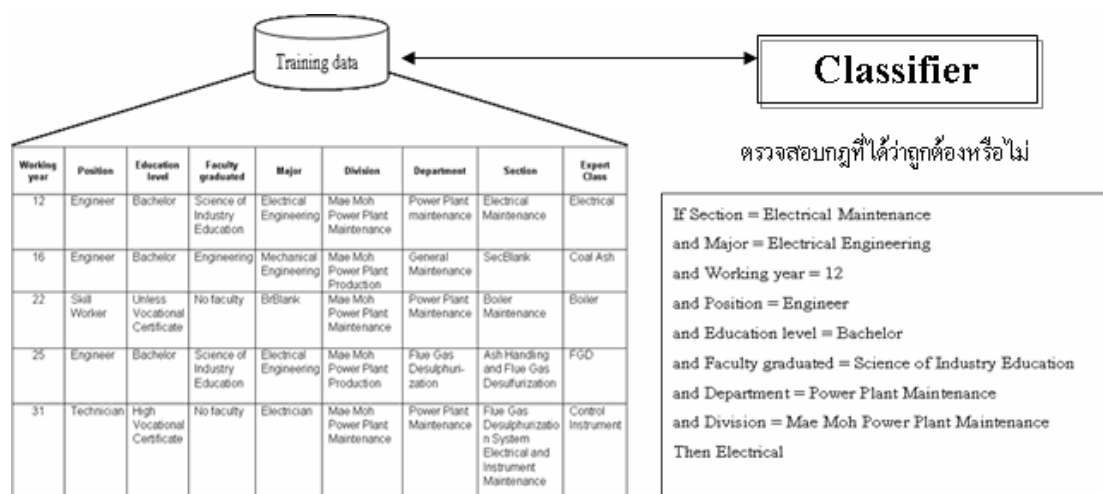


ภาพที่ 3 Classification Model ด้วยวิธี Decision Rules

ที่มา: Kantardzic (2003)

จากภาพที่ 3 ข้อมูลที่เป็นข้อมูลสอน จะถูกวิเคราะห์ด้วยอัลกอริทึมของการจำแนก จากนั้น โมเดลที่ได้เรียนรู้แล้วจะถูกแทนในรูปของกฎการจำแนก ซึ่งกฎที่ได้มาจะสามารถเรียนรู้ว่า ผู้ปฏิบัติงาน เชี่ยวชาญงานประเภทใด

2. Model evaluation (Accuracy) เป็นขั้นตอนการตรวจสอบความถูกต้องโดยจะมีข้อมูลที่ใช้สำหรับทดสอบ เรียกว่า Testing Data เป็นข้อมูลที่เราทราบอยู่แล้วว่าข้อมูลนั้นอยู่ในกลุ่มใด แล้วนำไปเปรียบเทียบกับ Learning Model เพื่อทดสอบว่าโมเดลที่สร้างไว้ สามารถจัดกลุ่มประเภทข้อมูลได้อย่างถูกต้องหรือไม่ โดยจะนำคลาสที่แท้จริงของข้อมูลเปรียบเทียบกับคลาสที่ได้จากโมเดล



ภาพที่ 4 ขั้นตอนการตรวจสอบความถูกต้องของโมเดล

ที่มา: Kantardzic (2003)

3. Model usage (Classification) เป็นขั้นตอนการนำโมเดลที่สร้างจากขั้นตอนแรกมาใช้กับข้อมูลที่ไม่เคยเห็นมาก่อน โดยจะใช้โมเดลทำการกำหนดคลาสให้กับข้อมูลที่ได้มา หรือทำนายค่าของข้อมูลนั้น ว่าควรอยู่คลาสใด

งานวิจัยฉบับนี้ ผู้วิจัยได้สร้าง Classifier Model แบบ Multiple Class โดยคลาสที่ได้จากการทำนายจะเป็นแบบ Multiple Class ตามประเภทความเชี่ยวชาญของผู้ปฏิบัติงานที่มีอยู่ 10 ประเภท

ชนิดของตัวแปรในการทำเหมืองข้อมูล

ในการทำเหมืองข้อมูลได้แบ่งตัวแปรออกเป็นแบบ Categorical และแบบ Quantitative

### ตัวแปรแบบ Categorical

1. Nominal Variable เป็นตัวแปรที่มีได้หลายค่า ไม่มีการอ้างอิงถึงลำดับของข้อมูล เช่น เพศ (ชาย, หญิง)
2. Ordinal Variable เป็นตัวแปรที่มีลำดับสำหรับค่าที่เป็นไปได้ เช่น ลำดับของ ลูกค้า (ดี, ปานกลาง, ไม่ดี) ระดับวุฒิการศึกษา (ปริญญาโท, ปริญญาตรี, ปวส., ปวช., ต่ำกว่า ปวช.)

### ตัวแปรแบบ Quantitative

1. Continuous เป็นตัวแปรที่มีค่าที่ต่อเนื่อง เช่น อายุงาน, รายได้
2. Discrete ตัวแปรที่มีค่าเป็นจำนวนเต็ม เช่น จำนวนพนักงาน

ตัวแปรที่ถูกเลือกสำหรับทำเหมืองข้อมูล จะถูกเรียกว่า “Active Variable” เพราะตัวแปรนั้นจะถูกใช้สร้างความแตกต่างของกลุ่มย่อยต่างๆ และสามารถนำมาใช้ในการทำนายผลได้ เพื่อให้มั่นใจว่าคุณภาพของข้อมูลที่ถูกเลือกนั้นเหมาะสม ข้อมูลมีคุณภาพ และมีความถูกต้อง ซึ่งสามารถวัดคุณภาพได้ด้วยวิธีทางสถิติ หรือสุ่มตัวอย่าง

ปัญหาที่มักพบในขั้นตอนการเตรียมข้อมูล ได้แก่ Noisy Data คือตัวแปรที่มีค่าเกินกว่าค่าที่เราคาดหวังไว้ อาจเกิดจากข้อมูลที่ไม่สมบูรณ์ จากความเผลอของมนุษย์ ตัวอย่างเช่น พนักงานป้อนข้อมูล ใส่อายุงานให้คนเป็น 300 ปี ค่าเหล่านี้ควรจะถูกแก้ไขให้ถูกต้อง หรือลบออกจากข้อมูลสำหรับการวิเคราะห์ ดังนั้น จึงกำหนดให้มีการตรวจสอบข้อมูลก่อนนำมาใช้

ค่าที่หายไป (Missing Value) คือค่าที่ไม่ได้แสดงในข้อมูลที่เราได้เลือกขึ้นมาแสดง หรือค่าที่ไม่สมบูรณ์ที่เราลบออกไป ระหว่างการทำ Noise Detection ค่าอาจจะหายไปเพราะเกิดจากความเผลอของมนุษย์ เพราะว่าไม่ได้ป้อนข้อมูล

## ทฤษฎีการจำแนกประเภท (Classification Theory)

การจำแนกข้อมูล คือ กระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ เพื่อแสดงให้เห็นความแตกต่างระหว่างคลาส หรือกลุ่มของข้อมูล เพื่อใช้ในการทำนายว่าข้อมูลนี้ควรจัดอยู่ในคลาสใด ซึ่งโมเดลที่ใช้จำแนกข้อมูลออกเป็นกลุ่มตามที่ได้กำหนดไว้ จะขึ้นอยู่กับกระบวนการหาค่าของข้อมูลทดลอง โดยใช้วิธีการนำชุดข้อมูลสอนมาให้ระบบเรียนรู้ ได้รับรู้ว่ามีข้อมูลใดอยู่ในคลาสเดียวกันบ้าง ผลลัพธ์ที่ได้จากการเรียนรู้ คือ โมเดลจำแนกประเภทข้อมูล (classifier model) โมเดลนี้ สามารถแทนได้ในหลายรูปแบบ เช่น ต้นไม้การตัดสินใจ กฎการตัดสินใจ (IF-THEN) สูตรทางคณิตศาสตร์ หรือโครงข่ายประสาทเทียม และจะนำข้อมูลส่วนที่เหลือจากข้อมูลที่ใช้ฝึกสอน เป็นข้อมูลที่ใช้ทดสอบ (Testing Data) กับกลุ่มที่หามาได้จากโมเดล เพื่อทดสอบความถูกต้องโดยเราจะปรับปรุงโมเดลจนกว่าจะได้ค่าความถูกต้องในระดับที่น่าพอใจ หลังจากนั้นเมื่อมีข้อมูลใหม่เข้ามา เราจะนำข้อมูลผ่านโมเดล โดยโมเดลจะสามารถทำนายกลุ่มของข้อมูลนี้ได้ (Tan *et al.*, 2006)

ข้อดีของการจำแนกประเภทข้อมูล (Classification) คือ เหมาะสำหรับการทำนายข้อมูลที่มีประเภทของข้อมูลน้อย และจะมีประสิทธิภาพลดลง เมื่อค่าของข้อมูลมีค่าที่ต่อเนื่อง ดังนั้น จึงต้องนำข้อมูลมาทำการแบ่งชั้น (Discretizing) โดยการแปลงตัวแปรเชิงปริมาณ (Quantitative) ให้เป็นแบบ Categorical โดยการแบ่งค่าของตัวแปรที่จะเป็นอินพุตให้เป็นช่วง ๆ เช่น การแปลงอายุงาน และอายุความเชี่ยวชาญ เป็นต้น

## อัลกอริทึมที่ใช้ในการจำแนกประเภทข้อมูล

เป็นลักษณะของการเรียนรู้ที่ต้องมีการสอน (Witten and Frank, 2005) ตัวอย่างของอัลกอริทึมนี้ได้แก่

1. ต้นไม้การตัดสินใจ (Decision tree) เป็นวิธีหนึ่งในเทคนิคการจำแนก โดยต้นไม้การตัดสินใจจะมีลักษณะคล้ายโครงสร้างต้นไม้ที่แต่ละโนดแสดงคุณลักษณะ (attribute) แต่ละกิ่งแสดงเงื่อนไขในการทดสอบ และโนดใบ แสดงกลุ่มที่กำหนดไว้

2. Decision rule กฎการตัดสินใจ สามารถสร้างขึ้นจากต้นไม้การตัดสินใจได้ง่าย โดยจะ

อยู่ในรูปของกฎ IF ... THEN โดยที่ IF {ส่วนของการกำหนดเงื่อนไข} THEN {ส่วนของการกระทำตามเงื่อนไข}

3. K-Nearest Neighbors เป็นเทคนิคที่เหมาะสมกับปัญหาการจำแนกประเภทข้อมูล โดยเทคนิคนี้จะแตกต่างจากเทคนิคอื่นที่ไม่ใช้ข้อมูลสอน ในการสร้างแบบจำลองแต่จะใช้ข้อมูลนั้นมาเป็นตัวแบบจำลอง ในการใช้งาน K-nearest neighbor (K-NN) นั้น จะต้องระบุค่าตัวเลขจำนวนเต็มบวกให้กับ k ด้วย โดยค่า k จะเป็นตัวบอกจำนวนของกรณี (case) ที่จะต้องค้นหาในการทำนายกรณีใหม่

4. Naïve Bayes เป็นอัลกอริทึมที่ใช้ทฤษฎี Bayes Theorem ในการคำนวณความน่าจะเป็นที่ใช้ในการแก้ปัญหาเรื่องการจำแนกประเภทข้อมูลที่ใช้แบบจำลองของการทำนาย โดยจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตาม

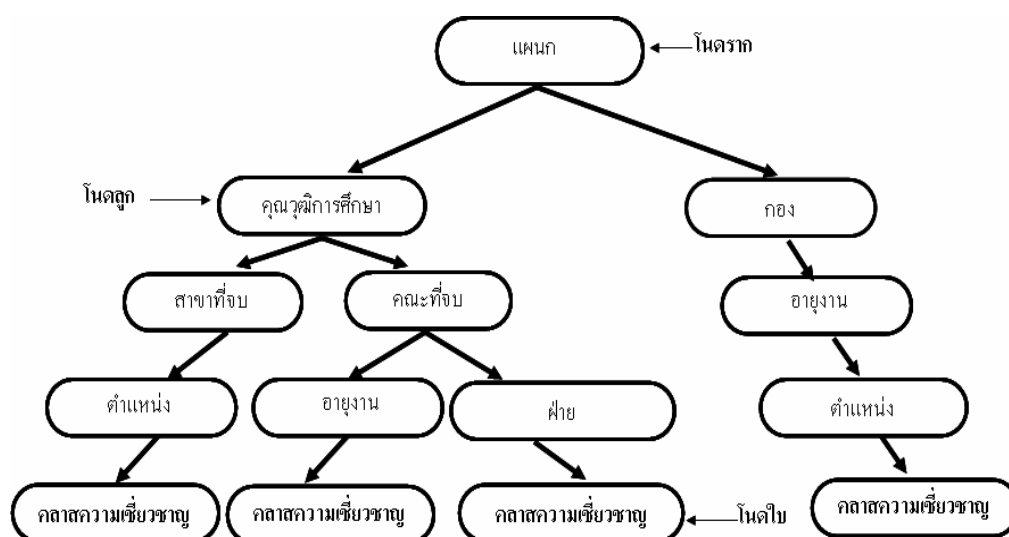
5. Bayesian Belief Network ข่ายงานความเชื่อเบย์ เรียกโดยย่อว่า ข่ายงานเบย์ (BayesNet) เป็นวิธีการเรียนรู้ที่ลดข้อจำกัดของการเรียนรู้แบบง่าย ในสมมติฐานของความไม่ขึ้นต่อกัน โดยใช้ข่ายงานความเชื่อเบย์ในการอธิบายความไม่ขึ้นต่อกันแบบมีเงื่อนไข (Condition Independence) ระหว่างตัวแปร

### ต้นไม้การตัดสินใจ (Decision Tree)

ต้นไม้การตัดสินใจนับว่าเป็นวิธีการเรียนรู้ที่ง่ายที่สุดแบบหนึ่งในการเรียนรู้ โดยการเรียนรู้แบบนี้เป็นการเรียนรู้โดยการจำแนกข้อมูลออกเป็นชั้น (class) ต่าง ๆ โดยใช้คุณสมบัติของข้อมูลในการจำแนกต้นไม้การตัดสินใจที่ได้จากการเรียนรู้ ทำให้ทราบว่าคุณสมบัติใดของข้อมูลที่เป็นตัวกำหนดการจำแนก โดยต้นไม้ที่ช่วยในการตัดสินใจจะมีลักษณะโครงสร้างคล้ายกับต้นไม้ที่แต่ละโนดจะแสดงคุณสมบัติที่ใช้ในการทดสอบ แต่ละกิ่งจะแสดงผลของการทดสอบ และใบของต้นไม้ (Leaves Node) จะแสดงคลาสหรือการแบ่งคลาส (ต้นไม้การตัดสินใจช่วยในการแปลงเป็นกฎในการแบ่งคลาส) ซึ่งการจำแนกประเภทใช้ในการทำนายประเภทของข้อมูล หรือข้อมูลที่ไม่มีทราบค่าซึ่งปกติแล้วค่าที่ทำนายได้จะเป็นลักษณะของตัวเลขหรือการทำนายที่เป็นจำเพาะ (บุญเสริม, 2546)

ต้นไม้การตัดสินใจ มีการทำงานแบบการเรียนรู้ที่ต้องมีการสอน คือ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่ได้กำหนดไว้ล่วงหน้า ที่เรียกว่า ชุดของข้อมูลที่ใช้สอน โดยอัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ที่เรียกว่า ชุดข้อมูลทดสอบ ได้ด้วย

รูปแบบของต้นไม้ จะประกอบด้วยโนดแรกสุดที่เรียกว่า โนดราก จากโนดราก จะแตกออกเป็นโนดลูก และที่โนดลูกก็จะมีลูกของตัวเองซึ่งโนดที่ระดับสุดท้ายจะเรียกว่าโนดใบ



ภาพที่ 5 การจำแนกประเภทผู้เชี่ยวชาญด้วยต้นไม้การตัดสินใจ

จะเห็นว่า จากโนดราก จนถึงโนดใบ จะมีหลายเส้นทาง ซึ่งเส้นทางนี้จะอธิบาย ถึงกฎที่ใช้สำหรับการจัดหมวดหมู่ของแต่ละคลาส ซึ่งในแต่ละโนดใบนั้นอาจเป็นคลาสเดียวกัน ซึ่งเกิดจากเหตุผลที่แตกต่างกันได้

## วิธีการเรียนรู้ของต้นไม้การตัดสินใจ (Decision Tree Learning)

การสร้างต้นไม้การตัดสินใจจะเป็นแบบการค้นหาลงล่างแบบละโมภหรือตะกราม (top-down greedy search) โดยเริ่มจากการเลือกคุณสมบัติที่ดีที่สุดมาสร้างเป็นโนดแรก เมื่อข้อมูลผ่านการแบ่งแยกที่โนดแรกตามค่าคุณสมบัติของโนดแรกแล้ว ก็จะหาคุณสมบัติที่ดีที่สุดของข้อมูลผ่านการแบ่งแยกนั้น มาสร้างเป็นโนดลูกของโนดแรกนั้นต่อไป และจะวนสร้างโนดลูกและต้นไม้ย่อยของแต่ละกิ่งไปเรื่อย ๆ จนกว่าข้อมูลที่ผ่านการแบ่งแยกนั้นจะจัดอยู่ในกลุ่มเดียวกัน หรือจำนวนข้อมูลที่ผ่านการแบ่งแยกในกิ่งหนึ่ง ๆ มีค่าน้อยกว่าค่าที่กำหนดไว้ (บุญเสริม, 2546)

## ลักษณะการเรียนรู้ของต้นไม้การตัดสินใจ

1. ผลการเรียนรู้แสดงอยู่ในรูปที่เข้าใจง่าย ทำให้ง่ายต่อการวิเคราะห์คุณสมบัติที่มีผลต่อการแยกแยะกลุ่มต่าง ๆ
2. แต่ละเส้นทางจากโนดแรกถึงโนดใบสามารถแสดงให้อยู่ในรูปกฎ IF-THEN ได้
3. มีความทนทานต่อข้อมูลมีสัญญาณรบกวน (noisy data) เช่น คุณสมบัติที่ไม่เกี่ยวข้องและค่าคุณสมบัติที่ผิดพลาดหรือขาดหาย
4. การเรียนรู้มีความรวดเร็ว เมื่อเทียบกับอัลกอริทึมสำหรับจำแนกชนิดอื่น
5. นำมาใช้ในการวิเคราะห์ความเสี่ยงของลูกค้า การวินิจฉัยทางการแพทย์ การวิเคราะห์กลุ่มดาว และงานทางด้านธุรกิจและวิทยาศาสตร์อื่น ๆ

## การแทนต้นไม้การตัดสินใจ (Decision Tree Representation)

1. โนดภายใน (internal node) คือ คุณสมบัติต่าง ๆ ของข้อมูล ซึ่งเมื่อข้อมูลใด ๆ ตกลงมาที่โนด จะใช้คุณสมบัตินี้เป็นตัวตัดสินใจว่าข้อมูลจะไปในทิศทางใด โดยโนดภายในที่เป็นจุดเริ่มต้นของต้นไม้เรียกว่าโนดแรก

2. กิ่ง (branch, link) เป็นค่าคุณสมบัติของคุณสมบัติในโนดภายในที่แตกกิ่งนี้ออกมา ซึ่งโนดภายในจะแตกกิ่งเป็นจำนวนเท่ากับจำนวนค่าคุณสมบัติของโนดภายในนั้น

3. โนดใบ (leaf node) คือ กลุ่มต่าง ๆ ซึ่งเป็นผลลัพธ์ในการแยกแยะข้อมูล

### วิธีการสร้างต้นไม้การตัดสินใจ

1. หากคุณลักษณะข้อมูล ที่สำคัญที่สุดมาแบ่งข้อมูลโดยคุณลักษณะข้อมูลนี้จะถูกนำมาสร้างเป็นโนดราก โดยจะมีคุณลักษณะข้อมูลที่เป็นเป้าหมาย เป็นผลลัพธ์ซึ่งเป็นโนดใบ ถูกกำหนดไว้ก่อน

2. นำค่าที่เป็นไปได้ในคุณลักษณะข้อมูลที่ถูกเลือกมาแตกออกเป็นกลุ่มของตัวเอง

3. แบ่งข้อมูลทั้งหมดตามกลุ่มที่แตกออกจากโนดราก

4. วนกลับไปทำที่ขั้นตอนแรก คือ หากคุณลักษณะของข้อมูล ที่สำคัญที่สุดจากข้อมูลที่เข้ามาเพื่อหาตัวแบ่งต่อไป

### ข้อจำกัดของต้นไม้การตัดสินใจ

1. การจำแนกข้อมูลโดยใช้ต้นไม้การตัดสินใจ กรณีของข้อมูลที่มีค่าต่อเนื่อง เช่น ข้อมูลอายุ ข้อมูลรายได้ ข้อมูลราคา ต้องทำการแปลงให้อยู่ในช่วงหรือตัดเป็นกลุ่มก่อน (Discrete)

2. เมื่ออัลกอริทึมเลือกจะใช้ค่าไหนเป็นตัวแบ่งกลุ่มแล้วก็จะไม่สนใจค่าอื่นที่อาจมีความสำคัญเช่นเดียวกัน

3. การจัดการกับข้อมูลที่ไม่ทราบค่า อาจมีผลกระทบกับความแม่นยำในการทำนายของต้นไม้การตัดสินใจ

3.1 ต้นไม้ที่มีระดับชั้นมากเกินไป จะทำให้ข้อมูลที่ผ่านโนด แยกออกเป็นชั้นเล็กชั้นน้อย ซึ่งข้อมูลเหล่านั้น จะไม่มีประโยชน์ในการนำมาทำการวิเคราะห์

3.2 ปัญหา Overfitting / Overtraining เกิดจากการที่แบบจำลองได้เรียนรู้ข้อมูลที่ใช้ในการสอนมากเกินไป จึงทำให้แบบจำลองนั้น จำข้อมูลที่สอนจนเกิดเป็นรูปแบบเฉพาะของข้อมูล ซึ่งโนดของต้นไม้การตัดสินใจที่ได้จะมีลักษณะเฉพาะตามข้อมูลที่สอน ดังนั้น เมื่อเรานำข้อมูลทดสอบเข้าสู่แบบจำลอง จึงทำให้แบบจำลองนั้น ทำนายค่าได้ไม่แม่นยำ วิธีการแก้ปัญหา คือ มีการตัดกิ่งของต้นไม้การตัดสินใจที่เกิดปัญหา Overfitting ออกไป

### อัลกอริทึม ID3

เป็นอัลกอริทึมในการสร้างต้นไม้การตัดสินใจ ที่ใช้หลักการของทฤษฎีข่าวสาร ค่าที่วัดได้จะนำมาใช้ในการตัดสินใจว่าจะใช้ตัวแปรใดเป็นตัวแบ่งข้อมูล โดยวิธีการกำหนดโครงสร้างต้นไม้การตัดสินใจจะเป็นลักษณะของการเลือกข้อมูลตามลำดับของตัวชี้วัดหรือค่าเกน (Gain) สูงที่สุดเป็นข้อมูลเริ่มต้น และข้อมูลถัดไปที่มีค่าลดหลั่นกันตามลำดับ โดยมีสูตรที่เกี่ยวข้องกับการสร้างต้นไม้การตัดสินใจ (Mitchell, 1997) ดังนี้

$$Entropy(S) \equiv -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus} \quad (1)$$

โดยที่

$S$  คือ จำนวนข้อมูลทั้งหมด

$p_{\oplus}$  คือ ข้อมูลที่มีค่าเป็นบวก

$p_{\ominus}$  คือ ข้อมูลที่มีค่าเป็นลบ

$Entropy(S_v)$  คือ จำนวนค่าของข้อมูลที่เป็นบวก

$$Gain(S,A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

โดยที่

$\sum_{v \in Values(A)}$  คือ ผลรวมของข้อมูลบวกและข้อมูลลบ

|         |                                                      |
|---------|------------------------------------------------------|
| $ S_v $ | คือ จำนวนค่าของข้อมูลหนึ่ง ๆ ที่มีค่าบวก รวมกับค่าลบ |
| $ S $   | คือ จำนวนข้อมูลทั้งหมด                               |

#### อัลกอริทึม C4.5

อัลกอริทึม C4.5 เป็นอัลกอริทึมในการจำแนกประเภทข้อมูล ได้พัฒนาต่อจากอัลกอริทึม ID3 ที่ใช้สำหรับการสร้างต้นไม้การตัดสินใจ นำเสนอโดย Ross Quinlan ในปี 1993 อัลกอริทึมนี้จะเยี่ยมแต่ละโนดแบบเรียกซ้ำ สามารถสร้างต้นไม้หลายรูปแบบโดยใช้แนวคิดของ Information Gain หรือ Entropy ในการเลือกโนดที่ดีที่สุดมาเป็นโนดเริ่มต้น นอกจากนี้ยังสามารถหลีกเลี่ยงการสร้างโครงสร้างต้นไม้ที่ใหญ่เกินไปได้ด้วย

#### จุดเด่นของอัลกอริทึม C4.5

1. ใช้หลักการทำงานแบบละโมภหรือตะกราม (Greedy approach) พื้นฐานของการสร้างต้นไม้การตัดสินใจ คืออัลกอริทึม greedy โดยจะสร้างต้นไม้จากบนลงล่างแบบเรียกซ้ำ (recursive) ด้วยวิธีการแบ่งแยกและเอาชนะ (divide-and-conquer) โดยแบ่งปัญหาใหญ่เป็นปัญหาย่อย
2. ใช้กับคุณลักษณะข้อมูลที่มีค่าไม่ต่อเนื่อง (Discrete)
3. คุณลักษณะข้อมูลถูกเลือกจากลำดับของตัววัด Gain ratio
4. ใช้เวลาในการเรียนรู้น้อยกว่าวิธีการอื่น
5. ผลลัพธ์ที่ได้เขียนเป็นกฎที่เข้าใจได้ง่าย

#### ข้อจำกัดของอัลกอริทึม C4.5

1. ไม่รองรับข้อมูลแบบต่อเนื่อง (Continuous data) จะต้องมีการแบ่งให้เป็นข้อมูลแบบไม่ต่อเนื่อง (discrete data) ก่อน

**อัลกอริทึมของ C4.5**

Pseudocode of the C4.5 Algorithm

FormTree(T)

- (1) ComputeClassFrequency(T);
- (2) if OneClass or FewCases  
    return a leaf;
- Create a decision node N;
- (3) ForEach Attribute A  
    ComputeGain (A);
- (4) N.test = AttributeWithBestGain;
- (5) if N.test is continuous  
    find Threshold;
- (6) ForEach T' in the splitting of T
- (7) if T' is Empty  
    Child of N is a leaf
- else
- (8) Child of N = FormTree (T');
- (9) ComputeErrors of N;
- return N

ภาพที่ 6 อัลกอริทึมต้นไม้การตัดสินใจ

ที่มา: Ruggieri (2002)

**กฎการตัดสินใจ (Decision Rules)**

กฎการตัดสินใจ จะนำเสนอรูปของ Production Rules ในลักษณะของ IF ... Then ที่ใช้ในการแก้ปัญหา สามารถสร้างกฎได้ง่ายและเป็นที่เข้าใจของผู้นำกฎนี้ไปใช้งาน ส่วนประกอบของกฎมีด้วยกัน 2 ส่วน ได้แก่ IF ที่จะเรียกว่า Antecedent เป็นส่วนของการใส่เงื่อนไขต่าง ๆ สำหรับ Then จะเรียกว่า Consequent ซึ่งจะเป็นส่วนของการกระทำตามเงื่อนไขที่กำหนด หรืออาจเป็นส่วนของการแสดงบทสรุปว่าจะเป็นอะไร หรือจะได้อะไร เป็นต้น

รูปแบบของกฎประกอบด้วย

**IF**      <antecedent>    **THEN**    <consequent>

ภาพที่ 7 รูปแบบของกฎการตัดสินใจ

เราสามารถเชื่อมกฎที่สร้างขึ้นด้วยคำเชื่อม AND , OR ซึ่งจะเรียกว่า Production rules งานวิจัยฉบับนี้ได้เลือกอัลกอริทึมที่ใช้ในการสร้างกฎ ได้แก่ OneR อัลกอริทึม

### อัลกอริทึม OneR

OneR เป็นอัลกอริทึมที่ง่ายที่สุดของอัลกอริทึมที่ใช้ในการจำแนกข้อมูล นำเสนอโดย Holte ในปี 1993 OneR อัลกอริทึมจะสร้างกฎให้กับแต่ละคุณลักษณะในช่วงของการใส่ข้อมูลสอน ซึ่งกฎที่ได้จะเป็นต้นไม้มาก่อนการตัดสินใจที่มีระดับเดียว แต่ละคุณลักษณะจะมีการกำหนดค่าไว้ก่อน โดยกฎที่เลือกใช้นั้นจะมีค่าผิดพลาดน้อยที่สุด

**ขั้นตอนวิธีของ OneR อัลกอริทึมที่จะนำมาสร้างกฎ สำหรับแต่ละคุณลักษณะข้อมูล**

#### Algorithm OneR

For each attribute,

{

For each value of that attribute, make a rule as follows:

