

การพัฒนาแบบจำลองภาระการทำงานของเครื่องจักรซีเอ็นซี จากการใช้กระแสไฟฟ้า 3 เฟส ด้วย
วิธีการแบบกลุ่ม

นายธนารักษ์ รักรธรรม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2554

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)

are the thesis authors' files submitted through the Graduate School.

DEVELOPMENT OF WORKLOAD MODELS FOR CNC MACHINES FROM 3 - PHASE
CURRENT CONSUMPTION USING ENSEMBLE METHOD

Mr. Thanarak Raktham

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

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2011

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การพัฒนาแบบจำลองภาระการทำงานของเครื่องจักรซีเอ็นซี
จากการใช้กระแสไฟฟ้า 3 เฟส ด้วยวิธีการแบบกลุ่ม

โดย

นายธนารักษ์ รักธรรม

สาขาวิชา

วิทยาศาสตร์คอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร. เกริก ภิรมย์โสภา

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็น
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศธีรบุญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.เฉลิมเอก อินทนาการวิวัฒน์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร. เกริก ภิรมย์โสภา)

..... กรรมการ
(อาจารย์ ดร.ณัฐวุฒิ หนูไพโรจน์)

..... กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.พงศ์วัช ชีพพิมลชัย)

ธนารักษ์ รักรธรรม : การพัฒนาแบบจำลองภาระการทำงานของเครื่องจักรซีเอ็นซี จากการใช้กระแสไฟฟ้า 3 เฟส ด้วยวิธีการแบบกลุ่ม. (DEVELOPMENT OF WORKLOAD MODELS FOR CNC MACHINES FROM 3 - PHASE CURRENT CONSUMPTION USING ENSEMBLE METHOD) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร. เกริก ภิรมย์ โสภา, 66 หน้า.

การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลสามารถพบได้ในหลายภาคส่วนของอุตสาหกรรมยานยนต์ และถูกใช้เป็นหนึ่งเครื่องมือสำคัญในหลายระบบของการผลิตแม่พิมพ์ฉีด ซึ่งเป็นหนึ่งในขั้นตอนสำคัญในการผลิตยานยนต์ ในการออกแบบแม่พิมพ์ฉีด จะเริ่มต้นที่แผนกออกแบบ (CAD) โดยการออกแบบจะใช้โปรแกรมคอมพิวเตอร์ช่วยออกแบบ จากนั้นจะส่งข้อมูลคอมพิวเตอร์ไปยังแผนกแคม (CAM) เพื่อแปลงข้อมูลทางคอมพิวเตอร์ไปเป็นทางเดินของเครื่องมือกัด และข้อมูลที่สำคัญอื่น ๆ เพื่อใช้ในการทำงานที่เครื่องจักร จากนั้นส่งข้อมูลไปยังแผนกเครื่องจักร โดยทั้งนี้ได้มีการพัฒนาระบบตรวจติดตามการทำงานของเครื่องจักรแบบอัตโนมัติ โดยการตรวจสอบการใช้กระแสไฟฟ้า 3 เฟส ของเครื่องจักรซีเอ็นซี เพื่อลดความผิดพลาดของมนุษย์จึงมีการนำเทคนิคการทำเหมืองข้อมูลมาใช้เพื่อจำแนกรูปแบบการใช้กระแสไฟฟ้า โดยสามารถระบุกระบวนการผลิตได้อย่างอัตโนมัติ ในงานวิจัยนี้เลือกใช้การสร้างตัวแบบจากวิธีการดังนี้คือ เบย์อย่างง่าย, เบย์เซียนเน็ตเวิร์ค, โครงข่ายประสาทเทียมชนิดแพร่กลับ, เคสตา, ตารางการตัดสินใจ และ ต้นไม้การตัดสินใจ ส่วนวิธีการแบบกลุ่มที่เลือกใช้คือ AdaBoostM1, Bagging, Stacking และ Vote ซึ่งการใช้วิธีแบบกลุ่มนั้นจะทำให้ตัวแบบมีความแม่นยำมากขึ้น ในการสร้างตัวแบบในขั้นตอนการฝึกและการทดสอบตัวแบบจะใช้วิธีการประเมินความแม่นยำด้วยวิธีการไขว้ข้าม 10 กลุ่ม จากผลการวิจัยจะพบว่าวิธีการสร้างตัวแบบกลุ่มโดยวิธี Bagging โดยใช้ตัวจำแนกต้นไม้การตัดสินใจซึ่งในขั้นตอนการเตรียมข้อมูลไม่ผ่านการทำให้เป็นแบบไม่ต่อเนื่องนั้นให้ค่า AUC เท่ากับ 0.946

ภาควิชาวิศวกรรมคอมพิวเตอร์.....ลายมือชื่อนิสิต.....
 สาขาวิชา.....วิทยาศาสตร์คอมพิวเตอร์.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
 ปีการศึกษา2554.....

5271425321 : MAJOR MASTER OF SCIENCE

KEYWORDS : DATA MINING / ENSEMBLE / BAGGING / CNC MACHINE

THANARAK RAKTHAM : DEVELOPMENT OF WORKLOAD MODELS FOR CNC MACHINES FROM 3 - PHASE CURRENT CONSUMPTION USING ENSEMBLE METHOD. ADVISOR : ASST. PROF. KRERK PIROMSOPA, Ph.D., 66 pp.

Applications using Data Mining Technique can be found in many facets of the automotive industry. In the injection mold manufacturing, a crucial step in automotive production, computational tools have been applied in several places. Starting from design section, injection mold is designed by the Computer Aided Design (CAD) software. The design is then sent to Computer Aided Manufacturing (CAM) section creating tool paths and the necessary information for machining. The final CAM data is later used for machining. In particular, we developed machine workload monitoring system to support the process planner by monitoring 3-phase electrical usage. To reduce human error, we applied data mining technique to electrical usage patterns for identifying current process running in CNC machines. In this paper, classifiers are created by applying 1) Naive Bayes 2) Bayes Net 3) Neural Network 4) KStar 5) Decision Table and 6) J48(C4.5) to electrical data. Later ensemble methods such as 1) AdaBoostM1, 2) Bagging, 3) Stacking, and 4) Vote are applied to each classifier to create more robust models. The models are trained and tested with 10-fold cross validation. Our preliminary result shows that bagging ensemble of J48 classifier with no discretization in the preprocessing step gives the best AUC = 0.946.

Department : Computer Engineering Student's Signature

Field of Study : Computer Science Advisor's Signature

Academic Year : 2011

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. เกริก ภริมย์โสภาอาจารย์ที่
ปรึกษาวิทยานิพนธ์ ที่ท่านได้กรุณาให้ความรู้ คำแนะนำ และคำปรึกษาต่างๆที่ทำให้วิทยานิพนธ์
ฉบับนี้สำเร็จลุล่วงได้ด้วยดี

ขอกราบขอบพระคุณ คุณพุมิโอะ อิวาซึกิ ประธานกรรมการบริหาร บริษัท ออโต
โมทีฟ โมเดล เทคโนโลยี จำกัด ที่ได้ให้ความอนุเคราะห์ในส่วนของข้อมูลจากระบบตรวจติดตาม
เครื่องจักรแบบอัตโนมัติเพื่อใช้ทำงานวิจัยในครั้งนี้

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.เฉลิมเอก อินทนากรวิวัฒน์
ประธานกรรมการ อาจารย์ ดร.ณัฐวุฒิ หนูไพโรจน์ กรรมการ อาจารย์ ดร.พงศ์วิรัช ชีพพิมลชัย
กรรมการภายนอกมหาวิทยาลัย ที่ได้ให้คำปรึกษาและช่วยเหลือตลอดระยะเวลาในการทำงาน
วิจัยนี้ ซึ่งทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากขึ้น

ขอกราบขอบพระคุณ บิดา มารดา ตลอดจนพี่น้องในครอบครัวที่คอยเป็นกำลังใจ
และช่วยเหลือผู้วิจัยมาโดยตลอด และขอบคุณเพื่อนๆทุกคนที่เป็นกำลังใจให้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญภาพ.....	ญ
บทที่	
1. บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.5 วิธีดำเนินการวิจัย.....	4
1.6 ลำดับขั้นตอนในการเสนอผลการวิจัย.....	5
2. เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
2.1 ทฤษฎีที่เกี่ยวข้อง.....	6
2.1.1 การทดสอบไคสแควร์.....	6
2.1.1.1 หลักการทดสอบไคสแควร์.....	6
2.1.1.2 การทดสอบความกลมกลืน (The goodness of fit test).....	7
2.1.2 วิธีการประเมินความแม่นยำด้วยวิธีการไขว้ข้าม k กลุ่ม (k – Fold Cross Validation).....	8
2.1.3 การจำแนกแบบเบย์อย่างง่าย (naive bayes).....	9
2.1.4 การเรียนรู้โครงสร้างเบย์เซียนเน็ตเวิร์ค.....	9
2.1.5 โครงข่ายประสาทเทียมชนิดแพร่กลับ.....	11
2.1.5.1 ขั้นตอนการเรียนรู้.....	12
2.1.5.2 ขั้นตอนการรู้จำ.....	14

บทที่	หน้า
2.1.6 ต้นไม้การตัดสินใจ (decision tree).....	14
2.1.7 การจำแนกแบบกลุ่มด้วยวิธี Bagging (Bootstrap Aggregating).....	15
2.1.8 การจำแนกแบบกลุ่มด้วยวิธี AdaBoost (Adaptive Boosting).....	15
2.1.9 การจำแนกแบบกลุ่มด้วยวิธี Stacking.....	17
2.1.10 การจำแนกแบบกลุ่มด้วยวิธีคะแนนเสียงข้างมาก (majority vote).....	17
2.2 เกณฑ์การประเมินประสิทธิภาพผลพยากรณ์ (Performance Statistic).....	18
2.2.1 เส้นโค้ง ROC (Receiver Operating Characteristic curve).....	18
2.2.2 พื้นที่ใต้โค้ง ROC (Area under the ROC Curve (AUC)).....	18
2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	19
3. วิธีดำเนินการวิจัย.....	21
3.1 ขั้นตอนในการทำงานวิจัย.....	21
3.2 รายละเอียดข้อมูลดิบ การสร้างและทดสอบชุดข้อมูล.....	24
3.2.1 ข้อมูลดิบ.....	24
3.2.2 การสร้างชุดข้อมูล.....	24
3.2.3 การทดสอบชุดข้อมูล.....	25
3.3 การสร้างตัวแบบด้วยโปรแกรม WEKA โดยใช้ KnowledgeFlow.....	25
4. ผลการวิเคราะห์ข้อมูล.....	30
4.1 ผลการสุ่มข้อมูลจากฐานข้อมูล.....	30
4.2 ผลความแม่นยำของตัวแบบแต่ละวิธี.....	30
5. สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	35
5.1 สรุปผลการวิจัย.....	35
5.2 ข้อเสนอแนะ.....	35
รายการอ้างอิง.....	37
ภาคผนวก.....	39
ประวัติผู้เขียนวิทยานิพนธ์.....	56

สารบัญตาราง

	หน้า
ตารางที่ 3.1	ข้อมูลดิบทั้งหมดแบ่งตามกระบวนการผลิต..... 24
ตารางที่ 3.2	จำนวนชุดข้อมูลที่สุ่มเลือกเพิ่มเติมเพื่อทดสอบกับตัวแบบแบบเดียว..... 25
ตารางที่ 4.1	ค่าโคสแควร์จากการสุ่มของแต่ละกระบวนการผลิต..... 30
ตารางที่ 4.2	ค่า AUC จากตัวแบบแบบเดียวโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง..... 30
ตารางที่ 4.3	ค่า AUC จากตัวแบบแบบเดียวโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง..... 31
ตารางที่ 4.4	ค่า AUC จากวิธีโครงข่ายประสาทเทียมโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง..... 31
ตารางที่ 4.5	ค่า AUC จากวิธีโครงข่ายประสาทเทียมโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง..... 32
ตารางที่ 4.6	ค่า AUC จากวิธีแบบกลุ่มโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง..... 32
ตารางที่ 4.7	ค่า AUC จากวิธีแบบกลุ่มโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง(ต่อ)..... 33
ตารางที่ 4.8	ค่า AUC จากวิธีแบบกลุ่มโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง. 33
ตารางที่ 4.9	ค่า AUC จากวิธีแบบกลุ่มโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง (ต่อ)..... 34

สารบัญภาพ

		หน้า
ภาพที่ 1.1	กระแสไฟฟ้าแบบ 3 เฟสในกระบวนการผลิตที่เรียกว่า SEMI FINISH CUTTING....	2
ภาพที่ 1.2	กระแสไฟฟ้าแบบ 3 เฟสในกระบวนการผลิตที่เรียกว่า UPPER SHAPE.....	2
ภาพที่ 1.3	ขั้นตอนกระบวนการวิจัย.....	4
ภาพที่ 2.1	การแบ่งข้อมูลสำหรับการเรียนรู้และทดสอบ.....	8
ภาพที่ 2.2	อัลกอริทึมการทำงานของการเรียนรู้แบบแพร่กลับ(Backpropagation).....	11
ภาพที่ 2.3	อัลกอริทึม Bagging.....	15
ภาพที่ 2.4	อัลกอริทึม AdaBoostM1.....	16
ภาพที่ 2.5	อัลกอริทึม AdaBoostM1 (ต่อ).....	17
ภาพที่ 2.6	รูปแบบความเห็นพ้องของคน 10 คน.....	17
ภาพที่ 2.7	ค่าไคสแควร์ที่ได้จากการทดลองของ Brian Connolly.....	19
ภาพที่ 3.1	ขั้นตอนในการทำงานวิจัย.....	21
ภาพที่ 3.2	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Stacking.....	25
ภาพที่ 3.3	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Majority Vote.....	26
ภาพที่ 3.4	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับ วิธีเบย์อย่างง่าย และเบย์เซียนเน็ตเวิร์ค.....	26
ภาพที่ 3.5	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธีเคสสตาร์ ตาราง ตัดสินใจ และต้นไม้การตัดสินใจ.....	27
ภาพที่ 3.6	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Bagging.....	28
ภาพที่ 3.7	การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี AdaBoostM1.....	29

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

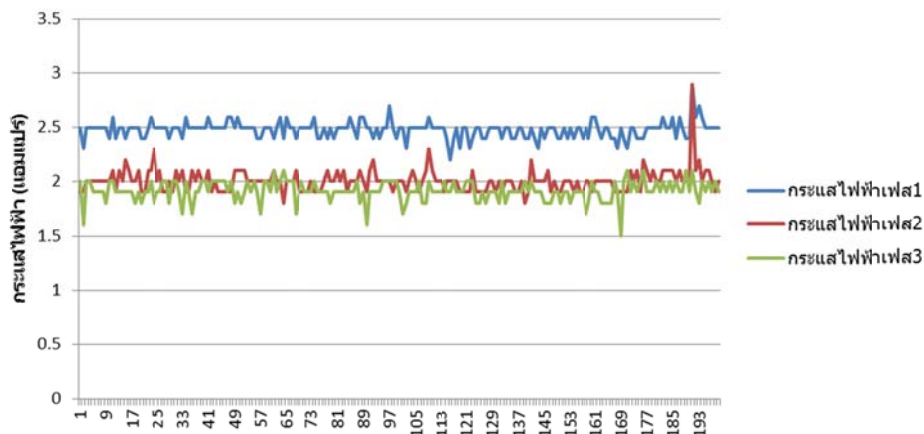
เนื่องจากปัจจุบันอุตสาหกรรมการผลิตมีการแข่งขันกันในระดับสูง ฉะนั้นจึงมีความพยายามที่จะลดต้นทุนในทุก ๆ ด้าน โดยผู้บริหารจึงจำเป็นต้องมีเครื่องมือที่สามารถช่วยในการตัดสินใจได้อย่างรวดเร็ว เพื่อความได้เปรียบทางธุรกิจ จึงมีแนวคิดที่จะนำเทคโนโลยีสารสนเทศมาใช้เพื่อเพิ่มขีดความสามารถในการแข่งขัน โดยหนึ่งในวัตถุประสงค์ อาจเพื่อให้สามารถทราบถึงสถานะของระบบการผลิตแบบปัจจุบัน ทำให้ผู้บริหารสามารถตอบสนอง โดยออกนโยบาย หรือปรับเปลี่ยนได้อย่างทันที่ อีกทั้งยังสามารถลดต้นทุนในส่วนของแรงงานคน และยังเพิ่มความแม่นยำให้กับระบบได้ดียิ่งขึ้น เพื่อให้การตัดสินใจดำเนินธุรกิจไม่ผิดพลาดหรือล่าช้ากว่าคู่แข่ง การที่มีเครื่องมือสนับสนุนเพื่อให้ได้รับข้อมูลที่ถูกต้อง แม่นยำและมีความเป็นปัจจุบันจึงมีความสำคัญ

การเหมืองข้อมูลได้เข้ามามีบทบาทในทางธุรกิจอย่างมาก เพราะการที่ฐานข้อมูลมีขนาดใหญ่ขึ้น ซึ่งยากแก่การค้นหาได้อย่างทันกาลด้วย DBMS (Database Management System) การทำเหมืองข้อมูลจึงเข้ามาช่วยแก้ปัญหาดังกล่าว ซึ่งทำให้ข้อมูลที่เป็นที่สนใจของผู้บริหารสามารถค้นหาได้ง่ายขึ้น สามารถนำมาเทียบเคียง ดูแนวโน้ม และนำข้อมูลที่จำเป็นของบริษัทส่งกลับให้ผู้บริหารตัดสินใจได้อย่างทันกาล

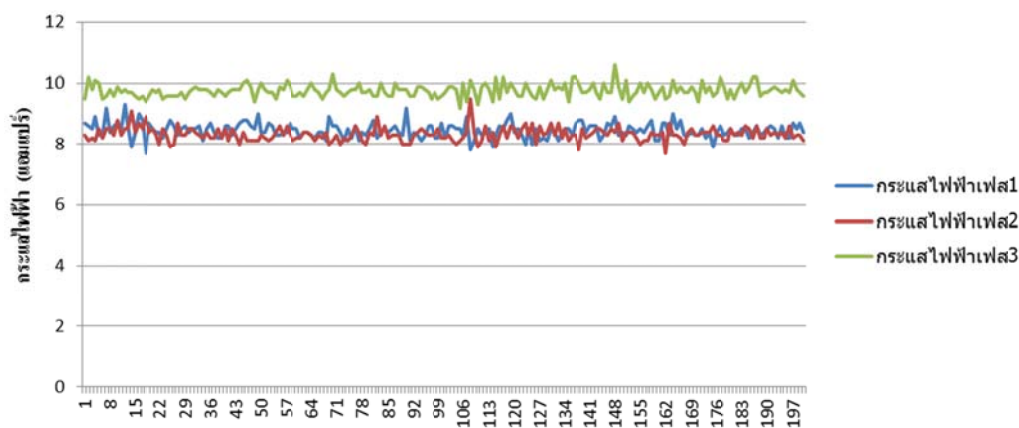
ในระบบการผลิตแม่พิมพ์นั้นขั้นตอนการผลิตจะเริ่มต้น ด้วยกระบวนการออกแบบด้วยโปรแกรมขึ้นรูปสามมิติ (CAD Software) และถูกเปลี่ยนเป็นค่าทางเทคนิคเพื่อป้อนให้กับเครื่องจักรด้วยโปรแกรมคำนวณทางเดินของเครื่องมือกัด (CAM Software) และค่าคุณสมบัติทางเทคนิคอื่น ๆ ที่จำเป็น เสร็จแล้วจึงส่งต่อให้แผนกเครื่องจักรเพื่อทำการผลิต

ในแผนกเครื่องจักรกระบวนการหนึ่งที่สำคัญคือ กระบวนการกัดชิ้นงานด้วยเครื่องมือกัด ซึ่งหัวกัดในเครื่องกัดขนาดกลางและใหญ่จะถูกส่งกำลังด้วยระบบไฟฟ้าแบบ 3 เฟส จึงมีแนวคิดที่จะนำค่าการใช้กระแสไฟฟ้าจากหัวกัดมาใช้วิเคราะห์ข้อมูล อย่างเช่นตรวจจับเวลาการทำงานเครื่องจักรเพื่อใช้คำนวณต้นทุนการผลิต และการวิเคราะห์และระบุชนิดของการกัดแบบอัตโนมัติ เพื่อใช้วิเคราะห์ประสิทธิภาพการทำงานของคนและเครื่องจักร

การใช้พลังงานไฟฟ้าของหัวกัดนั้นมีการใช้ไฟฟ้าแบบ 3 เฟส และมีรูปแบบที่ซับซ้อนเกินกว่าที่มนุษย์ หรือเครื่องมือแบบพื้นฐานจะแยกแยะได้



ภาพที่ 1.1 : กระแสไฟฟ้าแบบ 3 เฟสในกระบวนการผลิตที่เรียกว่า SEMI FINISH CUTTING



ภาพที่ 1.2 : กระแสไฟฟ้าแบบ 3 เฟสในกระบวนการผลิตที่เรียกว่า UPPER SHAPE

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อศึกษาขั้นตอนและวิธีการสร้างตัวแบบที่ถูกสร้างตัวด้วยวิธีการแบบเดี่ยว และแบบกลุ่ม

1.2.2 เพื่อทดลองและเปรียบเทียบประสิทธิภาพด้านความแม่นยำของตัวแบบ ที่ถูกสร้างด้วยวิธีการแบบเดี่ยวและแบบกลุ่ม ด้วยพื้นที่ได้คิงอาร์ไอซี

1.2.3 เพื่อสร้างตัวแบบที่มีประสิทธิภาพซึ่งถูกสร้างจากข้อมูลการใช้กระแสไฟฟ้า 3 เฟสที่เหมาะสมสำหรับนำไปใช้งานในการวิเคราะห์ขั้นตอนการทำงานของเครื่องจักรซีเอ็นซีแบบอัตโนมัติ

1.3 ขอบเขตของการวิจัย

1.3.1 งานวิจัยนี้ใช้ข้อมูลค่ากระแสไฟฟ้าแบบ 3 เฟส ของชุดหัวกัด (spindle) ของเครื่องจักรซีเอ็นซีชนิดเครื่องกัด จำนวนสองเครื่อง ซึ่งข้อมูลจะถูกเก็บทุก ๆ 3 วินาที ข้อมูลจะถูกเลือกมาแบบสุ่ม โดยสนใจเฉพาะตอนที่เครื่องจักรทำงานคือ ค่าผลรวมของค่ากระแสมีค่ามากกว่าหรือเท่ากับ 1 แอมป์แปร์

1.3.2 งานวิจัยนี้ทำการทดลองโดยเลือกข้อมูลจากฐานข้อมูล โดยสุ่มจากฐานข้อมูลด้วยฟังก์ชัน RAND ของ Microsoft SQL Server ตามวิธีการของ Brian Connolly (2004) แล้วทำการสร้างตัวแบบด้วยโปรแกรม Weka จากนั้นนำตัวแบบมาทดสอบหาค่าความแม่นยำ เพื่อทำการศึกษา หาตัวแบบที่เหมาะสมที่ใช้ในการแยกแยะขั้นตอนการผลิตของเครื่องจักรซีเอ็นซี

1.3.3 งานวิจัยนี้เลือกใช้ตัวเปรียบเทียบพื้นฐาน (Baseline) โดยใช้ตัวแยกแยะแบบเดี่ยวที่ให้ผลต่ำที่สุดซึ่งมีค่ามากกว่าการเลือกแบบสุ่ม

1.3.4 งานวิจัยนี้ทำการทดลองโดยเลือกวิธีการสร้างโมเดลสำหรับตัวแยกแยะแบบเดี่ยวและแบบกลุ่มดังนี้

1.3.4.1 การจำแนกแบบเบย์อย่างง่าย (naive bayes)

1.3.4.2 การจำแนกแบบเบย์เน็ตเวิร์คด้วย K2 Local Search

1.3.4.3 การจำแนกแบบเบย์เน็ตเวิร์คด้วย Hill Climber Local Search

1.3.4.4 การจำแนกแบบเบย์เน็ตเวิร์คด้วย Genetic Local Search

1.3.4.5 การจำแนกด้วยนิวรอลเน็ตเวิร์กแบบคพรอพาเกชัน

1.3.4.6 การจำแนกด้วยเคสตาร์

1.3.4.7 ต้นไม้การตัดสินใจ (decision tree)

1.3.4.8 การจำแนกแบบกลุ่มด้วยวิธี AdaBoostM1

1.3.4.9 การจำแนกแบบกลุ่มด้วยวิธี Bagging

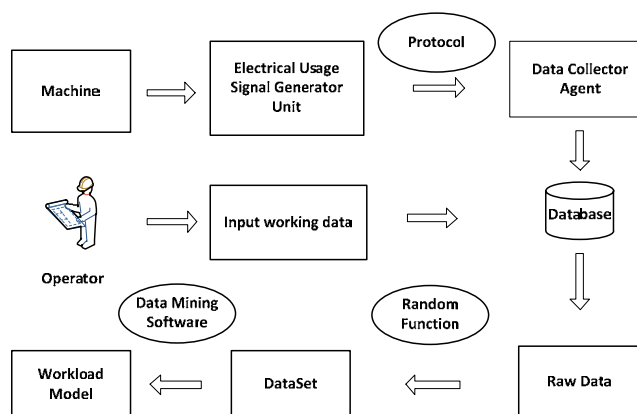
1.3.4.10 การจำแนกแบบกลุ่มด้วยวิธี Majority Vote

1.3.4.11 การจำแนกแบบกลุ่มด้วยวิธี Stacking

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 เข้าใจถึงวิธีที่ใช้ในการสร้างตัวแบบที่ใช้ตัวแยกแยะแบบเดี่ยวและแบบกลุ่ม
- 1.4.2 ได้รับความรู้เกี่ยวกับประสิทธิภาพความแม่นยำ ข้อดีและข้อเสีย ของการสร้างตัวแบบแต่ละวิธีเมื่อนำมาประยุกต์ใช้กับระบบการทำงานของเครื่องจักรซีเอ็นซีชนิดเครื่องกัด
- 1.4.3 ได้แบบจำลองที่สามารถนำมาประยุกต์ใช้กับระบบการทำงานของเครื่องจักรซีเอ็นซี

1.5 วิธีดำเนินการวิจัย



ภาพที่ 1.3 : ขั้นตอนกระบวนการวิจัย

- 1.6.1 ศึกษาผลงานเอกสารและรวบรวมข้อมูลที่เกี่ยวข้องได้แก่ข้อมูลเครื่องจักร, ข้อมูลโปรโตคอลตัวรับส่งข้อมูล, อุปกรณ์ที่ใช้จับกระแสไฟฟ้าแบบ 3 เฟส, เครื่องอ่านค่ากระแสไฟฟ้าแบบ 3 เฟส, ตัวเอเจนที่จะใช้รวบรวมข้อมูล, ฐานข้อมูลที่ใช้เก็บข้อมูล และระบบประมวลผลข้อมูล
- 1.6.2 จัดสร้างระบบตรวจติดตามเครื่องจักรแบบอัตโนมัติ (Machine Monitoring System)
- 1.6.3 ทำการสุ่มข้อมูลจากฐานข้อมูล ค่ากระแสไฟฟ้าแบบ 3 เฟสที่เก็บได้จากระบบตรวจติดตามเครื่องจักรแบบอัตโนมัติด้วยฟังก์ชัน RAND ของ Microsoft SQL Server จากนั้นคำนวณหาค่าไคสแควร์เพื่อตรวจสอบว่าเป็นการสุ่มที่ดีหรือไม่
- 1.6.4 ทำการสร้างตัวแบบจากวิธีแบบเดี่ยวและแบบกลุ่ม
- 1.6.5 เปรียบเทียบประสิทธิภาพในแต่ละวิธีด้วยค่าพื้นที่ใต้เส้นโค้ง ROC

ในการวิเคราะห์ข้อมูลข้างต้นจะใช้โปรแกรมสำเร็จรูปในการวิเคราะห์ข้อมูลได้แก่ WEKA และ Microsoft Excel

1.6 ลำดับขั้นตอนในการเสนอผลการวิจัย

งานวิจัยในครั้งนี้แบ่งเนื้อหาออกเป็น 5 บท ดังนี้

บทที่ 1 บทนำประกอบด้วยความเป็นมา และความสำคัญของปัญหา วัตถุประสงค์ของการวิจัย ขอบเขตของการวิจัย ประโยชน์ที่คาดว่าจะได้รับ และวิธีการดำเนินงานวิจัย

บทที่ 2 ทฤษฎีที่ใช้ในการศึกษา และ งานวิจัยที่เกี่ยวข้อง

บทที่ 3 วิธีการดำเนินการวิจัย ประกอบด้วย ขั้นตอนการจัดสร้างระบบตรวจติดตามเครื่องจักรแบบอัตโนมัติ และ วิธีในการสร้างตัวแบบวิธีต่างๆ

บทที่ 4 ผลการวิเคราะห์ประกอบด้วย

1) ผลการวิเคราะห์การสุม

2) ผลการคำนวณค่า ROC ของตัวแบบที่ได้จากวิธีการคัดเลือกค่าคุณลักษณะสำคัญวิธีต่างๆ

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ ในการศึกษาต่อในอนาคต

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้องกับการวิจัยในเรื่องการคัดเลือกค่าคุณลักษณะสำคัญ

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การทดสอบไคสแควร์

การทดสอบไคสแควร์ (Chi-Square Test) เป็นการทดสอบเพื่อเปรียบเทียบข้อมูลที่อยู่ในรูปของความถี่หรืออยู่ในรูปของสัดส่วน ดังเช่นการศึกษาเจตคติความคิดเห็น ความสนใจ หรือการยอมรับ เป็นต้น โดยไม่สามารถวัดค่าออกมาเป็นตัวเลขที่แน่นอน แต่สามารถจำแนกออกเป็นหมวดหมู่ได้ เช่น มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด ดี ไม่ดี เป็นต้นซึ่งเป็น ข้อมูลที่เกิดจากการเก็บรวบรวมจากตัวแปร ที่เกี่ยวข้องแล้วจำแนกออกมาเป็นความถี่หรือสัดส่วน ถ้าหากต้องการศึกษาว่าการแจกแจงความถี่ ของข้อมูลที่ได้จากตัวแปรหนึ่ง เป็นไปลักษณะใด หรือถ้าหากต้องการเปรียบเทียบตัวแปร 2 กลุ่มหรือมากกว่า 2 กลุ่มว่ามีความสัมพันธ์กันหรือไม่ การทดสอบไคสแควร์จะเหมาะสมกว่าการทดสอบด้วย z เนื่องจากการทดสอบด้วย z เหมาะสำหรับการทดสอบสัดส่วนของประชากรเพียงกลุ่มเดียว หรือการทดสอบความแตกต่างระหว่างสัดส่วนของสิ่ง ที่สนใจจากประชากร 2 กลุ่มเท่านั้น การทดสอบไคสแควร์จึงเป็นวิธีการทางสถิติที่นิยมใช้มากในการเปรียบเทียบ หรือทดสอบข้อมูลที่เป็นความถี่หรือข้อมูลที่อยู่ในรูปของสัดส่วน โดยเฉพาะ การใช้วิเคราะห์ข้อมูลจากการแบบสอบถามแบบมาตราส่วนประเมินค่า

2.1.1.1 หลักการทดสอบไคสแควร์

สมมติมีประชากรที่มีการแจกแจงปกติโดยมีค่าเฉลี่ย (μ) และความแปรปรวน (σ)

ถ้าหากสุ่มประชากรออกมา 1 คน แล้วนำมาแทนค่าในสมการที่ 2.1

$$Z^2 = \frac{(X-\mu)^2}{\sigma^2} \quad (2.1)$$

หลังจากนั้นจึงนำค่า ของ Z^2 ตั้งแต่ 0 จนถึง ∞ ไปเขียนกราฟแสดงการแจกแจงของ Z^2 จะพบว่า การแจกแจงของกราฟ Z^2 ที่ได้จะมีลักษณะเหมือนกับการแจกแจงของ

ไคสแควร์ (χ^2) ที่มีระดับของเสรีภาพเป็น 1 แต่ถ้าหากสุ่มประชากรออกมา N คน จำนวน 1 ครั้งและนำมาแทนค่าใน

สูตรเดียวกัน เพื่อหาค่า Z^2 และผลรวมของ $\sum Z^2$ กระทำลักษณะเช่นเดียวกันนี้เป็นจำนวน ∞ ครั้งแล้วนำไปเขียนกราฟเพื่อแสดงการแจกแจงของ $\sum Z^2$ จะพบว่าได้เส้นกราฟที่ได้จะมีลักษณะเหมือนกับการแจกแจงของไคสแควร์ ที่มีระดับของเสรีภาพเป็น N นั่นคือ

$$\chi^2(N) = \sum_{i=1}^N Z_i^2 \quad (2.2)$$

$$\chi^2(N) = \sum \frac{(X_i - \mu)^2}{\sigma^2} = \frac{(X - \mu)^2}{\sigma^2} \quad (2.3)$$