{

count how often each class appears

find the most frequent class

make the rule assign that class to this attribute-value

}

Calculate the error rate of the rules

}

Choose the rules with the smallest error rate

**ภาพที่ 8** ขั้นตอนการทำงานของอัลกอริทึม OneR

**ที่มา:** Witten and Frank (1999)

### หลักการสร้างกฎของอัลกอริทึม OneR ที่นำมาใช้ในการสร้างต้นไม้การตัดสินใจ

1. การสร้างกฎจาก บน-ล่าง เริ่มต้นจากการเลือกกฎทั่ว ๆ ไป ที่มีจำนวนมากที่สุดมาเป็นจุดเริ่มต้นของโนด จากนั้นให้เพิ่มกฎทีละกฎเข้าไป
2. การสร้างกฎจาก ล่าง - บน เริ่มต้นจากการเลือกกฎที่มีค่าเฉพาะเจาะจงมากที่สุดค่อย ๆ ลดกฎทีละกฎลงมา
3. แบบผสมร่วมกันระหว่างการสร้างกฎจากบน-ล่าง และจาก ล่าง-บน ใช้หลักการของ Candidate-elimination อัลกอริทึม กล่าวคือ จะทำการตัดกฎที่มีค่าผิดพลาดสูง ๆ ออก ให้เลือกเฉพาะกฎที่ต้องใช้งานจริง ๆ ที่มีค่าความถูกต้องสูงไว้

#### ข้อเด่นของอัลกอริทึม OneR

1. กฎที่เลือกมาใช้งานจะมีความถูกต้องสูงสุดจากบรรดากฎทั้งหลาย
2. มีอัตราความผิดพลาดต่ำ

### การเรียนรู้แบบเบย์

การเรียนรู้แบบเบย์ (Bayesian learning) เป็นวิธีการเรียนรู้ที่ใช้ทฤษฎีความน่าจะเป็น ซึ่งมีพื้นฐานมาจากทฤษฎีของเบย์ (Bayes theorem) เข้ามาช่วยในการเรียนรู้ จุดมุ่งหมายเพื่อต้องการสร้างโมเดลที่อยู่ในรูปของความน่าจะเป็น โดยสามารถใช้ข้อมูล และความรู้ก่อนหน้า (prior knowledge) เข้ามาช่วยในการเรียนรู้ ซึ่งความรู้ก่อนหน้า หมายถึง ความรู้ที่เรามีเกี่ยวกับสมมติฐานแต่ละตัวก่อนที่จะเก็บข้อมูล เมื่อใช้งานเราจะนำความน่าจะเป็นของข้อมูลที่เก็บได้มาปรับสมมติฐานซ้ำอีกครั้ง (บุญเสริม, 2546)

#### ทฤษฎีของเบย์

กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ความน่าจะเป็นของ A เมื่อรู้ B (ความน่าจะเป็นที่จะเกิดเหตุการณ์ A โดยมีเงื่อนไขว่าเหตุการณ์ B ได้เกิดขึ้นแล้ว) เขียนแทนด้วย  $P(A|B)$  สามารถคำนวณได้ด้วยทฤษฎีของเบย์ ดังนี้

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3)$$

โดยที่

$P(A|B)$  คือ ความน่าจะเป็นของเหตุการณ์ A เมื่อกำหนดเหตุการณ์ B

$P(B|A)$  คือ ความน่าจะเป็นของเหตุการณ์ B เมื่อกำหนดเหตุการณ์ A

$P(A)$  คือ ความน่าจะเป็นของเหตุการณ์ A

$P(B)$  คือ ความน่าจะเป็นของเหตุการณ์ B

กล่าวคือ ความน่าจะเป็นของ A เมื่อรู้ B (โดยมีเงื่อนไขว่า B เกิดขึ้นแล้ว) สามารถคำนวณได้จากผลคูณของความน่าจะเป็นของ B เมื่อรู้ A กับความน่าจะเป็นของ A หากด้วยความน่าจะเป็นของ B เราเรียก  $P(A)$  ว่าเป็นความน่าจะเป็นก่อน (prior probability) และเรียก  $P(A|B)$  ว่าเป็นความน่าจะเป็นภายหลัง (posterior probability) ความน่าจะเป็นก่อนเป็นค่าที่ได้จากข้อมูลเบื้องต้น ส่วนความน่าจะเป็นภายหลังเป็นค่าความน่าจะเป็นก่อนที่ถูกปรับด้วยข้อมูลที่เพิ่มขึ้น

การหาค่าความน่าจะเป็นที่สมมติฐาน (h) ที่สนใจว่ามีโอกาสจะเกิดขึ้นเท่าไร เราสามารถใช้ทฤษฎีของเบย์ในการคำนวณ ดังนี้

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (4)$$

โดยที่

$P(h|D)$  คือ เป็นความน่าจะเป็นภายหลังซึ่งเป็นความน่าจะเป็นที่

สมมติฐาน h จะเป็นจริงโดยมีเงื่อนไขว่า D เป็นจริง

$P(D|h)$  คือ ความน่าจะเป็นก่อนซึ่งเป็นความน่าจะเป็นที่สมมติฐาน h จะเป็นจริงโดยที่เรายังไม่ได้อ่านข้อมูลตัวอย่างสอน

$P(h)$  คือ สมมติฐาน

$P(D)$  คือ ข้อมูลตัวอย่างสอน

## ตัวจำแนกประเภทเบย์อย่างง่าย

ตัวจำแนกประเภทเบย์อย่างง่าย (Naïve Bayes Classifier) เป็นตัวจำแนกประเภทแบบหนึ่ง ที่ใช้งานได้ดี เหมาะกับกรณีของเซตตัวอย่างมีจำนวนมากและคุณสมบัติ (Attribute) ของตัวอย่างไม่ขึ้นต่อกัน มีการนำตัวจำแนกประเภทเบย์อย่างง่ายไปประยุกต์ใช้งานในด้านการจำแนกประเภทข้อความ (text classification) การวินิจฉัย (diagnosis) และพบว่าใช้งานได้ดีไม่ต่างจากการจำแนกประเภทวิธีการอื่น

## อัลกอริทึม Naïve Bayes

Naïve-Bayes เป็นเทคนิคที่ถูกตั้งชื่อตาม Thomas Bayes นำเสนอโดย Kohavi ในปี 1996 ได้ใช้ทฤษฎี Bayes Theorem ในการแก้ปัญหาเรื่องการจำแนกประเภทข้อมูลที่ใช้แบบจำลองของการทำนาย โดยจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัว กับตัวแปรตาม ทั้งนี้ เพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์ทางทฤษฎี การทำนายผลของ Naïve-Bayes จะมีความถูกต้อง เมื่อตัวแปรอิสระทั้งหมดมีความเป็นอิสระต่อกัน ซึ่งจะไม่ขึ้นกับตัวแปรอิสระตัวใดตัวหนึ่ง

## วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayesian Learning)

การเรียนรู้เบย์อย่างง่ายเป็นวิธีการจำแนกข้อมูลที่มีประสิทธิภาพ โดยผลลัพธ์ที่ได้นั้นเทียบได้กับผลลัพธ์จากอัลกอริทึมที่มีความซับซ้อนกว่า เช่น C4.5 [Dougherty, Kohavi & Sahami, 1995] การเรียนรู้เบย์อย่างง่ายมีพื้นฐานมาจากกฎของเบย์ แต่ละลดความซับซ้อนลงโดยจะเพิ่มสมมติฐานที่ว่าคุณสมบัติต่าง ๆ ของข้อมูลจะไม่ขึ้นต่อกัน หรือกล่าวได้ว่าความน่าจะเป็นของข้อมูลที่จะเป็นกลุ่ม  $C_i$  สำหรับข้อมูลที่มีคุณสมบัติ  $n$  ตัว  $X=\{A_1,...,A_n\}$  หรือใช้สัญลักษณ์ว่า  $P(C_i|A_1,...,A_n)$  คือ

จากกฎของเบย์

$$P(C_i | A_1,...,A_n) = \frac{P(A_1,...,A_n | C_i) \times P(C_i)}{P(A_1,...,A_n)} \quad (5)$$

โดยที่

$P(C_i | A_1, \dots, A_n)$  คือ กลุ่ม  $C_i$  ที่มีคุณสมบัติ  $n$  ตัว

$P(C_i)$  คือ กลุ่มหรือคลาสเบล

$P(A_i)$  คือ คุณสมบัติตัวที่  $i$  ของกลุ่ม  $C_i$

การนำวิธีการเรียนรู้แบบอย่างง่ายไปใช้ มีวิธีการดังต่อไปนี้ คือ

1. หา  $P(C_i | A_1, \dots, A_n)$  จากสมการด้านบนสำหรับทุก ๆ กลุ่ม  $i$
2. นำค่าที่ได้มาเปรียบเทียบกัน กลุ่มที่มีค่าความน่าจะเป็นสูงสุด คือคำตอบ

จากการที่ตัวจำแนกแบบอย่างง่ายขึ้นอยู่กับสมมติฐานที่ว่าคุณสมบัติแต่ละอย่างของข้อมูลไม่ขึ้นต่อกัน ทำให้ดูเหมือนจะนำไปใช้งานในทางปฏิบัติได้อย่างไม่มีประสิทธิภาพมากนัก แต่ผลจากการทดลองของ Domingos และ Pazzani (Domingos และ Pazzani, 1996) ชี้ว่า ตัวจำแนกแบบอย่างง่ายสามารถจำแนกข้อมูลได้อย่างมีประสิทธิภาพ แม้สมมติฐานไม่เป็นจริง ด้วยเหตุนี้จึงทำให้วิธีการนี้เป็นที่นิยมนำไปประยุกต์ใช้กันมาก

ข้อจำกัดของอัลกอริทึม Naïve Bayes

เทคนิคแบบ Naïve-Bayes จะไม่รองรับข้อมูลที่เป็นค่าต่อเนื่อง (Continuous Data) ดังนั้นตัวแปรอิสระ หรือตัวแปรตามที่มีค่าต่อเนื่อง จะต้องมีการแบ่งเป็นข้อมูลเป็นช่วง ๆ เช่น อายุงาน เป็นต้น จะเห็นได้ว่าถ้ามีการแบ่งช่วงของข้อมูลไม่เหมาะสม จะมีผลต่อคุณภาพของแบบจำลองที่สร้างขึ้นได้

### Algorithm Naïve-Bayes

Naïve\_Bayes\_Learn(examples)

FOR EACH target value  $v$  DO

$$\overline{P}(V_j) \leftarrow \text{estimate } P(V_j)$$

FOR EACH attribute value  $a$  of each attribute DO

$$\overline{P}(a_i | v_j) \leftarrow \text{estimate } P(a_i | V_j)$$

Classify\_New\_Example( $x$ )

$$V_{NB} = \frac{\arg \max_{V_j \in V} \overline{P}(V_j)}{\prod_{i=1}^n \overline{P}(a_i | v_j)}$$

ภาพที่ 9 ขั้นตอนการทำงานของอัลกอริทึม Naive Bayes

ที่มา: Salton *et al.* (1983)

### กฎความสัมพันธ์ (Association Rule)

กฎความสัมพันธ์ (Association Rule) เป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล มีหลักการทำงาน คือการค้นหาค่าความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มีอยู่เพื่อนำไปใช้ในการวิเคราะห์ หรือทำนายปรากฏการณ์ต่าง ๆ หรือวิเคราะห์การซื้อสินค้าของลูกค้าเรียกว่า “Market Basket Analysis” ซึ่งประเมินจากข้อมูลที่รวบรวมไว้ ผลการวิเคราะห์ที่ได้จะเป็นคำตอบของปัญหา ซึ่งการวิเคราะห์แบบนี้เป็นการใช้ “กฎความสัมพันธ์” (Association Rule) เพื่อหาความสัมพันธ์ของข้อมูล ตัวอย่างการนำเทคนิคนี้ไปประยุกต์ใช้กับงานจริง ได้แก่ การซื้อสินค้าของผู้บริโภคในซูเปอร์มาร์เก็ต ข้อมูลของลูกค้าจะถูกลำดับมาประมวลผล เพื่อหาความสัมพันธ์ของข้อมูล เช่น ลูกค้าซื้อนม มักจะซื้อสินค้าใดครบคู่กันด้วยเสมอ ความสัมพันธ์ที่ได้จากกระบวนการนี้ จะนำไปใช้คาดเดาว่าควรนำเสนอสินค้าตัวใดให้กับลูกค้าที่ซื้อนม เพิ่มเติม เช่น “Milk->Diaper” with  $D_{supp} = 20\%$ ,  $D_{conf} = 80\%$  ค่า  $D_{supp} = 20\%$  หมายความว่า มีลูกค้า 20 % ที่ซื้อนมและผ้าอ้อม สำหรับค่า  $D_{conf} = 80\%$  หมายความว่า เมื่อลูกค้าซื้อนมแล้ว มีโอกาสที่จะซื้อผ้าอ้อมด้วย 80% ดังนั้น ค่า Confidence จะเป็นตัววัดค่ากฎที่ได้นั้นดีหรือไม่ดี (Guoqing, 2005)

## การค้นหากฎความสัมพันธ์

การค้นหากฎความสัมพันธ์ (Association Rule) ในฐานข้อมูลขนาดใหญ่ถือเป็นงานหลักงานหนึ่งในการทำเหมืองข้อมูล กฎความสัมพันธ์สามารถเขียนได้ในรูปเซตไอเท็มที่เป็นเหตุไปสู่เซตไอเท็มที่เป็นผล ซึ่งมีรากฐานมาจากการวิเคราะห์ทางการตลาด เช่น ลูกค้าที่ซื้อผ้าอ้อมส่วนใหญ่มักจะซื้อจุกนมด้วย ก็สามารถเขียนกฎความสัมพันธ์ได้เป็น {ผ้าอ้อม}  $\rightarrow$  {จุกนม} เป็นต้น พื้นฐานของการค้นหากฎความสัมพันธ์ประกอบด้วยนิยามต่าง ๆ เหล่านี้

1. เซตไอเท็ม ( $I$ ) คือเซตที่มีไอเท็มทั้งหมดเป็นสมาชิก ซึ่งไอเท็มในที่นี้อาจเป็นชื่อสินค้าหรือชื่อใด ๆ ที่เป็นหน่วยพื้นฐานที่จะนำมาทำการเรียนรู้

2. ทรานแซกชัน ( $T$ ) เป็นเซตย่อยของเซตไอเท็ม โดยที่  $T \subseteq I$

3. เซตข้อมูล ( $D$ ) คือเซตที่มีทรานแซกชันทุกตัวเป็นสมาชิก

4. เรากล่าวว่าทรานแซกชัน  $T$  บรรจุเซตย่อยของไอเท็ม  $X$  ก็ต่อเมื่อ  $X \subseteq T$

ฉะนั้นจึงนิยามกฎความสัมพันธ์ได้ว่า กฎความสัมพันธ์ (association rule) คือ การอุปนัยในรูปแบบ  $X \rightarrow Y$  เมื่อ  $X \subset Y, Y \subset I$  และ  $X \cap Y = \emptyset$

นอกจากนี้ กฎความสัมพันธ์ทุกกฎจะประกอบไปด้วยค่าสนับสนุน (Support) และค่าความมั่นใจ (confidence) ซึ่งมีนิยาม ดังนี้

1. กฎความสัมพันธ์  $X \rightarrow Y$  มีค่าสนับสนุนเท่ากับ  $s$  ในเซตข้อมูล  $D$  ก็ต่อเมื่อ  $s\%$  ของทรานแซกชันใน  $D$  บรรจุ ใน  $X \cup Y$

2. กฎความสัมพันธ์  $X \rightarrow Y$  มีค่าความเชื่อมั่น เท่ากับ  $c$  ในเซตข้อมูล  $D$  ก็ต่อเมื่อ  $c\%$  ของทรานแซกชันใน  $D$  ที่บรรจุ ใน  $X \cup Y$  ด้วย

$$\text{Support, } s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \quad (6)$$

กฎความสัมพันธ์  $X \rightarrow Y$  มีค่าความเชื่อมั่นเท่ากับ  $c$  ในเซตข้อมูล  $D$  ก็ต่อเมื่อ  $c\%$  ของทรานแซกชันใน  $D$  ที่บรรจุ  $X \cup Y$  ด้วย

$$\text{Confidence, } c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \quad (7)$$

โดยที่

$$\begin{aligned} N &= \text{จำนวนทรานแซกชันทั้งหมด} \\ \sigma(X) &= \text{จำนวนครั้งที่เกิด } X \end{aligned}$$

### ปัญหาการค้นหากฎความสัมพันธ์

เป็นปัญหาทางคณิตศาสตร์ในการค้นหากฎความสัมพันธ์ กล่าวคือ การหากฎความสัมพันธ์ทั้งหมดในทรานแซกชันทุกตัวของเซตข้อมูลที่กำหนดให้ โดยกฎความสัมพันธ์ที่หาได้ทั้งหมดจะต้องมีค่าสนับสนุนมากกว่าค่าสนับสนุนน้อยสุดที่กำหนดไว้ และมีค่าความมั่นใจมากกว่าค่าความมั่นใจน้อยที่สุดที่กำหนดไว้เช่นกัน

### อัลกอริทึม Apriori

Apriori (Agrawal and Srikant, 1994) เป็นอัลกอริทึมพื้นฐานที่ใช้ในการหาความสัมพันธ์ของข้อมูล โดยใช้หลักการค้นหาแบบวงกว้างก่อนนับทรานแซกชัน ซึ่งจะสร้างและตรวจสอบเซตไอเท็มบอยที่ละชั้น โดยเริ่มจากเซตไอเท็มที่มีจำนวนสมาชิกเท่ากับหนึ่ง ถ้าเซตไอเท็มใดมีค่าสนับสนุนน้อยกว่าค่าสนับสนุนที่กำหนดก็จะตัดเซตไอเท็มนั้นออก ไม่นำไปสร้างเซตไอเท็มในชั้นต่อไป การทำงานของอัลกอริทึมจะวนไปเรื่อยๆ จนกระทั่งไล่ทุกระดับชั้น หรือไม่เหลือเซตไอเท็มที่จะสร้างเซตไอเท็มในชั้นต่อไป

ในการนับจำนวนทรานแซกชัน Apriori อัลกอริทึมจะไล่ทรานแซกชันครั้งเดียวในแต่ละระดับชั้น ในการตรวจดูว่าทรานแซกชันนั้นบรรจุเซตไอเท็มใดบ้าง เพื่อความรวดเร็วจะเก็บเซตไอเท็มในแต่ละระดับชั้นทั้งหมดไว้ในโครงสร้างต้นไม้แฮช (hash tree)

### ขั้นตอนการทำงานของอัลกอริทึม Apriori

1. การสร้างชุดของไอเท็ม (Frequent Itemset Generation) การค้นหาชุดของไอเท็มที่มีค่าสนับสนุนมากกว่าค่าสนับสนุนน้อยที่สุด เพื่อสร้างเป็นชุดของไอเท็ม (Frequency item)
2. การสร้างกฎความสัมพันธ์ (Rule Generation) เป็นการสร้างกฎ โดยจะดึงกฎที่มีค่าความเชื่อมั่นสูง ๆ จากชุดของไอเท็ม

### การวัดประสิทธิภาพการจำแนกประเภท

#### 1. Accuracy

ค่าความถูกต้องใช้ในการวัดความถูกต้องในการทำนายข้อมูล (prediction data for classification) โดยวัดค่าจากจำนวนข้อมูลที่ทำนายถูกต้องหารด้วยจำนวนข้อมูลทั้งหมด

$$\text{Accuracy} = \frac{\text{จำนวนข้อมูลที่ทำนายถูกต้อง}}{\text{จำนวนข้อมูลทั้งหมด}}$$

#### 2. Precision

การวัดค่าการค้นคืนคำตอบที่ได้จากการทำนาย (Prediction) ความเชี่ยวชาญของผู้ปฏิบัติงาน

$$\text{Precision (P)} = \frac{\text{จำนวนของเอกสารที่เกี่ยวข้องที่ค้นคืนได้}}{\text{จำนวนของเอกสารทั้งหมดที่ค้นคืน}}$$

#### 3. Recall

วัดค่าความระลึกในการตรวจพบความเชี่ยวชาญของผู้ปฏิบัติงาน

$$\text{Recall (R)} = \frac{\text{จำนวนของเอกสารที่เกี่ยวข้องที่ค้นคืนได้}}{\text{จำนวนเอกสารที่เกี่ยวข้องที่มีทั้งหมดในฐานข้อมูล}}$$

#### 4. F-measure

ค่าที่แสดงความสัมพันธ์ระหว่างค่า Precision และค่า Recall เพื่อหาความถูกต้องโดยรวม

$$\text{F-measure} = \frac{2 * P * R}{P + R}$$

โดยที่

P คือ Precision

R คือ Recall

#### 5. Root Mean-Squared Error (RMSE)

การคำนวณค่าความผิดพลาดที่ยอมรับได้

$$\text{Root Mean-Squared Error} = \sqrt{\frac{(a_1 - c_1)^2 + (a_2 - c_2)^2 + \dots + (a_n - c_n)^2}{n}}$$

โดยที่

$a_i$  คือ ค่าวัดได้จริง

$c_i$  คือ ค่าที่ทำนาย

$n$  คือ จำนวนข้อมูลทั้งหมด

### K-fold Cross-Validation

ในการทดสอบโมเดลที่สร้างขึ้นมานี้ ได้ทดสอบบนพื้นฐานวิธี k-fold cross-validation ซึ่งผู้วิจัยได้กำหนดค่า  $k = 10$  คือ ทำการแบ่งข้อมูลออกเป็น 10 ส่วนจำนวนเท่า ๆ กัน โดยจะมีข้อมูลจำนวน 9 ส่วนใช้สำหรับการสอน ส่วนที่เหลือจะเป็นข้อมูลที่ใช้สำหรับการทดสอบ ซึ่งในการทดลองนี้ได้กำหนดจำนวนรอบในการทำงานของแต่ละอัลกอริทึม จำนวน 100 รอบ

## งานวิจัยที่เกี่ยวข้อง

### งานวิจัยเกี่ยวกับการทำแผนที่ความรู้ (Knowledge Maps)

รายงานของ มหาวิทยาลัยเชียงใหม่ (2550) ได้กล่าวถึงแผนที่ความรู้ขององค์กรว่า เป็นการจัดทำแผนผังที่บอกถึงขอบเขตของความรู้ แหล่งความรู้ ผู้ที่มีความรู้ รวมถึงวิธีการที่จะเข้าถึงความรู้ วิลาวัณย์ ฉิมประเสริฐ (2548) ให้ความหมายของแผนที่ความรู้ว่าเป็นเข็มทิศที่ช่วยให้บุคคลสามารถเข้าถึงแหล่งความรู้ขององค์กร แผนที่ความรู้มักจะแสดงถึงความรู้ที่องค์กรมีอยู่ รวมทั้งบุคลากรที่มีความรู้นั้น และการไหลของความรู้ในองค์กร โดยแผนที่ความรู้จะช่วยให้องค์กรทราบถึงความรู้ความสามารถของบุคลากร

Eppler (2001) ได้ค้นหาเครื่องมือที่จะนำมาใช้ในการบริหารจัดการความรู้ในองค์กร ซึ่งได้ใช้แผนที่ความรู้เป็นเครื่องมือในการจัดการความรู้และได้แบ่งแผนที่ความรู้ ออกเป็น 5 ประเภท ได้แก่ แหล่งความรู้ (Knowledge –sources) สินทรัพย์ความรู้ (Knowledge assets) โครงสร้างความรู้ (Knowledge – structures) ความรู้เชิงประยุกต์ (Knowledge –applications) และการพัฒนาแผนที่ความรู้ (Knowledge –development maps)

Judith *et al.* (2004) ได้ศึกษาเรื่องการใช้ Concept maps ในการวัดความเข้าใจของนักศึกษา ในการเรียนตามหลักสูตรที่ทางมหาวิทยาลัยจัดให้ รวมถึงศึกษากฎเกณฑ์ที่ใช้ในการออกแบบ กระบวนการสร้างแผนที่ความรู้ โดยเชื่อมคำนามและคำกริยาเข้าด้วยกัน การสร้างแผนที่ในลักษณะ ผังเครือข่าย (Network diagrams) โดยโนดต่าง ๆ จะแทนด้วยคำนาม และเชื่อมความสัมพันธ์ระหว่างโนดด้วยคำกริยา เป็น directed graph และ acyclic graphs ผลการทดลอง ได้แผนที่ ที่มีความถูกต้อง 96% ความน่าเชื่อถือในการจัดเรียงลำดับอยู่ในช่วง 91-100 %

ต่อมาในปี 2006 Eppler (2006) นำเสนอการแบ่งประเภทของแผนที่ความรู้โดยใช้วิธี อนุกรมวิธาน วัดอุปสงค์เพื่อแบ่งประเภทของแผนที่ความรู้ให้เหมาะสมกับประเภทปัญหาของการจัดการความรู้ ด้านคุณภาพของการแบ่งประเภทของแผนที่พิจารณาจากความเสถียรและความสามารถในการบริหารจัดการ และองค์ประกอบสำคัญของแผนที่ความรู้ คือ Ground Layer และ Individual Elements เป็นการอธิบายถึงบริบทที่เกี่ยวข้องกับแผนที่ และรายการต่าง ๆ ที่จับคู่กับบริบทนั้น ซึ่งงานวิจัยนี้ได้แบ่งประเภทของแผนที่ความรู้ ตามวัตถุประสงค์ ตามรูปแบบทาง

กราฟฟิค ตามเนื้อหา ตามระดับของผู้ใช้งาน และตามวิธีการสร้างแผนที่ความรู้ ฉะนั้น การจัดทำแผนที่จะแตกต่างกันไปตามเนื้อหาของเรื่องนั้น ๆ งานวิจัยในอนาคตจะเป็นการพัฒนาต้นแบบของแผนที่ความรู้ประเภทต่าง ๆ ที่ระบุเงื่อนไขและข้อกำหนด รวมถึงเรื่องการติดตามประเมินผลการทำงานด้วย

Neshatian *et al.* (2006) ได้นำเสนอประสบการณ์ที่ได้จากหลายแหล่งความรู้ในองค์กร โดยใช้ออนโทโลยีในการนำเสนอประสบการณ์ของผู้ปฏิบัติงานในองค์กร ออนโทโลยีสามารถพิจารณาถึงแง่มุมของประสบการณ์ต่างๆ ที่จะนำมาใช้สร้างความรู้ในองค์กร โดยออนโทโลยีจะประกอบด้วย Local Domain Ontologies, Enterprise Ontology, Knowledge Map และ Research Roadmap โดยผู้วิจัยได้ใช้แผนที่ความรู้เป็นเครื่องมือที่บอกให้ทราบว่าแหล่งประสบการณ์ของความรู้ด้านต่าง ๆ อยู่ที่ไหน และใช้ออนโทโลยีเป็นตัวแทนรายละเอียดของความรู้ต่าง ๆ เป็นการจัดแบ่งหมวดหมู่ของความรู้ในขอบเขตความสนใจหนึ่ง ๆ

Richardson (2001) กล่าวถึงแผนที่ความรู้ว่าเป็นเครื่องมือที่ช่วยให้องค์กรทราบถึงความรู้ความสามารถของบุคลากร ช่วยในการวางแผนด้านความพร้อมและการแก้ปัญหาการขาดแคลนบุคลากร รวมทั้งสามารถพัฒนาบุคลากรให้มีความสามารถเพิ่มขึ้น

#### งานวิจัยเกี่ยวกับการแบ่งกลุ่ม (Clustering)

Lin และ Hsueh (2002) นำเสนอการสร้างแผนที่ความรู้เชิงเนื้อหา และการบำรุงรักษาความรู้ โดยใช้วิธีการแบ่งกลุ่มแบบ 2 ขั้นตอน ประกอบด้วย Hierarchical Clustering และ K-means Clustering ข้อมูลที่นำมาใช้ในการสร้างแผนที่ความรู้ได้นำมาจากเอกสารการสอนของอาจารย์ที่เก็บไว้บนอินเทอร์เน็ตที่เรียกว่า SCTNet และงานวิทยานิพนธ์ห้องสมุดแห่งชาติของประเทศไทย จากการทดลองสร้างแผนที่ความรู้และประเมินผลจากผู้เชี่ยวชาญ 16 กลุ่มโดยมีการปรับปรุงแผนที่ความรู้ใหม่ด้วยระบบการสร้างแผนที่ความรู้ ข้อมูลที่นำมาใช้ทดสอบมี 2 ชุด ชุดที่ 1 เรื่องการศึกษาวิทยาศาสตร์ธรรมชาติ และชุดที่ 2 เรื่องการบริหารจัดการสารสนเทศเกี่ยวกับบทคัดย่อในงานวิทยานิพนธ์ การวัดค่าความถูกต้องใช้ตัววัด precision และ recall การสร้างแผนที่ความรู้วัดค่า precision ได้ 91% recall ได้ 93%

Ong *et al.* (2005) นำเสนอการสร้างแผนที่ความรู้สำหรับบทความจากหนังสือพิมพ์ ซึ่งจะมีคอลัมน์การเงิน สุขภาพ ในลักษณะ Online โดยใช้วิธี Hierarchical Clustering ในการสร้าง Knowledge map สำหรับแหล่งความรู้ โดยเฉพาะแหล่งข่าวของหนังสือพิมพ์จีน นำเสนอความรู้ในรูปแบบของ PAT-Tree นอกจากนี้มีการสัปดาห์ การจัดกลุ่มข้อมูลโดยใช้ Self-Organizing Map วัดประสิทธิภาพของแผนที่ความรู้ด้วยค่า precision, recall