ในกรณีของการแจกแจงความแปรปรวน สมมติมีประชากรที่มีการแจกแจงปกติ และทราบค่าของความแปรปรวน (σ^2) ของประชากรกลุ่มนี้ สมมติต่อไปว่ามีการสุ่มประชากรออกมา N ค่า จำนวน ∞ ครั้งพร้อมทั้งหาค่าความแปรปรวน (S^2) ในแต่ละครั้ง หลังจากนั้นจึงนำค่ามาแจกแจงความแปรปรวนของประชากร จะได้ว่า

$$\chi^2(N - 1) = \frac{(N - 1)S^2}{\sigma^2} \quad (2.4)$$

$$\text{หรือ} \quad S^2 = \frac{\chi^2 \sigma^2}{N - 1} \quad (2.5)$$

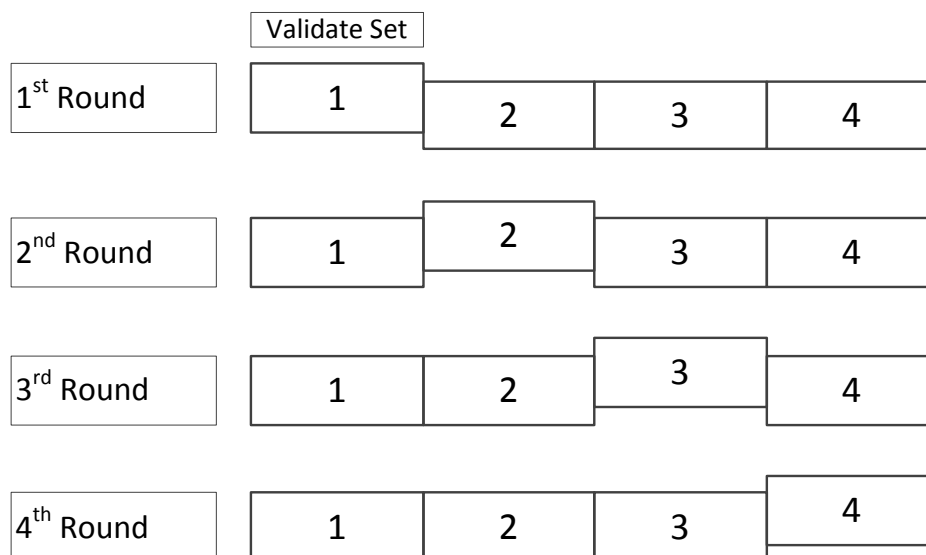
จะพบว่า $\frac{\sigma^2}{N - 1}$ คือค่าคงที่ของความแปรปรวนของประชากรและกลุ่มตัวอย่าง ดังนั้นการแจกแจงของ S^2 จึงขึ้นอยู่กับไคสแควร์จะมีความสัมพันธ์กับ z และความแปรปรวนตามความสัมพันธ์ของสูตรต่าง ๆ ที่กล่าวมาแล้ว

2.1.1.2 การทดสอบความกลมกลืน (The goodness of fit test)

การทดสอบความกลมกลืน (The goodness of fit test) เป็นการทดสอบไคสแควร์ เพื่อศึกษาว่าการแจกแจงความถี่ของตัวแปรเป็นไปตามรูปแบบที่กำหนดไว้หรือไม่ โดยศึกษาจากตัวแปรเพียงตัวเดียว โดยการเปรียบเทียบระหว่างข้อมูลจากตัวแปรกับข้อมูลที่ได้จากความคาดหมายหรือจากทฤษฎีใด ๆ ว่ามีความสอดคล้องกันหรือไม่

2.1.2 วิธีการประเมินความแม่นยำด้วยวิธีการไขว้ข้าม k กลุ่ม (k – Fold Cross Validation)

วิธีการนี้จะแบ่งข้อมูลออกเป็นกลุ่มจำนวน k กลุ่ม (k-Fold) โดยเรียกว่า “วิธีการประเมินตัวแบบด้วยวิธีการไขว้ข้าม k กลุ่ม” หรือ k Fold Cross – Validation เหมาะสำหรับชุดข้อมูลจำนวนไม่มาก โดยการแบ่งข้อมูลที่มีอยู่ออกเป็นส่วนๆ เรียกว่า โฟลด์ (Fold) ซึ่งข้อมูลจะถูกแบ่งออกเป็น k ส่วน ในตอนแรกจะเก็บส่วนแรกไว้ เพื่อเป็นชุดข้อมูลทดสอบ และส่วนที่ 2 ถึง k เป็นส่วนที่ใช้สำหรับเป็นชุดข้อมูลสำหรับการเรียนรู้ ทำเช่นนี้ไป จนทุกๆ ส่วนถูกใช้สำหรับทดสอบจนครบ จากนั้นหาค่าเฉลี่ยของความผิดพลาดของแต่ละตัวแบบที่ได้ ซึ่งสามารถทำให้มองเห็นภาพรวมของความผิดพลาดที่เกิดขึ้นได้ หรือเป็นประสิทธิภาพของตัวแบบที่ได้ ข้อมูลทุกตัวอย่างจะได้เป็นทั้งชุดทดสอบและชุดสอน ตัวอย่างดังแสดงในภาพที่ 2.1 เป็นการประเมินตัวแบบด้วยวิธีการไขว้ข้าม 4 กลุ่ม โดยการจัดเตรียมข้อมูลทั้งหมด เพื่อใช้ในการเรียนรู้โดยวิธีการจำแนกข้อมูลนั้น ข้อมูลได้ถูกแบ่งออกเป็น 2 ชุดโดย ชุดแรกสำหรับเป็นชุดข้อมูลทดสอบ มีจำนวน 10 % ของจำนวนข้อมูลทั้งหมด และชุดที่ 2 ถึง 4 เป็นชุดข้อมูลสำหรับการเรียนรู้ คิดเป็น 90 % ข้อมูลทั้งหมด



ภาพที่ 2.1 : การแบ่งข้อมูลสำหรับการเรียนรู้และทดสอบ

2.1.3 การจำแนกแบบเบย์อย่างง่าย (naive bayes)

เป็นเทคนิคที่ถูกตั้งชื่อตาม Thomas Bayes (1702-1761) โดยใช้หลักความน่าจะเป็นซึ่งอยู่บนพื้นฐานของ ทฤษฎีของเบย์ (Bayes' theorem) และสมมติฐานที่กำหนดการเกิดของเหตุการณ์ต่างๆ ที่ใช้ในการจัดกลุ่มเป็นอิสระต่อกัน (independence) การจำแนกแบบเบย์อย่างง่าย จะใช้การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตาม เพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็น ในทางทฤษฎี การทำนายผลของการจำแนกแบบเบย์อย่างง่ายจะถูกตัดง ถ้าตัวแปรอิสระทั้งหมดเป็นอิสระต่อกัน ไม่ขึ้นกับตัวแปรอิสระตัวใดตัวหนึ่ง การจำแนกแบบเบย์อย่างง่ายนั้นไม่รองรับข้อมูลที่เป็นข้อมูลต่อเนื่อง ดังนั้นตัวแปรอิสระ หรือตัวแปรตามที่เป็นค่าต่อเนื่องควรถูกแบ่งเป็นช่วง การจำแนกแบบเบย์อย่างง่ายสามารถให้ผลลัพธ์ที่รวดเร็วและเป็นเครื่องมือที่ดีในการพัฒนาตัวแบบ

2.1.4 การเรียนรู้โครงสร้างเบย์เซียนเน็ตเวิร์ค

การเรียนรู้เบย์เซียนเน็ตเวิร์คแบ่งได้เป็น 2 ส่วนคือ เรียนรู้โครงสร้างซึ่งจะพิจารณาถึงโครงสร้างหรือ โทโปโลยีของเบย์เซียนเน็ตเวิร์ค และอีกส่วนจะเป็นการเรียนรู้พารามิเตอร์คือ Conditional Probabilities ในแต่ละโหนดของเน็ตเวิร์ค

อัลกอริทึมแรกๆที่ใช้เรียนรู้โครงสร้างของเบย์เซียนเน็ตเวิร์คจากข้อมูลได้แก่ K2 Algorithm ของ Cooper และ Herskovits (Cooper และ Herkivit, 1992) อัลกอริทึมนี้ต้องทราบ Topological Order ของเบย์เซียนเน็ตเวิร์คล่วงหน้าและจะทำการสร้างโครงสร้างของเบย์เซียนเน็ตเวิร์ค B_s จากฐานข้อมูล D ที่ทำให้ $P(B_s, D)$ มีค่าสูงสุด (Heckerman, 1995; Heckerman et al., 1995) สำหรับ $P(B_s, D)$ สามารถคำนวณได้จาก

$$P(B_s, D) = P(B_s) \prod_{i=1}^n f(i, \pi) \quad (2.6)$$

โดยที่ $f(i, \pi)$ คือค่าความน่าจะเป็นของฐานข้อมูลซึ่งมี Parent ของ x_i คือ π_i สามารถคำนวณได้จาก

$$f(i, \pi) = \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} \alpha_{ijk}! \quad (2.7)$$

D = ฐานข้อมูลสำหรับหาโครงสร้างของเบย์เซียนเน็ตเวิร์ค

B_s = โครงสร้างเน็ตเวิร์คที่กำลังพิจารณา

n = จำนวนโหนดในเน็ตเวิร์ค

π_i = เซตของ Parent ของโหนด x_i

\emptyset_i = รายการของค่า Possible Instantiations จาก Parent ของ x_i ในฐานข้อมูล D ซึ่งคือ ถ้า p_1, \dots, p_s คือ Parent ของ x_i แล้ว \emptyset_i คือ Cartesian Product

$$\{v_1^{p_1}, \dots, v_{r_{p_1}}^{p_1}\} \times \dots \times \{v_1^{p_s}, \dots, v_{r_{p_s}}^{p_s}\} \quad (2.8)$$

ของค่าที่เป็นไปได้ทั้งหมดจากแอททริบิวต์ของ p_1 ถึง p_s

$$q_i = |\emptyset_i|$$

$$r_i = |V_i|$$

α_{ijk} = จำนวนกรณีที่เกิดขึ้นใน D ซึ่งแอททริบิวต์เป็น Instantiation กับค่าของ k^{th} และ Parent ของ x_i ใน π_i เป็น Instantiations กับค่า j^{th} Instantiation ใน \emptyset_i

$N_{ij} = \sum_{k=1}^{r_i} \alpha_{ijk}$ ซึ่งคือจำนวนกรณีในฐานข้อมูลซึ่ง Parent ของ x_i เป็น Instantiations กับค่า j^{th} Instantiation ใน \emptyset_i

อัลกอริทึม K2 จะมีข้อจำกัดซึ่งจะต้องมีการกำหนด Topology Order ของเบย์เซียนเน็ตเวิร์คล่วงหน้าซึ่งอาจทำไม่ได้ในทางปฏิบัติ นอกจากนี้ถ้าอัลกอริทึมมีลักษณะเป็น Exhaustive Search จะทำให้เสียเวลาในการทำงานมาก

Larranaga et al. (1996) และ Larranaga et al. (1997) ได้นำเสนอรูปแบบโครงโมไซม์ที่เรียกว่า Connectivity Matrix (c_{ij}) เพื่อใช้แทนโครงสร้างเบย์เซียนเน็ตเวิร์คโดยมีรูปแบบโครงโมไซม์ที่เรียกว่า Connectivity Matrix (C_{ij}) เพื่อใช้แทนโครงสร้างเบย์เซียนเน็ตเวิร์คโดยมีรูปแบบโครงโมไซม์เป็นเมตริกซ์ C ขนาด $n \times n$ โดย n คือจำนวนตัวแปรหรือโหนดแต่ละ c_{ij} กำหนดได้ดังนี้

$$C_{ij} = \begin{cases} 1, & \text{if } j \text{ is a parent of } i \\ 0, & \text{otherwise} \end{cases} \quad (2.9)$$

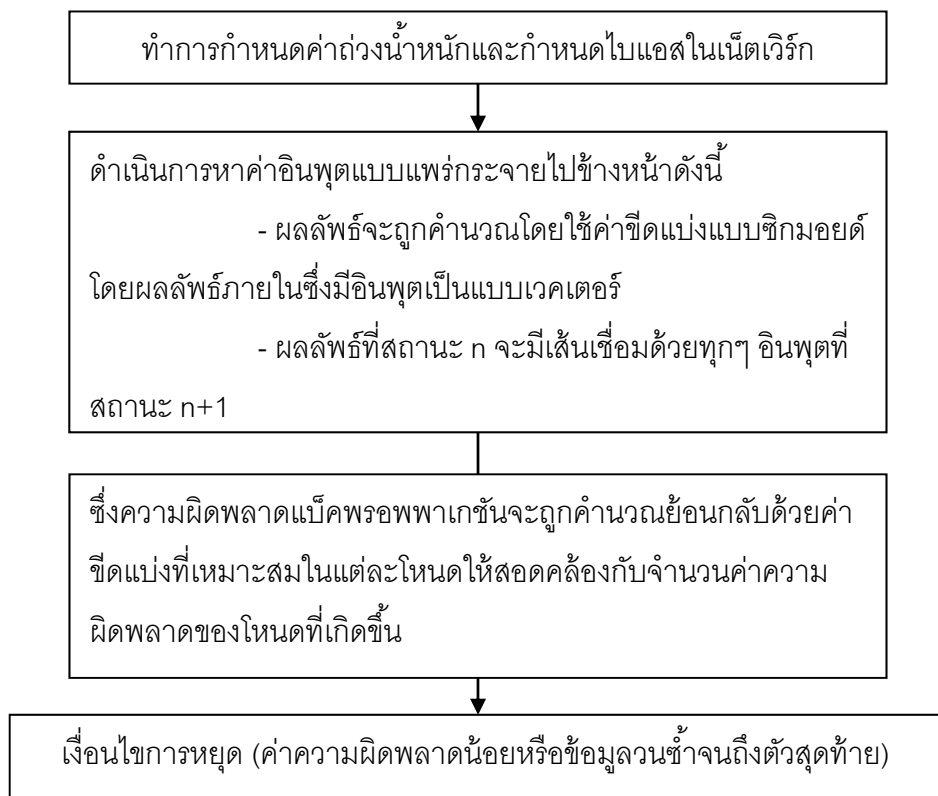
แต่ละประชากรแทนด้วย String:

$$C_{11}C_{21} \dots C_{n1}C_{12}C_{22} \dots C_{n2} \dots C_{1n}C_{2n} \dots C_{nn}$$

โดยมีความยาวของ String คือ n^2

2.1.5 โครงข่ายประสาทเทียมชนิดแพร่กลับ

โครงข่ายประสาทเทียมชนิดแพร่กลับ ซึ่งเป็น Feed Forward Network โดยในขั้นตอนการทำงานจะไม่มีป้อนผลลัพธ์ที่ได้ในแต่ละโหนดย้อนกลับไปยังโหนดที่ส่งข้อมูลมาให้ (ในขั้นตอนการเรียนรู้จะมีการปรับปรุ้ค่าน้ำหนักของแต่ละโหนดเมื่อตอบคำตอบผิด โดยส่งข้อมูลย้อนกลับไปจึงเรียกว่า Backpropagation) ค่าน้ำหนักที่ได้จะเป็นค่าน้ำหนักที่ทำให้ค่าผิดพลาดกำลังสองเฉลี่ย (Mean Square Error) มีค่าต่ำสุด โครงสร้างของโครงข่ายประสาทเทียมชนิดแพร่กลับ ประกอบด้วย ชั้นอินพุต (input layer) ชั้นซ่อน(hidden layer) ซึ่งอาจมีได้มากกว่า 1 ชั้น และชั้นเอาต์พุต (output layer) ซึ่งแต่ละโหนดจะมีการเชื่อมต่อกันแบบเชื่อมต่อทั้งหมด (Fully Connected)



ภาพที่ 2.2 : อัลกอริทึมการทำงานของการเรียนรู้แบบแพร่กลับ(Backpropagation)

2.1.5.1 ขั้นตอนการเรียนรู้

ก่อนที่จะนำโครงข่ายประสาทเทียมไปใช้งานในการรู้จำนั้น จะต้องสอนโครงข่ายประสาทเทียม (train) โดยการหาสามารถหาความสัมพันธ์ของ อินพุต และ เอาต์พุต ให้ถูกต้อง หรือผิดพลาดให้น้อยที่สุดเสียก่อนซึ่งเรียกว่าขั้นตอนการเรียนรู้ (learning mode, training mode) และมีวิธีการทำงานดังนี้

2.1.5.1.1 กำหนดค่าเริ่มต้นแก่ค่าน้ำหนัก (weight) และ ค่า bias โดยการสุ่มตัวเลขที่มีค่าน้อยๆ เช่นระหว่าง - 0.1 ถึง 0.1

2.1.5.1.2 ส่งข้อมูลที่จะสอนนิรลเน็ตเวิร์ก เข้าสู่อินพุตโนด แล้วดำเนินการเหมือนในขั้นตอนการรู้จำ คือ คำนวณค่าผลลัพธ์ของแต่ละโนด (ยกเว้นอินพุตโนด) ตามสมการที่ 2.10 และ 2.11

$$O_j = F(\theta_j + \sum_{i=1}^n w_{ji} o_i) \quad (2.10)$$

$$F(a) = \frac{1}{1+e^{-a}} \quad (2.11)$$

เมื่อ o_j คือผลลัพธ์ของโนดที่ j

θ_j คือค่า bias ของโนดที่ j

w_{ji} คือค่าน้ำหนักของโนดที่ i ซึ่งส่งข้อมูล o_i มาให้โนดที่ j

o_i คือผลลัพธ์ของโนดที่ i

$F(a)$ คือ ฟังก์ชันกระตุ้น (activation function) ซึ่งใช้ฟังก์ชันซิกมอยด์ (sigmoid function)

2.1.5.1.3 ทำการปรับค่าน้ำหนัก เริ่มต้นที่เอาต์พุตโนด แล้วค่อยทำการย้อนกลับไปยังฮิดเดนโนด ในชั้นซ้อน โดยคำนวณค่าน้ำหนักใหม่ ($W_{ji}(t+1)$) ได้ตามสมการที่ 2.12

$$W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji} \quad (2.12)$$

เมื่อ $W_{ji}(t)$ คือค่าน้ำหนักของโนด ที่ i ซึ่งเชื่อมกับโนดที่ j เวลาที่ t (หรือ iteration ที่ t)

ΔW_{ji} คือค่าน้ำหนักที่ต้องการปรับ

ค่าน้ำหนักที่ต้องการปรับคำนวณได้จากสมการที่ 2.13

$$\Delta W_{ji} = \eta \delta_j O_i \quad (2.13)$$

เมื่อ η คือค่า Learning Rate ($0 < \eta < 1$)

δ_j คือค่าความผิดพลาดของผลลัพธ์ของโนดที่ j

แต่ในกรณีที่ต้องการที่จะให้โครงข่ายประสาทเทียม เข้า (convergence) เร็วขึ้น ก็จะสามารถทำได้โดยเพิ่มค่า inertia หรือ momentum ตามสมการที่ 2.14

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_j O_i + \alpha (W_{ji}(t) - W_{ji}(t-1)) \quad (2.14)$$

เมื่อ α คือค่า Inertia หรือ momentum ($0 < \alpha < 1$)

2.1.5.1.4 ค่าความผิดพลาดของผลลัพธ์หาได้ดังนี้

หากเป็นเอาต์พุตโนดใช้สมการที่ 2.15

$$\delta_j = O_j(1 - O_j)(T_j - O_j) \quad (2.15)$$

เมื่อ T_j คือผลลัพธ์ที่ต้องการได้ (target) จากเอาต์พุตโนดที่ j

O_j คือผลลัพธ์ที่คำนวณได้ (actual) จากเอาต์พุตโนดที่ j

2.1.5.1.5 หากเป็นฮิดเดนโนด ใช้สมการที่ 2.16

$$\delta_j = O_j(1 - O_j) \sum_k \delta_k W_{kj} \quad (2.16)$$

เมื่อ δ_k คือค่าความผิดพลาดของผลลัพธ์ของโนดที่ k (ซึ่งเชื่อมต่อกับโนดที่ j)

W_{kj} คือค่าน้ำหนักของโนดที่ j ซึ่งเชื่อมกับโนดที่ k

ทำการปรับค่าน้ำหนักในโนดอื่นจนครบ แล้วทำการสอนนิเวรอลเน็ตเวิร์กในตัวอย่างถัดไป (iteration ถัดไป) จนกระทั่งเน็ตเวิร์กนั้นลู่เข้า

2.1.5.2 ขั้นตอนการรู้จำ

ขั้นตอนการรู้จำ (recognition mode) จะมีการรับข้อมูล (สมาชิกของเวกเตอร์รูปแบบ) เข้าที่อินพุตโนด (ในชั้นอินพุต) จากนั้นแต่ละอินพุตโนดจะส่งข้อมูลไปยังทุกๆ ฮิดเดนโนด (ในชั้นซ่อน) แล้วแต่ละฮิดเดนโนด จะทำการคำนวณค่าที่ได้รับมา ตามสมการ 2.17 และ 2.18

$$O_j = F(\theta_j + \sum_{i=1}^n w_{ji} o_i) \quad (2.17)$$

$$F(a) = \frac{1}{1+e^{-a}} \quad (2.18)$$

เมื่อ O_j คือผลลัพธ์ของโนดที่ j

θ_j คือค่า bias ของโนดที่ j

w_{ji} คือค่าน้ำหนักของโนดที่ i ซึ่งส่งข้อมูล o_i มาให้โนดที่ j

o_i คือผลลัพธ์ของโนดที่ i

$F(a)$ คือ ฟังก์ชันกระตุ้น (activation function) ซึ่งใช้ฟังก์ชันซิกมอยด์ (sigmoid function)

จากนั้นฮิดเดนโนดจะส่งผลลัพธ์ของตัวเองต่อไปยังทุกๆ เอาต์พุตโนด (ในชั้นเอาต์พุต) ซึ่งแต่ละเอาต์พุตโนด ก็จะสามารถคำนวณค่าที่ได้รับมาตามสมการข้างบน แล้วตอบผลลัพธ์ของโนดตัวเอง ซึ่งเราสามารถพิจารณาผลลัพธ์ของทุกๆ เอาต์พุตโนด แล้วหาผลลัพธ์สุดท้ายได้ เช่น ถ้าผลลัพธ์ของ เอาต์พุตโนดที่ 1 มีค่าเท่ากับ 1 และผลลัพธ์ของเอาต์พุตโนดตัวที่เหลือเท่ากับ 0 ทั้งหมด แสดงว่า ข้อมูลที่ได้รับคือ ข้อมูลของตัวอักษร ก เป็นต้น

2.1.6 ต้นไม้การตัดสินใจ (decision tree)

เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ วิธีต้นไม้การตัดสินใจ ประกอบด้วยโนด สำหรับเงื่อนไขในการตัดสินใจ กิ่งแสดงถึงผลลัพธ์ที่ได้จากการพิจารณาเงื่อนไขที่โนด โดยแต่ละกิ่งจะนำไปสู่ผลลัพธ์สุดท้ายคือใบ (leaf node หรือ decision node) วิธีการต้นไม้การตัดสินใจจะจำกัดข้อมูลที่เป็นตัวแปรตาม หนึ่งตัวต่อหนึ่งต้น ถ้าต้องการทำนายตัวแปรตามหลายตัว ในการวิเคราะห์ต้องสร้างตัวแบบสำหรับตัวแปรตามแต่ละตัว ขั้นตอนวิธีของวิธีการต้นไม้การตัดสินใจส่วนใหญ่ไม่รองรับข้อมูลแบบต่อเนื่อง (Continuous data) จะต้องมีการแบ่งให้เป็นข้อมูลแบบไม่ต่อเนื่อง (discretized data) เสียก่อน จึงใช้ต้นไม้การตัดสินใจได้ เป็นวิธีการที่ค่อนข้างแพร่หลายเนื่องจากความไม่ซับซ้อนของขั้นตอนการทำงาน และ

สามารถตีความและเข้าใจลักษณะของรูปแบบข้อมูลได้ง่าย เพราะมีการแยกออกเป็นกฎ หรือข้อกำหนดได้

2.1.7 การจำแนกแบบกลุ่มด้วยวิธี Bagging (Bootstrap Aggregating)

Bagging (Bootstrap Aggregating) (Breiman,1996) เป็นวิธีสร้างตัวแยกแยะหลายๆตัวโดยที่แตกต่างกันและใช้ตัวแยกแยะเหล่านั้นเพื่อหาคำตอบ โดยเป็นสร้างชุดข้อมูลสอนชุดใหม่ จากชุดข้อมูลสอนเดิม โดยการสุ่มเอาข้อมูลจากชุดข้อมูลสอนเดิมมาสร้างชุดข้อมูลสอนใหม่หลายๆชุด โดยให้จำนวนของข้อมูลในแต่ละชุดเท่าเดิม (อาจมีข้อมูลบางตัวซ้ำหลายครั้งได้) ซึ่งชุดข้อมูลสอนใหม่แต่ละชุดนี้จะถูกนำไปใช้สร้างตัวแยกแยะตัวใหม่ที่ละตัว วิธี Bagging จะเหมาะสำหรับนิรवलเน็ตเวิร์กและตัวแยกแยะอื่นๆที่การเปลี่ยนแปลงเพียงเล็กน้อยของชุดข้อมูลสอน จะทำให้เกิดการเปลี่ยนแปลงอย่างมากของผลลัพธ์ที่ได้

1. Initialize the parameter
 - $D = \emptyset$, the ensemble.
 - L , the number of classifiers to train.
2. For $k = 1, \dots, L$
 - Take a bootstrap sample S_k from Z .
 - Build a classifier D_k using S_k as the training set.
 - Add the classifier to the current ensemble, $D = D \cup D_k$.
3. Return D .
4. Run D_1, \dots, D_L on the input x .
5. The class with the maximum number of votes is chosen as the label for x .

ภาพที่ 2.3 : อัลกอริทึม Bagging

แหล่งที่มา : Ludmila I. Kuncheva, 2004

2.1.8 การจำแนกแบบกลุ่มด้วยวิธี AdaBoost (Adaptive Boosting)

AdaBoost (Adaptive Boosting) (Freund และ Schapire,1995) จะเป็นอีกวิธีการหนึ่งซึ่งมีความคล้ายกับ Bagging แต่ AdaBoost จะสร้างตัวแยกแยะที่ต่างกันโดยการจัดการกับชุดข้อมูลสอน (ไม่ใช้การสุ่มข้อมูลใหม่) โดย AdaBoost จะทำการกำหนดน้ำหนัก (weight) ให้กับข้อมูลสอนแต่ละตัวให้มีน้ำหนักแตกต่างกันไป ในแต่ละรอบที่มีการสร้างตัว

แยกแยะใหม่ค่าน้ำหนักจะถูกเปลี่ยนไปตามความผิดพลาดของผลลัพธ์ที่ตัวแยกแยะในรอบนั้นๆ กระทำต่อข้อมูลถ้าตัวแยกแยะตอบผลลัพธ์ถูกสำหรับข้อมูลสอนตัวใด ข้อมูลสอนตัวนั้นก็ถูกลดค่าน้ำหนัก และในทางตรงข้ามถ้าตัวแยกแยะตอบผลลัพธ์ผิดสำหรับข้อมูลสอนตัวใด ข้อมูลสอนนั้นก็ถูกเพิ่มค่าน้ำหนัก โดยค่าน้ำหนักที่ได้จะใช้สำหรับกำหนดความน่าจะเป็นที่ข้อมูลตัวนั้นจะถูกเลือกให้อยู่ในชุดข้อมูลสอนชุดต่อไป ซึ่งในตอนแรกข้อมูลทุกตัวจะถูกกำหนดให้มีค่าความน่าจะเป็นที่จะถูกเลือกเท่ากัน สำหรับจุดมุ่งหมายของ AdaBoost ก็คือจะให้ความสนใจกับข้อมูลที่ทำนายผิดพลาดมากขึ้นสำหรับการสร้างตัวแยกแยะในชุดต่อไป

1) Initialize the parameters

- Set the weights $w^1 = [w_1, \dots, w_N], w_j^1 \in [0, 1], \sum_{j=1}^N w_j^1 = 1$
- Initially the ensemble $D = \emptyset$.
- Pick L , the number of classifier to train.

2) For $k = 1, \dots, L$

- Take a sample S_k from Z using distribution w^k .
- Build a classifier D_k using S_k as the training set.
- Calculate the weighted ensemble error at step k by

$$\epsilon_k = \sum_{i=1}^N w_j^k l_k^j$$

($l_k^j = 1$ if D_k misclassifies z_j and $l_k^j = 0$ otherwise.)

- If $\epsilon_k = 0$ or $\epsilon_k \geq 0.5$, ignore D_k , reinitialize the weights w_j^k to $\frac{1}{N}$ and continue.
- Else, calculate

$$\beta_k = \frac{\epsilon_k}{1 - \epsilon_k}, \text{ where } \epsilon_k \in (0, 0.5),$$

- Update the individual weights

$$w_j^{k+1} = \frac{w_j^k \beta_k^{(1-l_k^j)}}{\sum_{i=1}^N w_i^k \beta_k^{(1-l_k^i)}}, \quad j = 1, \dots, N.$$

ภาพที่ 2.4 : อัลกอริทึม AdaBoostM1

แหล่งที่มา : Ludmila I. Kuncheva, 2004

3) Return D and β_1, \dots, β_L .

4) Calculate the support for class ω_t by

$$\mu_t(x) = \sum_{D_k(x)=\omega_t} \ln\left(\frac{1}{\beta_k}\right)$$

The class with the maximum support is chosen label for x .

ภาพที่ 2.5 : อัลกอริทึม AdaBoostM1 (ต่อ)

แหล่งที่มา : Ludmila I. Kuncheva, 2004

2.1.9 การจำแนกแบบกลุ่มด้วยวิธี Stacking

Stacking (Wolpert D., 1992) วิธีนี้เป็นวิธีที่ใช้โมเดลระดับสูง (high-level model) ในการรวมผลลัพธ์ของโมเดลระดับต่ำ (low-level model) เพื่อให้ได้ค่าความถูกต้องที่สูงขึ้น สมมุติว่ามีกลุ่มตัวอย่างที่ใช้สอน $\{(x_1, y_1), \dots, (x_m, y_m)\}$ และมี L ขั้นตอนวิธี $A_1, A_2, A_3, \dots, A_L$ จากนั้นจะใช้ วิธีเหล่านี้เพื่อสร้างสมมติฐาน h_1, \dots, h_L จุดมุ่งหมายของวิธีนี้ก็คือการหาตัวแยกแยะ ที่สามารถรวมผลลัพธ์จาก h_1 ถึง h_L ได้ ค่าตอบของวิธีนี้หาได้จาก $h^*(h_1(x), \dots, h_L(x))$

2.1.10 การจำแนกแบบกลุ่มด้วยวิธีคะแนนเสียงข้างมาก (majority vote)

เป็นหนึ่งในยุทธศาสตร์ที่ถูกใช้มาตั้งแต่สมัยโบราณสำหรับการเลือกตั้ง (W. H. E. Day, 1988) แบ่งได้เป็น 3 รูปแบบดังนี้

2.1.10.1 ความเห็นพ้องเป็นเอกฉันท์ (unanimity)

2.1.10.2 ความเห็นพ้องของเสียงส่วนใหญ่ที่เกิดขึ้นครั้งหนึ่ง (simple majority)

2.1.10.3 ความเห็นพ้องของเสียงส่วนใหญ่แต่ไม่เกิดขึ้นครั้งหนึ่ง (plurality)



ภาพที่ 2.6 : รูปแบบความเห็นพ้องของคน 10 คน

แหล่งที่มา : Ludmila I. Kuncheva, 2004

2.2 เกณฑ์การประเมินประสิทธิภาพผลพยากรณ์ (Performance Statistic)

2.2.1 เส้นโค้ง ROC (Receiver Operating Characteristic curve)

เส้นโค้ง ROC เป็นกราฟที่พล็อตระหว่างค่า sensitivity (SN) หรือสัดส่วนของการพยากรณ์ได้ถูกต้องของการเกิดเหตุการณ์ที่สนใจ และค่า $1 - \text{specificity}$ (SP) หรือสัดส่วนของการพยากรณ์ผิดของการเกิดเหตุการณ์ที่สนใจ (Thomas A. Lasko และคณะ, 2005)

2.2.2 พื้นที่ใต้โค้ง ROC (Area under the ROC Curve (AUC))

พื้นที่ใต้โค้ง ROC ใช้เป็นดัชนีในการบ่งชี้ถึงความถูกต้องของการพยากรณ์หรือความเชื่อถือได้ของตัวแบบ จะแสดงให้เห็นถึงความสามารถในการจำแนกกลุ่มเหตุการณ์ที่สนใจ ออกจากกลุ่มเหตุการณ์ที่ไม่สนใจ ซึ่งในการคำนวณค่าประมาณพื้นที่ใต้โค้ง ROC (AUC) สามารถคำนวณได้หลายวิธี ในงานวิจัยนี้ขอกล่าวอ้างถึงระเบียบวิธีที่ไม่ใช้พารามิเตอร์ (Nonparametric Method) โดยในการสร้างเส้นโค้ง ROC จะเชื่อมต่อด้านละจุดของข้อมูลซึ่งมีค่า $1 - \text{specificity}$ อยู่ในแนวแกน x และ sensitivity อยู่ในแนวแกน y ($1 - SP$, SN) เป็นเส้นตรง จากนั้นทำการประมาณค่าพื้นที่ใต้โค้ง ROC ซึ่งอาศัยหลักเกณฑ์เชิงสี่เหลี่ยมคางหมู (Trapezoidal Rule) ซึ่งวิธีการดังกล่าวเป็นการอ้างระเบียบขั้นตอนทางสถิติที่ไม่ใช้พารามิเตอร์ โดยค่าประมาณของพื้นที่ใต้โค้ง ROC หลังจากวิธีการนี้จะมีค่าเท่ากับผลการทดสอบข้อมูลชุดเดิมด้วยการทดสอบของแมน-วิทนี (Mann-Whitney Test) เป็นการทดสอบโดยการคำนวณค่าจากจำนวนคู่ที่มีความเป็นไปได้ของเหตุการณ์ที่สนใจกับเหตุการณ์ที่ไม่สนใจ ซึ่งเป็นแนวความคิดเดียวกับการทดสอบผลรวมอันดับของวิลคอกซัน (Wilcoxon Rank-Sum Test) และการทดสอบ C-index (Thomas A. Lasko และคณะ, 2005)