Liao *et al.* (2007) นำเทคนิคการทำเหมืองข้อมูลมาใช้ในตลาดหลักทรัพย์ของประเทศไต้หวัน โดยทำการจัดกลุ่มข้อมูลแบบ 2 ขั้นตอน ขั้นตอนแรกใช้เทคนิค Association Rules โดยใช้ อัลกอริทึม Apriori ในการหาความสัมพันธ์ของ stock category ขั้นตอนที่ 2 ใช้ K-means อัลกอริทึม ในการจัดกลุ่ม stock category เพื่อใช้เป็นข้อมูลประกอบการลงทุนในตลาดหลักทรัพย์ ผลของงานวิจัยนี้สามารถค้นหา portfolio ของ stock categories ที่มีความเป็นไปได้ เรื่องการลงทุนในตลาดหลักทรัพย์ของไต้หวัน

#### งานวิจัยเกี่ยวกับการจำแนก (Classification)

Wang *et al.* (2003) นำเสนอการจำแนกเอกสารในห้องสมุด โดยใช้ทฤษฎีเบย์ ซึ่งใช้ Naïve Bayes อัลกอริทึม และได้สร้างแบบจำลองขึ้นมา 3 แบบ ได้แก่ 1. Multi-variate Bernoulli Model (MB) และ Multinomial Model (MN) 2. Smoothing methods 3. Feature selection methods ซึ่งได้วัดประสิทธิภาพของแบบจำลองด้วยตัววัด Precision Recall และ F-measure มีการเปรียบเทียบกับผลงานวิจัยก่อนหน้า ผลการทดลองปรากฏว่า แบบจำลองที่ 2 ให้ค่าความถูกต้องในการจำแนกข้อมูลสูงกว่าแบบจำลองอื่น โดยมีค่า F-measure เพิ่มขึ้นกว่าเดิม 20 % ในปีเดียวกัน Huang และคณะ (Huang *et al.*, 2003) เปรียบเทียบความสามารถในการจำแนกประเภทข้อมูลของ 3 อัลกอริทึม ประกอบด้วย Naïve Bayes อัลกอริทึม C4.5 อัลกอริทึม และ SVM (Support Vector Machine) ด้วยข้อมูลทดสอบจาก UCI Repository จำนวน Data Set 18 ชุด ผลการทดลองปรากฏว่า C4.5 อัลกอริทึม ให้ค่าความถูกต้องโดยเฉลี่ยสูงสุด ในขณะที่ Naïve Bayes อัลกอริทึม และ SVM (Support Vector Machine) มีค่าความถูกต้องโดยเฉลี่ยเท่ากัน

Soman และ Bobbie (2005) นำเสนอการจำแนกประเภทผู้ป่วยที่เป็นโรคหัวใจเต้นผิดปกติ ด้วยอัลกอริทึม OneR C4.5 และ Naïve Bayes โดยแบ่งข้อมูลที่ใช้ในการฝึกสอน และข้อมูลทดสอบ เป็น 4 ส่วนประกอบ 1. ใช้ข้อมูลฝึกสอน 100% แล้วนำมาเป็นข้อมูลทดสอบ ผลการทดลองปรากฏว่า J48 มีความถูกต้องในการทำนายสูงสุด (91.81%)

2. ข้อมูลทดสอบ 50% และชุดข้อมูลสอน 50% ผลการทดลองปรากฏว่า Naïve Bayes มีความถูกต้องในการทำนายสูงสุด (70.80%) 3. ข้อมูลฝึกสอน 70% ข้อมูลทดสอบ 30% ผลการทดลองปรากฏว่า Naïve Bayes มีความถูกต้องในการทำนายสูงสุด (75%) 4. ข้อมูลฝึกสอน 80% ข้อมูลทดสอบ 20% ผลการทดลองปรากฏว่า Naïve Bayes มีความถูกต้องในการทำนายสูงสุดในการแบ่งข้อมูลเป็นชุดทดสอบและชุดฝึกสอน สำหรับ C4.5 ให้ผลการทำนายสูงสุดในกรณีที่ใช้ข้อมูลฝึกสอน 100% แล้วนำมาทดสอบ

Ordonez (2006) ทำการเปรียบเทียบการทำนายผลของผู้ป่วยที่เป็นโรคหลอดเลือดแดง ด้วยเทคนิค Association Rules และ Decision Trees ผลการทดลองพบว่า Decision Trees แสดงผลการทำนายไม่ครอบคลุม เหมือนกับ Association Rules ข้อสรุปของการทดลองว่า Decision Trees มักใช้กับการค้นหากฎที่ง่าย ๆ ไม่ซับซ้อน และในการแบ่งว่าโรคใดควรจะนำมาสร้างเป็นโนดราก Decision Trees จะมีการแบ่งโนดที่ต่างไปจากทางการแพทย์ และกฎส่วนมากที่ได้จะเป็นเซตเล็ก ๆ ของผู้ป่วย แต่สำหรับ Association Rules โดยทั่วไปจะรวมกฎที่เกี่ยวกับการทำนายโรคของผู้ป่วย กฎที่ได้จะมีความน่าเชื่อถือสูง และครอบคลุมกลุ่มของผู้ป่วย

Wang *et al.* (2006) นำเสนออัลกอริทึมใหม่ ได้แก่ Self-adaptive NBTree (S-NBTree) ซึ่งเป็น Hybrid decision tree ใช้อยู่ร่วมกับ Naïve Bayes โดย Wang เปรียบเทียบประสิทธิภาพการจำแนกของ 3 อัลกอริทึม ได้แก่ C4.5 อัลกอริทึม Naïve Bayes และ S-NBTree ผลการทดลองปรากฏว่า Naïve Bayes มีประสิทธิภาพน้อยกว่า C4.5 ส่วน S-NBTree มีความถูกต้องในการทำนายสูงสุด

### สรุปงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับเรื่อง อัลกอริทึมในการจำแนกบุคลากรในองค์กร สำหรับการสร้างแผนที่ความรู้ ประกอบด้วงานวิจัยในเรื่องการสร้างแผนที่ความรู้ (Knowledge Maps) งานวิจัยเกี่ยวกับการแบ่งกลุ่ม (Clustering) และงานวิจัยเกี่ยวกับการจำแนกข้อมูล (Classification)

งานวิจัยที่เกี่ยวข้องในเรื่องการสร้างแผนที่ความรู้ (Knowledge Maps) เทคนิคที่นำมาใช้ในการสร้างแผนที่ความรู้ เช่น 1) Concept Maps 2) อนุกรมวิธาน 3) ออนโทโลยี และสามารถแบ่งแผนที่ความรู้เป็น 5 ประเภท ได้แก่ 1) แหล่งความรู้ (Knowledge –sources) 2) สินทรัพย์ความรู้ (Knowledge assets) 3) โครงสร้างความรู้ (Knowledge – structures) 4) ความรู้เชิงประยุกต์

(Knowledge –applications) 5) การพัฒนาแผนที่ความรู้ (Knowledge –development maps) โดยงานวิจัยฉบับนี้ ผู้วิจัยได้สร้างแผนที่ความรู้ประเภทแหล่งความรู้ ที่จะสามารถระบุแหล่งของภูมิปัญญาในระดับบุคคลและองค์กร

งานวิจัยที่เกี่ยวข้องกับการแบ่งกลุ่ม (Clustering) มีการใช้เทคนิคการแบ่งกลุ่มแบบ 2 ขั้นตอน ประกอบด้วย Hierarchical Clustering, K-means Clustering และ Self-Organizing Map และมีการวัดประสิทธิภาพด้วยค่า Precision และ Recall

งานวิจัยที่เกี่ยวข้องกับการจำแนกข้อมูล (Classification) เทคนิคที่นำมาใช้ในการจำแนกประเภทข้อมูล เช่น อัลกอริทึม Naïve Bayes อัลกอริทึม C4.5 SVM (Support Vector Machine) อัลกอริทึม OneR และ S-NBTree ซึ่งเป็นแบบ Hybrid decision tree โดยใช้ต้นไม้การตัดสินใจร่วมกับ Naïve Bayes สำหรับเรื่องการหาความสัมพันธ์ พบว่างานวิจัยที่เกี่ยวข้องใช้ Association Rules ด้วยวิธีอัลกอริทึม Apriori และใช้ตัววัดประสิทธิภาพด้วยค่า Precision, Recall และ F-measure งานวิจัยฉบับนี้ใช้อัลกอริทึม 3 แบบ ได้แก่ อัลกอริทึม C4.5 อัลกอริทึม OneR และ อัลกอริทึม Naïve Bayes ในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานในองค์กร เนื่องจากเป็นอัลกอริทึมที่สามารถสร้างต้นไม้การตัดสินใจที่มีความถูกต้องสูง ผู้สนใจศึกษาสามารถเรียนรู้และทำความเข้าใจได้ง่าย ไม่ซับซ้อน สำหรับงานวิจัยที่เกี่ยวข้องกับการศึกษาครั้งนี้ ได้นำมาสรุปรวบรวมไว้ในตารางที่ 1, 2 และ 3

ตารางที่ 1 ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับแผนที่ความรู้ (Knowledge Maps)

| งานวิจัย                    | เทคนิค                                                                                                       | ข้อมูล                                                                                                                                                                            | ผล                                                                                                                                                     |
|-----------------------------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Eppler (2001)               | ได้ค้นหาเครื่องมือที่จะนำมาใช้ในการบริหารจัดการความรู้ในองค์กร ได้แก่ แผนที่ความรู้                          | <ul style="list-style-type: none"> <li>- ตัวแทนด้านมัลติมีเดีย</li> <li>- กลุ่มที่ปรึกษา</li> <li>- บริษัทวิจัยทางการตลาด</li> <li>- บริษัทดำเนินการด้านบริการขนาดเล็ก</li> </ul> | แบ่งแผนที่ความรู้เป็น 5 ประเภท ได้แก่ Knowledge-sources, Knowledge-assets, Knowledge-structures, Knowledge-applications และ Knowledge-development maps |
| Eppler (2006)               | นำเสนอการแบ่งประเภทแผนที่ความรู้ด้วยวิธีอนุกรมวิธาน                                                          | ชุมชนด้านวิทยาศาสตร์ :<br>ศึกษาด้านการศึกษา ศึกษาด้านองค์กร และวิศวกรความต้องการ                                                                                                  | แบ่งประเภทความรู้ตามวัตถุประสงค์ตามรูปแบบทางกราฟิก ตามเนื้อหา ตามระดับของผู้ใช้งาน และตามวิธีการสร้างแผนที่ความรู้                                     |
| Judith <i>et al.</i> (2004) | ศึกษาความเข้าใจของนักศึกษาในการสร้าง Concept maps โดยใช้ Network diagrams, directed graph และ acyclic graphs | ใช้ข้อมูลของนิสิตชั้นปีที่ 3 ที่เรียนสาขาวิศวกรรมศาสตร์                                                                                                                           | Concept maps ที่สร้างมีความถูกต้องชัดเจน และจัดเรียงตามรูปแบบที่กำหนด 96% ความน่าเชื่อถือในการจัดเรียงอยู่ในช่วง 91-100 %                              |

ตารางที่ 1 (ต่อ)

| งานวิจัย                          | เทคนิค    | ข้อมูล                                                              | ผล                                                                                                                  |
|-----------------------------------|-----------|---------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| Neshatian <i>et al.</i><br>(2006) | ออนโทโลยี | ประสบการณ์ของผู้ปฏิบัติงานใน<br>องค์กร โดยแบ่งเป็นประเภท<br>ความรู้ | ออนโทโลยีสามารถพิจารณาได้ถึง<br>แง่มุมของประสบการณ์ต่างๆ ที่จะ<br>นำมาใช้สร้างความรู้ และแผนที่ความรู้<br>ขององค์กร |

**ตารางที่ 2** ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มข้อมูล (Clustering)

| งานวิจัย                  | เทคนิค                                                                                                                                                | ข้อมูล                                                                                           | ผล                                                                                                                                                                             |
|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Lin และ Hsueh (2002)      | นำเสนอการสร้างแผนที่ความรู้และการบำรุงรักษาความรู้ โดยใช้ two-stage clustering ประกอบด้วย hierarchical clustering และ K-means clustering              | 1. ด้านวิทยาศาสตร์ธรรมชาติ<br>2. สารสนเทศในระดับปริญญาโท<br>ในส่วนที่เป็นบทคัดย่อ<br>วิทยานิพนธ์ | ค่า precision ได้ 91% recall ได้ 93%                                                                                                                                           |
| Ong <i>et al.</i> (2005)  | นำเสนอการสร้างแผนที่ความรู้ด้วยวิธี Hierarchical Clustering และจัดกลุ่มข้อมูลด้วยวิธี Self-Organizing Map                                             | หนังสือพิมพ์คอลัมน์การเงิน และ<br>สุขภาพ                                                         | มีการจัดกลุ่มข้อมูลที่ต้องกว่าที่ใช้<br>บุคลากรในหน่วยงานจัดเอง ใช้ตัววัด<br>ประสิทธิภาพ Precision และ Recall<br>นอกจากนี้ยังช่วยลดความซับซ้อนของ<br>ข้อมูล (High dimensional) |
| Liao <i>et al.</i> (2007) | จัดกลุ่มข้อมูลแบบ 2 ขั้นตอน ขั้นตอนที่ 1 ใช้ Apriori อัลกอริทึมในการหาความสัมพันธ์ของ stock category ขั้นตอนที่ 2 ใช้ K-means อัลกอริทึมในการจัดกลุ่ม | ข้อมูลในตลาดหลักทรัพย์ของ<br>ไต้หวัน                                                             | สามารถค้นหา portfolio ของ stock<br>categories ที่มีความเป็นไปได้เรื่องการ<br>ลงทุนในตลาดหลักทรัพย์ของไต้หวัน                                                                   |

ตารางที่ 2 (ต่อ)

| งานวิจัย | เทคนิค                                                              | ข้อมูล | ผล |
|----------|---------------------------------------------------------------------|--------|----|
|          | stock category เพื่อใช้เป็นข้อมูลประกอบการลงทุน<br>ในตลาดหลักทรัพย์ |        |    |

ตารางที่ 3 ตารางสรุปผลงานวิจัยที่เกี่ยวข้องกับการจำแนกประเภทข้อมูล (Classification)

| งานวิจัย                  | เทคนิค                                  | ข้อมูล                     | ผล                                                                                                     |
|---------------------------|-----------------------------------------|----------------------------|--------------------------------------------------------------------------------------------------------|
| Wang <i>et al.</i> (2003) | จำแนกเอกสาร ด้วย Naïve Bayes อัลกอริทึม | เอกสารที่ห้องสมุดของรัฐสภา | เปรียบเทียบกับผลงานวิจัยก่อนหน้า<br>การทดลองครั้งนี้มีค่า F-measure<br>เพิ่มขึ้นกว่าเดิม 20 % (67.19%) |

ตารางที่ 3 (ต่อ)

| งานวิจัย                   | เทคนิค                                                                                                                         | ข้อมูล                                   | ผล                                                                                                                                                                                                                                                                                     |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Huang <i>et al.</i> (2003) | นำเสนอการทำนายความถูกต้องในการจำแนกประเภทข้อมูล โดยใช้ Naïve Bayes อัลกอริทึม C4.5 อัลกอริทึม และ SVM (Support Vector Machine) | ข้อมูลจาก UCI repository<br>จำนวน 18 ชุด | C4.5 อัลกอริทึมให้ค่าความถูกต้องโดยเฉลี่ยสูงสุด 86.7%                                                                                                                                                                                                                                  |
| Soman และ Bobbie (2005)    | นำเสนอการจำแนกประเภทผู้ป่วยที่เป็นโรคหัวใจ ต้นฝึกปกติ ด้วย OneR, Decision Tree J48 (C4.5) และ Naïve Bayes                      | ข้อมูลผู้ป่วยที่เป็นโรคหัวใจต้นฝึกปกติ   | การทดลองแบ่งเป็น 4 เงื่อนไข<br>1. ข้อมูลสอนและทดสอบ 100% พบว่า Decision Tree J48(C4.5) มีความถูกต้องในการทำนายสูงสุด (91.81%) 2. ข้อมูลสอน 50% ทดสอบ 50% Naïve Bayes มีความถูกต้องในการทำนายสูงสุด (70.80%) 3. ข้อมูลสอน 70% ทดสอบ 30% Naïve Bayes มีความถูกต้องในการทำนายสูงสุด (75%) |

ตารางที่ 3 (ต่อ)

| งานวิจัย                  | เทคนิค                                                                                                         | ข้อมูล                                        | ผล                                                                                            |
|---------------------------|----------------------------------------------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------------------------------------------------------|
|                           |                                                                                                                |                                               | 4. ข้อมูลสอน 80% ทดสอบ 20% Naïve Bayes ให้ค่าความถูกต้องในการทำนายสูงสุด (74.73%)             |
| Ordonez (2006)            | เปรียบเทียบการทำงานของอัลกอริทึม Association Rules กับ Decision Trees ในการจำแนกข้อมูล                         | ผู้ป่วยที่เป็นโรคหัวใจ                        | ผลการทดลองพบว่า Decision Trees ไม่สามารถแสดงผลการทำนายได้ครอบคลุม เหมือนกับ Association Rules |
| Wang <i>et al.</i> (2006) | เปรียบเทียบประสิทธิภาพของอัลกอริทึม C4.5, Naïve Bayes และ S-NBTree (Hybride decision tree ร่วมกับ Naïve Bayes) | กลุ่มข้อมูลเกี่ยวกับธรรมชาติ (natural domain) | Naïve Bayes มีประสิทธิภาพน้อยกว่า C4.5 และ S-NBTree มีความถูกต้องสูงสุด                       |

## อุปกรณ์และวิธีการ

### อุปกรณ์

#### 1. Hardware

เครื่องคอมพิวเตอร์ส่วนบุคคล มีส่วนประกอบ ดังนี้ CPU Intel 2 Duo Processor T5600 1.83 GHz, 667 MHz FSB, 2MB L2 cache, 120GB HDD และ 512 MB DDR2

#### 2. Software

2.1 MatLab เวอร์ชัน 7.0.1

2.2 Waikato Environment for Knowledge Analysis เวอร์ชัน 3.5.5

#### 3. Database

3.1 MySQL-front เวอร์ชัน 2.5

3.2 Microsoft Access 2003

#### 4. Operating System

4.1 Window XP 5.1 Service Pack 2

### วิธีการ

#### ข้อมูลสำหรับการทดลอง (Data set)

ข้อมูลที่ใช้ในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงาน ได้มาจากหน่วยงาน การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง ซึ่งส่วนบริหารคุณภาพและเพิ่มผลผลิตสายงานผลิตไฟฟ้า

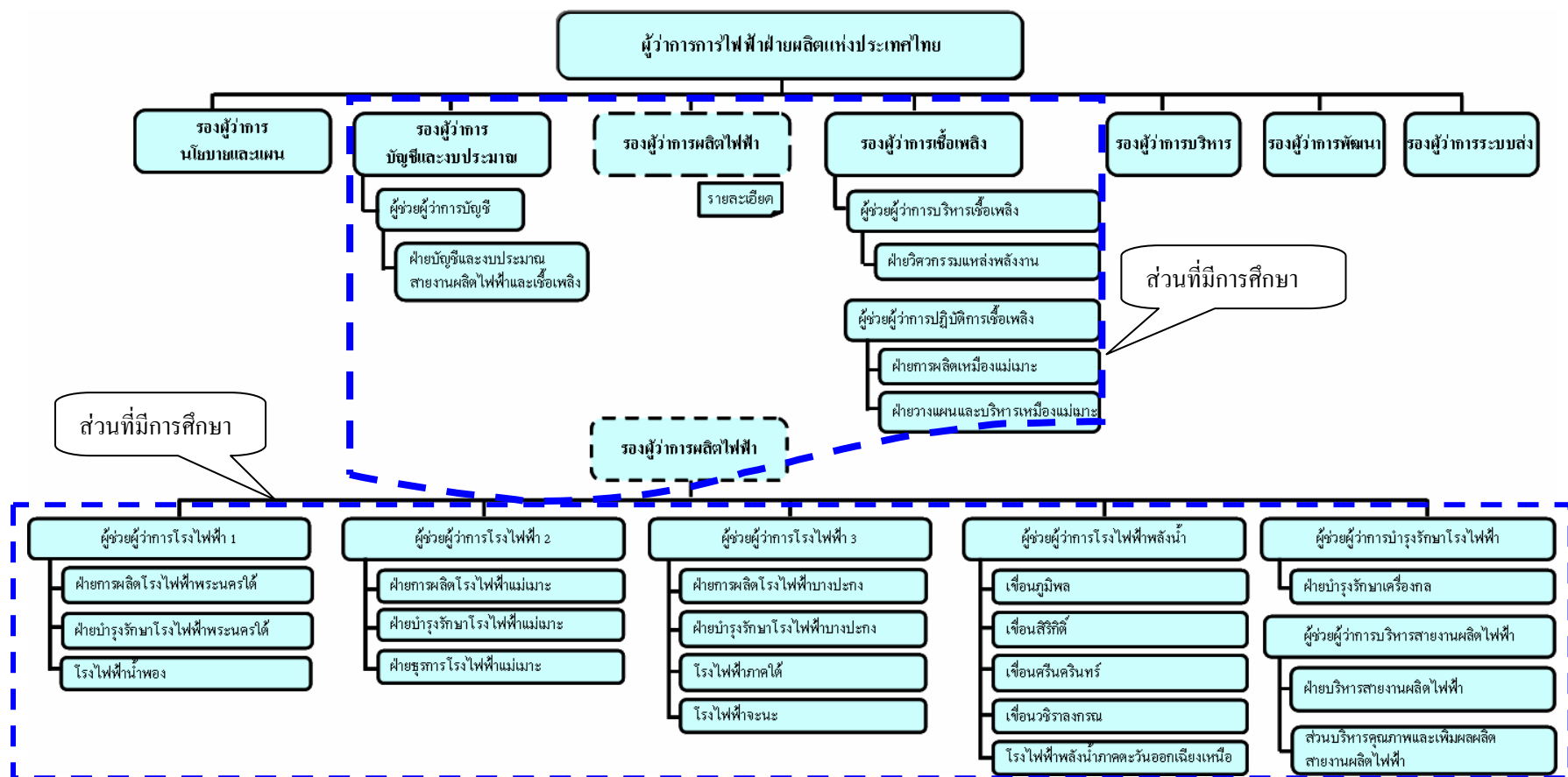
(สพฟ.) เป็นหน่วยงานหลักที่รับผิดชอบในเรื่อง กำหนดกลยุทธ์และแผนงานที่ทำให้ทุกหน่วยงานมีการจัดการความรู้อย่างเป็นระบบ และพัฒนาไปสู่องค์กรแห่งการเรียนรู้ ข้อมูลที่หน่วยงานต่าง ๆ บันทึกความรู้ไว้จะถูกบันทึกผ่านทางเว็บไซต์ที่ชื่อว่า “เว็บภูมิปัญญา” ซึ่งเปิดตัวเว็บไซต์ เมื่อปี 2549 โดยฝ่ายต่าง ๆ ที่สังกัดรองผู้ว่าการผลิตไฟฟ้าจะเป็นหน่วยงานที่มีความพร้อมมากที่สุด ได้นำข้อมูลจัดเก็บลงในเว็บไซต์ สำหรับสายงานอื่น ๆ เริ่มมีการตื่นตัวในการนำความรู้มาบันทึกบนเว็บไซต์เพิ่มขึ้น ดังนั้น ข้อมูลที่นำมาใช้ในการทดลองมีจำนวน 2 ชุด คือ ข้อมูลชุดที่ 1 จำนวน 343 ระเบียบ และข้อมูลชุดที่ 2 จำนวน 805 ระเบียบ ในปี 2550 ข้อมูลแบ่งเป็น 4 ส่วน คือ

### ส่วนที่ 1 ข้อมูลบุคคล

ข้อมูลบุคคล หมายถึง ชื่อ – นามสกุล ของผู้เปิดเผยตัวเองว่าเป็นผู้เชี่ยวชาญ/ผู้รู้ ในงานด้านใด คุณลักษณะหลักในเรื่องของบุคคล ประกอบด้วย 3 คุณลักษณะ ได้แก่ 1) รหัสพนักงาน 2) อายุ งาน และ 3) ตำแหน่งงาน ข้อมูลที่นำมาทดลองผู้ปฏิบัติงานจะมีตำแหน่งเป็นช่าง 71.93% ตำแหน่งวิศวกร 16.40% และตำแหน่งอื่น ๆ 11.40%

### ส่วนที่ 2 ข้อมูลหน่วยงาน

ข้อมูลหน่วยงาน หมายถึง หน่วยงานต่าง ๆ ในระดับผู้ช่วยผู้ว่าการ และฝ่ายที่อยู่ภายใต้การบังคับบัญชา จำนวน 28 หน่วยงานจาก 3 สายงาน ได้แก่ 1) สายงานรองผู้ว่าการบัญชีและการเงิน 2) สายงานรองผู้ว่าการผลิตไฟฟ้า 3) สายงานรองผู้ว่าการเชื้อเพลิง โดยโครงสร้างหน่วยงานเริ่มจากระดับผู้ว่าการ รองผู้ว่าการ ผู้ช่วยผู้ว่าการ และหน่วยงานระดับฝ่ายที่สังกัดผู้ช่วยผู้ว่าการ ตามสายงานต่าง ๆ ซึ่งหน่วยงานระดับฝ่ายได้แบ่งผู้เชี่ยวชาญตามหน่วยงานที่ปฏิบัติได้ 3 คุณลักษณะ คือ 1) หน่วยงานระดับฝ่าย 2) ระดับกอง 3) ระดับแผนก



ภาพที่ 10 โครงสร้างการบังคับบัญชาของ กฟผ. เฉพาะหน่วยงานตัวอย่างที่นำข้อมูลมาใช้ในการทดลอง ปี พ.ศ. 2550

### ส่วนที่ 3 ข้อมูลการศึกษา

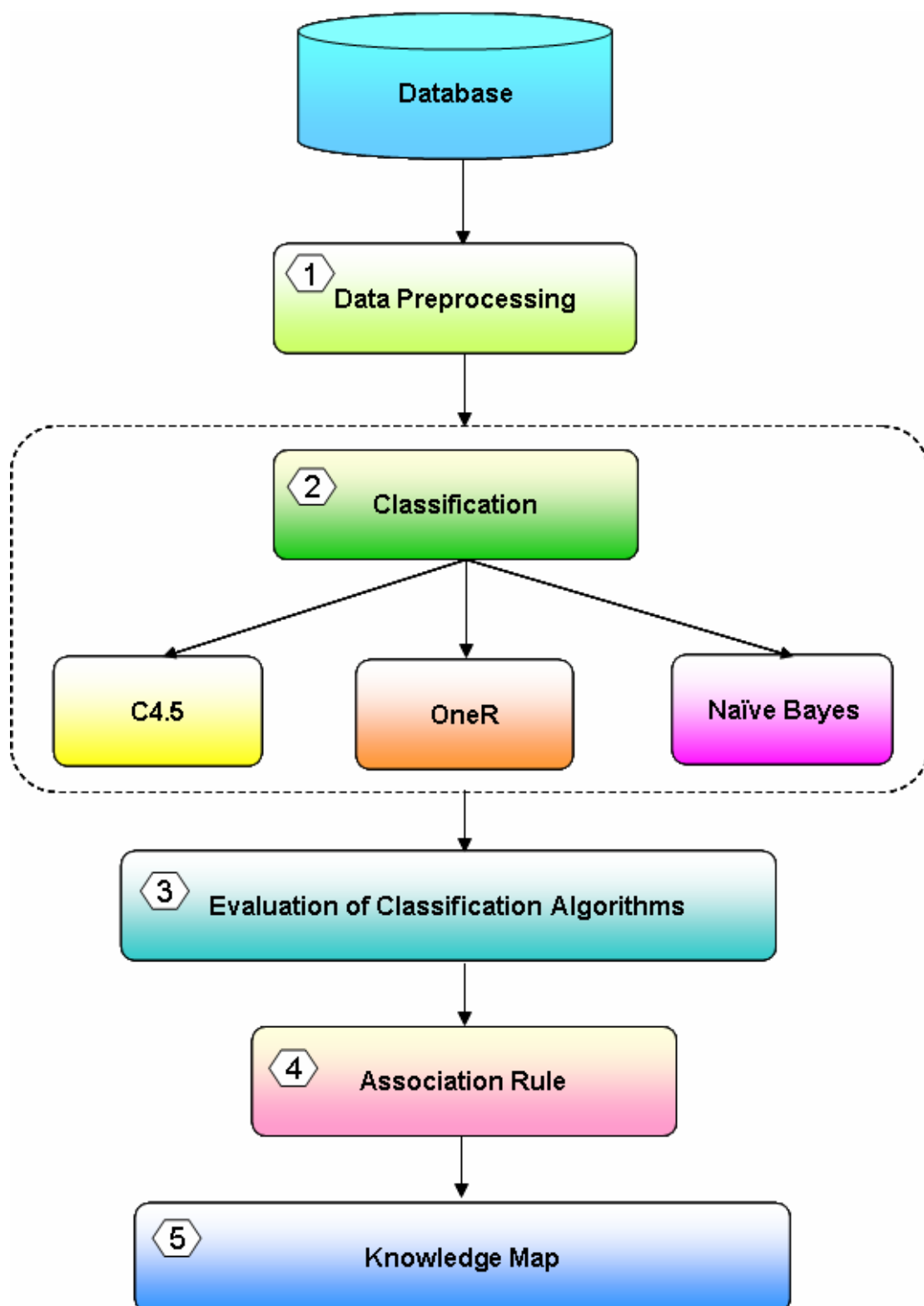
ข้อมูลการศึกษา หมายถึง การศึกษาของผู้ปฏิบัติงานที่เป็นผู้เชี่ยวชาญ/ผู้รู้ ในงานด้านต่าง ๆ ประกอบด้วย 3 คุณลักษณะ ได้แก่ 1) คุณวุฒิการศึกษา 2) คณะที่จบการศึกษา และ 3) สาขาที่จบการศึกษา สำหรับคุณลักษณะข้อมูลด้านคุณวุฒิการศึกษาจะมีระดับ ปวส. 44.47% ระดับ ปวช. 23.48% ระดับปริญญาตรี 18.14% อื่น ๆ 13.91%