โดยกำหนดให้

d_1, d_2, \dots, d_{n_D} เป็นค่าพยากรณ์สำหรับกลุ่มเหตุการณ์ที่สนใจ

h_1, h_2, \dots, h_{n_H} เป็นค่าพยากรณ์สำหรับกลุ่มเหตุการณ์ที่ไม่สนใจ

n_D เป็นจำนวนค่าพยากรณ์ตัวอย่างจากกลุ่มเหตุการณ์ที่สนใจ

n_H เป็นจำนวนค่าพยากรณ์ตัวอย่างจากกลุ่มเหตุการณ์ที่ไม่สนใจ

และ $C(d_i, h_j)$ เป็นฟังก์ชันที่ใช้ในการเปรียบเทียบ เมื่อ

$$C(d_i, h_j) = \begin{cases} 1, & \text{ถ้า } d_i > h_j \\ 0.5, & \text{ถ้า } d_i = h_j \\ 0, & \text{ถ้า } d_i < h_j \end{cases} \quad (2.19)$$

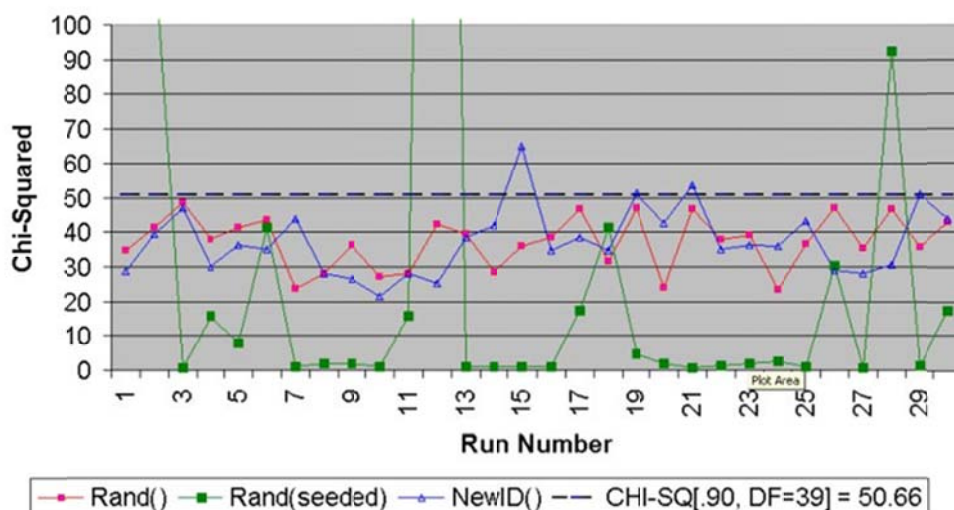
ค่าประมาณของพื้นที่ใต้โค้ง ROC ($A\hat{U}C$) คือค่าเฉลี่ยของผลรวมที่เกิดจากฟังก์ชันที่ใช้ในการเปรียบเทียบ ($C(d_i, h_j)$) ของคู่เหตุการณ์ที่สนใจกับเหตุการณ์ที่ไม่สนใจทุกคู่ ดังนี้

$$(A\hat{U}C) = \frac{1}{n_D n_H} \sum_{i=1}^{n_D} \sum_{j=1}^{n_H} C(d_i, h_j) \quad (2.20)$$

จากระเบียบวิธีที่ไม่ใช้พารามิเตอร์จะเห็นได้ว่าวิธีการดังกล่าวไม่มีข้อสมมติเกี่ยวกับข้อมูล จึงส่งผลให้ระเบียบวิธีการนี้เป็นที่นิยมและถูกนำไปประยุกต์ใช้ในงานด้านต่าง ๆ อย่างกว้างขวาง

2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง

การสุ่มด้วยฟังก์ชัน RAND (Brian C. , 2004) ทำการสุ่มข้อมูลมาทั้งหมด 30 ครั้ง จากจำนวนข้อมูลดิบ 100,000 ชุด มา 1 % โดยใช้การแบ่งช่วงข้อมูลออกเป็น 40 ช่วง (องศาความเป็นอิสระเท่ากับ 39) โดยแต่ละช่วงมีระยะห่าง 2500 ซึ่งในแต่ละช่วงควรมีความถี่ของข้อมูลที่ถูสุ่มเท่ากับ 25 โดยที่ค่าความเชื่อมั่นที่ 90 % จะสามารถหาค่า χ^2 จากตาราง Chi-square function table ได้ 50.660 ซึ่งถ้าคำนวณค่า χ^2 ออกมาได้น้อยกว่า 50.660 จะถือว่าข้อมูลชุดดังกล่าวข้อมูลแบบสุ่มที่ดี ซึ่งจากเทคนิคต่าง ๆ ได้ผลดังนี้



ภาพที่ 2.7 : ค่าไคสแควร์ที่ได้จากการทดลองของ Brian Connolly

แหล่งที่มา : Brian Connolly, 2004

Kolen และ Pollack (1991) เสนอวิธีพื้นฐานสำหรับการสร้างตัวแยกแยะ สำหรับนิเวศเน็ตเวิร์กแบบแบคพรอพาเกชัน (Backpropagation Neural Network) โดยแสดงให้เห็นว่าเพียงแค่การกำหนดค่าน้ำหนักเริ่มต้นเปลี่ยนไปก็สามารถสร้างตัวแยกแยะที่ต่างไปจากเดิมได้

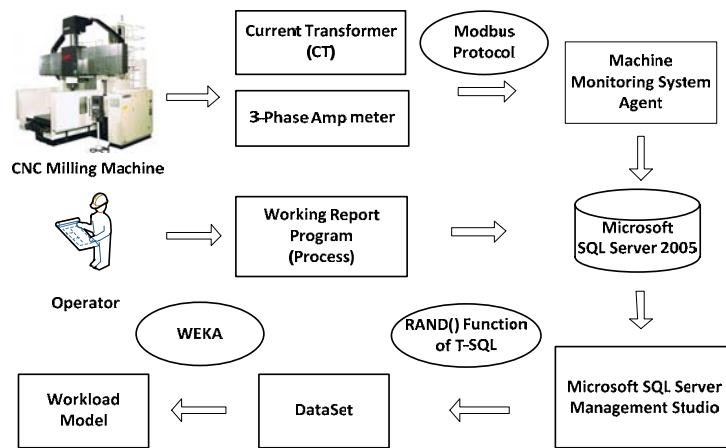
fBoost (Harries M. , 1999) วิธีนี้คล้ายกับ Boosting แต่ fBoost ไม่ได้ใช้ค่าความผิดพลาดในการปรับน้ำหนักของแต่ละตัวอย่าง fBoost ใช้ข้อมูลที่ได้จากสมมติฐาน (hypothesis) วิธีนี้แสดงให้เห็นว่าการปรับค่าน้ำหนักไม่จำเป็นที่จะต้องมีการตอบผิด

Unweighted voting (Clemen, R. T. ,1989) วิธีนี้เป็นวิธีที่ง่ายที่สุดในการรวมผลลัพธ์จากตัวแยกแยะแต่ละตัว อย่างไรก็ตามการไม่ใช้น้ำหนักถ่วงน้ำหนักนั้นมิใช่ข้อดีหลายข้อ วิธีนี้ถูกใช้ในหลายขั้นตอนวิธี (algorithm) เช่น Bagging และ ECOC การรวมแบบไม่ถ่วงน้ำหนักนั้นเหมาะสำหรับตัวแยกแยะที่ให้คำตอบเป็นค่าความน่าจะเป็นที่จะเป็นประเภทนั้นๆ (probability estimate) มากกว่าตัวแยกแยะที่ให้คำตอบเป็นคลาส (simple classification decision) คำตอบของวิธีนี้ได้มาจากคลาส (class) ที่มีความน่าจะเป็นสูงสุด

บทที่ 3

วิธีดำเนินการวิจัย

ในงานวิจัยนี้จะทำการศึกษาเพื่อที่จะคัดเลือกหาตัวแบบที่ดีที่สุดเพื่อนำไปใช้ใน
ระบบการผลิตแม่พิมพ์โดยใช้ค่าพื้นที่ใต้โค้ง ROC เป็นตัวเปรียบเทียบ ทั้งตัวแบบแบบเดี่ยวและ
แบบกลุ่ม



ภาพที่ 3.1 : ขั้นตอนในการทำงานวิจัย

3.1 ขั้นตอนในการทำงานวิจัย

3.1.1 ทำการสร้างระบบตรวจติดตามเครื่องจักรแบบอัตโนมัติ (Machine Monitoring System) ซึ่งระบบประกอบไปด้วย

3.1.1.1 ใช้ Current Transformer เป็นตัวคล้องจับกระแสไฟฟ้า

3.1.1.2 ใช้เครื่องวัดกระแสไฟฟ้าแบบ 3 เฟส เป็นตัวอ่านค่ากระแสไฟฟ้าแบบมีพอร์ตอนุกรม

3.1.1.3 ใช้โปรโตคอลมอดบัสสำหรับรับส่งข้อมูลจากเครื่องวัดมายังคอมพิวเตอร์ผ่านทางพอร์ตอนุกรม

3.1.1.4 ใช้โปรแกรมสำหรับบันทึกข้อมูลค่ากระแสไฟฟ้าลงฐานข้อมูล พัฒนาโดยใช้ Microsoft Visual Studio 2010 โดยใช้ภาษาเบสิก

3.1.1.5 ใช้โปรแกรมสำหรับบันทึกข้อมูลการทำงานลงฐานข้อมูล พัฒนาโดยใช้ Microsoft Visual Studio 2010 โดยใช้ภาษาเบสิก

- 3.1.2 ทำการสุ่มข้อมูลจากฐานข้อมูลโดยใช้ฟังก์ชันRAND (Brian Connolly, 2004) ของ Microsoft SQL Server แล้วคำนวณหาค่าโคสแควร์ด้วยกับโปรแกรม Microsoft Excel 2010 เพื่อตรวจสอบว่าเป็นการสุ่มที่ดีหรือไม่
- 3.1.3 นำค่าที่ได้จากการสุ่มมาสร้างชุดข้อมูล โดยในขั้นตอนการเตรียมข้อมูล (preprocess) ชุดข้อมูลจะถูกแบ่งเป็น 2 แบบด้วยกันคือ
 - 3.1.3.1 ผ่านกระบวนการทำให้เป็นแบบไม่ต่อเนื่อง
 - 3.1.3.2 ไม่ผ่านกระบวนการทำให้เป็นแบบไม่ต่อเนื่อง
- 3.1.4 นำชุดข้อมูลไปทำการสร้างตัวแบบด้วยกับโปรแกรม WEKA โดยแบ่งตามชื่อคลาส และคุณสมบัติที่เลือกใช้ดังนี้
 - 3.1.4.1 Naive Bayes
 - 3.1.4.2 Bayes Net K2 local search
 - 3.1.4.3 Bayes Net HillClimber local search
 - 3.1.4.4 Bayes Net GeneticSearch local search
 - 3.1.4.5 Back propagation Multilayer Perceptron Neural Network
 - 3.1.4.5.1 Learning Rate = 0.1 Training Time = 200
 - 3.1.4.5.2 Learning Rate = 0.2 Training Time = 200
 - 3.1.4.5.3 Learning Rate = 0.3 Training Time = 200
 - 3.1.4.5.4 Learning Rate = 0.1 Training Time = 500
 - 3.1.4.5.5 Learning Rate = 0.2 Training Time = 500
 - 3.1.4.5.6 Learning Rate = 0.3 Training Time = 500
 - 3.1.4.5.7 Learning Rate = 0.1 Training Time = 1000
 - 3.1.4.5.8 Learning Rate = 0.2 Training Time = 1000
 - 3.1.4.5.9 Learning Rate = 0.3 Training Time = 1000
 - 3.1.4.6 KStar
 - 3.1.4.7 AdaBoostM1 โดยเลือกใช้ Classifier ดังต่อไปนี้
 - 3.1.4.7.1 Naive Bayes
 - 3.1.4.7.2 Bayes Net GeneticSearch

- 3.1.4.7.3 Back propagation Multilayer Perceptron Neural Network
Learning Rate = 0.1 Training Time = 200
- 3.1.4.7.4 KStar
- 3.1.4.7.5 DecisionTable
- 3.1.4.7.6 J48
- 3.1.4.8 Bagging โดยเลือกใช้ Classifier ดังต่อไปนี้
 - 3.1.4.8.1 Naive Bayes
 - 3.1.4.8.2 Bayes Net GeneticSearch
 - 3.1.4.8.3 Back propagation Multilayer Perceptron Neural Network
Learning Rate = 0.1 Training Time = 200
 - 3.1.4.8.4 KStar
 - 3.1.4.8.5 DecisionTable
 - 3.1.4.8.6 J48
- 3.1.4.9 Stacking โดยใช้ metaClassifier = J48 และใช้ Classifier 6 ตัวซึ่งประกอบไปด้วย
 - 3.1.4.9.1 Naive Bayes
 - 3.1.4.9.2 Bayes Net GeneticSearch
 - 3.1.4.9.3 Back propagation Multilayer Perceptron Neural Network
Learning Rate = 0.1 Training Time = 200
 - 3.1.4.9.4 KStar
 - 3.1.4.9.5 DecisionTable
 - 3.1.4.9.6 J48
- 3.1.4.10 Vote โดยใช้ combinationRule = Majority Voting และใช้ Classifier 6 ตัวซึ่งประกอบไปด้วย
 - 3.1.4.10.1 Naive Bayes
 - 3.1.4.10.2 Bayes Net GeneticSearch

3.1.4.10.3 Back propagation Multilayer Perceptron Neural Network

Learning Rate = 0.1 Training Time = 200

3.1.4.10.4 KStar

3.1.4.10.5 DecisionTable

3.1.4.10.6 J48

3.1.4.11 Rules

3.1.4.11.1 DecisionTable BestFirst Search

3.1.4.11.2 DecisionTable GreedyStepwise Search

3.1.4.12 Trees

3.1.4.12.1 J48

3.1.5 เปรียบเทียบประสิทธิภาพของตัวแบบในแต่ละวิธีด้วยค่า AUC

3.2 รายละเอียดข้อมูลดิบ การสร้างและทดสอบชุดข้อมูล

3.2.1 ข้อมูลดิบ

โดยข้อมูลดิบที่ทำการศึกษามีดังนี้

ตารางที่ 3.1 : ข้อมูลดิบทั้งหมดแบ่งตามกระบวนการผลิต

กระบวนการผลิต	จำนวนข้อมูลดิบ
Masking	7,000
Rough Cut	258,000
Drill	518,000
Semi Finish	633,000

3.2.2 การสร้างชุดข้อมูล

ทำการสุ่มข้อมูลในแต่ละโพรเซสจำนวน 400, 1200, 2400 และ 4000 โดยการสุ่มใช้วิธีการสร้าง แถวข้อมูลเพิ่มโดยฟังก์ชัน RAND() ของ Microsoft SQL Server ตามวิธีการของ Brian Connolly (2004)

นำข้อมูลที่สุ่มได้ มาสร้างชุดข้อมูลโดยรวมแต่ละโพรเซสที่จำนวนการสุ่มเท่ากัน เข้าด้วยกัน ซึ่งจะได้ชุดข้อมูลทั้งหมด 4 ชุด ที่มีจำนวนชุดข้อมูล เท่ากับ 1600, 4800, 9600 และ 16000

เฉพาะตัวแบบแบบเดียวให้ทดสอบกับชุดข้อมูลเพิ่มเติมอีก 3 ชุดดังนี้
ตารางที่ 3.2 : จำนวนชุดข้อมูลที่สุ่มเลือกเพิ่มเติมเพื่อทดสอบกับตัวแบบแบบเดียว

กระบวนการผลิต		จำนวนข้อมูลดิบ	จำนวนข้อมูล		
			ชุดที่1	ชุดที่2	ชุดที่3
Masking		7,000	7,000	7,000	7,000
Rough Cut		258,000	10,000	20,000	50,000
Drill		518,000	10,000	20,000	50,000
Semi Finish		633,000	10,000	20,000	50,000
รวม			37,000	67,000	157,000

3.2.3 การทดสอบชุดข้อมูล

โดยใช้การแบ่งช่วงข้อมูลดิบออกเป็น 40 ช่วง ทำให้ได้ค่าองศาความเป็นอิสระ = 39 โดยในแต่ละช่วงมีระยะห่างเท่ากับ จำนวนข้อมูลดิบในแต่ละกระบวนการผลิตหารด้วย 40 โดยที่ค่าความเชื่อมั่นที่ 90 % จะสามารถหาค่าไคสแควร์จากตาราง Chi-square function table ได้เท่ากับ 50.660