### ส่วนที่ 4 ความเชี่ยวชาญ

ความเชี่ยวชาญ หมายถึง งานที่ผู้ปฏิบัติงานมีความเชี่ยวชาญ หรือเป็นผู้รู้ในงานด้านนั้น ที่จะสามารถแก้ปัญหาในงานที่ปฏิบัติ หรือสามารถที่จะช่วยเหลือผู้อื่นในยามที่เกิดปัญหา เป็นการแก้ไขปัญหาที่ทันกาล หรือสามารถที่จะสอนงานให้ผู้ปฏิบัติงานท่านอื่นได้ และผู้อื่นสามารถนำความรู้ที่ได้รับไปใช้ในการปฏิบัติงานได้อย่างมีประสิทธิภาพ รายละเอียดข้อมูลความเชี่ยวชาญปรากฏในตารางที่ 14

### ขั้นตอนการทดลอง

โครงสร้างการทำงานของระบบการทำเหมืองข้อมูลงานทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง



ภาพที่ 11 โครงสร้างของงานวิจัยสำหรับทำนายความเชื่อหาญของผู้ปฏิบัติงาน

## 1. การเตรียมข้อมูล (Preprocessing)

ขั้นตอนของการเตรียมข้อมูล เริ่มจากการรวบรวมข้อมูลจากเว็บไซต์ภูมิปัญญาของสายงานผลิตไฟฟ้า การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย ที่เก็บไว้ในฐานข้อมูล MySQL รวบรวมได้ทั้งหมด 805 รายการ (transaction) โดยทำการเลือกข้อมูลจากแหล่งข้อมูล และทำการแยกข้อมูลที่ไม่มีค่า ข้อมูล ที่บันทึกไม่สมบูรณ์ออก และทำการรวบรวมข้อมูลที่ต้องการจากหลายตารางข้อมูล เช่น ข้อมูลบุคคล ข้อมูลหน่วยงาน ข้อมูลการศึกษา ข้อมูลความเชี่ยวชาญ และนำคุณลักษณะจากตารางมาใช้งาน โดยข้อมูลที่ใช้ในการทดลองมีจำนวน 10 คุณลักษณะ ประกอบด้วย 1) อายุงาน 2) ตำแหน่ง 3) คุณวุฒิการศึกษา 4) คณะที่จบการศึกษา 5) สาขาที่จบ 6) ฝ่าย 7) กอง 8) แผนก 9) ระดับความเชี่ยวชาญ และ 10) ประเภทความเชี่ยวชาญ

### 1.1 การแปลงข้อมูล (Data Transformation)

ขั้นตอนของการแปลงข้อมูลจะทำการแปลงให้อยู่ในรูปแบบที่พร้อมจะนำมาใช้ในการวิเคราะห์ตามอัลกอริทึม โดยเริ่มจากขั้นตอนที่ 1 การแปลงข้อมูลที่เก็บรวบรวมมาที่ผ่านการทำความสะอาดข้อมูลแล้ว จัดข้อมูลให้อยู่ในรูปแบบของไฟล์ Excel (\*.xls) ขั้นตอนที่ 2 บันทึกไฟล์ Excel เป็นนามสกุล \*.csv เพื่อนำไปใช้งานในโปรแกรม WEKA (Waikato Environment for Knowledge Analysis) เวอร์ชัน 3.5.5 ที่เป็นเครื่องมือใช้ในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน

## 2. การจำแนกประเภทข้อมูล

ในขั้นตอนการจำแนกประเภทข้อมูล (Data Classification) เป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล ผู้วิจัยได้แบ่งคลาสความเชี่ยวชาญของผู้ปฏิบัติงาน เป็น 10 ประเภท ตามคุณลักษณะที่กำหนดโดยหน่วยงาน กฟผ. ดังนี้ 1) Turbine 2) Control Instrument 3) Water Treatment 4) ICT 5) Civil 6) Boiler 7) Electrical 8) Fuel Gas Desulfurization (FGD) 9) Coal Ash 10) Management และเลือกอัลกอริทึมที่ใช้ในการจำแนกประเภทข้อมูล ได้แก่ อัลกอริทึม Decision Tree C4.5 อัลกอริทึม OneR และ อัลกอริทึม Naïve Bayes โดยทำการเปรียบเทียบระหว่างอัลกอริทึมในประเด็นต่าง ๆ ดังนี้

1. ด้านความสามารถของอัลกอริทึมในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานได้ถูกต้อง

2. ด้านความสามารถของอัลกอริทึมในการค้นคืนคำตอบได้ถูกต้อง

3. ด้านความผิดพลาดของความสามารถของแต่ละอัลกอริทึมที่จะยอมรับได้

3. เปรียบเทียบประสิทธิภาพการทำงานของอัลกอริทึมที่ใช้ในการจำแนกข้อมูล

นำข้อมูลที่ได้จากขั้นตอนข้างต้น มาเปรียบเทียบว่าอัลกอริทึมใดให้ค่าความถูกต้องในการทำนายสูงสุด ระหว่าง Decision Tree (C4.5) Decision Rules (OneR) และ Bayes (Naïve Bayes) เพื่อเลือกอัลกอริทึมที่ดีที่สุดนำไปใช้สร้างแผนที่ความรู้ (Knowledge map) ขององค์กรต่อไป

เหตุผลประกอบในการเลือกแต่ละอัลกอริทึมมาใช้ในการวิจัย มีดังนี้

Decision Tree: อัลกอริทึม C4.5 เป็นอัลกอริทึมที่ใช้ในการสร้างต้นไม้การตัดสินใจที่พัฒนามาจากอัลกอริทึม ID3 เพราะต้นไม้ที่สร้างได้นั้น ได้ผ่านการพรวนให้ต้นไม้มีความถูกต้องและสมบูรณ์มากกว่าเดิม

Decision Rules: อัลกอริทึม OneR เป็นอัลกอริทึมที่ใช้ในการสร้างกฎที่สามารถแปลงมาเป็นต้นไม้การตัดสินใจได้ง่าย ซึ่งจะสามารถนำผลที่ได้จากการทดลองมาเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึมได้ เนื่องจากอยู่ในรูปของต้นไม้การตัดสินใจ และจากการตรวจสอบเอกสารยังไม่พบผู้วิจัยท่านใดนำ OneR อัลกอริทึมมาใช้ในการจำแนกข้อมูลด้านความเชี่ยวชาญต่างๆ ของผู้ปฏิบัติงาน

อัลกอริทึม Naïve Bayes เป็นอัลกอริทึมที่ใช้ในการหาความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ ให้ผลลัพธ์ที่ดีและรวดเร็ว จึงได้นำอัลกอริทึมนี้มาเป็นเครื่องมือที่ใช้ในการสร้างแบบจำลองและหารูปแบบความสัมพันธ์ที่ไม่ซับซ้อน และจะได้นำมาเปรียบเทียบกับ 2 อัลกอริทึมที่

กล่าวมาแล้วในข้างต้น ว่าอัลกอริทึมใดที่ให้ค่าความถูกต้องในการทำนายสูงสุด ในจำนวน 3 เทคนิคที่เลือกนำมาใช้ในงานวิจัย

ขั้นตอนที่ 3 นำผลการทดลองที่ได้จากโปรแกรม WEKA มาประมวลผลที่โปรแกรม MathLab และวิเคราะห์ข้อมูล ตามตัววัดประสิทธิภาพ ที่ผู้วิจัยได้เลือกมาใช้วัดผลการทดลอง จำนวน 4 ตัววัด ประกอบด้วย (1) วัดค่าความถูกต้องในการจำแนกด้วยค่า Accuracy (2) วัดค่าความถูกต้องในการค้นคืนด้วยค่า Precision, Recall (3) วัดค่าความถูกต้องโดยรวม ด้วยค่า F-measure (4) วัดค่าความผิดพลาดที่ยอมรับได้ ด้วยค่า Root Means-Squared Error จึงนับเป็นขั้นตอนที่สำคัญมาก เนื่องจากความถูกต้อง และความสมบูรณ์ของผลลัพธ์ขึ้นอยู่กับขั้นตอนนี้

#### 4. การหาความสัมพันธ์ของคุณลักษณะสำคัญของข้อมูล

นำข้อมูลที่ใช้ในการทดลองมาหาความสัมพันธ์คุณลักษณะของข้อมูลในชุดที่ 2 (จำนวน 805 ระเบียบ) เท่านั้น โดยใช้เทคนิค Association Rule เพื่อการค้นพบความสัมพันธ์ที่น่าสนใจในเซตของ Items ในแต่ละ Transaction ของฐานข้อมูล กระบวนการหา Association Rules ประกอบด้วย 2 ส่วนหลัก คือ

1. การหา Frequent Itemsets ที่มีค่าสนับสนุน (Support) มากกว่าหรือเท่ากับค่าความถี่ขั้นต่ำ (Minimum Support)

2. การสร้าง Association Rules จาก Frequent Itemsets ที่หาได้จากขั้นตอนที่ 1 และจะยอมรับ Association Rules ที่สร้างขึ้นได้ก็ต่อเมื่อกฎนี้มีค่าความเชื่อมั่น (Confidence) มากกว่าหรือเท่ากับค่าความเชื่อมั่นขั้นต่ำ (Minimum Confidence)

#### 5. แผนที่ความรู้

การสร้างแผนที่ความรู้องค์กร สามารถสร้างได้จากการหาความสัมพันธ์ของ คุณลักษณะข้อมูลที่เกิดขึ้นร่วมกันมาใช้ในการสร้างแผนที่ความรู้ เพื่อจำแนกความเชี่ยวชาญของ ผู้ปฏิบัติงาน

### ข้อมูลที่ใช้ในการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงาน

ตารางที่ 4 คุณลักษณะของข้อมูลที่ใช้ในการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงาน

| ลำดับที่ | คุณลักษณะข้อมูล     | ชื่อตัวแปร        | จำนวน (item) |
|----------|---------------------|-------------------|--------------|
| 1        | อายุงาน             | Working year      | 9            |
| 2        | ตำแหน่งงาน          | Position          | 12           |
| 3        | คุณวุฒิการศึกษา     | Education level   | 6            |
| 4        | คณะที่จบการศึกษา    | Faculty graduated | 12           |
| 5        | สาขาที่จบ           | Major             | 49           |
| 6        | ฝ่าย                | Division          | 28           |
| 7        | กอง                 | Department        | 45           |
| 8        | แผนก                | Section           | 78           |
| 9        | ระดับความเชี่ยวชาญ  | Expert Level      | 4            |
| 10       | ประเภทความเชี่ยวชาญ | Expert Class      | 10           |

ตารางที่ 5 คุณลักษณะของข้อมูลอายุงาน (Working years)

| ลำดับที่ | อายุงาน (ปี) |
|----------|--------------|
| 1        | 00 – 05      |
| 2        | 06 – 10      |
| 3        | 11 – 15      |
| 4        | 16 – 20      |
| 5        | 21 – 25      |
| 6        | 26 – 30      |
| 7        | 31 – 35      |
| 8        | 36 – 40      |
| 9        | 41 – 45      |

ตารางที่ 6 คุณลักษณะของข้อมูลตำแหน่งงาน (Position)

| ลำดับที่ | รหัสตำแหน่งงาน | ตำแหน่งงาน              |
|----------|----------------|-------------------------|
| 1        | 49800          | สถาปนิก                 |
| 2        | 49900          | วิศวกร                  |
| 3        | 44900          | วิทยาการ                |
| 4        | 33800          | พนักงานวิชาชีพ          |
| 5        | 22900          | พนักงานขับเครื่องจักรกล |
| 6        | 46900          | นิติกร                  |
| 7        | 48900          | นักวิทยาศาสตร์          |
| 8        | 45900          | นักบัญชี                |
| 9        | 49700          | นักธรณีวิทยา            |
| 10       | 44800          | นักคอมพิวเตอร์          |
| 11       | 23000          | ช่างชำนาญการ            |
| 12       | 33900          | ช่าง                    |

ตารางที่ 7 คุณลักษณะของข้อมูลคุณวุฒิการศึกษา (Education Level)

| ลำดับที่ | รหัสคุณวุฒิการศึกษา | คุณวุฒิการศึกษา |
|----------|---------------------|-----------------|
| 1        | 1                   | ปริญญาโท        |
| 2        | 2                   | ปริญญาตรี       |
| 3        | 3                   | ปวส.            |
| 4        | 4                   | ปวช.            |
| 5        | 5                   | ต่ำกว่า ปวช.    |
| 6        | 6                   | ไม่ระบุ         |

ตารางที่ 8 คุณลักษณะของข้อมูลคณะที่จบการศึกษา (Faculty graduated)

| ลำดับที่ | รหัสคณะที่จบการศึกษา | คณะที่จบการศึกษา  |
|----------|----------------------|-------------------|
| 1        | 052                  | บริหารธุรกิจ      |
| 2        | 057                  | บัญชี             |
| 3        | 301                  | อุตสาหกรรมศาสตร์  |
| 4        | 401                  | ครุศาสตร์         |
| 5        | 501                  | นิติศาสตร์        |
| 6        | 526                  | วิทยาศาสตร์       |
| 7        | 560                  | วิศวกรรมศาสตร์    |
| 8        | 589                  | ศิลปศาสตร์        |
| 9        | 601                  | เศรษฐศาสตร์       |
| 10       | 701                  | สถาปัตยกรรมศาสตร์ |
| 12       | 000                  | ไม่ระบุ           |

ตารางที่ 9 คุณลักษณะของข้อมูลสาขาที่จบการศึกษา (Major)

| ลำดับที่ | รหัสสาขาที่จบการศึกษา | สาขาที่จบการศึกษา   |
|----------|-----------------------|---------------------|
| 1        | 051                   | การบริหารทั่วไป     |
| 2        | 057                   | การบัญชี            |
| 3        | 157                   | คอมพิวเตอร์ศึกษา    |
| 4        | 161                   | คอมพิวเตอร์ธุรกิจ   |
| 5        | 168                   | เคมีปฏิบัติการ      |
| 6        | 170                   | เคมีอุตสาหกรรม      |
| 7        | 173                   | เครื่องกลอุตสาหกรรม |
| 8        | 195                   | ช่างกลโรงงาน        |
| 9        | 196                   | ช่างกลโลหะ          |
| 10       | 200                   | ช่างก่อสร้าง        |

ตารางที่ 9 (ต่อ)

| ลำดับที่ | รหัสสาขาที่จบการศึกษา | สาขาที่จบการศึกษา            |
|----------|-----------------------|------------------------------|
| 11       | 214                   | ช่างเครื่องกล                |
| 12       | 226                   | ช่างจักรกลหนัก               |
| 13       | 228                   | ช่างจิ๊กและทำเครื่องมือ      |
| 14       | 231                   | ช่างเชื่อมและโลหะแผ่น        |
| 15       | 247                   | ช่างท่อและประสาน             |
| 16       | 248                   | ช่างเทคนิคการผลิต            |
| 17       | 249                   | เทคนิคการผลิต                |
| 18       | 250                   | ช่างเคมี                     |
| 19       | 251                   | ช่างเทคนิคสถาปัตยกรรม        |
| 20       | 252                   | เทคนิคอุตสาหกรรม             |
| 21       | 253                   | ช่างเทคนิคอุตสาหกรรม         |
| 22       | 260                   | ช่างไฟฟ้า                    |
| 23       | 261                   | ช่างไฟฟ้ากำลัง               |
| 24       | 262                   | ช่างไฟฟ้าอุตสาหกรรม          |
| 25       | 269                   | ช่างยนต์                     |
| 26       | 273                   | ช่างโยธา                     |
| 27       | 277                   | ช่างโลหะ                     |
| 28       | 282                   | ช่างวิทยุและโทรคมนาคม        |
| 29       | 287                   | ช่างสำรวจ                    |
| 30       | 291                   | ช่างอิเล็กทรอนิกส์           |
| 31       | 295                   | เคมี                         |
| 32       | 327                   | เทคนิควิศวกรรมอิเล็กทรอนิกส์ |
| 33       | 346                   | เทคโนโลยีไฟฟ้าอุตสาหกรรม     |
| 34       | 347                   | เทคโนโลยีอิเล็กทรอนิกส์      |
| 35       | 363                   | เทคโนโลยีโยธา                |
| 36       | 392                   | บริหารธุรกิจ                 |

ตารางที่ 9 (ต่อ)

| ลำดับที่ | รหัสสาขาที่จบการศึกษา | สาขาที่จบการศึกษา    |
|----------|-----------------------|----------------------|
| 37       | 465                   | ไฟฟ้า                |
| 38       | 466                   | ไฟฟ้ากำลัง           |
| 39       | 516                   | เลขานุการ            |
| 40       | 529                   | คอมพิวเตอร์          |
| 41       | 546                   | วิศวกรรมเคมี         |
| 42       | 547                   | วิศวกรรมเครื่องกล    |
| 43       | 561                   | วิศวกรรมไฟฟ้า        |
| 44       | 566                   | วิศวกรรมไฟฟ้าสื่อสาร |
| 45       | 585                   | วิศวกรรมอุตสาหการ    |
| 46       | 602                   | เศรษฐศาสตร์การคลัง   |
| 47       | 697                   | ช่างเทคนิคโลหะ       |
| 48       | 706                   | เทคโนโลยีโครงสร้าง   |
| 49       | 000                   | ไม่ระบุ              |

ตารางที่ 10 คุณลักษณะของข้อมูลระดับฝ่าย (Division)

| ลำดับที่ | รหัสฝ่าย | ชื่อฝ่าย                                             |
|----------|----------|------------------------------------------------------|
| 1        | 9224000  | ฝ่ายบัญชีและงบประมาณสายงาน<br>ผลิตไฟฟ้าและเชื้อเพลิง |
| 2        | 9510000  | ผู้ช่วยผู้ว่าการ โรงไฟฟ้า 1                          |
| 3        | 9511000  | ฝ่ายการผลิต โรงไฟฟ้าพระนครใต้                        |
| 4        | 9512000  | ฝ่ายบำรุงรักษา โรงไฟฟ้าพระนครใต้                     |
| 5        | 9514000  | โรงไฟฟ้าน้ำพอง                                       |
| 6        | 9520000  | ผู้ช่วยผู้ว่าการ โรงไฟฟ้า 2                          |
| 7        | 9521000  | ฝ่ายการผลิต โรงไฟฟ้าแม่เมาะ                          |
| 8        | 9522000  | ฝ่ายบำรุงรักษา โรงไฟฟ้าแม่เมาะ                       |
| 9        | 9523000  | ฝ่ายธุรการ โรงไฟฟ้าแม่เมาะ                           |

ตารางที่ 10 (ต่อ)

| ลำดับที่ | รหัสฝ่าย | ชื่อฝ่าย                                       |
|----------|----------|------------------------------------------------|
| 10       | 9531000  | ฝ่ายการผลิตโรงไฟฟ้าบางปะกง                     |
| 11       | 9553000  | ฝ่ายจัดการธุรกิจบำรุงรักษา                     |
| 12       | 9532000  | ฝ่ายบำรุงรักษาโรงไฟฟ้าบางปะกง                  |
| 13       | 9533000  | โรงไฟฟ้าภาคใต้                                 |
| 14       | 9534000  | โรงไฟฟ้าจะนะ                                   |
| 15       | 9541000  | เขื่อนภูมิพล                                   |
| 16       | 9542000  | เขื่อนสิริกิติ์                                |
| 17       | 9543000  | เขื่อนศรีนครินทร์                              |
| 18       | 9544000  | เขื่อนวชิราลงกรณ์                              |
| 19       | 9545000  | เขื่อนรัชชประภา                                |
| 20       | 9546000  | โรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ           |
| 21       | 9550000  | ผู้ช่วยผู้ว่าการธุรกิจเดินเครื่องและบำรุงรักษา |
| 22       | 9561000  | ฝ่ายบำรุงรักษาเครื่องกล                        |
| 23       | 9572000  | ฝ่ายบริหารสายงานผลิตไฟฟ้า                      |
| 24       | 9574000  | ส่วนบริหารคุณภาพและเพิ่มผลผลิตสายงานผลิตไฟฟ้า  |
| 25       | 9612000  | ฝ่ายวิศวกรรมแหล่งพลังงาน                       |
| 26       | 9620000  | ผู้ช่วยผู้ว่าการปฏิบัติการเชื้อเพลิง           |
| 27       | 9622000  | ฝ่ายการผลิตเหมืองแม่เมาะ                       |
| 28       | 9621000  | ฝ่ายวางแผนและบริหารเหมืองแม่เมาะ               |

ตารางที่ 11 คุณลักษณะของข้อมูลระดับกอง (Department)

| ลำดับที่ | รหัสกอง | ชื่อกอง                                 |
|----------|---------|-----------------------------------------|
| 1        | 9224110 | กองงบประมาณสายงานผลิตไฟฟ้าและเชื้อเพลิง |
| 2        | 9324020 | กองบริการทั่วไป                         |
| 3        | 9421020 | กองโยธา                                 |
| 4        | 9511050 | กองการผลิตโรงไฟฟ้าลานกระบือ             |

ตารางที่ 11 (ต่อ)

| ลำดับที่ | รหัสกอง | ชื่อกอง                                     |
|----------|---------|---------------------------------------------|
| 5        | 9224110 | กองงบประมาณสายงานผลิตไฟฟ้าและเชื้อเพลิง     |
| 6        | 9324020 | กองบริการทั่วไป                             |
| 7        | 9421020 | กองโยธา                                     |
| 8        | 9511050 | กองการผลิตโรงไฟฟ้าลานกระบือ                 |
| 9        | 9512010 | กองบำรุงรักษาโรงไฟฟ้าพลังความร้อน           |
| 10       | 9512020 | กองบำรุงรักษาโรงไฟฟ้าพลังความร้อนร่วม 1     |
| 11       | 9512030 | กองบำรุงรักษาโรงไฟฟ้าพลังความร้อนร่วม 2     |
| 12       | 9513010 | กองการผลิตโรงไฟฟ้าพลังความร้อนร่วม          |
| 13       | 9521010 | กองการผลิต 1                                |
| 14       | 9521020 | กองการผลิต 2                                |
| 15       | 9521030 | กองปฏิบัติการระบบลำเลียงถ่านหินและยิปซัม    |
| 16       | 9521040 | กองปฏิบัติการระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ |
| 17       | 9521050 | กองเชื้อเพลิง ถ่าน และน้ำ                   |
| 18       | 9522010 | กองบำรุงรักษาโรงไฟฟ้า 1                     |
| 19       | 9522020 | กองบำรุงรักษาโรงไฟฟ้า 2                     |
| 20       | 9522030 | กองบำรุงรักษาโรงไฟฟ้า 3                     |
| 21       | 9522040 | กองบำรุงรักษาโรงไฟฟ้า 4                     |
| 22       | 9522050 | กองบำรุงรักษากลาง                           |
| 23       | 9531010 | กองการผลิตโรงไฟฟ้าพลังความร้อน              |
| 24       | 9531020 | กองการผลิตโรงไฟฟ้าพลังความร้อนร่วม 1        |
| 25       | 9531030 | กองการผลิตโรงไฟฟ้าพลังความร้อนร่วม 2        |
| 26       | 9532020 | กองบำรุงรักษาโรงไฟฟ้าพลังความร้อนร่วม       |
| 27       | 9532030 | กองบำรุงรักษาตามแผน                         |
| 28       | 9532040 | กองบำรุงรักษาอุปกรณ์สนับสนุนโรงไฟฟ้า        |
| 29       | 9532050 | กองพัฒนาระบบงานบำรุงรักษาและธุรกิจ          |
| 30       | 9533020 | กองการผลิตโรงไฟฟ้ากระบี่                    |
| 31       | 9534010 | กองเดินเครื่อง                              |

ตารางที่ 11 (ต่อ)

| ลำดับที่ | รหัสกอง | ชื่อกอง                                  |
|----------|---------|------------------------------------------|
| 32       | 9534020 | กองวางแผนการผลิตและบำรุงรักษา            |
| 33       | 9541020 | กองบำรุงรักษาโรงไฟฟ้า                    |
| 34       | 9541030 | กองบำรุงรักษาโยธา                        |
| 35       | 9545040 | กองโรงไฟฟ้าเขื่อนบางลาง                  |
| 36       | 9546040 | กองโรงไฟฟ้าเขื่อนอุบลรัตน์               |
| 37       | 9546050 | กองโรงไฟฟ้าเขื่อนสิรินธร                 |
| 38       | 9546060 | กองโรงไฟฟ้าเขื่อนจุฬาภรณ์                |
| 39       | 9561090 | กองแผนงานและบริการ                       |
| 40       | 9572060 | กองพัฒนาบุคลากรและระบบงานสายงานผลิตไฟฟ้า |
| 41       | 9572090 | กองเทคโนโลยีสารสนเทศสายงานผลิตไฟฟ้า      |
| 42       | 9612020 | กองวิศวกรรมเครื่องกลและไฟฟ้า             |
| 43       | 9621030 | กองวิศวกรรมธรณี                          |
| 44       | 9621040 | กองประมวลผลและวิเคราะห์                  |
| 45       | 9622040 | กองบำรุงรักษาไฟฟ้า                       |

ตารางที่ 12 คุณลักษณะของข้อมูลระดับแผนก (Section)

| ลำดับที่ | รหัสแผนก | ชื่อแผนก                                 |
|----------|----------|------------------------------------------|
| 1        | 9511011  | แผนกเดินเครื่องกะ 1                      |
| 2        | 9511012  | แผนกเดินเครื่องกะ 2                      |
| 3        | 9511013  | แผนกเดินเครื่องกะ 3                      |
| 4        | 9511014  | แผนกเดินเครื่องกะ 4                      |
| 5        | 9512012  | แผนกบำรุงรักษาหม้อน้ำ                    |
| 6        | 9512015  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัด |
| 7        | 9513021  | แผนกวางแผนบำรุงรักษา                     |
| 8        | 9521001  | แผนกสิ่งแวดล้อม                          |
| 9        | 9521015  | แผนกประสิทธิภาพ 1                        |

ตารางที่ 12 (ต่อ)

| ลำดับที่ | รหัสแผนก | ชื่อแผนก                                                            |
|----------|----------|---------------------------------------------------------------------|
| 10       | 9521016  | แผนกเคมีการผลิต                                                     |
| 11       | 9521025  | แผนกประสิทธิภาพ 2                                                   |
| 12       | 9521032  | แผนกปฏิบัติการระบบลำเลียงเถ้านหินและยิปซัม 2                        |
| 13       | 9521033  | แผนกปฏิบัติการระบบลำเลียงเถ้านหินและยิปซัม 3                        |
| 14       | 9521034  | แผนกปฏิบัติการระบบลำเลียงเถ้านหินและยิปซัม 4                        |
| 15       | 9521035  | แผนกวิเคราะห์ระบบลำเลียงเถ้านหินและยิปซัม                           |
| 16       | 9521041  | แผนกเดินเครื่องกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 1                         |
| 17       | 9521042  | แผนกเดินเครื่องกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 2                         |
| 18       | 9521043  | แผนกเดินเครื่องกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 3                         |
| 19       | 9521044  | แผนกเดินเครื่องกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 4                         |
| 20       | 9521046  | แผนกเคมีระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์                              |
| 21       | 9521052  | แผนกเดินอุปกรณ์สายพานและจัดหน้า 2                                   |
| 22       | 9521053  | แผนกเดินอุปกรณ์สายพานและจัดหน้า 3                                   |
| 23       | 9521054  | แผนกเดินอุปกรณ์สายพานและจัดหน้า 4                                   |
| 24       | 9522011  | แผนกบำรุงรักษาหม้อน้ำ 1                                             |
| 25       | 9522012  | แผนกบำรุงรักษาถังหิน 1                                              |
| 26       | 9522013  | แผนกบำรุงรักษาไฟฟ้า 1                                               |
| 27       | 9522014  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัด 1                          |
| 28       | 9522015  | แผนกบำรุงรักษาเครื่องกลระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 1             |
| 29       | 9522016  | แผนกบำรุงรักษาไฟฟ้าและอุปกรณ์ควบคุมระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 1 |
| 30       | 9522021  | แผนกบำรุงรักษาหม้อน้ำ 2                                             |
| 31       | 9522022  | แผนกบำรุงรักษาถังหิน 2                                              |
| 32       | 9522023  | แผนกบำรุงรักษาไฟฟ้า 2                                               |
| 33       | 9522024  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัด 2                          |

ตารางที่ 12 (ต่อ)

| ลำดับที่ | รหัสแผนก | ชื่อแผนก                                                            |
|----------|----------|---------------------------------------------------------------------|
| 34       | 9522025  | แผนกบำรุงรักษาเครื่องกลระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 2             |
| 35       | 9522026  | แผนกบำรุงรักษาไฟฟ้าและอุปกรณ์ควบคุมระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 2 |
| 36       | 9522031  | แผนกบำรุงรักษาหม้อน้ำ 3                                             |
| 37       | 9522032  | แผนกบำรุงรักษาถังหั่น 3                                             |
| 38       | 9522033  | แผนกบำรุงรักษาไฟฟ้า 3                                               |
| 39       | 9522034  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัด 3                          |
| 40       | 9522035  | แผนกบำรุงรักษาเครื่องกลระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 3             |
| 41       | 9522036  | แผนกบำรุงรักษาไฟฟ้าและอุปกรณ์ควบคุมระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ 3 |
| 42       | 9522042  | แผนกบำรุงรักษาเครื่องกลระบบลำเลียงขี้เถ้า                           |
| 43       | 9522044  | แผนกบำรุงรักษาไฟฟ้า 4                                               |
| 44       | 9522045  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัด 4                          |
| 45       | 9522051  | แผนกวางแผนบำรุงรักษากลาง                                            |
| 46       | 9522052  | แผนกบำรุงรักษาหม้อน้ำกลาง                                           |
| 47       | 9522053  | แผนกบำรุงรักษาถังหั่นกลาง                                           |
| 48       | 9522054  | แผนกบำรุงรักษาไฟฟ้ากลาง                                             |
| 49       | 9522055  | แผนกบำรุงรักษาอุปกรณ์ควบคุมและเครื่องวัดกลาง                        |
| 50       | 9522056  | แผนกโรงงานเครื่องกล                                                 |
| 51       | 9530002  | แผนกเทคโนโลยีสารสนเทศ                                               |
| 52       | 9531016  | แผนกเดินเครื่องกะ 6                                                 |
| 53       | 9531017  | แผนกเดินเครื่องกะ 7                                                 |
| 54       | 9531019  | แผนกประสิทธิภาพ                                                     |
| 55       | 9531041  | แผนกงานอาคาร                                                        |
| 56       | 9532011  | แผนกวางแผนบำรุงรักษาโรงไฟฟ้าพลังความร้อน                            |

ตารางที่ 12 (ต่อ)

| ลำดับที่ | รหัสแผนก | ชื่อแผนก                                |
|----------|----------|-----------------------------------------|
| 57       | 9532012  | แผนกบำรุงรักษาเครื่องกล                 |
| 58       | 9532013  | แผนกบำรุงรักษาไฟฟ้าและเครื่องวัดคุม     |
| 59       | 9532033  | แผนกบำรุงรักษาถังหั่น                   |
| 60       | 9532034  | แผนกบำรุงรักษาไฟฟ้า                     |
| 61       | 9532035  | แผนกบำรุงรักษาอุปกรณ์เครื่องวัดคุม      |
| 62       | 9532052  | แผนกวิศวกรรมบำรุงรักษา                  |
| 63       | 9532053  | แผนกพัฒนาบุคลากรและธุรกิจ               |
| 64       | 9533016  | แผนกบำรุงรักษาไฟฟ้าและอุปกรณ์ควบคุม     |
| 65       | 9533018  | แผนกประสิทธิภาพและวางแผน                |
| 66       | 9533019  | แผนกบำรุงรักษาโยธา                      |
| 67       | 9541031  | แผนกบำรุงรักษาอาคาร                     |
| 68       | 9543042  | แผนกบำรุงรักษาโรงไฟฟ้า                  |
| 69       | 9546041  | แผนกโรงไฟฟ้าเขื่อนอุบลรัตน์             |
| 70       | 9546042  | แผนกโรงไฟฟ้าเขื่อนน้ำพุง                |
| 71       | 9561092  | แผนกระบบงานและข้อมูล                    |
| 72       | 9564017  | แผนกโรงงานไฟฟ้า                         |
| 73       | 9572062  | แผนกอบรมและพัฒนาการจัดการ               |
| 74       | 9572092  | แผนกพัฒนาเครือข่ายระบบงานและศูนย์ข้อมูล |
| 75       | 9612022  | แผนกวิศวกรรมไฟฟ้า                       |
| 76       | 9621042  | แผนกควบคุมคอมพิวเตอร์                   |
| 77       | 9621043  | แผนกคอมพิวเตอร์สื่อสาร                  |
| 78       | 9622041  | แผนกบำรุงรักษาไฟฟ้าระบบขนส่งวัสดุ       |

ตารางที่ 13 คุณลักษณะของอายุความเชี่ยวชาญในงานที่ปฏิบัติ (Expert Level)

| ลำดับที่ | ระดับความเชี่ยวชาญ | ชื่อระดับความเชี่ยวชาญ |
|----------|--------------------|------------------------|
| 1        | 01 – 05 ปี         | Beginner               |
| 2        | 06 – 10 ปี         | Experience             |
| 3        | 11 – 15 ปี         | Skill                  |
| 4        | 16 – 20 ปีขึ้นไป   | Expert                 |

ตารางที่ 14 คุณลักษณะของข้อมูลประเภทความเชี่ยวชาญ (Expert Class)

| ลำดับที่ | ความเชี่ยวชาญ                              | รายละเอียด                                                                                                     |
|----------|--------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| 1        | Turbine                                    | งานด้านกังหันและอุปกรณ์ เช่น Condensor, ESP และ BFP                                                            |
| 2        | Control Instrument                         | งานด้านเครื่องมือวัดและควบคุม เช่น DCS, BMS และ ABC                                                            |
| 3        | Water Treatment                            | งานด้านระบบผลิตน้ำ งานด้านเคมี งานด้านสิ่งแวดล้อม                                                              |
| 4        | Information Communication Technology (ICT) | งานด้านระบบข้อมูลคอมพิวเตอร์ ระบบเครือข่ายและระบบสื่อสาร                                                       |
| 5        | Civil                                      | งานด้านก่อสร้าง บำรุงรักษาอาคาร บริเวณ เขื่อน อ่างเก็บน้ำ และคลองส่งน้ำ                                        |
| 6        | Boiler                                     | งานด้านหม้อน้ำและอุปกรณ์ เช่น PULV, SSC และ Soot                                                               |
| 7        | Electrical                                 | งานด้านระบบไฟฟ้า เช่น Generator, Motor และ EHG                                                                 |
| 8        | Fuel Gas Desulfurization (FGD)             | งานด้านเครื่องกำจัดกำมะถันเฟอร์ร่า และอุปกรณ์ เช่น Booster Fan, Absorber และ Slurry Pump                       |
| 9        | Coal Ash                                   | งานด้านระบบสายพานลำเลียงถ่าน จี้ถ่าน ยิปซัม งานจำหน่ายวัตถุพลอยได้                                             |
| 10       | Management                                 | งานด้านระบบบริหารจัดการ วางแผนบำรุงรักษา บริหารการผลิต ระบบมาตรฐานการจัดการ บริหารความปลอดภัย การจัดการความรู้ |

## ผลและวิจารณ์

### ผล

จากการทดลองทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน ในเบื้องต้นได้ทำการทดลองกับข้อมูลชุดที่ 1 จำนวน 343 ระเบียบ แสดงผลการทดลองในตารางที่ 15

**ตารางที่ 15** การวัดประสิทธิภาพของอัลกอริทึม C4.5 อัลกอริทึม OneR อัลกอริทึม Naïve Bayes แบบเฉลี่ยโดยการแบ่งข้อมูลแบบ 10-Fold Cross-Validation ของข้อมูลชุดที่ 1

| ตัววัดประสิทธิภาพ | C4.5   | OneR   | Naïve Bayes |
|-------------------|--------|--------|-------------|
| Percent correct   | 82.893 | 80.117 | 73.380      |
| Percent incorrect | 17.107 | 19.883 | 26.620      |
| Precision         | 0.904  | 0.636  | 0.739       |
| Recall            | 0.691  | 0.737  | 0.635       |
| F-measure         | 0.758  | 0.658  | 0.653       |
| RMSE              | 0.275  | 0.276  | 0.203       |

จากนั้นได้ทำการเพิ่มจำนวนข้อมูลที่มีขนาดใหญ่ขึ้น เป็นข้อมูลชุดที่ 2 ที่มีคุณลักษณะข้อมูลเหมือนกับข้อมูลชุดที่ 1 จำนวนข้อมูล 805 ระเบียบ จากการทดลองทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน ด้วย 3 อัลกอริทึม ประกอบด้วย Decision Tree (อัลกอริทึม C4.5) Decision Rules (อัลกอริทึม OneR) และอัลกอริทึม Naïve Bayes ในการวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม 3 แบบ ในชุดข้อมูลที่ 2 ใช้วิธีการแบ่งข้อมูลแบบ 10-fold Cross-Validation จำนวนรอบการทำงานเท่ากับ 100 รอบ สามารถแสดงผลที่ได้จากการทดลอง ดังตารางที่ 16-18

ตารางที่ 16 การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม C4.5 โดยกำหนด (n = 805)

| อัลกอริทึม C4.5 |                 |                   |              |              |              |              |
|-----------------|-----------------|-------------------|--------------|--------------|--------------|--------------|
| Fold            | Percent correct | Percent incorrect | Precision    | Recall       | F-measure    | RMSE         |
| 1               | 95.074          | 4.926             | 0.952        | <b>1.000</b> | 0.975        | 0.099        |
| 2               | 95.000          | 5.000             | 0.951        | 0.999        | 0.975        | 0.099        |
| 3               | 95.037          | 4.963             | 0.951        | <b>1.000</b> | 0.975        | 0.099        |
| 4               | 94.988          | 5.012             | 0.951        | 0.999        | 0.974        | 0.099        |
| 5               | 96.150          | 3.850             | 0.963        | 0.999        | 0.980        | 0.088        |
| 6               | 96.213          | 3.788             | 0.963        | 0.999        | <b>0.981</b> | <b>0.087</b> |
| 7               | 96.213          | 3.788             | 0.963        | <b>1.000</b> | <b>0.981</b> | <b>0.087</b> |
| 8               | 96.138          | 3.863             | 0.963        | 0.999        | 0.980        | 0.088        |
| 9               | <b>96.225</b>   | <b>3.775</b>      | <b>0.964</b> | 0.999        | <b>0.981</b> | <b>0.087</b> |
| 10              | 95.025          | 4.975             | 0.951        | 0.999        | 0.975        | 0.099        |
| เฉลี่ย          | 95.606          | 4.394             | 0.957        | 0.999        | 0.978        | 0.093        |

ตารางที่ 16 แสดงผลการทดลองในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน ด้วยอัลกอริทึม C4.5 ปรากฏว่า Fold ที่ 9 ให้ค่าความถูกต้องในการทำนายสูงสุด เท่ากับ 96.225% และมีค่าการทำนายไม่ถูกต้องต่ำสุด เท่ากับ 3.775% สำหรับค่า Precision จะเห็นว่าที่ Fold เท่ากับ 9 จะสามารถค้นคืนคำตอบได้มากที่สุด เท่ากับ 0.964 และ Recall ค่าที่เหมาะสมอยู่ที่ Fold 1, 3 และ 7 ค่าเท่ากับ 1.000 ค่า F-measure ซึ่งเป็นค่าที่แสดงความสัมพันธ์ระหว่างค่า precision และค่า recall เพื่อหาความถูกต้องโดยรวมมีค่ามากที่สุดอยู่ที่ Fold 6, 7 และ 9 ค่าเท่ากับ 0.981 ด้านค่าความผิดพลาดในการทำนายที่สามารถยอมรับได้ (RMSE) มีค่าต่ำสุดอยู่ที่ Fold 6, 7 และ 9 เท่ากับ 0.087

ตารางที่ 17 การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม OneR โดยกำหนด (n = 805)

| อัลกอริทึม OneR |                 |                   |              |              |              |              |
|-----------------|-----------------|-------------------|--------------|--------------|--------------|--------------|
| Fold            | Percent correct | Percent incorrect | Precision    | Recall       | F-measure    | RMSE         |
| 1               | 94.988          | 5.012             | 0.953        | <b>0.999</b> | 0.976        | 0.100        |
| 2               | 94.889          | 5.111             | 0.953        | 0.998        | 0.975        | 0.101        |
| 3               | 94.926          | 5.074             | 0.953        | <b>0.999</b> | 0.975        | 0.101        |
| 4               | 94.926          | 5.074             | 0.952        | <b>0.999</b> | 0.975        | 0.101        |
| 5               | <b>96.150</b>   | 3.850             | 0.964        | <b>0.999</b> | 0.981        | <b>0.088</b> |
| 6               | 96.075          | 3.925             | <b>0.965</b> | 0.998        | 0.981        | <b>0.088</b> |
| 7               | 96.138          | <b>3.863</b>      | <b>0.965</b> | <b>0.999</b> | <b>0.982</b> | <b>0.088</b> |
| 8               | 96.013          | 3.988             | 0.964        | 0.998        | 0.980        | 0.089        |
| 9               | 96.100          | 3.900             | <b>0.965</b> | 0.998        | 0.981        | <b>0.088</b> |
| 10              | 94.925          | 5.075             | 0.954        | <b>0.999</b> | 0.976        | 0.101        |
| เฉลี่ย          | 95.513          | 4.487             | 0.959        | <b>0.999</b> | 0.978        | 0.094        |

ตารางที่ 17 แสดงผลการทดลองในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม OneR ปรากฏว่า Fold ที่ 5 ให้ค่าความถูกต้องในการทำนายสูงสุด เท่ากับ 96.150% Fold ที่ 7 มีค่าการทำนายไม่ถูกต้องต่ำสุด เท่ากับ 3.863% สำหรับค่า Precision และ Recall จะเห็นว่าที่ Fold 6, 7 และ 9 จะสามารถค้นคืนคำตอบได้มากที่สุด เท่ากับ 0.965 และ Recall ค่าที่เหมาะสมอยู่ที่ Fold 1, 3, 4, 5, 7 และ 10 ค่าเท่ากับ 0.999 ค่า F-measure มีค่าสูงสุดที่ Fold 7 เท่ากับ 0.982 ด้านค่าความผิดพลาดในการทำนายที่สามารถยอมรับได้ (RMSE) มีค่าต่ำสุดอยู่ที่ Fold 5, 6, 7 และ 9 เท่ากับ 0.088

**ตารางที่ 18** การวัดประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม Naïve Bayes โดยกำหนด (n = 805)

| อัลกอริทึม Naïve Bayes |                 |                   |              |              |              |              |
|------------------------|-----------------|-------------------|--------------|--------------|--------------|--------------|
| Fold                   | Percent correct | Percent incorrect | Precision    | Recall       | F-measure    | RMSE         |
| 1                      | 91.235          | 8.765             | 0.953        | <b>0.958</b> | 0.955        | 0.122        |
| 2                      | 90.346          | 9.654             | 0.953        | 0.948        | 0.951        | 0.127        |
| 3                      | 90.420          | 9.580             | 0.952        | 0.950        | 0.951        | 0.125        |
| 4                      | 90.543          | 9.457             | 0.950        | 0.952        | 0.951        | 0.126        |
| 5                      | 91.412          | 8.588             | 0.964        | 0.949        | 0.957        | 0.117        |
| 6                      | 91.450          | 8.550             | 0.964        | 0.949        | 0.956        | 0.117        |
| 7                      | 91.450          | 8.550             | <b>0.965</b> | 0.949        | 0.956        | 0.118        |
| 8                      | 91.188          | 8.813             | 0.964        | 0.946        | 0.955        | 0.119        |
| 9                      | <b>91.800</b>   | <b>8.200</b>      | <b>0.965</b> | 0.952        | <b>0.958</b> | <b>0.116</b> |
| 10                     | 90.537          | 9.463             | 0.953        | 0.951        | 0.952        | 0.125        |
| เฉลี่ย                 | 91.038          | 8.962             | 0.958        | 0.950        | 0.954        | 0.121        |

ตารางที่ 18 แสดงผลการทดลองในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงาน ด้วย Naïve Bayes อัลกอริทึม ปรากฏว่า Fold ที่ 9 ให้ค่าความถูกต้องในการทำนายสูงสุด เท่ากับ 91.800% และมีค่าการทำนายไม่ถูกต้องต่ำสุด เท่ากับ 8.200% ค่า Precision ที่แสดงถึงการค้นคืนคำตอบที่ถูกต้อง จะเห็นได้ว่า Fold ที่ 9 สามารถค้นคืนได้มากที่สุด เท่ากับ 0.965 Recall ที่ Fold 1 มีค่าสูงสุด เท่ากับ 0.958 สำหรับค่าของ F-measure เพื่อหาความถูกต้องโดยรวมมีค่าสูงสุดที่ Fold 9 เท่ากับ 0.958 ด้านค่าความผิดพลาดในการทำนายที่สามารถยอมรับได้ (RMSE) มีค่าต่ำสุดอยู่ที่ Fold 9 เท่ากับ 0.116

**ตารางที่ 19** การวัดประสิทธิภาพของอัลกอริทึม C4.5 อัลกอริทึม OneR อัลกอริทึม Naïve Bayes  
แบบเฉลี่ยโดยการแบ่งข้อมูลแบบ 10-Fold Cross-Validation

| ตัววัดประสิทธิภาพ | C4.5          | OneR         | Naïve Bayes  |
|-------------------|---------------|--------------|--------------|
| Percent correct   | <b>95.606</b> | 95.513       | 91.038       |
| Percent incorrect | <b>4.394</b>  | 4.487        | <b>8.962</b> |
| Precision         | 0.957         | <b>0.959</b> | 0.958        |
| Recall            | <b>0.999</b>  | <b>0.999</b> | 0.950        |
| F-measure         | <b>0.978</b>  | <b>0.978</b> | 0.954        |
| RMSE              | <b>0.093</b>  | 0.094        | 0.121        |

ตารางที่ 19 แสดงผลการเปรียบเทียบประสิทธิภาพของทั้ง 3 อัลกอริทึม ได้แก่ C4.5 อัลกอริทึม OneR อัลกอริทึมและ Naïve Bayes อัลกอริทึม โดยเป็นการนำเสนอในลักษณะของการหาค่าเฉลี่ยของการแบ่งข้อมูลแบบ 10-Fold Cross-Validation ผลปรากฏว่าเปอร์เซ็นต์ความถูกต้อง (Percent correct) ในการทำนายสูงสุด ได้แก่ อัลกอริทึมของ Decision Trees คือ C4.5 เท่ากับ 95.606% รองลงมาได้แก่ Decision Rules คือ OneR อัลกอริทึม เท่ากับ 95.513% อัลกอริทึมที่ให้ค่าความถูกต้องในการทำนายต่ำสุด ได้แก่ Naïve Bayes เท่ากับ 91.038%

เปอร์เซ็นต์ของการทำนายไม่ถูกต้อง (Percent incorrect) สูงสุด ได้แก่ Naïve Bayes เท่ากับ 8.962% รองลงมาได้แก่ Decision Rules คือ อัลกอริทึม OneR เท่ากับ 4.487% อัลกอริทึมที่ให้ค่าความไม่ถูกต้องในการทำนายต่ำสุด ได้แก่ อัลกอริทึมของ Decision Trees คือ C4.5 เท่ากับ 4.394%

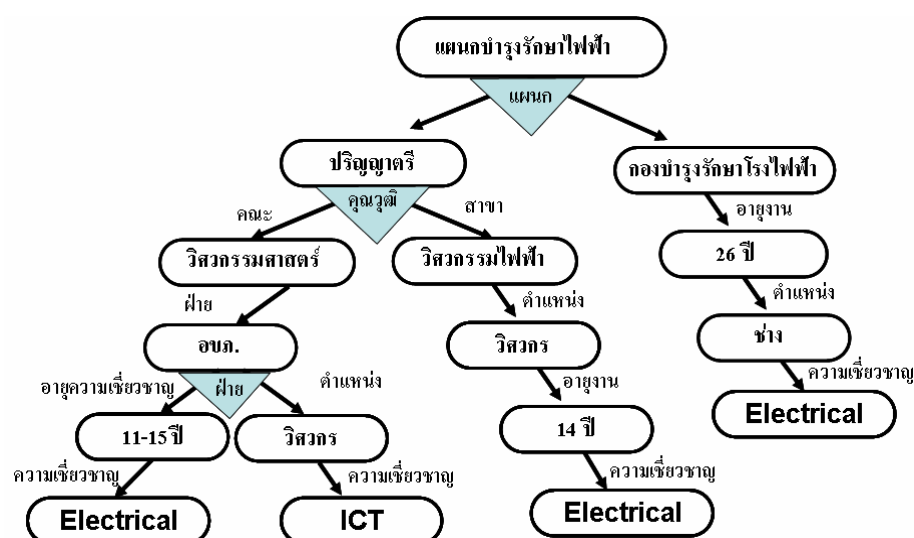
ค่า Precision ของอัลกอริทึมทั้ง 3 มีค่าใกล้เคียงกันมากจนไม่สามารถบอกความแตกต่างกันได้ ที่สามารถค้นคืนคำตอบ โดยอัลกอริทึม C4.5 ได้เท่ากับ 0.957 อัลกอริทึม OneR ได้เท่ากับ 0.959 และอัลกอริทึม Naïve Bayes ได้เท่ากับ 0.958

ค่า Recall ที่ได้จากการตรวจคืนข้อมูล ด้วยอัลกอริทึม Naïve Bayes มีค่าต่ำสุด เท่ากับ 0.950 สำหรับอัลกอริทึม C4.5 และอัลกอริทึม OneR มีค่าเท่ากัน คือ 0.999

ค่า F-measure เป็นค่าที่แสดงความสัมพันธ์ระหว่างค่า Precision และค่า Recall เพื่อหาความถูกต้องโดยรวมพบว่า Decision Trees C4.5 และ OneR อัลกอริทึม ให้ค่าสูงสุด เท่ากับ 0.978 และสุดท้าย คือ อัลกอริทึมของ Bayes (Naïve Bayes) ซึ่งเป็นอัลกอริทึมที่ให้ค่าความถูกต้องโดยรวมต่ำสุดในจำนวน 3 อัลกอริทึม เท่ากับ 0.954

ค่าความผิดพลาดในการทำนายที่สามารถยอมรับได้สูงสุด (RMSE) ได้แก่ อัลกอริทึมของ C4.5 มีค่าเท่ากับ 0.093 ตามมาด้วย OneR อัลกอริทึม มีค่าเท่ากับ 0.094 สุดท้าย คือ Bayes (Naïve Bayes) อัลกอริทึม เท่ากับ 0.121 นับเป็นอัลกอริทึมที่มีค่าความผิดพลาดสูงเป็นอันดับหนึ่ง

จากผลการทดลองด้วยข้อมูลชุดที่ 2 ขนาดข้อมูล 805 ระเบียบ สามารถแสดงภาพของต้นไม้การตัดสินใจที่สร้างได้จากอัลกอริทึม C4.5 ที่มีความถูกต้องในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานได้สูงสุด



ภาพที่ 12 ต้นไม้การตัดสินใจสำหรับความเชี่ยวชาญด้าน Electrical และ ICT ด้วยอัลกอริทึม C4.5

ภาพที่ 12 แสดงต้นไม้การตัดสินใจสำหรับความเชี่ยวชาญด้าน Electrical และ ICT ที่สร้างจาก C4.5 อัลกอริทึม จะเห็นได้ว่า โหนดราก เป็นคุณลักษณะข้อมูลที่เป็นสังกัดของผู้ปฏิบัติงาน ระดับแผนก และมีกิ่ง 2 กิ่ง ได้แก่ กิ่งด้านซ้ายเป็นคุณลักษณะข้อมูลของคุณวุฒิการศึกษา กิ่งทางด้านขวาเป็นคุณลักษณะข้อมูลของสังกัดผู้ปฏิบัติงานระดับ กอง ในแต่ละกิ่งจะแตกเป็นกิ่งย่อย ๆ ต่อลงไปจนถึงโหนดใบ ที่แสดงถึงคลาสความเชี่ยวชาญของผู้ปฏิบัติงาน นั่นเอง

**ตารางที่ 20** เปรียบเทียบจำนวนข้อมูลที่ใช้ในการทดลองด้วยตัววัดประสิทธิภาพของอัลกอริทึม  
C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes แบบเฉลี่ยโดยการแบ่งข้อมูลแบบ  
10-Fold Cross-Validation โดยใช้ข้อมูล 2 ชุด

| ตัววัด<br>ประสิทธิภาพ | อัลกอริทึม<br>C4.5 |                | อัลกอริทึม<br>OneR |                | อัลกอริทึม<br>Naïve Bayes |                |
|-----------------------|--------------------|----------------|--------------------|----------------|---------------------------|----------------|
|                       | ชุดที่ 1           | ชุดที่ 2       | ชุดที่ 1           | ชุดที่ 2       | ชุดที่ 1                  | ชุดที่ 2       |
|                       | 343<br>ระเบียน     | 805<br>ระเบียน | 343<br>ระเบียน     | 805<br>ระเบียน | 343<br>ระเบียน            | 805<br>ระเบียน |
| Percent correct       | 82.893             | <b>95.606</b>  | 80.117             | <b>95.513</b>  | 73.380                    | <b>91.038</b>  |
| Percent incorrect     | 17.107             | <b>4.394</b>   | 19.883             | <b>4.487</b>   | 26.620                    | <b>8.962</b>   |
| Precision             | 0.904              | <b>0.957</b>   | 0.636              | <b>0.959</b>   | 0.739                     | <b>0.958</b>   |
| Recall                | 0.691              | <b>0.999</b>   | 0.737              | <b>0.999</b>   | 0.635                     | <b>0.950</b>   |
| F-measure             | 0.758              | <b>0.978</b>   | 0.658              | <b>0.978</b>   | 0.653                     | <b>0.954</b>   |
| RMSE                  | 0.275              | <b>0.093</b>   | 0.276              | <b>0.094</b>   | 0.203                     | <b>0.121</b>   |

ตารางที่ 20 แสดงผลการเปรียบเทียบข้อมูล 2 ชุด ที่มีขนาดแตกต่างกันในการทดลองความถูกต้องในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม C4.5 พบว่า ชุดข้อมูลที่มีขนาดใหญ่ (ชุดข้อมูลที่ 2 ข้อมูล 805 ระเบียน) จะมีค่าความถูกต้องในการทำนายที่สูงกว่าชุดข้อมูลขนาดเล็กกว่า (ชุดข้อมูลที่ 1 ข้อมูล 343 ระเบียน) และค่าการทำนายไม่ถูกต้องในชุดข้อมูล 2 ชุดมีค่าที่แปรผกผันกัน คือ ข้อมูลชุดที่ 1 มีข้อมูลขนาดเล็ก จะมีค่าความไม่ถูกต้องในการทำนายที่สูงกว่าชุดข้อมูลที่ 2 ค่าเท่ากับ 17.107

การทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม OneR พบว่า ชุดข้อมูลที่มีขนาดใหญ่ (ชุดข้อมูลที่ 2 ข้อมูล 805 รายการ) จะมีค่าความถูกต้องในการทำนาย เท่ากับ 95.513 และค่าการทำนายไม่ถูกต้องในชุดข้อมูล 2 ชุด มีค่าที่แปรผกผันกัน คือ ข้อมูลชุดที่ 1 มีขนาดข้อมูลเล็ก จะมีค่าความไม่ถูกต้องในการทำนายที่สูงกว่าชุดข้อมูลที่ 2 ค่าเท่ากับ 19.883

การทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม Naive Bayes พบว่า ชุดข้อมูลที่มีขนาดใหญ่ (ชุดข้อมูลที่ 2 ข้อมูล 805 รายการ) จะมีค่าความถูกต้องในการทำนาย เท่ากับ 91.038 และค่าการทำนายไม่ถูกต้องในชุดข้อมูล 2 ชุดมีค่าที่แปรผกผันกัน คือ ข้อมูลชุดที่ 1 มีขนาดข้อมูลเล็ก จะมีค่าความไม่ถูกต้องในการทำนายที่สูงกว่าชุดข้อมูลที่ 2 ค่าเท่ากับ 26.620

การเปรียบเทียบประสิทธิภาพการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานด้วยอัลกอริทึม 3 แบบ คือ อัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes ด้วยชุดข้อมูล 2 ชุด จากการทดลองพบว่า ตัววัดประสิทธิภาพ Precision และ Recall ของข้อมูลชุดที่ 2 สามารถทำนายคำตอบได้ถูกต้องมากกว่าข้อมูลชุดที่ 1 ในทุกอัลกอริทึม ซึ่งแสดงถึงแนวโน้มของความถูกต้องในการทำนายคลาสข้อมูล ถ้าข้อมูลมีจำนวนมากขึ้น ความแม่นยำในการทำนายคำตอบจะมีประสิทธิภาพสูงขึ้น