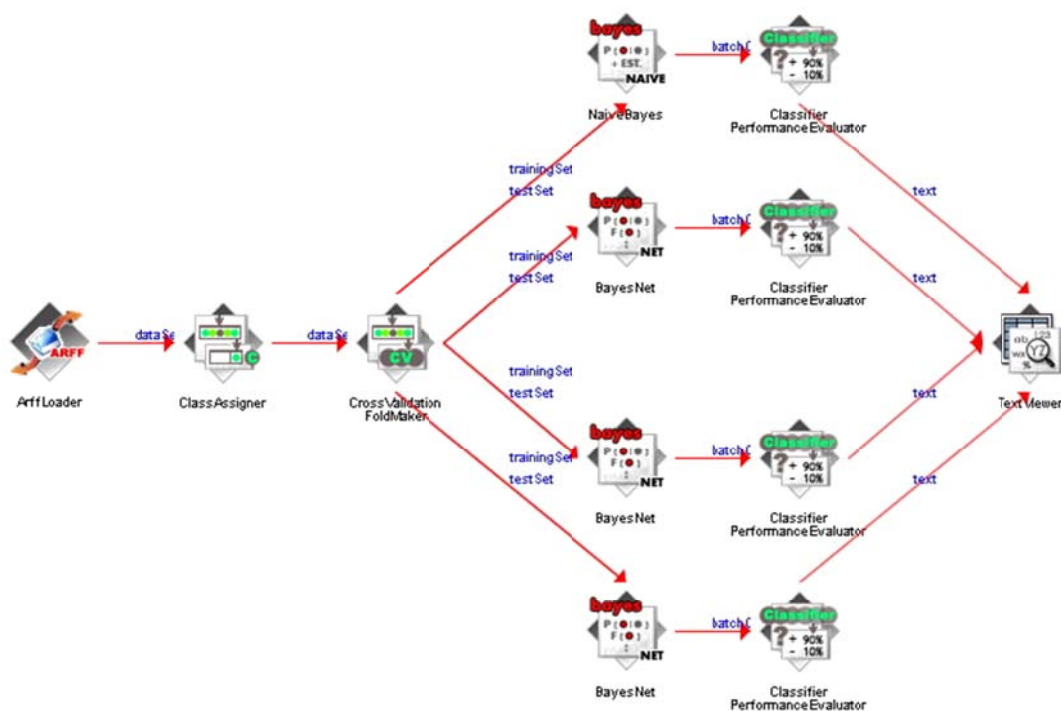
3.3 การสร้างตัวแบบด้วยโปรแกรม WEKA โดยใช้ KnowledgeFlow



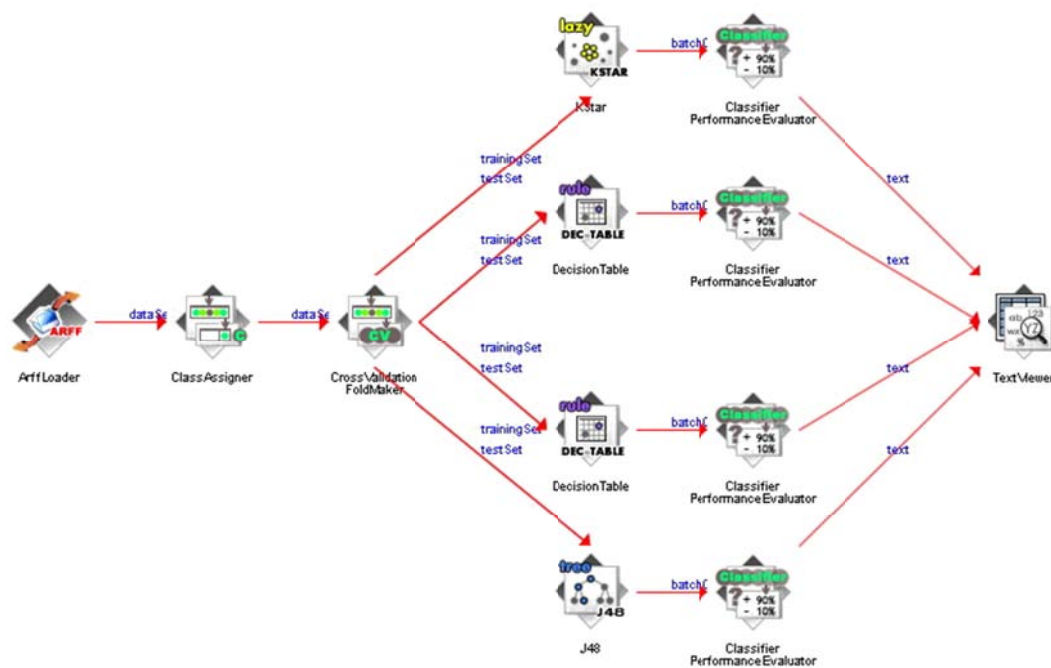
ภาพที่ 3.2 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Stacking



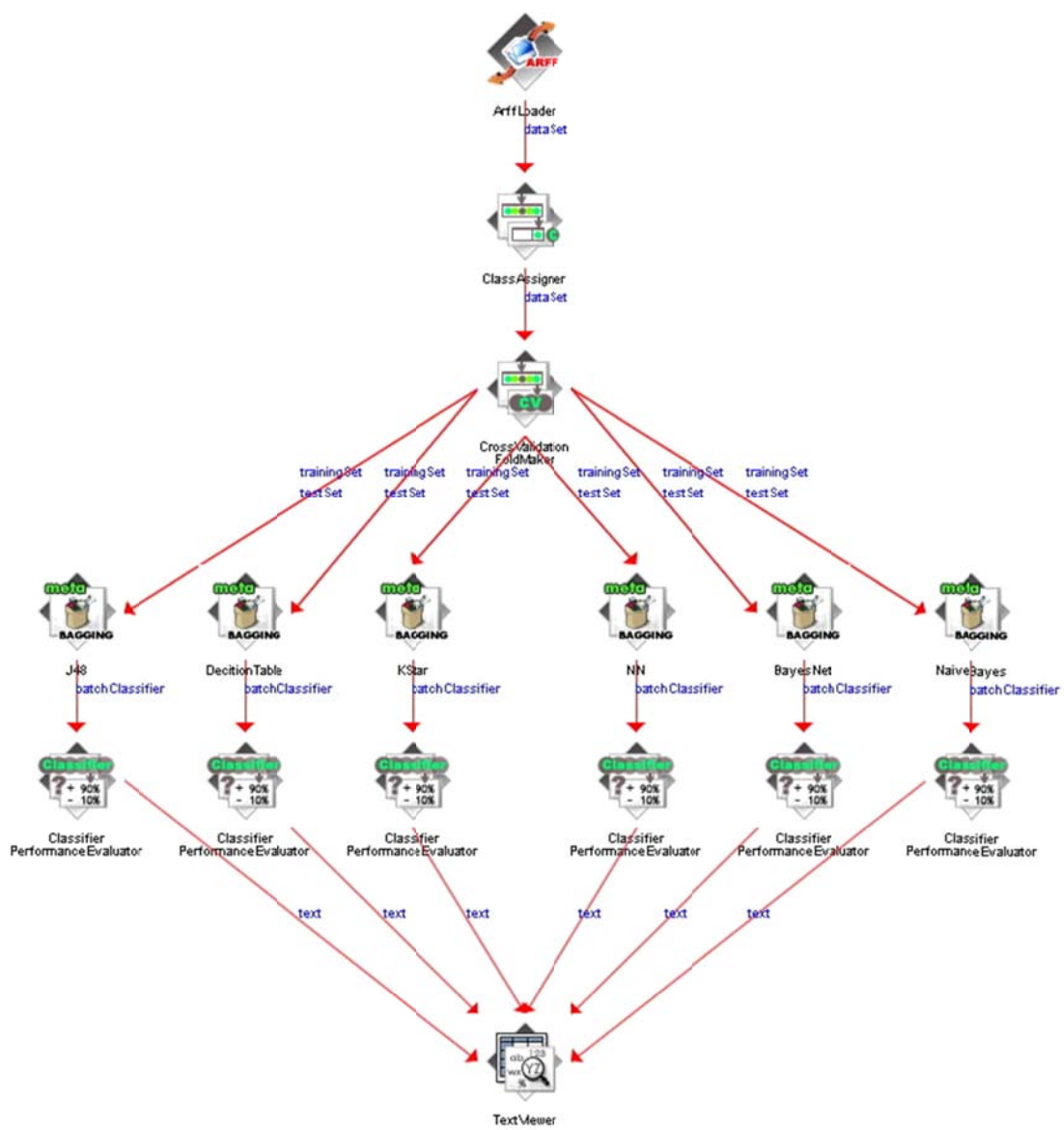
ภาพที่ 3.3 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Majority Vote



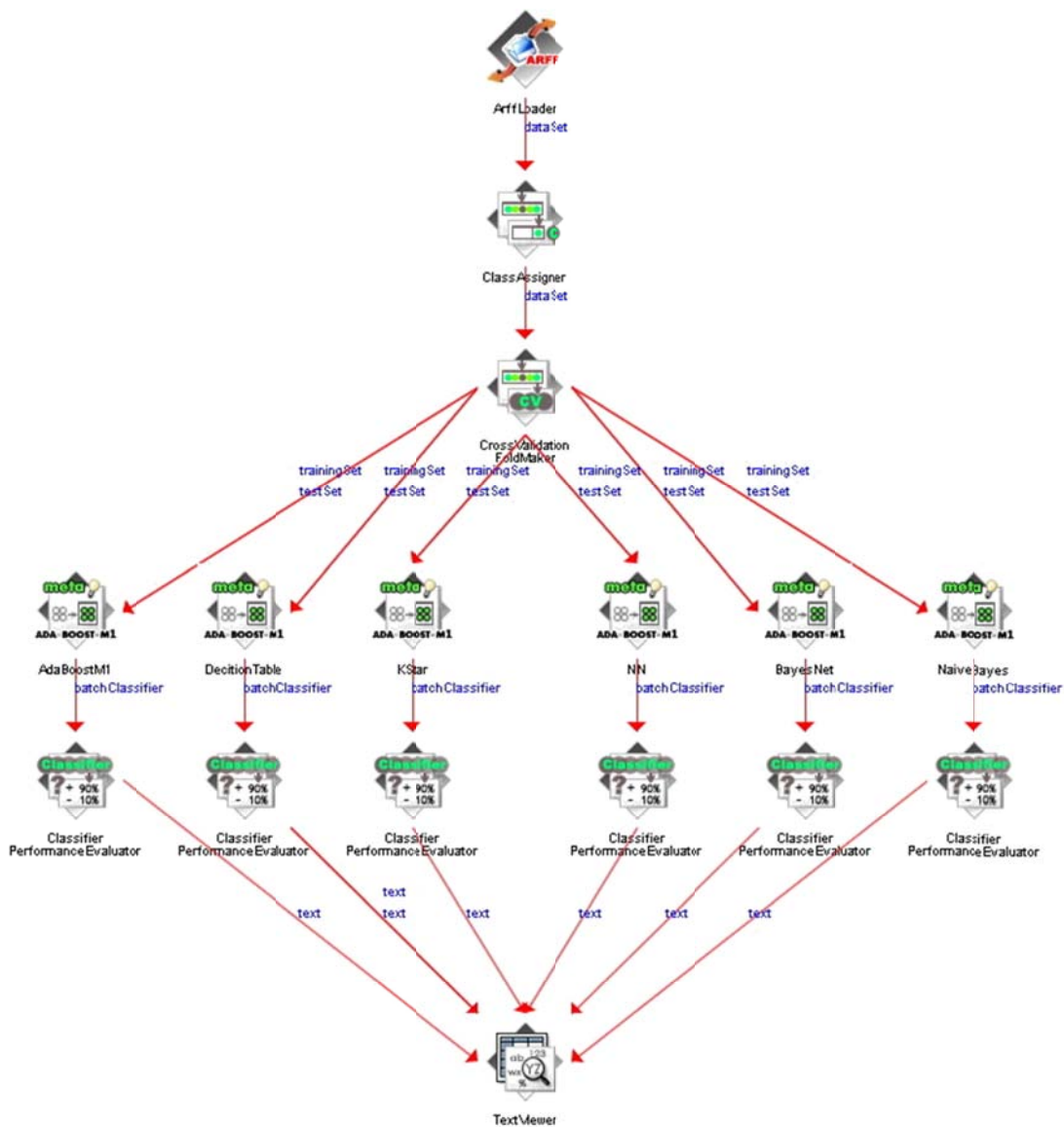
ภาพที่ 3.4 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับ วิธีเบย์อย่างง่ายและเบย์เซียนเน็ตเวิร์ค



ภาพที่ 3.5 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธีเคสสตาร์ ตารางตัดสินใจ และต้นไม้การตัดสินใจ



ภาพที่ 3.6 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี Bagging



ภาพที่ 3.7 : การสร้างตัวแบบด้วย WEKA knowledgeFlow สำหรับวิธี AdaBoostM1

บทที่ 4

ผลการวิเคราะห์ข้อมูล

งานวิจัยนี้มีจุดประสงค์เพื่อ หาตัวแบบที่ดีที่สุดจากวิธีการแบบเดี่ยวและแบบกลุ่ม ดังนั้นในบทนี้จะกล่าวถึงการศึกษาให้ครอบคลุมประเด็นต่างๆที่ได้กล่าวมาข้างต้น

4.1 ผลการสุ่มข้อมูลจากฐานข้อมูล

ตารางที่ 4.1 : ค่าไคสแควร์จากการสุ่มของแต่ละกระบวนการผลิต

ชื่อกระบวนการผลิต	จำนวนข้อมูลสุ่ม			
	400	1200	2400	4000
Masking	24.4	32.06	33.23	17.66
Rough Cut	50.4	32.46	34.83	47.12
Drill	31.2	39.6	49.6	39.33
Semi Finish	33	40.46	34.58	35.42

จากตารางจะเห็นว่าค่าไคสแควร์ที่ได้มีค่าไม่เกิน 50.660 (ค่าจากตาราง Chi-square function table) จึงสรุปได้ว่า ชุดข้อมูลที่สุ่มมาในทุกกระบวนการผลิตเป็นการสุ่มที่ดี

4.2 ผลความแม่นยำของตัวแบบแต่ละวิธี

ตารางที่ 4.2 : ค่า AUC จากตัวแบบแบบเดี่ยวโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีแบบเดี่ยว	ค่า AUC			
	1600	4800	9600	16000
NaiveBayes	0.716	0.716	0.713	0.715
BayesNet K2	0.876	0.891	0.897	0.903
BayesNet HillClimber	0.876	0.891	0.897	0.903
<i>BayesNet GeneticSearch</i>	0.888	0.912	0.927	0.935
KStar	0.890	0.895	0.896	0.897
DecisionTable BestFirst	0.874	0.900	0.913	0.921
DecisionTable GreedyStepwise	0.873	0.900	0.913	0.920
J48	0.869	0.903	0.912	0.921

จากตารางที่ 4.2 จะเห็นได้ว่าวิธีการสร้างตัวแบบแบบเดียวกับ BayesNet Genetic Search นั้นให้ค่า AUC ดีที่สุดเท่ากับ 0.935 และดีที่สุดในเกือบทุกชุดข้อมูลทดสอบ

ตารางที่ 4.3 : ค่า AUC จากตัวแบบแบบเดียวโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีแบบเดียว	ค่า AUC						
	1600	4800	9600	16000	37000	67000	157000
NaiveBayes	0.883	0.895	0.901	0.905	0.892	0.877	0.865
BayesNet K2	0.884	0.896	0.901	0.906	0.892	0.877	0.865
BayesNet HillClimber	0.884	0.896	0.901	0.906	0.892	0.877	0.865
BayesNet GeneticSearch	0.896	0.918	0.932	0.938	0.932	0.925	0.923
KStar	0.897	0.919	0.931	0.936	0.928	0.920	0.916
DecisionTable BestFirst	0.881	0.901	0.919	0.922	0.915	0.905	0.907
DecisionTable Greedy	0.881	0.902	0.923	0.921	0.915	0.905	0.919
J48	0.873	0.895	0.916	0.922	0.920	0.916	0.914

จากตารางที่ 4.3 จะเห็นได้ว่าวิธีการสร้างตัวแบบแบบเดียวกับ BayesNet Genetic Search ที่จำนวนชุดข้อมูลเท่ากับ 16,000 ชุด นั้นให้ค่า AUC เท่ากับ 0.938 ซึ่งให้ค่าดีที่สุดในเกือบทุกชุดข้อมูล และหากผ่านกระบวนการเตรียมข้อมูลโดยการทำให้เป็นข้อมูลแบบไม่ต่อเนื่องแล้วจะเห็นได้ว่าเกือบทุกการทดลองให้ค่า AUC สูงขึ้นเกือบทั้งหมด

ตารางที่ 4.4 : ค่า AUC จากวิธีโครงข่ายประสาทเทียมโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีโครงข่ายประสาทเทียม L (อัตราการเรียนรู้) T (จำนวนรอบการเรียนรู้)	ค่า AUC			
	1600	4800	9600	16000
L = 0.1 T= 200	0.755	0.757	0.758	0.760
L = 0.2 T= 200	0.758	0.761	0.757	0.777
L = 0.3 T= 200	0.767	0.760	0.767	0.776
L = 0.1 T= 500	0.763	0.760	0.768	0.761
L = 0.2 T= 500	0.768	0.774	0.771	0.782
L = 0.3 T= 500	0.778	0.772	0.776	0.782
L = 0.1 T= 1000	0.768	0.760	0.774	0.762
L = 0.2 T= 1000	0.774	0.780	0.775	0.785
L = 0.3 T= 1000	0.783	0.777	0.779	0.784