สำหรับตัววัดประสิทธิภาพ F-Measure และ RMSE พบว่า ข้อมูลชุดที่ 2 สามารถค้นคืนคำตอบได้ถูกต้องมากกว่าข้อมูลชุดที่ 1 ในทุกอัลกอริทึม และตัววัดประสิทธิภาพ ค่าความผิดพลาดที่ยอมรับได้ (RMSE) จากการทดลองด้วยอัลกอริทึม 3 แบบ พบว่าชุดข้อมูลที่ 2 ที่มีขนาดข้อมูลใหญ่กว่าข้อมูลชุดที่ 1 จะมีค่าความผิดพลาดที่ยอมรับได้มากกว่า โดยผลการทดลองของชุดข้อมูลที่ 2 ด้วยอัลกอริทึม C4.5 และอัลกอริทึม OneR จะเห็นว่ามีค่า RMSE ที่ใกล้เคียงกันมาก คือ 0.093 และ 0.094 สำหรับอัลกอริทึม Naïve Bayes จะมีค่า RMSE ที่สูงกว่าอัลกอริทึมอื่น นั้นแสดงให้เห็นว่า อัลกอริทึม Naïve Bayes มีค่าความผิดพลาดที่ยอมรับได้น้อยกว่าอัลกอริทึม C4.5 และอัลกอริทึม OneR

### ผลการวิเคราะห์ความสัมพันธ์ของคุณลักษณะ ด้วย Apriori อัลกอริทึม ของ Association Rules

ในขั้นตอนการหาความสัมพันธ์ของการทำนายความถูกต้องในการจำแนกความเชี่ยวชาญในงานด้านต่าง ๆ จากคุณลักษณะข้อมูล จำนวน 10 คุณลักษณะ ตัววัดคือ ค่าสนับสนุน (Support) ค่าความเชื่อมั่น (Confidence) และกำหนดหน่วยวัดความสัมพันธ์ (Lift) มีค่ามากกว่า 1 ขึ้นไป ซึ่งสามารถหาความสัมพันธ์ของแต่ละคุณลักษณะ (Attributes) ในแต่ละประเภทความเชี่ยวชาญได้ดังต่อไปนี้

**ตารางที่ 21** กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Boiler (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.8)

| ที่ | เหตุการณ์                                                                  | สิ่งที่เกิดขึ้น<br>(Consequence)                                                      | ค่าสนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น | หน่วยวัด<br>ความสัมพันธ์ |
|-----|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------|----------------------|--------------------------|
| 1   | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์                                            | มีความเชี่ยวชาญ<br>ด้าน Boiler                                                        | 0.8                      | 0.98                 | 1.01                     |
| 2   | มีความเชี่ยวชาญด้าน<br>Boiler                                              | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์                                                       | 0.8                      | 0.86                 | 1.01                     |
| 3   | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์ และ<br>จบการศึกษาต่ำกว่า<br>ระดับปริญญาตรี | มีความเชี่ยวชาญ<br>ด้าน Boiler                                                        | 0.8                      | 1                    | 1                        |
| 4   | มีความเชี่ยวชาญด้าน<br>Boiler                                              | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์<br>และสาขาอื่นที่<br>คุณวุฒิต่ำกว่า<br>ระดับปริญญาตรี | 0.8                      | 0.86                 | 1                        |

ตารางที่ 21 แสดงให้เห็นถึงตัวอย่างกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Boiler จำนวน 4 กฎ จากจำนวนกฎทั้งสิ้น 10 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับคณะที่จบ

การศึกษาของผู้ปฏิบัติงาน คุณวุฒิการศึกษาและความเชี่ยวชาญในงานด้าน Boiler ดังเช่น กฎความสัมพันธ์ที่ 4 คือเมื่อผู้ปฏิบัติงานจบการศึกษาระดับวิศวกรรมศาสตร์และมีผู้ปฏิบัติงานที่จบการศึกษาต่ำกว่าระดับปริญญาตรี ถึงร้อยละ 80 จะมีความเชี่ยวชาญด้าน Boiler ตามกฎความสัมพันธ์ที่ 4 อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

ตารางที่ 22 กฎความสัมพันธ์ของความเชี่ยวชาญด้านโยธา Civil (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.9)

| ที่ | เหตุการณ์<br>(Antecedent)                                                             | สิ่งที่เกิดขึ้น<br>(Consequence)        | ค่าสนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------------------------------------|-----------------------------------------|--------------------------|--------------------------------------|------------------------------------|
| 1   | มีความเชี่ยวชาญด้าน<br>โยธา (Civil)                                                   | มีความ<br>เชี่ยวชาญงาน<br>ด้านเดียว     | 0.9                      | 1                                    | 1                                  |
| 2   | ผู้ปฏิบัติงานสังกัดแผนก<br>บำรุงรักษาอาคาร                                            | มีความ<br>เชี่ยวชาญด้าน<br>โยธา (Civil) | 0.9                      | 1                                    | 1                                  |
| 3   | แผนกบำรุงรักษาอาคาร                                                                   | มีความ<br>เชี่ยวชาญด้าน<br>โยธา (Civil) | 0.9                      | 1                                    | 1                                  |
| 4   | ผู้ปฏิบัติงานสังกัดแผนก<br>บำรุงรักษาอาคาร และมี<br>ความเชี่ยวชาญด้าน<br>โยธา (Civil) | เป็นความ<br>เชี่ยวชาญงาน<br>ด้านเดียว   | 0.9                      | 1                                    | 1                                  |

ตารางที่ 22 แสดงให้เห็นถึงตัวอย่างกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Civil จำนวน 4 กฎ จากจำนวนกฎทั้งสิ้น 10 กฎ ซึ่งกฎโดยส่วนใหญ่จะเกี่ยวข้องกับหน่วยงานที่ผู้ปฏิบัติงานสังกัดระดับแผนก และความเชี่ยวชาญด้าน Civil ดังเช่น กฎความสัมพันธ์ที่ 2 คือผู้ปฏิบัติงานสังกัดแผนกบำรุงรักษาอาคาร ร้อยละ 90 จะมีความเชี่ยวชาญด้านโยธา (Civil) ซึ่งมีโอกาสที่เป็นไปได้อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

ตารางที่ 23 กฎความสัมพันธ์ของความเชื่อชาญด้าน Coal Ash (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.95 )

| ที่ | เหตุการณ์<br>(Antecedent)                      | สิ่งที่เกิดขึ้น<br>(Consequence)                                                      | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|------------------------------------------------|---------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | จบการศึกษาคณะ<br>วิทยาศาสตร์                   | สังกัดผู้ช่วยผู้ว่าการ<br>ปฏิบัติการเชื้อเพลิง                                        | 0.95                         | 1                                    | 1.03                               |
| 2   | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี           | สังกัดฝ่ายการผลิต<br>โรงไฟฟ้า<br>แม่เมาะ                                              | 0.95                         | 1                                    | 1.03                               |
| 3   | จบการศึกษาคณะ<br>วิทยาศาสตร์                   | สังกัดผู้ช่วยผู้ว่าการ<br>ปฏิบัติการเชื้อเพลิง<br>และมีความเชี่ยวชาญ<br>ด้าน Coal Ash | 0.95                         | 1                                    | 1.03                               |
| 4   | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี           | สังกัดฝ่ายการผลิต<br>โรงไฟฟ้า<br>แม่เมาะ และมีความ<br>เชี่ยวชาญด้าน Coal<br>Ash       | 0.95                         | 1                                    | 1.03                               |
| 5   | สังกัดผู้ช่วยผู้ว่าการ<br>ปฏิบัติการเชื้อเพลิง | จบการศึกษาคณะ<br>วิทยาศาสตร์ และมี<br>ความเชี่ยวชาญด้าน<br>Coal Ash                   | 0.95                         | 0.97                                 | 1.03                               |
| 6   | สังกัดฝ่ายการผลิต<br>โรงไฟฟ้าแม่เมาะ           | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี และมี<br>ความเชี่ยวชาญด้าน<br>Coal Ash           | 0.95                         | 0.97                                 | 1.03                               |

ตารางที่ 23 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                                | สิ่งที่เกิดขึ้น<br>(Consequence)                   | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|--------------------------------------------------------------------------|----------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 7   | จบการศึกษาคณะ<br>วิทยาศาสตร์ และมี<br>ความเชี่ยวชาญด้าน<br>Coal Ash      | สังกัดผู้ช่วยผู้ว่า<br>การปฏิบัติการ<br>เชื้อเพลิง | 0.95                         | 1                                    | 1.03                               |
| 8   | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี และมีความ<br>เชี่ยวชาญด้าน Coal Ash | สังกัดฝ่ายการ<br>ผลิตโรงไฟฟ้า<br>แม่เมาะ           | 0.95                         | 1                                    | 1.03                               |

ตารางที่ 23 แสดงให้เห็นกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Coal Ash จำนวน 8 กฎ จากจำนวนกฎทั้งสิ้น 10 กฎ ซึ่งกฎโดยส่วนใหญ่จะเกี่ยวข้องกับคณะที่ผู้ปฏิบัติงานจบการศึกษา หน่วยงานที่ผู้ปฏิบัติงานสังกัดอยู่ ระดับฝ่าย และความเชี่ยวชาญด้าน Coal Ash กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานจบการศึกษาจากคณะวิทยาศาสตร์ จะสังกัดผู้ช่วยผู้ว่าการปฏิบัติการเชื้อเพลิง และมีผู้ปฏิบัติงานที่มีคุณสมบัติต่ำกว่าระดับปริญญาตรี สังกัดฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ กฎความสัมพันธ์ที่ 6 คือ ผู้ปฏิบัติงานสังกัดฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ จะมีคุณสมบัติต่ำกว่าระดับปริญญาตรี และมีความเชี่ยวชาญด้าน Coal Ash ร้อยละ 95 ซึ่งมีโอกาสที่เป็นไปได้ค่อนข้างแน่นอน โดยสามารถพิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 0.97

ตารางที่ 24 กฎความสัมพันธ์ของความเชื่อชาวยุคด้าน Control Instrument (กำหนดค่าสนับสนุนขั้นต่ำ เท่ากับ 0.75)

| ที่ | เหตุการณ์<br>(Antecedent)                   | สิ่งที่เกิดขึ้น<br>(Consequence)                                                               | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------|------------------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ตำแหน่งวิศวกร                               | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์<br>และมีความเชื่อชาวยุค<br>งานด้าน Control<br>Instrument       | 0.75                         | 1                                    | 1.31                               |
| 2   | ตำแหน่งช่าง                                 | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี                                                           | 0.75                         | 1                                    | 1.31                               |
| 3   | ตำแหน่งวิศวกร                               | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์                                                                | 0.75                         | 1                                    | 1.31                               |
| 4   | ตำแหน่งช่าง                                 | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี และมี<br>ความเชื่อชาวยุคงาน<br>ด้าน Control<br>Instrument | 0.75                         | 1                                    | 1.31                               |
| 5   | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์             | ตำแหน่งวิศวกร และ<br>มีความเชื่อชาวยุคด้าน<br>Control Instrument                               | 0.75                         | 1                                    | 1.31                               |
| 6   | มีคุณสมบัติต่ำกว่าระดับ<br>ปริญญาตรี        | ตำแหน่งช่าง และมี<br>ความเชื่อชาวยุคด้าน<br>Control Instrument                                 | 0.75                         | 1                                    | 1.31                               |
| 7   | มีความเชื่อชาวยุคด้าน<br>Control Instrument | ตำแหน่งเป็นวิศวกร<br>และช่าง                                                                   | 0.75                         | 0.77                                 | 1                                  |

ตารางที่ 24 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                | สิ่งที่เกิดขึ้น<br>(Consequence)                                                        | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|----------------------------------------------------------|-----------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 8   | มีความเชี่ยวชาญด้าน<br>Control Instrument<br>และ Turbine | จบการศึกษาคณะ<br>วิศวกรรมศาสตร์<br>และสาขาอื่นที่มี<br>คุณวุฒิต่ำกว่าระดับ<br>ปริญญาตรี | 0.75                         | 0.77                                 | 1                                  |

ตารางที่ 24 แสดงให้เห็นกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Control Instrument จำนวน 8 กฎ จากจำนวนกฎทั้งสิ้น 12 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับ ตำแหน่งงาน คุณวุฒิต่ำกว่าระดับปริญญาตรี คณะที่จบการศึกษาและความเชี่ยวชาญด้าน Control Instrument ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานมีตำแหน่งเป็นวิศวกร จบการศึกษาคณะ วิศวกรรมศาสตร์ และมีความเชี่ยวชาญด้าน Control Instrument ร้อยละ 75 โอกาสที่ผู้ปฏิบัติงานมี ตำแหน่งเป็นวิศวกรจะจบการศึกษาคณะวิศวกรรมศาสตร์และมีความเชี่ยวชาญด้าน Control Instrument มีความเป็นไปได้อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

ตารางที่ 25 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Electrical (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.75)

| ที่ | เหตุการณ์<br>(Antecedent)                               | สิ่งที่เกิดขึ้น<br>(Consequence)               | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------|------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ตำแหน่งเป็นช่างและมี<br>ความเชี่ยวชาญด้าน<br>Electrical | มีคุณวุฒิ<br>การศึกษาต่ำกว่า<br>ระดับปริญญาตรี | 0.75                         | 1                                    | 1.27                               |

ตารางที่ 25 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                                       | สิ่งที่เกิดขึ้น<br>(Consequence)                                                             | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 2   | มีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี                                          | ตำแหน่งเป็นช่าง                                                                              | 0.75                         | 0.98                                 | 1.27                               |
| 3   | ตำแหน่งเป็นวิศวกร                                                               | มีคุณวุฒิ<br>การศึกษาระดับ<br>ปริญญาตรี และมี<br>ความเชี่ยวชาญ<br>ด้าน Electrical<br>และ ICT | 0.75                         | 1                                    | 1.27                               |
| 4   | มีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี<br>และมีความเชี่ยวชาญ<br>ด้าน Electrical | ตำแหน่งเป็นช่าง                                                                              | 0.75                         | 0.98                                 | 1.27                               |
| 5   | มีคุณวุฒิการศึกษาระดับ<br>ปริญญาตรี                                             | ตำแหน่งเป็น<br>วิศวกร และมี<br>ความเชี่ยวชาญ<br>ด้าน ICT                                     | 0.75                         | 0.98                                 | 1.27                               |
| 6   | ตำแหน่งเป็นช่าง                                                                 | มีคุณวุฒิ<br>การศึกษาต่ำกว่า<br>ระดับปริญญาตรี<br>และมีความ<br>เชี่ยวชาญด้าน<br>Electrical   | 0.75                         | 1                                    | 1.27                               |
| 7   | มีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี                                          | มีความเชี่ยวชาญ<br>ด้าน Electrical                                                           | 0.75                         | 1                                    | 1.02                               |

ตารางที่ 25 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                                | สิ่งที่เกิดขึ้น<br>(Consequence) | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|--------------------------------------------------------------------------|----------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 8   | มีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี และมีความเชี่ยวชาญด้าน Electrical | มีความเชี่ยวชาญด้านเดียว         | 0.75                         | 1                                    | 1.02                               |

ตารางที่ 25 แสดงให้เห็นกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Electrical จำนวน 8 กฎ จากจำนวนกฎทั้งสิ้น 14 กฎ ซึ่งกฎโดยส่วนใหญ่จะเกี่ยวข้องกับตำแหน่งงาน คุณวุฒิต่ำกว่าระดับปริญญาตรี และความเชี่ยวชาญด้าน Electrical ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานมีตำแหน่งเป็นช่าง มีความเชี่ยวชาญด้าน Electrical และมีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี ร้อยละ 75 มีโอกาสที่ผู้ปฏิบัติงานที่มีตำแหน่งเป็นช่าง มีความเชี่ยวชาญด้าน Electrical จะมีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

ตารางที่ 26 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Fuel Gas Desulfurization (FGD) (กำหนดค่าสนับสนุนขั้นต่ำ เท่ากับ 0.8)

| ที่ | เหตุการณ์<br>(Antecedent)      | สิ่งที่เกิดขึ้น<br>(Consequence) | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|--------------------------------|----------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ผู้ปฏิบัติงานมีตำแหน่งเป็นช่าง | มีคุณวุฒิต่ำกว่าระดับปริญญาตรี   | 0.8                          | 1                                    | 1.15                               |
| 2   | มีคุณวุฒิต่ำกว่าระดับปริญญาตรี | ผู้ปฏิบัติงานมีตำแหน่งเป็นช่าง   | 0.8                          | 0.97                                 | 1.15                               |

ตารางที่ 26 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                           | สิ่งที่เกิดขึ้น<br>(Consequence)                                        | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------------------|-------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 3   | ผู้ปฏิบัติงานมีตำแหน่ง<br>เป็นช่าง                                  | มีคุณวุฒิต่ำกว่า<br>ระดับปริญญาตรี<br>และมีความ<br>เชี่ยวชาญด้าน<br>FGD | 0.8                          | 1                                    | 1.15                               |
| 4   | ผู้ปฏิบัติงานมีตำแหน่ง<br>เป็นช่าง และมีความ<br>เชี่ยวชาญ FGD       | มีคุณวุฒิต่ำกว่า<br>ระดับปริญญาตรี                                      | 0.8                          | 1                                    | 1.15                               |
| 5   | ผู้ปฏิบัติงานมีตำแหน่ง<br>เป็นช่าง                                  | มีความเชี่ยวชาญ<br>ด้าน FGD                                             | 0.8                          | 1                                    | 1                                  |
| 6   | ผู้ปฏิบัติงานมีตำแหน่ง<br>เป็นช่าง และมีคุณวุฒิต่ำ<br>กว่าปริญญาตรี | มีความเชี่ยวชาญ<br>ด้าน FGD                                             | 0.8                          | 1                                    | 1                                  |

ตารางที่ 26 แสดงให้เห็นกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน FGD ที่เกี่ยวกับงานด้านการจำกัดก๊าซซัลเฟอร์ไดออกไซด์ จำนวน 6 กฎ จากจำนวนกฎทั้งสิ้น 15 กฎ ซึ่งกฎโดยส่วนใหญ่จะเกี่ยวข้องกับตำแหน่งงาน คุณวุฒิการศึกษา และความเชี่ยวชาญด้าน FGD กฎความสัมพันธ์ที่ 4 คือ ผู้ปฏิบัติงานมีตำแหน่งเป็นช่าง และมีความเชี่ยวชาญด้าน FGD จะมีคุณวุฒิต่ำกว่าระดับปริญญาตรี ร้อยละ 80 มีโอกาสที่ผู้ปฏิบัติงานที่มีตำแหน่งเป็นช่าง และมีความเชี่ยวชาญด้าน FGD จะมีคุณวุฒิต่ำกว่าระดับปริญญาตรีอย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ

ตารางที่ 27 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน ICT (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.65)

| ที่ | เหตุการณ์<br>(Antecedent)                | สิ่งที่เกิดขึ้น<br>(Consequence)      | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|------------------------------------------|---------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | กองบำรุงรักษา<br>โรงไฟฟ้า 1              | มีความเชี่ยวชาญ<br>ด้าน ICT           | 0.65                         | 1                                    | 1                                  |
| 2   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน ICT | สังกัดกอง<br>บำรุงรักษา<br>โรงไฟฟ้า 1 | 0.65                         | 0.71                                 | 1                                  |

ตารางที่ 27 แสดงให้เห็นกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน ICT จำนวน 2 กฎ จากจำนวนกฎทั้งสิ้น 4 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับหน่วยงานที่ผู้ปฏิบัติงานสังกัดอยู่ระดับกอง และความเชี่ยวชาญด้าน ICT ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานที่สังกัดกองบำรุงรักษาโรงไฟฟ้าจะมีความเชี่ยวชาญด้าน ICT ร้อยละ 65 กฎนี้มีความเป็นไปได้อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่น เท่ากับ 1

ตารางที่ 28 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Management (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.5)

| ที่ | เหตุการณ์<br>(Antecedent)                             | สิ่งที่เกิดขึ้น<br>(Consequence)                   | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|-------------------------------------------------------|----------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ผู้ปฏิบัติงานจบ<br>การศึกษาระดับ ปวช.<br>และปริญญาตรี | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน<br>Management | 0.5                          | 1                                    | 1                                  |

ตารางที่ 28 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                                                   | สิ่งที่เกิดขึ้น<br>(Consequence)                                        | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 2   | ผู้ปฏิบัติงานจบ<br>การศึกษาระดับ ปวช.<br>และปริญญาตรี                                       | มีความเชี่ยวชาญด้าน<br>Management และ<br>เป็นความเชี่ยวชาญ<br>ด้านเดียว | 0.5                          | 0.92                                 | 1.01                               |
| 3   | ผู้ปฏิบัติงานจบ<br>การศึกษาระดับ ปวช.<br>และปริญญาตรี มี<br>ความเชี่ยวชาญด้าน<br>Management | มีความเชี่ยวชาญด้าน<br>เดียว                                            | 0.5                          | 0.92                                 | 1.01                               |
| 4   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน<br>Management เป็น<br>ความเชี่ยวชาญด้าน<br>เดียว       | ผู้ปฏิบัติงานจบ<br>การศึกษาระดับ<br>ปวช. และปริญญา<br>ตรี               | 0.5                          | 0.58                                 | 1.01                               |
| 5   | ผู้ปฏิบัติงานจบ<br>การศึกษาระดับ ปวช.<br>และปริญญาตรี                                       | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน<br>Management                      | 0.5                          | 1                                    | 1                                  |

ตารางที่ 28 แสดงให้เห็นถึงกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Management จำนวน 5 กฎ จากจำนวนกฎทั้งสิ้น 10 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับคุณวุฒิการศึกษา และความเชี่ยวชาญด้าน Management ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานมีคุณวุฒิ การศึกษาระดับ ปวช. และปริญญาตรี โดยจะมีความเชี่ยวชาญด้าน Management ร้อยละ 50 และมี โอกาสที่ผู้ปฏิบัติงานที่มีคุณวุฒิการศึกษาระดับ ปวช. และปริญญาตรี จะมีความเชี่ยวชาญด้าน Management มีโอกาสความเป็นไปได้ โดยพิจารณาจากค่าความเชื่อมั่น เท่ากับ 1

ตารางที่ 29 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Turbine (กำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.8 )

| ที่ | เหตุการณ์<br>(Antecedent)                                                                 | สิ่งที่เกิดขึ้น<br>(Consequence)                                                     | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน Control<br>Instrument                                | มีความเชี่ยวชาญ<br>ด้าน Turbine                                                      | 0.8                          | 1                                    | 1                                  |
| 2   | ผู้ปฏิบัติงานมีคุณสมบัติ<br>ต่ำกว่าระดับปริญญาตรี                                         | มีความเชี่ยวชาญ<br>ด้าน Turbine                                                      | 0.8                          | 1                                    | 1                                  |
| 3   | ผู้ปฏิบัติงานมีคุณสมบัติ<br>ต่ำกว่าระดับปริญญาตรี                                         | มีความเชี่ยวชาญ<br>ด้าน Control<br>Instrument เป็น<br>ความเชี่ยวชาญ<br>ด้านเดียว     | 0.8                          | 0.98                                 | 1                                  |
| 4   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน Control<br>Instrument เป็นความ<br>เชี่ยวชาญด้านเดียว | ผู้ปฏิบัติงานมี<br>คุณสมบัติต่ำกว่า<br>ระดับปริญญาตรี                                | 0.8                          | 0.83                                 | 1                                  |
| 5   | ผู้ปฏิบัติงานมีคุณสมบัติ<br>ต่ำกว่าระดับปริญญาตรี                                         | มีความเชี่ยวชาญ<br>ด้าน Turbine และ<br>มีความเชี่ยวชาญ<br>ด้าน Control<br>Instrument | 0.8                          | 0.98                                 | 1                                  |
| 6   | ผู้ปฏิบัติงานมีคุณสมบัติ<br>ต่ำกว่าระดับปริญญาตรี<br>และมีความเชี่ยวชาญ<br>ด้าน Turbine   | มีความเชี่ยวชาญ<br>ด้าน Control<br>Instrument                                        | 0.8                          | 0.98                                 | 1                                  |

ตารางที่ 29 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                    | สิ่งที่เกิดขึ้น<br>(Consequence)                                                 | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|----------------------------------------------|----------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 7   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน Turbine | มีความเชี่ยวชาญ<br>ด้าน Control<br>Instrument เป็น<br>ความเชี่ยวชาญด้าน<br>เดียว | 0.8                          | 0.98                                 | 1                                  |

ตารางที่ 29 แสดงให้เห็นถึงกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Turbine จำนวน 7 กฎ จากจำนวนกฎทั้งสิ้น 12 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับคุณวุฒิการศึกษา และความเชี่ยวชาญด้าน Turbine และ Control Instrument ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานมีความเชี่ยวชาญด้าน Control Instrument และจะมีความเชี่ยวชาญด้าน Turbine ด้วย ร้อยละ 80 และมีโอกาสความเป็นไปได้ โดยพิจารณาจากค่าความเชื่อมั่น เท่ากับ 1

ตารางที่ 30 กฎความสัมพันธ์ของความเชี่ยวชาญด้าน Water Treatment (กำหนดค่าสนับสนุนขั้นต่ำ เท่ากับ 0.9)

| ที่ | เหตุการณ์<br>(Antecedent)                             | สิ่งที่เกิดขึ้น<br>(Consequence)           | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|-------------------------------------------------------|--------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 1   | ผู้ปฏิบัติงานสังกัดฝ่าย<br>การผลิตโรงไฟฟ้า<br>แม่เมาะ | มีความเชี่ยวชาญ<br>ด้าน Water<br>Treatment | 0.9                          | 1                                    | 1                                  |

ตารางที่ 30 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                                                 | สิ่งที่เกิดขึ้น<br>(Consequence)                               | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------------------------|----------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 2   | ผู้ปฏิบัติงานมีคุณสมบัติต่ำกว่าระดับปริญญาตรี                             | มีความเชี่ยวชาญด้าน Water Treatment                            | 0.9                          | 1                                    | 1                                  |
| 3   | ผู้ปฏิบัติงานสังกัดฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ และมีความเชี่ยวชาญด้านเดียว | มีความเชี่ยวชาญด้าน Water Treatment                            | 0.9                          | 1                                    | 1                                  |
| 4   | ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ                                                | มีความเชี่ยวชาญด้านเดียว                                       | 0.9                          | 0.93                                 | 1.04                               |
| 5   | ผู้ปฏิบัติงานมีความเชี่ยวชาญด้านเดียว                                     | สังกัดฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ                               | 0.9                          | 1                                    | 1.04                               |
| 6   | ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ                                                | มีความเชี่ยวชาญด้าน Water Treatment เป็นความเชี่ยวชาญด้านเดียว | 0.9                          | 0.93                                 | 1.04                               |
| 7   | ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ และมีความเชี่ยวชาญด้าน Water Treatment         | มีความเชี่ยวชาญด้านเดียว                                       | 0.9                          | 0.93                                 | 1.04                               |

ตารางที่ 30 (ต่อ)

| ที่ | เหตุการณ์<br>(Antecedent)                               | สิ่งที่เกิดขึ้น<br>(Consequence)                                                       | ค่า<br>สนับสนุน<br>(Support) | ค่าความ<br>เชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|---------------------------------------------------------|----------------------------------------------------------------------------------------|------------------------------|--------------------------------------|------------------------------------|
| 8   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้านเดียว               | สังกัดฝ่ายการ<br>ผลิตโรงไฟฟ้าแม่<br>เมาะ และมีความ<br>เชี่ยวชาญด้าน<br>Water Treatment | 0.9                          | 1                                    | 1.04                               |
| 9   | ผู้ปฏิบัติงานมีความ<br>เชี่ยวชาญด้าน Water<br>Treatment | สังกัดฝ่ายการ<br>ผลิตโรงไฟฟ้า<br>แม่เมาะ                                               | 0.9                          | 1                                    | 1.04                               |

ตารางที่ 30 แสดงให้เห็นถึงกฎความสัมพันธ์ของประเภทความเชี่ยวชาญด้าน Water Treatment จำนวน 9 กฎ จากจำนวนกฎทั้งสิ้น 14 กฎ ซึ่งกฎโดยส่วนใหญ่นั้นจะเกี่ยวข้องกับหน่วยงานที่ผู้ปฏิบัติงานสังกัดระดับฝ่าย คุณวุฒิการศึกษาและความเชี่ยวชาญด้าน Water Treatment ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานที่สังกัดฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ จะมีความเชี่ยวชาญด้าน Water Treatment ร้อยละ 90 โอกาสที่ผู้ปฏิบัติงานจะมีความเป็นด้าน Water Treatment มีความเป็นไปได้อย่างแน่นอน พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

ตารางที่ 31 กฎความสัมพันธ์ของความเชี่ยวชาญทุกประเภท (กำหนดค่าสนับสนุนขั้นต่ำ เท่ากับ 0.15)

| ที่ | เหตุการณ์<br>(Antecedent)          | สิ่งที่เกิดขึ้น<br>(Consequence)                                          | ค่า<br>สนับสนุน<br>(Support) | ค่าความเชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|------------------------------------|---------------------------------------------------------------------------|------------------------------|----------------------------------|------------------------------------|
| 1   | ผู้ปฏิบัติงานมีอายุงาน<br>21-25 ปี | จะมีตำแหน่งเป็น<br>ช่าง และมีคุณวุฒิ<br>การศึกษาต่ำกว่า<br>ระดับปริญญาตรี | 0.3                          | 1                                | 1.27                               |

ตารางที่ 31 (ต่อ)

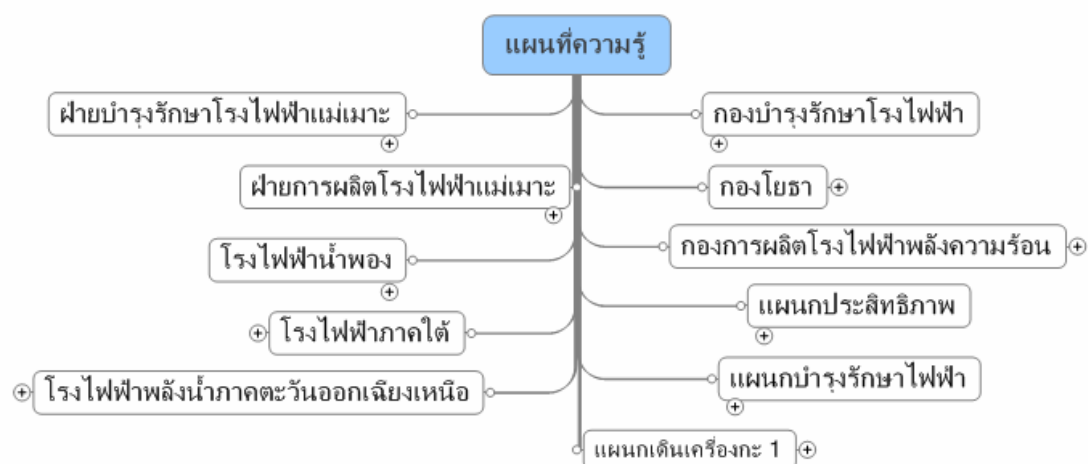
| ที่ | เหตุการณ์<br>(Antecedent)                                                                             | สิ่งที่เกิดขึ้น<br>(Consequence)                                                       | ค่า<br>สนับสนุน<br>(Support) | ค่าความเชื่อมั่น<br>(Confidence) | หน่วยวัด<br>ความสัมพันธ์<br>(Lift) |
|-----|-------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|------------------------------|----------------------------------|------------------------------------|
| 2   | ผู้ปฏิบัติงานมีคุณวุฒิ<br>การศึกษาระดับ ปวส.<br>และปฏิบัติงานที่ฝ่าย<br>บำรุงรักษาโรงไฟฟ้า<br>แม่เมาะ | จะมีตำแหน่งเป็น<br>ช่าง                                                                | 0.2                          | 1                                | 1.41                               |
| 3   | ผู้ปฏิบัติงานมีคุณวุฒิ<br>การศึกษาระดับ ปวส.                                                          | จะมีตำแหน่งเป็น<br>ช่าง                                                                | 0.4                          | 0.98                             | 1.38                               |
| 4   | ผู้ปฏิบัติงานมีคุณวุฒิ<br>การศึกษาระดับ ปวช.                                                          | จะมีตำแหน่งเป็น<br>ช่าง                                                                | 0.15                         | 0.96                             | 1.35                               |
| 5   | ผู้ปฏิบัติงานมีอายุงาน<br>21-25 ปี                                                                    | เป็นผู้มีคุณวุฒิ<br>การศึกษาระดับ<br>ปวส.                                              | 0.15                         | 0.87                             | 1.94                               |
| 6   | ผู้ปฏิบัติงานมีตำแหน่ง<br>เป็นช่าง                                                                    | จะมีอายุงาน 21-25<br>ปี                                                                | 0.3                          | 0.45                             | 1.23                               |
| 7   | ผู้ปฏิบัติงานมีคุณวุฒิ<br>การศึกษาต่ำกว่าระดับ<br>ปริญญาตรี                                           | จะตำแหน่งเป็น<br>ช่างและปฏิบัติงาน<br>ที่หน่วยงานฝ่าย<br>บำรุงรักษา<br>โรงไฟฟ้าแม่เมาะ | 0.3                          | 0.44                             | 1.27                               |

ตารางที่ 31 แสดงให้เห็นถึงกฎความสัมพันธ์ของความเชี่ยวชาญทุกประเภท จำนวน 8 กฎ โดยกฎที่ได้ส่วนใหญ่จะเกี่ยวข้องกับอายุงาน ตำแหน่งงาน คุณวุฒิการศึกษา หน่วยงานที่ผู้ปฏิบัติงานสังกัดระดับฝ่าย ดังเช่น กฎความสัมพันธ์ที่ 1 คือ ผู้ปฏิบัติงานมีอายุงาน 21-25 ปี จะมีตำแหน่งเป็นช่าง และมีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี ร้อยละ 30 โอกาสที่ผู้ปฏิบัติงานที่มี

ตำแหน่งเป็นช่าง และมีคุณวุฒิการศึกษาต่ำกว่าระดับปริญญาตรี มีความเป็นไปได้อย่างแน่นอน  
พิจารณาได้จากค่าความเชื่อมั่นที่เท่ากับ 1

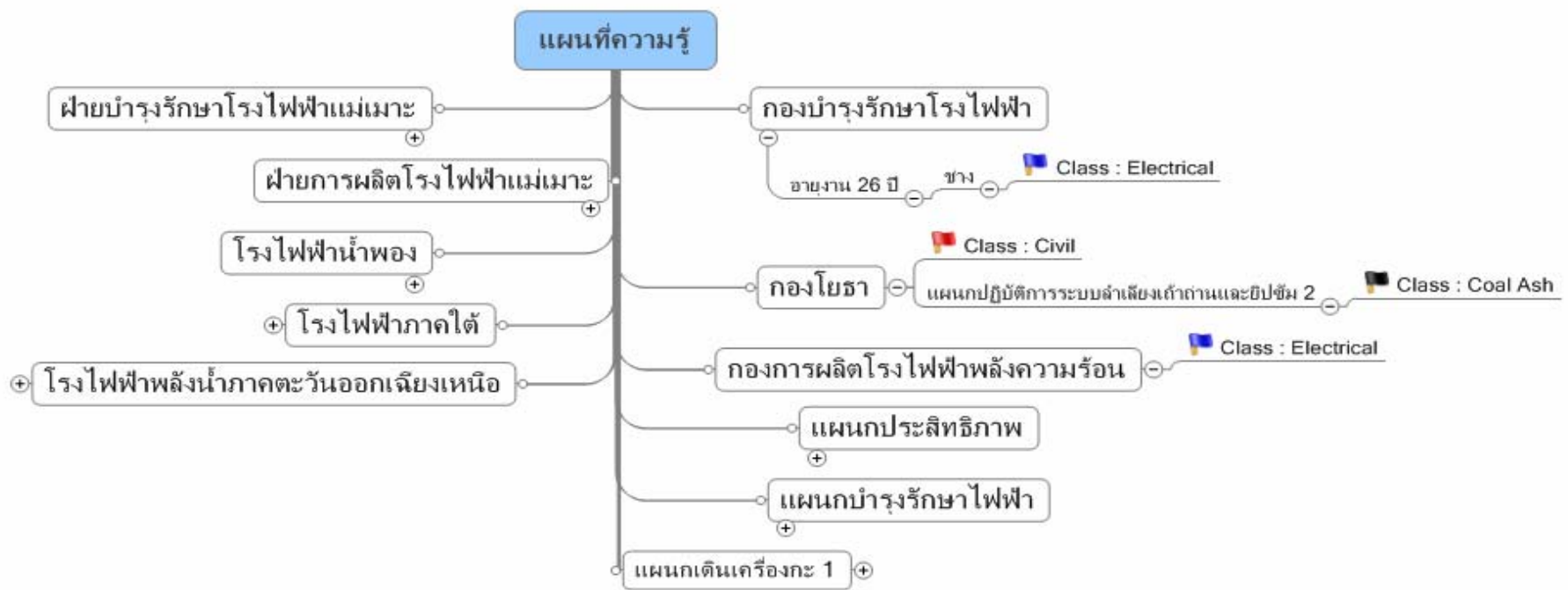
### แผนที่ความรู้

จากการทดลองในขั้นตอนแรก ใช้วิธีการจำแนกประเภทความเชี่ยวชาญเชิงเปรียบเทียบ  
ของ 3 อัลกอริทึม ได้แก่ อัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes เพื่อหา  
อัลกอริทึมที่มีค่าความถูกต้องในการจำแนกประเภทความเชี่ยวชาญได้สูงสุด ผลการทดลองด้วย  
โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) เวอร์ชัน 3.5.5 พบว่า  
อัลกอริทึม C4.5 มีค่าความถูกต้องในการทำนายความเชี่ยวชาญของผู้ปฏิบัติงานได้สูงสุด โดยใช้ตัว  
วัดประสิทธิภาพ 4 ตัว ได้แก่ 1) Accuracy 2) Precision/Recall 3) F-measure และ 4) Root Mean-  
Squared Error (RMSE) และจากการทดลองในขั้นตอนที่สอง คือ การหาความสัมพันธ์ของ  
คุณลักษณะที่มีผลต่อความเชี่ยวชาญของผู้ปฏิบัติงาน โดยใช้ Association Rule Mining (อัลกอริทึม  
Apriori) ทำให้สามารถหาความสัมพันธ์ของคุณลักษณะข้อมูลแต่ละคลาสได้ ผู้วิจัยจึงได้นำผลการ  
ทดลองที่ได้มาใช้ในการสร้างแผนที่ความรู้ขององค์กร ทั้งนี้ ได้ใช้ต้นไม้การตัดสินใจที่ได้จากการ  
ทดลองโดยอัลกอริทึม C4.5 มาใช้ในการสร้างแผนที่ความรู้ขององค์กร ทำให้ทราบได้ว่าองค์กรมี  
ความรู้อะไรบ้าง ความรู้ต่าง ๆ มีอยู่ที่ไหน และความรู้ในเรื่องต่าง ๆ มีความซ้ำซ้อนกับหน่วยงาน  
อื่นหรือไม่ ทั้งนี้ได้แสดงรายละเอียดของแผนที่ความรู้ของ กฟผ. ดังนี้

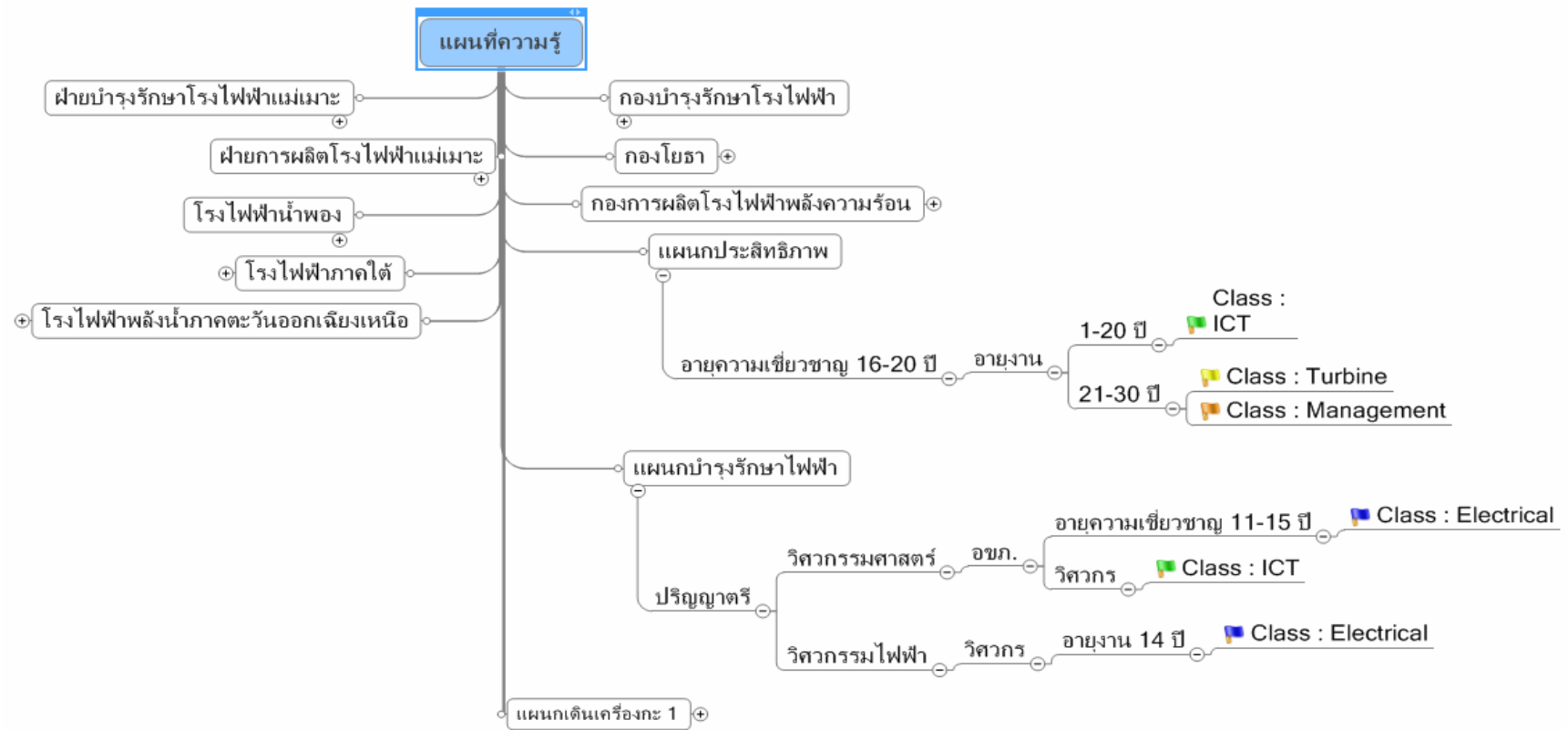


ภาพที่ 13 ภาพรวมของแผนที่ความรู้ขององค์กร กฟผ.

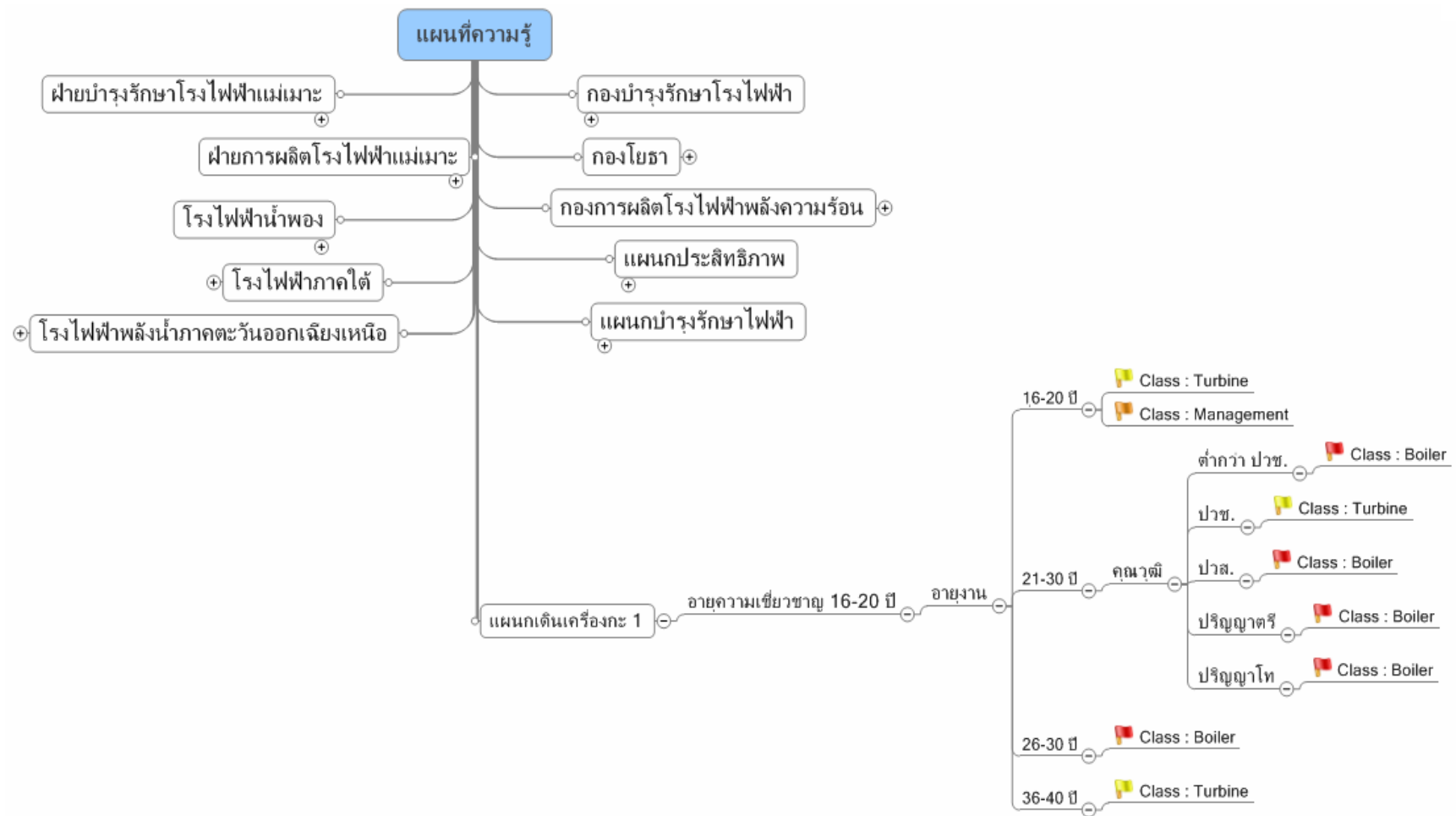
ภาพที่ 13 แสดงภาพรวมของแผนที่ความรู้ขององค์กร กฟผ. แยกเป็นหน่วยงานระดับฝ่าย ระดับกอง และระดับแผนก เริ่มจากระดับฝ่าย ประกอบด้วย 1) ฝ่ายบำรุงรักษาโรงไฟฟ้าแม่เมาะ 2) ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ 3) โรงไฟฟ้าน้ำพอง 4) โรงไฟฟ้าภาคใต้ 5) โรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ หน่วยงานระดับกอง ประกอบด้วย 1) กองบำรุงรักษาโรงไฟฟ้า 2) กองโยธา 3) กองการผลิตโรงไฟฟ้าพลังความร้อน หน่วยงานระดับแผนก ประกอบด้วย 1) แผนกประสิทธิภาพ 2) แผนกบำรุงรักษาไฟฟ้า 3) แผนกเดินเครื่องกะ 1



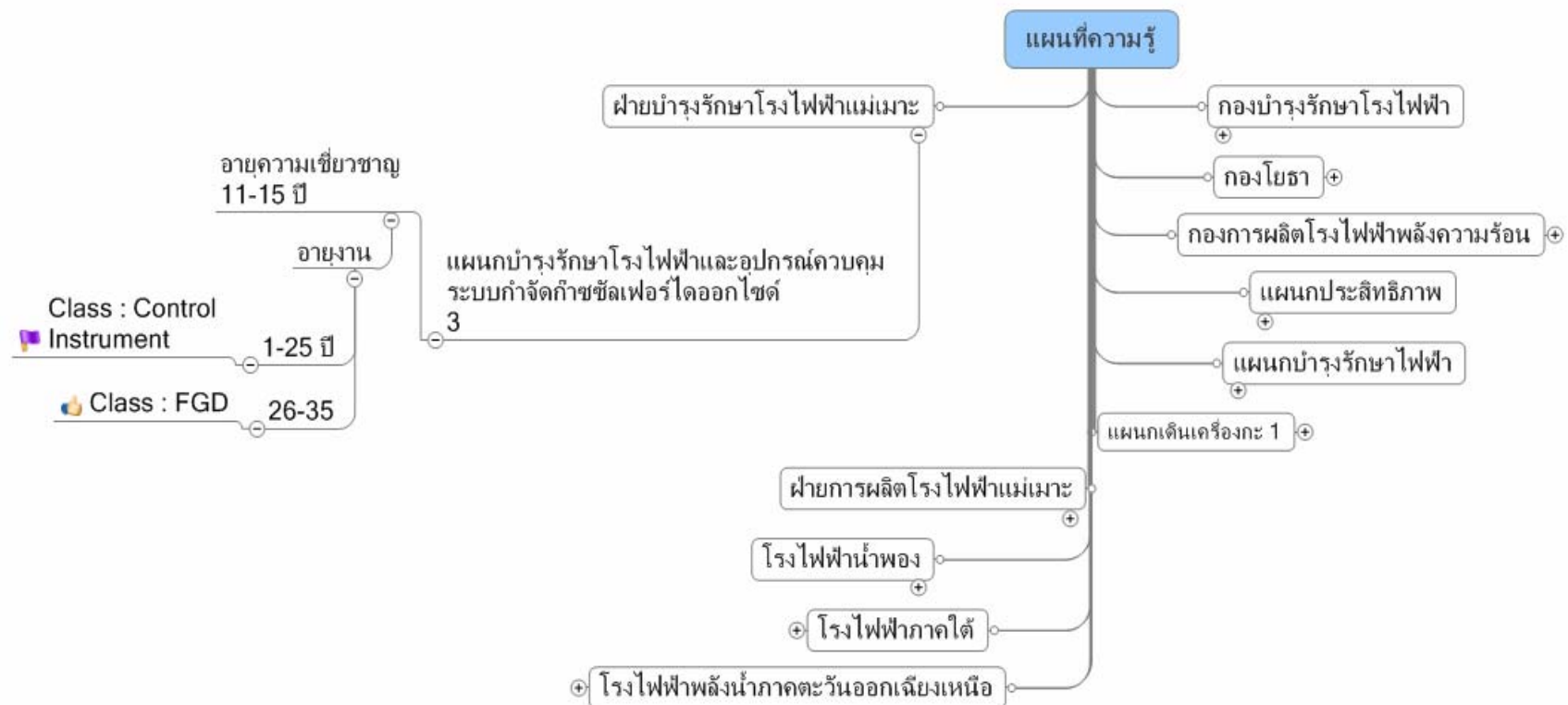
ภาพที่ 14 แผนที่ความรู้โครงสร้าง กฟผ. ระดับกอง



ภาพที่ 15 แผนที่ความรู้ขององค์กร กฟผ. แผนกประสิทธิภาพและแผนกบำรุงรักษาไฟฟ้า



ภาพที่ 16 แผนที่ความรู้องค์กร กฟผ. แผนกเดินเครื่องกะ 1



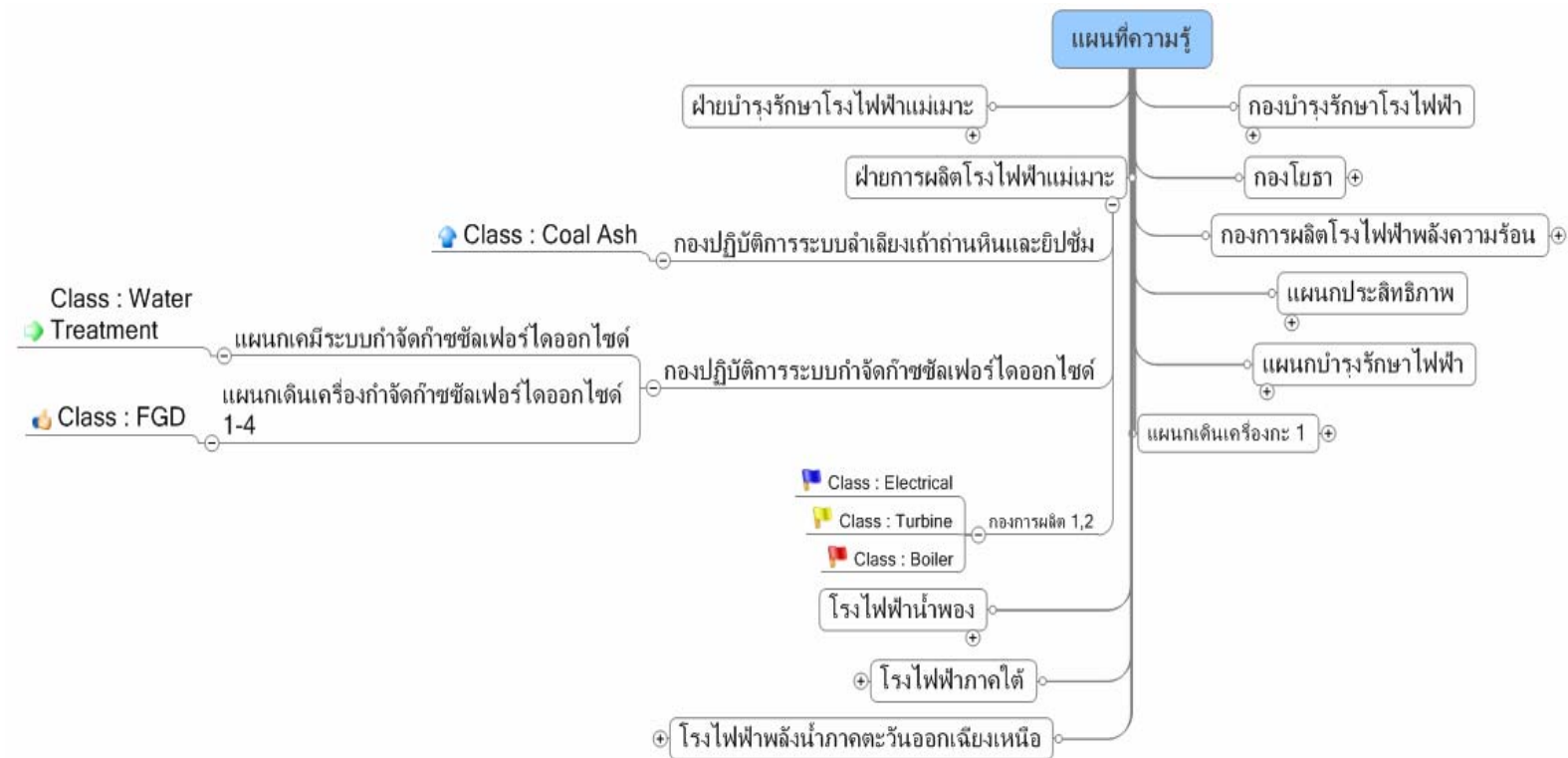
ภาพที่ 17 แผนที่ความรู้ขององค์กร กฟผ. ฝ่ายบำรุงรักษาโรงไฟฟ้าแม่เมาะ

ภาพที่ 14 แสดงแผนที่ความรู้องค์กร กฟผ. ระดับกอง ประกอบด้วย กองบำรุงรักษา โรงไฟฟ้า ที่ผู้ปฏิบัติงานมีอายุงาน 26 ปี และมีตำแหน่งเป็นช่าง จะมีความเชี่ยวชาญด้าน Electrical สำหรับกองโยธา จะมีความเชี่ยวชาญด้านโยธา (Civil) และแผนกในกองโยธา ได้แก่ แผนก ปฏิบัติการระบบลำเลียงถ่านและยิปซัม 2 จะมีความเชี่ยวชาญด้าน Coal Ash

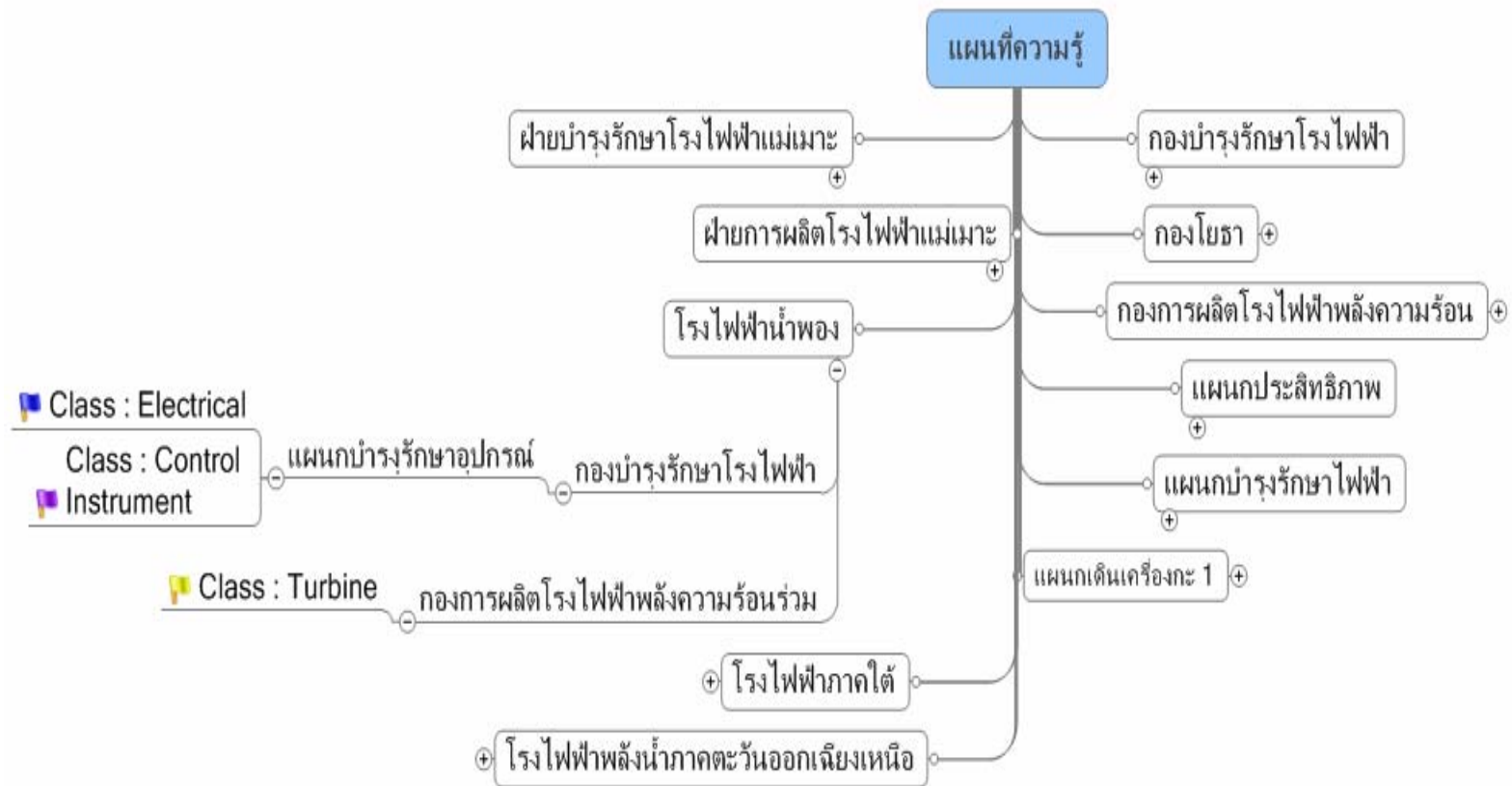
ภาพที่ 15 แสดงแผนที่ความรู้องค์กร กฟผ. แผนกประสิทธิภาพ มีคุณลักษณะข้อมูลด้าน อายุความเชี่ยวชาญในงานที่ปฏิบัติ อยู่ในช่วง 16-20 ปี และผู้ปฏิบัติงานมีอายุงาน 1-20 ปี จะมีความเชี่ยวชาญด้าน ICT สำหรับผู้ปฏิบัติงานที่มีอายุงาน 21-30 ปี จะมีความเชี่ยวชาญด้าน Turbine และ Management สำหรับแผนกบำรุงรักษาไฟฟ้า ผู้ปฏิบัติงานจะมีคุณวุฒิการศึกษาระดับปริญญาตรีจบ การศึกษาจากคณะวิศวกรรมศาสตร์ ปฏิบัติงานที่ เขื่อนภูมิพล (อบก.) ถ้ามีอายุความเชี่ยวชาญใน งานอยู่ในช่วง 11-15 ปี จะมีความเชี่ยวชาญด้าน Electrical แต่ถ้าผู้ปฏิบัติงานมีตำแหน่งเป็นวิศวกร จะมีความเชี่ยวชาญด้าน ICT ด้วย และถ้ามีคุณลักษณะข้อมูลในเรื่องสาขาที่จบการศึกษา ด้าน วิศวกรรมไฟฟ้า จะมีตำแหน่งเป็นวิศวกร ที่มีอายุงาน 14 ปี จะมีความเชี่ยวชาญด้าน Electrical