จากตารางที่ 4.4 จะเห็นได้ว่าวิธีการสร้างตัวแบบแบบเดียวด้วยวิธีโครงข่ายประสาทเทียมที่ไม่ผ่านการเตรียมข้อมูลโดยการทำให้เป็นค่าไม่ต่อเนื่องที่จำนวนรอบการเรียนรู้เท่ากับ 1000 และอัตราการเรียนรู้เท่ากับ 0.2 ให้ค่า AUC ดีที่สุดเท่ากับ 0.785

ตารางที่ 4.5 : ค่า AUC จากวิธีโครงข่ายประสาทเทียมโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีโครงข่ายประสาทเทียม L (อัตราการเรียนรู้) T (จำนวนรอบการเรียนรู้)	ค่า AUC			
	1600	4800	9600	16000
L = 0.1 T= 200	0.894	0.914	0.920	0.928
L = 0.2 T= 200	0.895	0.911	0.919	0.914
L = 0.3 T= 200	0.892	0.911	0.893	0.894
L = 0.1 T= 500	0.893	0.915	0.920	0.926
L = 0.2 T= 500	0.891	0.911	0.918	0.922
L = 0.3 T= 500	0.890	0.910	0.905	0.889
L = 0.1 T= 1000	0.892	0.916	0.921	0.926
L = 0.2 T= 1000	0.889	0.912	0.919	0.918
L = 0.3 T= 1000	0.889	0.911	0.905	0.900

จากตารางที่ 4.5 จะเห็นได้ว่าวิธีการสร้างตัวแบบแบบเดียวด้วยวิธีโครงข่ายประสาทเทียมที่ผ่านการเตรียมข้อมูลโดยการทำให้เป็นค่าไม่ต่อเนื่องที่จำนวนรอบการเรียนรู้เท่ากับ 200 และอัตราการเรียนรู้เท่ากับ 0.1 ที่จำนวนชุดข้อมูลเท่ากับ 16,000 ชุด ให้ค่า AUC ดีที่สุดเท่ากับ 0.928 และในทุกการทดลองให้ค่า AUC ที่สูงขึ้นเมื่อเทียบกับวิธีการที่ไม่ผ่านการทำให้เป็นค่าแบบไม่ต่อเนื่อง

ตารางที่ 4.6 : ค่า AUC จากวิธีแบบกลุ่มโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีแบบกลุ่ม	ค่า AUC			
	1600	4800	9600	16000
AdaBoostM1(NaiveBayes)	0.716	0.716	0.713	0.715
AdaBoostM1(BayesNet GeneticSearch)	0.852	0.884	0.903	0.916
AdaBoostM1(KStar)	0.908	0.915	0.910	0.908

ตารางที่ 4.7 : ค่า AUC จากวิธีแบบกลุ่มโดยไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง (ต่อ)

วิธีแบบกลุ่ม	ค่า AUC			
	1600	4800	9600	16000
AdaBoostM1(DecisionTable BestFirst)	0.809	0.896	0.916	0.923
AdaBoostM1(Neural Network)	0.749	0.747	0.751	0.734
AdaBoostM1(J48)	0.899	0.922	0.932	0.934
Bagging(NaiveBayes)	0.716	0.715	0.712	0.715
Bagging(BayesNet GeneticSearch)	0.904	0.924	0.936	0.941
Bagging(KStar)	0.889	0.895	0.896	0.897
Bagging(DecisionTable BestFirst)	0.903	0.924	0.935	0.938
Bagging(Neural Network)	0.761	0.767	0.772	0.777
Bagging(J48)	0.918	0.938	0.943	0.946
Stacking	0.805	0.842	0.859	0.875
Vote	0.788	0.812	0.829	0.839

จากตารางที่ 4.6 และ 4.7 วิธีการแบบกลุ่มด้วยวิธี Bagging ไม่ผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง ที่จำนวนชุดข้อมูลเท่ากับ 16,000 ชุด โดยใช้ตัวแยกแยะด้วยวิธี J48 ให้ค่า AUC ดีที่สุดเท่ากับ 0.946

ตารางที่ 4.8 : ค่า AUC จากวิธีแบบกลุ่มโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง

วิธีแบบกลุ่ม	ค่า AUC			
	1600	4800	9600	16000
AdaBoostM1(NaiveBayes)	0.832	0.855	0.859	0.869
AdaBoostM1(BayesNet GeneticSearch)	0.851	0.886	0.898	0.908
AdaBoostM1(KStar)	0.878	0.905	0.919	0.923
AdaBoostM1(DecisionTable BestFirst)	0.873	0.906	0.919	0.925
AdaBoostM1(Neural Network)	0.875	0.909	0.920	0.922
AdaBoostM1(J48)	0.879	0.909	0.924	0.927
Bagging(NaiveBayes)	0.883	0.896	0.901	0.905
Bagging(BayesNet GeneticSearch)	0.896	0.917	0.932	0.939
Bagging(KStar)	0.897	0.919	0.931	0.935

ตารางที่ 4.9 : ค่า AUC จากวิธีแบบกลุ่มโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง (ต่อ)

วิธีแบบกลุ่ม	ค่า AUC			
	1600	4800	9600	16000
Bagging(DecisionTable BestFirst)	0.896	0.915	0.929	0.934
Bagging(Neural Network)	0.898	0.922	0.934	0.940
Bagging(J48)	0.889	0.911	0.925	0.932
Stacking	0.859	0.876	0.889	0.896
Vote	0.802	0.823	0.842	0.845

จากตาราง 4.8 และ 4.9 เป็นวิธีการแบบกลุ่มโดยผ่านกระบวนการทำให้เป็นข้อมูลแบบไม่ต่อเนื่องวิธีการ Bagging โดยใช้ตัวแยกแยะโครงข่ายประสาทเทียม ที่จำนวนรอบการเรียนรู้เท่ากับ 200 อัตราการเรียนรู้เท่ากับ 0.1 และจำนวนชุดข้อมูลเท่ากับ 16,000 ชุด ให้ค่า AUC ดีที่สุดเท่ากับ 0.940

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การวิจัยนี้มุ่งเน้นที่การหาวิธีการสร้างตัวแบบที่ดีที่สุดโดยการเปรียบเทียบค่าพื้นที่ใต้โค้ง ROC (AUC) โดยใช้ข้อมูลในการสร้างตัวแบบ ทั้งขั้นตอนการฝึกตัวแบบและการทดสอบตัวแบบ จากข้อมูลการใช้กระแสไฟฟ้าแบบ 3 เฟส ที่ถูกเก็บจากเครื่องจักร 14 เครื่อง โดยในงานวิจัยนี้ มุ่งเน้นศึกษาที่กระบวนการผลิต 4 กระบวนการด้วยกัน ประกอบไปด้วย Masking, Rough Cut, Drill และ Semi Finish จำนวนชุดข้อมูลทั้งหมดมี 7 ชุดด้วยกันคือ 1600, 4800, 9600, 16000, 37000, 67000 และ 157000 ซึ่งในขั้นตอนการเตรียมข้อมูลจะแบ่งชุดข้อมูลออกเป็น 2 แบบคือแบบที่ผ่านการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง กับแบบที่ไม่ผ่านการทำให้เป็นข้อมูลแบบไม่ต่อเนื่อง ซึ่งผลของ AUC ที่ได้เป็นดังนี้ ในส่วนของตัวแบบแบบเดี่ยว ตัวแบบที่สร้างขึ้นด้วยวิธีเบย์เซียนเน็ตเวิร์คด้วยวิธีการค้นหาแบบพันธุกรรมซึ่งผ่านการทำให้เป็นข้อมูลแบบไม่ต่อเนื่องให้ค่า AUC ที่ดีที่สุดคือ 0.938 ที่ชุดทดสอบที่มีจำนวนข้อมูลเท่ากับ 1600 แต่เมื่อเพิ่มจำนวนชุดทดสอบมากขึ้นค่า AUC จะเริ่มลดลง ในส่วนของวิธีการแบบกลุ่มตัวแบบที่สร้างด้วยวิธีการ Bagging โดยใช้ตัวจำแนกด้วยวิธีต้นไม้ตัดสินใจ ซึ่งในขั้นตอนการเตรียมข้อมูลไม่ผ่านการทำให้เป็นข้อมูลแบบไม่ต่อเนื่องให้ค่า AUC ที่ดีที่สุดคือ 0.946 ที่ชุดทดสอบที่มีจำนวนข้อมูลเท่ากับ 1600 จะเห็นได้ว่าการสร้างตัวแบบด้วยวิธีการแบบกลุ่มเมื่อเทียบกับวิธีการสร้างตัวแบบแบบเดี่ยวแล้วจะทำให้ตัวแบบมีความแม่นยำสูงขึ้น ซึ่งวิธีการต้นไม้ตัดสินใจโดยใช้เทคนิคการตัดแต่ง (pruning) นั้น เหมาะสมที่สุดที่จะนำมาประยุกต์ใช้กับข้อมูลการใช้กระแสไฟฟ้าแบบ 3 เฟส จากเครื่องจักรประเภทเครื่องกัด

5.2 ข้อเสนอแนะ

5.2.1 ในงานวิจัยนี้ใช้ช่วงเวลาเก็บข้อมูลเพียง 4 เดือน ซึ่งอาจยังไม่ครอบคลุมกับรูปของของกระบวนการผลิตทั้งหมด หากใช้เวลาในการเก็บข้อมูลยาวนานขึ้นอาจได้ตัวแบบที่มีความแม่นยำสูงขึ้น

5.2.2 ในการสร้างตัวแบบโดยใช้วิธีโครงข่ายประสาทเทียมอาจมีวิธีการที่จะใช้ในการช่วยหาค่าพารามิเตอร์ที่ดีที่สุด เช่น อัตราการเรียนรู้, จำนวนรอบการเรียนรู้

5.2.3 ในการสร้างตัวแบบในบางวิธีมีการใช้งานหน่วยความจำเป็นจำนวนมาก เช่น วิธีโครงข่ายประสาทเทียมผู้วิจัยสามารถสร้างตัวแบบได้สูงสุดที่ชุดข้อมูลที่มีจำนวนเท่ากับ 1600 เท่านั้น หากใช้ชุดข้อมูลที่มากกว่านี้จะทำให้หน่วยความจำไม่เพียงพอ ซึ่งอาจแก้ปัญหาด้วยการใช้โปรแกรมที่สามารถทำงานแบบขนานได้

5.2.4 ในการสร้างตัวแบบด้วยวิธีแบบกลุ่มนั้นในบางวิธีใช้เวลาในการสร้างค่อนข้างมากแต่กลับให้ผลของความแม่นยำที่ต่ำ วิธีแบบกลุ่มนั้นใช้ได้กับตัวแบบบางวิธีเท่านั้น

รายการอ้างอิง

ภาษาอังกฤษ

Breiman, L. Bagging predictors. Machine Learning, 26(1996): 123–140.

Connolly B. Random Sampling in T-SQL [online]. 2004. Available from : <http://msdn.microsoft.com/en-us/library/aa175776%28v=sql.80%29.aspx> [2011,Jan].

Efron B., and Tibshirani R. An Introduction to the Bootstrap. NY:Chapman & Hall, 1993.

Fawcett T. ROC Graphs: Notes and Practical Considerations for Researchers [online]. 2004. Available from : <http://citeseerx.ist.psu.edu> [2011,Jan].

Freund, Y., and Schapire R. E. A decision–theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55(1997): 119–139.

Kuncheva, L. Combining Pattern Classifiers Methods and Algorithms. Wiley-Interscience, 2004.

Modbus organization, Modbus over serial line specification and implementation guide [online]. 2002. Available from : http://www.modbus.org/docs/Modbus_over_serial_line_V1.pdf [2011,Jan].

Nowlan S. J., and Hinton, G. E. Evaluation of adaptive mixtures of competing experts. Advances in Neural Information Processing Systems (1991): 774–780.

Usama M., and Keki B. Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning. Proceedings of the International Joint Conference on Uncertainty in AI, pp.1022-1027. Morgan Kaufmann, 1993.

Witten I., and Frank E., Data Mining practical machine learning tools and techniques with Java implementations. Morgan Kaufmann, 2000.

Wolpert D. Stacked Generalization. Neural Network 5(1992): 241-259.

Xu L., Krzyzak A. and Suen C. Y. Methods of combining multiple classifiers and their application to handwriting recognition. IEEE Transactions on Systems, Man, and Cybernetics 22(1992): 418-435.

ภาคผนวก

1. โพรโตคอล Modbus

โพรโตคอล Modbus เป็นโพรโตคอลที่ถูกพัฒนาโดยบริษัทผู้ผลิต PLC (Programmable Logic Controller) ชื่อบริษัท Modicon ของประเทศสหรัฐอเมริกาซึ่งปัจจุบันได้ถูกควบรวมกิจการเป็นส่วนหนึ่งของ บริษัท Schneider Electric

โพรโตคอล Modbus เป็นโพรโตคอลที่มีโครงสร้างค่อนข้างง่าย และอ้างอิงถึง OSI เพียงไม่กี่ชั้น ดังนั้นมันจึงเป็นที่นิยม โดยเฉพาะในอุปกรณ์ประเภท PLC และดิจิตอลมิเตอร์ รวมทั้งระบบที่ต้องการความเร็วสูงในการตอบสนองเช่น โรงผลิตไฟฟ้า เป็นต้น

ตัวมาตรฐาน Modbus ประกอบด้วยโพรโตคอลในระดับแอปพลิเคชัน (OSI ชั้นที่ 7) ที่ใช้การสื่อสารแบบไคลเอนต์/เซิร์ฟเวอร์ (Client/Server) ระหว่างอุปกรณ์ที่ถูกเชื่อมต่อกับระบบบัสของเครือข่าย

โพรโตคอล Modbus ใช้โครงสร้างการสื่อสารแบบไคลเอนต์/เซิร์ฟเวอร์ โดยมีโหมดการทำงานแบบ รีควีสต์/เรสปอนส์ (Request/Response) หรือการร้องขอและการตอบสนอง โดยไม่มีวิธีการควบคุมการเข้าถึงสื่อ (Media Access Control) ที่ถูกใช้ในเลเยอร์ที่ 2 โมเดลไคลเอนต์/เซิร์ฟเวอร์มีรูปแบบเมสเสจ (Message) อยู่ 4 ชนิดหลัก จากมุมมองของไคลเอนต์และเซิร์ฟเวอร์

การส่งข้อมูลแบบอนุกรม

มีสองโหมดใน การส่งข้อมูลแบบอนุกรมที่ถูกนิยามใน Modbus นั่นคือ โหมด RTU และ ASCII แต่ละโหมดได้ถูกนิยามการเข้ารหัสในเมสเสจและการส่งบิตข้อมูลบนสายสื่อสารแบบอนุกรม โหมดการส่งข้อมูลต้องเหมือนกันทั้งสองด้านของการสื่อสารแบบอนุกรม

Modbus RTU สมควรเป็นโหมดที่ทุกอุปกรณ์ที่สนับสนุนโพรโตคอล Modbus ต้องสนับสนุน ส่วน Modbus ASCII เป็นเพียงโหมดทางเลือกที่ใช้ในระบบงานบางระบบเท่านั้น อุปกรณ์ควรสามารถเลือกโหมดตามที่ใช้ต้องการไม่ว่าจะเป็น RTU และ ASCII โดยที่โหมดดีฟอลต์ต้องเป็นโหมด RTU

โหมต RTU

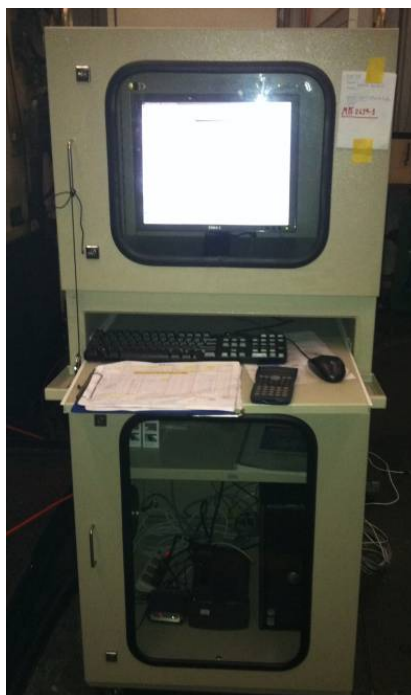
เมื่ออุปกรณ์สื่อสารใช้โหมต Modbus RTU แต่ละ 8 บิตในเมสเสจจะถูกแบ่งเป็น 2 ส่วน ส่วนละ 4 บิตเพื่อแทนที่หรือแสดงด้วยอักขระ ASCII 2 อักขระ ข้อดีของโหมต RTU คือสามารถลดจำนวนไบต์ที่ใช้ส่งข้อมูลได้เกือบครึ่งหนึ่งเมื่อเทียบกับโหมต ASCII แต่ละเมสเสจต้องส่งข้อมูลเป็นไบต์ต่อเนื่องกัน

2. เครื่องจักรซีเอ็นซีชนิดเครื่องกัด



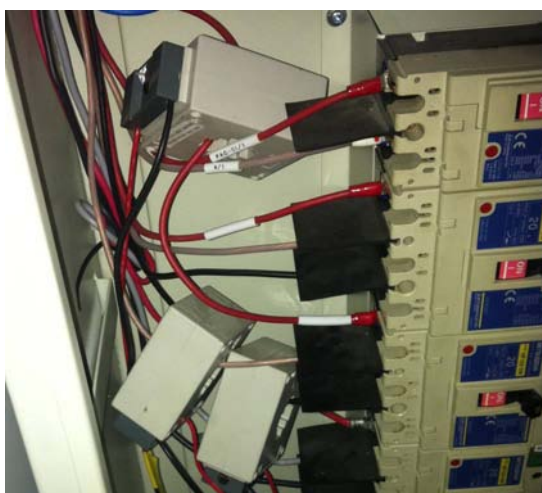
ภาพที่ 1 เครื่องจักรซีเอ็นซีชนิดเครื่องกัด

3. ตู้ควบคุม



ภาพที่ 2 ตู้ควบคุม

4. ชุดขดลวดเหนี่ยวนำ (Current Transformer)



ภาพที่ 3 ชุดขดลวดเหนี่ยวนำ (Current Transformer)

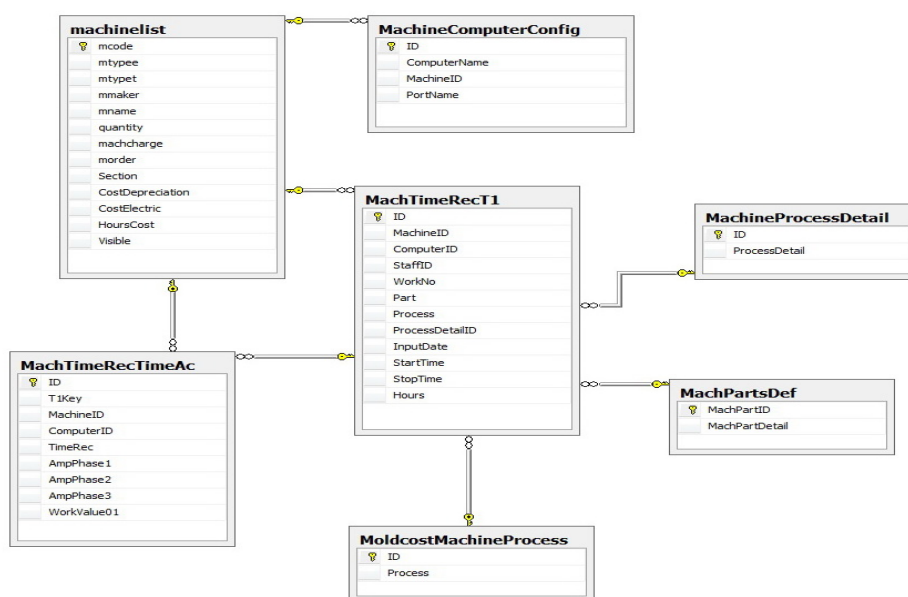
5. เครื่องวัดกระแสไฟฟ้าชนิด 3 เฟส



ภาพที่ 4 เครื่องวัดกระแสไฟฟ้าชนิด 3 เฟส

6. ฐานข้อมูลที่ใช้เก็บข้อมูล

โดยใช้ Microsoft SQL Server 2005 เป็นตัวเก็บข้อมูล



ภาพที่ 5 แผนภาพของฐานข้อมูลที่ใช้เก็บข้อมูล

7. โปรแกรมสำหรับเก็บค่ากระแสไฟฟ้าจากเครื่องจักร

System Status :

System is OK
Host Name : NCM-MACHINE01
Machine ID : 25 Name: MPF-2614DS No.1
Port : COM7

System is OK
Host Name : NCM-MACHINE01
Machine ID : 37 Name: MPF-2614DS No.2
Port : COM8

14/01/2554 17:22:43	A1=9.00	A2=8.40	A3=8.40
14/01/2554 17:22:46	A1=9.50	A2=8.20	A3=8.20
14/01/2554 17:22:49	A1=9.00	A2=8.30	A3=8.30
14/01/2554 17:22:52	A1=8.90	A2=8.30	A3=8.30
14/01/2554 17:22:55	A1=8.60	A2=8.70	A3=8.70
14/01/2554 17:22:58	A1=8.30	A2=8.70	A3=8.70
14/01/2554 17:23:01	A1=8.60	A2=8.70	A3=8.70
14/01/2554 17:23:04	A1=8.60	A2=8.40	A3=8.40

14/01/2554 17:22:43	A1=9.80	A2=8.50	A3=11.10
14/01/2554 17:22:47	A1=9.10	A2=9.10	A3=11.60
14/01/2554 17:22:49	A1=8.00	A2=10.20	A3=11.30
14/01/2554 17:22:53	A1=9.00	A2=10.00	A3=10.50
14/01/2554 17:22:56	A1=10.10	A2=8.60	A3=10.70
14/01/2554 17:22:58	A1=10.50	A2=7.80	A3=11.40
14/01/2554 17:23:02	A1=10.00	A2=8.20	A3=12.20
14/01/2554 17:23:04	A1=8.80	A2=9.40	A3=11.90

Last Update : 14/01/2554 17:23:04

Copyright 2010, Automotive Mold Technology Co., Ltd.
Version Beta 2.0 update 30/11/2010 Developed By Thanarak Raktham

ภาพที่ 6 โปรแกรมเก็บข้อมูลกระแสไฟฟ้าจากเครื่องจักรซีเอ็นซี

8. โปรแกรมสำหรับเก็บข้อมูลรายละเอียดงานหน้าเครื่องจักร

» Home » Daily Report System » Machine Report System » ระบบข้อมูลโรงงาน » Tool Control System » ระบบบริหารงานบุคคล » Mold Cost System »

Machine Time Record System : Start New Job

1. เลือกเครื่องจักร

2. เลือกวันที่ เวลา เวลา

3. เลือกรหัสงาน

4. เลือก Parts

5. เลือก Process (แผนก NCM)

6. เลือก Process ย่อ

รหัสข้อมูล	Machine No.	Name	Maker	Job Code	Part Name	Process	Process Detail	Start Time	End Time
10220	25	MPF-2614DS No.1	Toshiba	I151	Cavity	SEMI FINISH CUTTING	เครื่องจักรทำงาน	14/01/2011 15:18	No Data
10218	37	MPF-2614DS No.2	Toshiba	I151	Core	SEMI FINISH CUTTING	เครื่องจักรทำงาน	14/01/2011 10:19	No Data

ภาพที่ 7 โปรแกรมสำหรับเก็บข้อมูลรายละเอียดงานหน้าเครื่องจักร

9. ผลการคำนวณจากการสร้างตัวแบบ

วิธีการ NaiveBay โดยไม่ผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 9,600

Correctly Classified Instances	4010	41.7708 %
Incorrectly Classified Instances	5590	58.2292 %
Kappa statistic	0.2236	
Mean absolute error	0.2951	
Root mean squared error	0.4098	
Relative absolute error	78.6822 %	
Root relative squared error	94.6363 %	
Coverage of cases (0.95 level)	83.5104 %	
Mean rel. region size (0.95 level)	64.0156 %	
Total Number of Instances	9600	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.467	0.317	0.329	0.467	0.386	0.683	DRILL
	0.697	0.170	0.577	0.697	0.631	0.806	MARKING
	0.283	0.017	0.849	0.283	0.425	0.681	ROUGH CUT
	0.224	0.273	0.215	0.224	0.219	0.682	SEMI FINISH CUTTING
W Avg.	0.418	0.194	0.493	0.418	0.415	0.713	

วิธีการ BayesNet Genetic Search โดยไม่ผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000

Correctly Classified Instances	12112	75.7 %
Incorrectly Classified Instances	3888	24.3 %
Kappa statistic	0.676	
Mean absolute error	0.158	
Root mean squared error	0.2858	
Relative absolute error	42.1317 %	
Root relative squared error	65.9988 %	
Coverage of cases (0.95 level)	97.75 %	
Mean rel. region size (0.95 level)	53.425 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.689	0.128	0.642	0.689	0.665	0.902	DRILL
	0.893	0.04	0.882	0.893	0.887	0.982	MARKING
	0.679	0.088	0.72	0.679	0.699	0.917	ROUGH CUT
	0.766	0.068	0.791	0.766	0.778	0.938	SEMI FINISH CUTTING
W Avg.	0.757	0.081	0.758	0.757	0.757	0.935	

วิธีการ NaiveBayes โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 1,600

Correctly Classified Instances	1109	69.3125 %
Incorrectly Classified Instances	491	30.6875 %
Kappa statistic	0.5908	
Mean absolute error	0.1716	
Root mean squared error	0.3488	
Relative absolute error	45.7675 %	
Root relative squared error	80.5582 %	
Coverage of cases (0.95 level)	89.1875 %	
Mean rel. region size (0.95 level)	46.1563 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.7	0.105	0.69	0.7	0.695	0.877	SEMI FINISH CUTTING
	0.523	0.068	0.718	0.523	0.605	0.857	ROUGH CUT
	0.748	0.201	0.554	0.748	0.636	0.839	DRILL
	0.803	0.035	0.884	0.803	0.841	0.96	MARKING
W Avg.	0.693	0.102	0.711	0.693	0.694	0.883	

วิธีการ BayesNet Genetic Search โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000

Correctly Classified Instances	12205	76.2813 %
Incorrectly Classified Instances	3795	23.7188 %
Kappa statistic	0.6838	
Mean absolute error	0.1492	
Root mean squared error	0.2832	
Relative absolute error	39.7941 %	
Root relative squared error	65.4041 %	
Coverage of cases (0.95 level)	97.2688 %	
Mean rel. region size (0.95 level)	48.3688 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.663	0.108	0.672	0.663	0.668	0.903	DRILL
	0.883	0.039	0.882	0.883	0.883	0.984	MARKING
	0.711	0.09	0.725	0.711	0.718	0.921	ROUGH CUT
	0.794	0.079	0.77	0.794	0.782	0.943	SEMI FINISH CUTTING
W Avg.	0.763	0.079	0.762	0.763	0.762	0.938	

วิธีการโครงข่ายประสาทเทียม โดยไม่ผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 1,600 จำนวน
รอบการเรียนรู้เท่ากับ 200 และอัตราการเรียนรู้เท่ากับ 0.1

Correctly Classified Instances	826	51.625 %
Incorrectly Classified Instances	774	48.375 %
Kappa statistic	0.355	
Mean absolute error	0.2921	
Root mean squared error	0.3805	
Relative absolute error	77.8845 %	
Root relative squared error	87.8642 %	
Coverage of cases (0.95 level)	98.75 %	
Mean rel. region size (0.95 level)	90.7031 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.53	0.249	0.415	0.53	0.465	0.724	SEMI FINISH CUTTING
	0.435	0.102	0.588	0.435	0.5	0.769	ROUGH CUT
	0.428	0.235	0.377	0.428	0.401	0.676	DRILL
	0.673	0.059	0.791	0.673	0.727	0.852	MARKING
W Avg.	0.516	0.161	0.543	0.516	0.523	0.755	

วิธีการโครงข่ายประสาทเทียม โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000 จำนวน
รอบการเรียนรู้เท่ากับ 500 และอัตราการเรียนรู้เท่ากับ 0.3

Correctly Classified Instances	11715	73.2188 %
Incorrectly Classified Instances	4285	26.7813 %
Kappa statistic	0.6429	
Mean absolute error	0.1445	
Root mean squared error	0.3377	
Relative absolute error	38.537 %	
Root relative squared error	77.987 %	
Coverage of cases (0.95 level)	82.8375 %	
Mean rel. region size (0.95 level)	34.1313 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.632	0.125	0.628	0.632	0.63	0.842	DRILL
	0.843	0.031	0.902	0.843	0.871	0.953	MARKING
	0.686	0.096	0.705	0.686	0.695	0.873	ROUGH CUT
	0.768	0.106	0.708	0.768	0.737	0.888	SEMI FINISH CUTTING
W Avg.	0.732	0.089	0.735	0.732	0.733	0.889	

วิธีการโครงข่ายประสาทเทียม โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000 จำนวน
รอบการเรียนรู้เท่ากับ 200 และอัตราการเรียนรู้เท่ากับ 0.1

Correctly Classified Instances	12210	76.3125 %
Incorrectly Classified Instances	3790	23.6875 %
Kappa statistic	0.6842	
Mean absolute error	0.1447	
Root mean squared error	0.2916	
Relative absolute error	38.5789 %	
Root relative squared error	67.3346 %	
Coverage of cases (0.95 level)	94.9063 %	
Mean rel. region size (0.95 level)	44.5063 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.646	0.1	0.683	0.646	0.664	0.892	DRILL
	0.897	0.037	0.889	0.897	0.893	0.978	MARKING
	0.73	0.104	0.701	0.73	0.715	0.91	ROUGH CUT
	0.78	0.075	0.777	0.78	0.778	0.934	SEMI FINISH CUTTING
W Avg.	0.763	0.079	0.762	0.763	0.763	0.928	

วิธีการ Bagging ด้วยกับ NaiveBayes โดยไม่ผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 9,600

Correctly Classified Instances	4103	42.7396 %
Incorrectly Classified Instances	5497	57.2604 %
Kappa statistic	0.2365	
Mean absolute error	0.295	
Root mean squared error	0.4098	
Relative absolute error	78.6721 %	
Root relative squared error	94.6459 %	
Coverage of cases (0.95 level)	83.3125 %	
Mean rel. region size (0.95 level)	64.0234 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.489	0.323	0.335	0.489	0.398	0.683	DRILL
	0.697	0.168	0.58	0.697	0.633	0.807	MARKING
	0.285	0.019	0.833	0.285	0.425	0.681	ROUGH CUT
	0.239	0.254	0.239	0.239	0.239	0.68	SEMI FINISH CUTTING
W Avg.	0.427	0.191	0.497	0.427	0.424	0.712	

วิธีการ Bagging ด้วยกับ J48 โดยไม่ผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000

Correctly Classified Instances	12705	79.4063 %
Incorrectly Classified Instances	3295	20.5938 %
Kappa statistic	0.7254	
Mean absolute error	0.1351	
Root mean squared error	0.2691	
Relative absolute error	36.022 %	
Root relative squared error	62.1552 %	
Coverage of cases (0.95 level)	97.475 %	
Mean rel. region size (0.95 level)	47.3922 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.704	0.09	0.723	0.704	0.713	0.919	DRILL
	0.91	0.035	0.896	0.91	0.903	0.984	MARKING
	0.753	0.091	0.735	0.753	0.744	0.929	ROUGH CUT
	0.81	0.059	0.82	0.81	0.815	0.952	SEMI FINISH CUTTING
Weighted Avg.	0.794	0.069	0.794	0.794	0.794	0.946	

วิธีการ Vote โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 1,600

Correctly Classified Instances	1125	70.3125 %
Incorrectly Classified Instances	475	29.6875 %
Kappa statistic	0.6042	
Mean absolute error	0.1484	
Root mean squared error	0.3853	
Relative absolute error	39.5833 %	
Root relative squared error	88.9757 %	
Coverage of cases (0.95 level)	70.3125 %	
Mean rel. region size (0.95 level)	25 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.713	0.104	0.695	0.713	0.704	0.804	SEMI FINISH CUTTING
	0.585	0.09	0.684	0.585	0.631	0.748	ROUGH CUT
	0.713	0.183	0.565	0.713	0.631	0.765	DRILL
	0.803	0.019	0.933	0.803	0.863	0.892	MARKING
W Avg.	0.703	0.099	0.719	0.703	0.707	0.802	

วิธีการ Bagging ด้วยกับโครงข่ายประสาทเทียม โดยผ่านการ discretize ที่จำนวนชุดทดสอบเท่ากับ 16,000

Correctly Classified Instances	12426	77.6625 %
Incorrectly Classified Instances	3574	22.3375 %
Kappa statistic	0.7022	
Mean absolute error	0.1441	
Root mean squared error	0.2779	
Relative absolute error	38.4217 %	
Root relative squared error	64.1884 %	
Coverage of cases (0.95 level)	96.8563 %	
Mean rel. region size (0.95 level)	47.4953 %	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.686	0.106	0.683	0.686	0.685	0.908	DRILL
	0.902	0.036	0.894	0.902	0.898	0.984	MARKING
	0.724	0.088	0.734	0.724	0.729	0.923	ROUGH CUT
	0.795	0.069	0.795	0.795	0.795	0.945	SEMI FINISH CUTTING
W Avg.	0.777	0.074	0.776	0.777	0.776	0.940	

ประวัติผู้เขียนวิทยานิพนธ์

นายธนารักษ์ รักรธรรม เกิดเมื่อวันที่ 29 มกราคม พุทธศักราช 2524 สำเร็จ การศึกษาระดับปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมการผลิต จากภาควิชาวิศวกรรม การผลิต คณะวิศวกรรมศาสตร์ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ เมื่อปีการศึกษา 2546 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชา วิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2552