ภาพที่ 16 แสดงแผนที่ความรู้องค์กร กฟผ. แผนกเดินเครื่องกะ 1 ผู้ปฏิบัติงานมีอายุความ เชี่ยวชาญในงานที่ปฏิบัติ อยู่ในช่วง 16-20 ปี และมีอายุงาน 16-20 ปี จะมีความเชี่ยวชาญด้าน Turbine และ Management ถ้ามีอายุงาน 21-30 ปี และมีคุณวุฒิการศึกษาต่ำกว่า ปวช. จะมีความ เชี่ยวชาญด้าน Boiler ถ้ามีคุณวุฒิการศึกษา ปวช. จะมีความเชี่ยวชาญด้าน Turbine มีคุณวุฒิ การศึกษา ปวส. จะมีความเชี่ยวชาญด้าน Boiler และถ้ามีคุณวุฒิการศึกษาปริญญาตรี ปริญญาโท จะมีความเชี่ยวชาญด้าน Boiler และผู้ที่มีอายุงาน 26-30 ปี จะมีความเชี่ยวชาญด้าน Boiler สำหรับ อายุงาน 36-40 ปี จะมีความเชี่ยวชาญด้าน Turbine

ภาพที่ 17 แสดงแผนที่ความรู้องค์กร กฟผ. ฝ่ายบำรุงรักษาโรงไฟฟ้าแม่เมาะ ผู้ปฏิบัติงาน สังกัดแผนกบำรุงรักษาโรงไฟฟ้าและอุปกรณ์ควบคุมระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ ผู้ปฏิบัติงานที่มีอายุความเชี่ยวชาญ อยู่ในช่วง 11-15 ปี และมีอายุงาน 1-25 ปี จะมีความเชี่ยวชาญ ด้าน Control Instrument และถ้ามีอายุงาน 26-35 ปี จะมีความเชี่ยวชาญด้าน FGD



ภาพที่ 18 แผนที่ความรู้องค์กร กฟผ. ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ



ภาพที่ 19 แผนที่ความรู้องค์กร กฟผ. โรงไฟฟ้าน้ำพอง



ภาพที่ 20 แผนที่ความรู้องค์กร กฟผ. โรงไฟฟ้าภาคใต้และโรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ

ภาพที่ 18 แสดงแผนที่ความรู้องค์กร กฟผ. ฝ่ายการผลิต โรงไฟฟ้าแม่เมาะ ผู้ปฏิบัติงาน สังกัดกองปฏิบัติการระบบลำเลียงถ่านหินและยิปซั่ม จะมีความเชี่ยวชาญด้าน Coal Ash และ ผู้ปฏิบัติงานสังกัดกองปฏิบัติการระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ ปฏิบัติงานที่แผนกเคมีระบบกำจัดก๊าซซัลเฟอร์ไดออกไซด์ จะมีความเชี่ยวชาญด้าน Water Treatment และผู้ปฏิบัติงานที่สังกัด แผนกเดินเครื่องกำจัดก๊าซซัลเฟอร์ไดออกไซด์ จะมีความเชี่ยวชาญด้าน FGD สำหรับกองการผลิต 1 และ 2 ผู้ปฏิบัติงานจะมีความเชี่ยวชาญด้าน Electrical, Turbine และ Boiler

ภาพที่ 19 แสดงแผนที่ความรู้องค์กร กฟผ. โรงไฟฟ้าน้ำพอง ผู้ปฏิบัติงานกองบำรุงรักษา โรงไฟฟ้า สังกัดแผนกบำรุงรักษาอุปกรณ์ จะมีความเชี่ยวชาญด้าน Electrical และ Control Instrument สำหรับกองการผลิตโรงไฟฟ้าพลังความร้อนร่วม จะมีความเชี่ยวชาญด้าน Turbine

ภาพที่ 20 แสดงแผนที่ความรู้องค์กร กฟผ. โรงไฟฟ้าภาคใต้ และโรงไฟฟ้าพลังน้ำภาค ตะวันออกเฉียงเหนือ โดยผู้ปฏิบัติงานโรงไฟฟ้าภาคใต้จะมีความเชี่ยวชาญในงานด้าน Control Instrument และผู้ปฏิบัติงานโรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ จะมีความเชี่ยวชาญในงาน ด้าน Control Instrument เช่นกัน

## วิจารณ์

งานวิจัยฉบับนี้ได้นำหลักการของการทำเหมืองข้อมูลมาใช้ในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงาน โดยใช้อัลกอริทึมประเภทการสร้างแบบจำลองแบบทำนาย (Predictive Data Mining) มาทำการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงานแยกตามคลาสที่กำหนด ซึ่งได้แบ่งความเชี่ยวชาญในงานออกเป็น 10 ประเภท และทำการทดลองด้วย อัลกอริทึม 3 แบบ ได้แก่ อัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes จากการทดลองพบว่า อัลกอริทึมที่ให้ผลในการทำนายที่มีประสิทธิภาพดีที่สุด เป็นอัลกอริทึมที่ใช้หลักการของการสร้างต้นไม้การตัดสินใจ คือ อัลกอริทึม C4.5 โดยวิเคราะห์ในแต่ละรอบและแบบเฉลี่ย 10 รอบ (10-fold) ให้ผลการทำงานดีเป็นอันดับหนึ่ง ส่วนอัลกอริทึมที่ให้ผลในการทำนายแม่นยำน้อยที่สุด ได้แก่ Naïve Bayes อัลกอริทึม ซึ่งเป็นอัลกอริทึมที่ใช้หลักการทางสถิติ จากนั้นใช้เทคนิค Association Rules ในการหาความสัมพันธ์ของแต่ละคุณลักษณะ (Attributes) ทั้งนี้ เพื่อนำไปใช้สร้างแผนที่ความรู้ในองค์กรต่อไป

จากขั้นตอนการเปรียบเทียบอัลกอริทึมของการจำแนกความเชี่ยวชาญของอัลกอริทึม 3 แบบ จะเห็นได้ว่า C4.5 อัลกอริทึมให้ค่าความถูกต้องในการทำนายสูงสุด ซึ่งเป็นไปตามงานวิจัยของ Huang และคณะปี 2003 และงานวิจัยของ Soman และ Bobbie ในปี 2005 ที่ได้นำเสนอการจำแนกข้อมูลเชิงเปรียบเทียบอัลกอริทึมไว้ ผลที่ได้พบว่า C4.5 อัลกอริทึมให้ค่าความถูกต้องสูงสุด

สำหรับขั้นตอนการหาความสัมพันธ์ของคุณลักษณะข้อมูลโดย Association Rule Mining (อัลกอริทึม Apriori) จะเห็นได้ว่าผู้ปฏิบัติงานส่วนใหญ่จะมีคุณสมบัติต่ำกว่าระดับปริญญาตรี โดยความเชี่ยวชาญด้านต่าง ๆ สามารถระบุหน่วยงานของผู้เชี่ยวชาญได้ว่าสังกัดอยู่ที่ใด เช่น ผู้ปฏิบัติงานที่มีตำแหน่งเป็นช่าง จะมีคุณสมบัติต่ำกว่าระดับปริญญาตรี และมีความเชี่ยวชาญด้าน Coal Ash จะสังกัด ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ เป็นต้น สำหรับผู้ปฏิบัติงานที่มีตำแหน่งเป็นวิศวกร จะจบการศึกษาคณะวิศวกรรมศาสตร์ นอกจากนี้งานที่มีส่วนเกี่ยวข้องกันจะช่วยให้ผู้ปฏิบัติงานมีความเชี่ยวชาญในงานอื่นเพิ่มขึ้นด้วย และความสัมพันธ์ของคุณลักษณะข้อมูลด้านความเชี่ยวชาญที่ผู้ปฏิบัติงานแต่ละคนมี อาจจะมีมีความเชี่ยวชาญได้หลายด้าน อาทิ ผู้ปฏิบัติงานที่มีความเชี่ยวชาญด้าน Control Instrument แล้ว จะมีความเชี่ยวชาญด้าน Turbine ด้วย เป็นต้น คุณลักษณะของข้อมูล

หน่วยงานที่ผู้ปฏิบัติงานสังกัด ระดับฝ่าย กอง แผนก สามารถที่จะระบุความเชี่ยวชาญของ  
ผู้ปฏิบัติงานได้เช่นกัน

แผนที่ความรู้ที่สร้างได้จากงานวิจัยฉบับนี้ จะเป็นเครื่องมือที่บอกให้ทราบว่าภายในองค์กร  
มีความรู้ ความเชี่ยวชาญในงานด้านใด ความรู้ต่าง ๆ อยู่ที่หน่วยงานไหน และมีความรู้ด้านใดที่  
ซ้ำซ้อนกันระหว่างหน่วยงาน เป็นต้น แต่กระนั้น แผนที่ความรู้ที่สร้างได้ก็ยังไม่สามารถครอบคลุม  
ความรู้ทุกประเภทได้อย่างชัดเจน เนื่องจากจำนวนข้อมูลที่ใช้ในการทดลองมีจำนวนไม่มากพอ  
หากต้องการพัฒนาแผนที่ความรู้ที่มีความละเอียดและชัดเจนมากยิ่งขึ้น จะต้องมีการเพิ่มปริมาณ  
ข้อมูลที่ใช้ในการทดลองให้มากกว่านี้ ดังนั้น แผนที่ความรู้ฉบับนี้จึงนับเป็นต้นแบบให้กับผู้ที่สนใจ  
จะสร้างแผนที่ความรู้ขององค์กร ของหน่วยงาน

## สรุปและข้อเสนอแนะ

### สรุป

งานวิจัยนี้มีจุดมุ่งหมายในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานตามคลาสเบลของความเชี่ยวชาญ และการหาความสัมพันธ์ของคุณลักษณะข้อมูลที่เกี่ยวข้องกับความเชี่ยวชาญในสาขาอาชีพที่ทำการศึกษา โดยใช้หลักการของการทำเหมืองข้อมูล (Data Mining) มาใช้ในการจำแนกประเภทข้อมูล ข้อมูลที่ใช้ในการทดลองได้จากหน่วยงาน กฟผ. สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง จำนวนกลุ่มตัวอย่าง 28 หน่วยงาน และได้แบ่งคุณลักษณะของข้อมูลที่ใช้ในการจำแนกประเภทความเชี่ยวชาญของผู้ปฏิบัติงานเป็น 10 คุณลักษณะ ประกอบด้วย 1) อายุงาน 2) ตำแหน่ง 3) คุณวุฒิการศึกษา 4) คณะที่จบการศึกษา 5) สาขาที่จบ 6) ฝ่าย 7) กอง 8) แผนก 9) ระดับความเชี่ยวชาญ และ 10) ประเภทความเชี่ยวชาญ

ขั้นตอนการทดลองในงานวิจัยนี้แบ่งเป็น 2 ส่วน 1) การทดลองขั้นตอนแรกใช้วิธีการจำแนกประเภทความเชี่ยวชาญเชิงเปรียบเทียบด้วยอัลกอริทึม 3 แบบ คือ 1) อัลกอริทึม C4.5 2) อัลกอริทึม OneR 3) อัลกอริทึม Naïve Bayes เพื่อหาอัลกอริทึมที่มีความแม่นยำที่สุดในการทำนายความเชี่ยวชาญ และใช้ตัววัดประสิทธิภาพ 4 ตัว 1) Accuracy 2) Precision/Recall 3) F-measure และ 4) Root Mean-Squared Error (RMSE) และการทดลองขั้นตอน 2 คือ การหาความสัมพันธ์ของคุณลักษณะที่มีผลต่อความเชี่ยวชาญของผู้ปฏิบัติงาน โดยใช้ Association Rule Mining (อัลกอริทึม Apriori)

ผลการทดลองพบว่า อัลกอริทึม C4.5 ให้ค่าความถูกต้องในการทำนายคลาสของความเชี่ยวชาญได้สูงสุด สามารถแสดงผลที่ได้จากตัววัดประสิทธิภาพ ดังนี้ 1) Accuracy 95.61% 2) ค่า Precision เท่ากับ 0.957 ค่า Recall เท่ากับ 0.999 3) ค่า F-measure เท่ากับ 0.978 และ 4) ค่า RMSE เท่ากับ 0.093 รองลงมาคือ อัลกอริทึม OneR ผลที่ได้จากตัววัดประสิทธิภาพ ดังนี้ 1) Accuracy 95.513% 2) ค่า Precision เท่ากับ 0.959 ค่า Recall เท่ากับ 0.999 3) ค่า F-measure เท่ากับ 0.978 และ 4) ค่า RMSE เท่ากับ 0.094 สุดท้ายอัลกอริทึม Naïve Bayes ให้ค่าความถูกต้องในการทำนายค่าสุดท้าย ผลที่ได้จากตัววัดประสิทธิภาพ ดังนี้ 1) Accuracy 91.038% 2) ค่า Precision เท่ากับ 0.958 ค่า Recall เท่ากับ 0.950 3) ค่า F-measure เท่ากับ 0.954 4) ค่า RMSE เท่ากับ 0.121 จากผลการทดลอง

ในขั้นตอนแรกช่วยให้สามารถบรรลุวัตถุประสงค์ของงานวิจัยข้อที่ 1 ที่ต้องการศึกษาและเปรียบเทียบประสิทธิภาพการทำงานนายความเชี่ยวชาญของผู้ปฏิบัติงาน และวัตถุประสงค์ข้อที่ 2 ได้ อัลกอริทึมที่เหมาะสมกับการสร้างแผนที่ความรู้ นั่นคือ อัลกอริทึม C4.5

สำหรับการหาความสัมพันธ์ของคุณลักษณะที่มีผลต่อความเชี่ยวชาญของผู้ปฏิบัติงาน โดยใช้ Association Rule Mining (อัลกอริทึม Apriori) คุณลักษณะข้อมูลที่ศึกษา ได้แก่ ข้อมูลของคุณวุฒิการศึกษา คณะที่จบการศึกษา ตำแหน่งงาน หน่วยงานที่ผู้ปฏิบัติงานสังกัดและอายุความเชี่ยวชาญ ผลจากการวิเคราะห์ คือ 1) ตำแหน่งงานช่าง จะมีคุณวุฒิต่ำกว่าระดับปริญญาตรี 2) หน่วยงาน และอายุความเชี่ยวชาญในงานมีความสัมพันธ์กับความเชี่ยวชาญของผู้ปฏิบัติงาน 3) งานที่มีส่วนเกี่ยวข้องกันจะช่วยให้ผู้ปฏิบัติงานมีความเชี่ยวชาญในงานเพิ่มขึ้น กล่าวคือ การหาความสัมพันธ์ของคุณลักษณะข้อมูลช่วยให้สามารถวิเคราะห์ความสัมพันธ์ของคุณลักษณะข้อมูลที่เกิดขึ้นร่วมกัน ในงานด้านต่าง ๆ ได้เป็นอย่างดี

จากการทดลองขั้นตอนแรก และขั้นตอนที่ 2 โดยใช้โปรแกรม WEKA เวอร์ชัน 3.5.5 ผู้วิจัยได้เลือกอัลกอริทึมที่มีความเหมาะสมมาใช้ในการสร้างแผนที่ความรู้ นั่นคือ อัลกอริทึม C4.5 ที่เป็นรูปแบบของต้นไม้การตัดสินใจ ร่วมกับการหาความสัมพันธ์ของคุณลักษณะข้อมูลที่จำแนกตามความเชี่ยวชาญ มาใช้ในการสร้างแผนที่ความรู้ขององค์กร กฟผ. โดยโครงสร้างของแผนที่ความรู้ประกอบด้วย หน่วยงานระดับฝ่าย ระดับกอง และระดับแผนก เริ่มจากระดับฝ่าย ประกอบด้วย 1) ฝ่ายบำรุงรักษาโรงไฟฟ้าแม่เมาะ 2) ฝ่ายการผลิตโรงไฟฟ้าแม่เมาะ 3) โรงไฟฟ้าน้ำพอง 4) โรงไฟฟ้าภาคใต้ 5) โรงไฟฟ้าพลังน้ำภาคตะวันออกเฉียงเหนือ หน่วยงานระดับกอง ประกอบด้วย 1) กองบำรุงรักษาโรงไฟฟ้า 2) กองโยธา 3) กองการผลิตโรงไฟฟ้าพลังความร้อน หน่วยงานระดับแผนก ประกอบด้วย 1) แผนกประสิทธิภาพ 2) แผนกบำรุงรักษาโรงไฟฟ้า 3) แผนกเดินเครื่องกะ 1 แต่ละหน่วยงานจะมีคุณลักษณะข้อมูลที่สามารถบ่งชี้ความเชี่ยวชาญของผู้ปฏิบัติงานในหน่วยงาน

สำหรับการเปรียบเทียบประสิทธิภาพการทำงานของอัลกอริทึม C4.5 อัลกอริทึม OneR และอัลกอริทึม Naïve Bayes เมื่อทดสอบด้วยข้อมูลที่มีขนาดต่างกัน คือ ข้อมูลชุดที่ 1 จำนวน 343 ระเบียน และข้อมูลชุดที่ 2 จำนวน 805 ระเบียนพบว่า ข้อมูลที่มีขนาดที่ใหญ่ขึ้น จะมีค่าความแม่นยำในการทำนายคลาสความเชี่ยวชาญสูงกว่าชุดข้อมูลที่มีขนาดเล็ก

### ข้อเสนอแนะ

งานวิจัยฉบับนี้ใช้เทคนิคการทำเหมืองข้อมูลมาใช้ในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานในองค์กร โดยจะเห็นได้ว่าปัจจุบันความรู้ต่าง ๆ ในองค์กรมีลักษณะการกระจาย ขาดการรวบรวมให้เป็นศูนย์กลางความรู้ที่จะบอกถึงความรู้ที่มีอยู่ในองค์กร และความรู้ที่จำเป็นสำหรับองค์กรใครเป็นผู้รู้ ผู้เชี่ยวชาญในงานด้านใด ซึ่งเรื่องดังกล่าวมีความจำเป็นอย่างยิ่งที่จะต้องอาศัยการจัดการความรู้ที่เหมาะสม สำหรับผู้วิจัยที่สนใจในเรื่องการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงาน มีความจำเป็นจะต้องเข้าใจในเรื่องจัดการความรู้ และเห็นประโยชน์ของการจัดการความรู้ขององค์กร ที่เรา ๆ ท่าน ๆ สามารถรวบรวมองค์ความรู้ได้จากผู้ปฏิบัติงาน และเอกสารความรู้ต่าง ๆ มาพัฒนาต่อให้เป็นระบบ เพื่อให้ทุกคนในองค์กรสามารถเข้าถึงความรู้ ผู้รู้และพัฒนาตนเองให้เป็นผู้เชี่ยวชาญในงาน โดยสามารถนำขั้นตอนวิธีการวิจัยนี้ ไปใช้เป็นแนวทางในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานสำหรับหน่วยงานได้ จุดเด่นของงานวิจัยฉบับนี้ที่แตกต่างจากงานวิจัยอื่น ๆ ในการจำแนกข้อมูล คือ กระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในคลาส หรือกลุ่มข้อมูลที่กำหนด โดยใช้ทฤษฎีการจำแนกประเภทข้อมูล (Classification Theory) และอัลกอริทึมที่นำมาใช้ในงานวิจัยนี้ยังไม่พบว่างานวิจัยอื่นนำมาใช้ในการจำแนกความเชี่ยวชาญของผู้ปฏิบัติงานในองค์กร

ผลที่ได้จากงานวิจัยฉบับนี้ ผู้สนใจสามารถนำไปพัฒนาต่อให้เกิดประสิทธิภาพมากยิ่งขึ้น โดยการใช้ข้อมูลให้มีปริมาณมากขึ้นและ ครอบคลุมความรู้ทั้งหมดในองค์กร เพิ่มจำนวนคุณลักษณะของข้อมูล เช่น รางวัลที่ได้รับ งานวิจัย จากนั้นเลือกใช้อัลกอริทึมในการจำแนกความเชี่ยวชาญ งานวิจัยที่จะพัฒนาต่อในอนาคตผู้วิจัยจะใช้อัลกอริทึม K-Nearest Neighbors และ RandomForest ในการจำแนกความเชี่ยวชาญให้มีความแม่นยำเพิ่มขึ้น และจะสร้างแผนที่ความรู้ขององค์กรด้วย Concept Maps โดยใช้ Network Diagrams ทั้งนี้ เพื่อให้สามารถแสดงแผนที่ความรู้ได้อย่างชัดเจน และเห็นความเชื่อมโยงของแหล่งความรู้ต่าง ๆ ในองค์กรได้เป็นอย่างดี

## เอกสารและสิ่งอ้างอิง

มหาวิทยาลัยเชียงใหม่ คณะวิทยาศาสตร์. 2550. การทำแผนที่ความรู้บุคลากรคณะวิทย์ ม.ช.  
แผนที่ความรู้ของคณะวิทยาศาสตร์. แหล่งที่มา: <http://gotoknow.org/blog/map>, 23 ,  
มีนาคม 2551.

บุญเสริม กิจศิริกุล. 2546. โครงการวิจัยร่วมภาครัฐและเอกชน โครงการย่อยที่ 7 อัลกอริทึมการทำ  
เหมืองข้อมูล.

\_\_\_\_\_ 2546. ปัญญาประดิษฐ์. ภาควิชาวิศวกรรมคอมพิวเตอร์, จุฬาลงกรณ์มหาวิทยาลัย.

วิลาวัลย์ นิมประเสริฐ. 2548. การวิจัยเกี่ยวกับการจัดการความรู้. แหล่งที่มา: [http://blog-for-thai-km.blogspot.com/2005/02/blog-post\\_06.html](http://blog-for-thai-km.blogspot.com/2005/02/blog-post_06.html), 23 มีนาคม 2551.

สำนักคณะกรรมการพัฒนาระบบราชการ [ก.พ.ร.]. 2548. การจัดการความรู้. : 21-25.

Agrawal, R. and R. Srikant. 1994. Fast Algorithms for Mining Association Rules, pp. 487-499.  
*In Proceedings of the VLDB Conference 20*. ed. IBM Almaden Research Center,  
Santiago, Chile.

Chen, G., H. Liu, L. Yu, Q. Wei and X Zhang. 2005. A new approach to classification based on  
association rule mining, pp. 674-689. *In Decision Support Systems 42*. ed.

Eppler, M. J. 2001. Making Knowledge Visible Through Intranet Knowledge Maps: Concepts,  
Elements, Cases, System Sciences, pp. 225-444. *In Proceedings of the Annual  
Hawaii International Conference. 34*. ed.

- Eppler, M.J. 2006. Toward a Pragmatic Taxonomy of Knowledge Maps : Classification Principles, Sample Typologies, and Application Examples, pp. 195-204. *In* **Proceedings of the Information Visualization (IV'06)**.
- Fayyad, U., G. S. Piatetsky and P. Smyth. 1996. Knowledge Discovery and Data Mining : Towards a Unifying Framework, pp. 156-164. *In* **The Association for the Advancement of Artificial Intelligence Press**.
- Huang, J., J. Lu and C.X. Ling. 2003. Comparing Naïve Bayes, Decision Trees, and SVM with AUC and Accuracy., *In* **Proceedings of the Third IEEE International Conference on Data Mining (ICDM'03)**.
- Kantardzic, M. 2003. **Data Mining : Concepts, Models, Methods and Algorithms**. A John Wiley & Sons Inc.
- Larose, D. T. 2005. **Discovering Knowledge in Data**. A John Wiley & Sons, Inc.
- Liao, S.H., H.H. Ho and H.W. Lin. 2007. Mining Stock Category Association and Cluster on Taiwan Stock Market, *In* **Expert System with Application**.
- Lin, F.R. and C.M. Hsueh. 2006. Knowledge Map Creation and Maintenance for Virtual Communities of Practice. **Information Processing and Management** 2 (42): 551-568.
- Nanoka, I. and H. Takeuchi. 1995. **The Knowledge Creating Company**. Oxford University.
- Neshatian, K., M. Kharrat and S.B. Khamaneh. 2006. Organizational Experience Management Through Knowledge Maps-An Ontological Approach, pp. 1-8. *In* **World Automation Congress 6. ed.** Iran Telecom Research Center, Budapest.

- Ong, T.H., H. Chen, W.K. Sung and B. Zhu. 2005. Newsmap : a knowledge map for online news, *In Decision Support Systems* 4 (39): 583-597.
- Richardson, D. 2004. **The Practical Reality of Knowledge Management within Development Initiatives.** Available Source: <http://tumi.lamolina.edu.pe/ipps/ict/ict11.pdf>, February 21, 2008.
- Ruggieri, S. 2002. Efficient C4.5. *IEEE Transactions* 2 (14): 438-444.
- Soman, T. and P.O. Bobbie. 2003. **Classification of Arrhythmia Using Machine Learning Techniques.** *In School of Computing and Software Engineering Southern Polytechnic State University (SPSU), USA.*
- Tan, P.N., M. Steinbach and V. Kumar. 2006. **Introduction to Data Mining.** Pearson Education, Inc., USA.
- Wang, Y., J. Hodges and B. Tang. 2003. Classification of Web Documents Using a Naive Bayes Method, pp. 560-564. *In Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'03).*
- Wang, L.M., X.L. Li, C.H. Cao and S.M. Yuan. 2006. Combining decision tree and Naïve Bayes for classification. *In Knowledge Based Systems* 7 (19): 511-515.
- Weka. **Data Mining with Open Source Machine Learning Software in Java.** Available Source: <http://www.cs.waikato.ac.nz/ml/weka>, May 2008.
- Witten, I.H. and E. Frank. 2005. **Data Mining: Practical Machine Learning Tools and Techniques with Java Implementation.** Morgan Kaufmann, San Francisco.

## ประวัติการศึกษา และการทำงาน

|                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ชื่อ –นามสกุล                  | นางสาวบุษกร สุคนธวงศาโรจน์                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| วัน เดือน ปี ที่เกิด           | วันที่ 23 พฤศจิกายน 2513                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| สถานที่เกิด                    | กรุงเทพมหานคร                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| ประวัติการศึกษา                | วท.บ.(วิทยาการคอมพิวเตอร์) ราชภัฏสวนสุนันทา                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| ตำแหน่งหน้าที่การงานปัจจุบัน   | นักคอมพิวเตอร์ระดับ 7                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| สถานที่ทำงานปัจจุบัน           | การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| ผลงานดีเด่นและรางวัลทางวิชาการ | <ol style="list-style-type: none"> <li>นำเสนอ และตีพิมพ์บทความวิชาการเรื่อง<br/>           “Classification algorithms for knowledge mapping of expert in Energy industry” ในงานประชุมวิชาการ The 9th International DSI Conference in conjunction with The 12th APDSI Conference 2007 กรุงเทพมหานคร ในวันที่ 11-15 กรกฎาคม 2550</li> <li>นำเสนอ และตีพิมพ์บทความวิชาการเรื่อง<br/>           “Comparison of Classification Algorithm Performances in Knowledge Mapping for Organization Experts” ในงานประชุมวิชาการ The 7th International Conference on Electronic Business 2007 ประเทศไต้หวัน ในวันที่ 2-6 ธันวาคม 2550</li> </ol> |

