

บทที่ 3

วิธีการดำเนินการวิจัย

การศึกษาค้นคว้าครั้งนี้เป็นการศึกษาในรูปแบบงานวิจัยพื้นฐานและการวิจัยเชิงประยุกต์ (Basic Research and Applied Research) โดยนำข้อมูลจากฐานข้อมูลที่มีอยู่มาทำเหมืองข้อมูล เพื่อค้นหาความรู้ (Knowledge Discovery) รูปแบบของการเกิดรพุนแบบต่างๆในกระบวนการคิดค้นรูปโปรเตอร์ด้วยอคูมิเนียม และนำไปสู่การพัฒนาสร้างตัวแบบจำลองพารามิเตอร์ที่ช่วยลดปัญหาการเกิดรพุนแบบต่างๆ ที่ทำการศึกษได้ ในการดำเนินการวิจัยนี้จะใช้กระบวนการค้นความรู้แบบ CRISP-DM ร่วมกับทฤษฎีการทำเหมืองข้อมูล (Data mining theory) ในการสร้างตัวแบบจำลอง

3.1 ขั้นตอนในการสร้างตัวแบบ

3.1.1 การทำความเข้าใจในเรื่องธุรกิจที่จะทำ (Business understanding)

3.1.1.1 การกำหนดวัตถุประสงค์ทางธุรกิจ (Determine business Objectives) เป็นเป้าหมายอย่างแรกสำหรับการวิเคราะห์ที่จะต้องเข้าใจถึงวัตถุประสงค์ทางธุรกิจที่ชัดเจน เป้าหมายของการวิเคราะห์จะไม่ครอบคลุมถึงปัจจัยที่สำคัญตอนเริ่มโครงการ แต่จะมีผลกระทบต่อผลลัพธ์ในขั้นสุดท้าย กล่าวคือ ผลลัพธ์ที่ผิดพลาดในขั้นตอนนี้จะทำให้ได้คำตอบที่ผิดหลังจากที่โครงการเสร็จ

3.1.1.2 การประเมินสถานการณ์ปัจจุบัน (Assess situation) ในขั้นตอนนี้จะเกี่ยวข้องกับรายละเอียดที่เป็นจริงของทรัพยากรที่มีอยู่ ข้อจำกัด สมมติฐาน หรือปัจจัยอื่นๆที่ควรนำมาพิจารณาเพื่อช่วยในการตัดสินใจเกี่ยวกับเป้าหมายของการ วิเคราะห์ข้อมูลและโครงการ

3.1.1.3 กำหนดเป้าหมายของการทำเหมืองข้อมูล (Determine data mining goals) เป้าหมายของการทำเหมืองข้อมูลจะหมายถึงวัตถุประสงค์ทางด้านเทคนิค เช่น การสร้างตัวแบบพยากรณ์ว่าจะใช้ Algorithmประเภทใดเพื่อวิเคราะห์รูปแบบการเกิดของเสียประเภทต่างๆ

3.1.1.4 สร้างแผนโครงการ (Produce project plan) อธิบายจุดมุ่งหมายของแผนสำหรับทำให้เป้าหมายของการทำเหมืองข้อมูลเพื่อสร้างตัวแบบประสบความสำเร็จ

3.1.2 การทำความเข้าใจกับข้อมูลที่จัดเก็บ (Data understanding)

3.1.2.1 การเก็บข้อมูลขั้นต้น (Collect initial data) ความต้องการข้อมูลของโครงการ/ การเข้าถึงข้อมูลซึ่งแสดงอยู่ในรายการทรัพยากรของโครงการ การเก็บข้อมูลเบื้องต้นรวมถึงการดึงข้อมูล ถ้าจำเป็นเพื่อให้เข้าใจได้ง่ายขึ้น เช่น ถ้าหากเราสนใจที่จะใช้เครื่องมือบางอย่างในการทำ ความเข้าใจถึงข้อมูลจำเป็นที่จะต้องดึงข้อมูลเหล่านั้นเข้าสู่เครื่องมือนั้นก่อน

3.1.2.2 การอธิบายข้อมูล (Describe data) คือ การที่สามารถอธิบายข้อมูลและ ทดสอบข้อมูลทั้งหมดได้

3.1.2.3 การขยายข้อมูล (Explore data) ในส่วนนี้จะเกี่ยวข้องโดยตรงกับคำถามของ เหมืองข้อมูล ซึ่งอาจต้องใช้ในการสืบค้น สร้างภาพ และรายงาน การวิเคราะห์เหล่านี้อาจจะหมายถึง เป้าหมายของเหมืองข้อมูล อย่างไรก็ตามผู้ใช้นั้นมักจะให้ปรับเปลี่ยนคำอธิบายของข้อมูลและ คุณภาพของรายงาน อาจจะรวมไปถึงการแปลงและการจัดเตรียมข้อมูลอื่นๆ เพื่อการวิเคราะห์ใน อนาคต

3.1.2.4 ตรวจสอบคุณภาพของข้อมูล (Verify data quality) ตรวจสอบคุณภาพของ ข้อมูล เช่น ข้อมูลสมบูรณ์หรือยัง (ครอบคลุมความต้องการทั้งหมดหรือไม่) ข้อมูลถูกต้องหรือไม่ และมีข้อผิดพลาดหรือไม่ ถ้ามีมันคืออะไร มีค่าที่หายไปของข้อมูลหรือไม่ ถ้ามีจะกู้คืนอย่างไร มัน เกิดขึ้นที่ไหนและจะแก้ไขอย่างไร

3.1.3 การเตรียมข้อมูล (Data preparation)

3.1.3.1 การคัดเลือกข้อมูล (Select data) ตัดสินใจว่าข้อมูลใดถูกใช้ในการวิเคราะห์ บรรทัดฐานที่เกี่ยวข้องกับเป้าหมายของเหมืองข้อมูล คุณภาพและข้อจำกัดทางด้านเทคนิค เช่น จำนวนข้อมูลที่จำกัด หรือประเภทของข้อมูล

3.1.3.2 การทำความสะอาดข้อมูล (Data Cleansing) คือการเพิ่มคุณภาพของข้อมูลสู่ ระดับที่ต้องการโดยเลือกใช้เทคนิคในการวิเคราะห์ซึ่งอาจจะเกี่ยวข้องกับการเลือกข้อมูลย่อย บางส่วน การเพิ่มเติมค่าคงที่ หรือ การใช้เทคนิคขั้นสูง เช่นประมาณการข้อมูลที่หายไปโดยใช้ตัว แบบ

3.1.3.3 การสร้างข้อมูลใหม่ (Construct data) งานส่วนนี้จะรวมไปถึงวิธีการดำเนินการ ในการจัดเตรียมข้อมูล เช่น การสร้างลักษณะประจำใหม่ การทำข้อมูลชุดใหม่ให้สมบูรณ์ หรือการ เปลี่ยนแปลงค่าของลักษณะประจำเดิมที่มีอยู่

3.1.3.4 การรวมข้อมูล (Integrate data) เป็นขั้นตอนในการรวมข้อมูลจากหลายๆ ตาราง หรือหลายๆ แหล่งที่มาเพื่อสร้างข้อมูลชุดใหม่

3.1.3.5 การเปลี่ยนรูปแบบของข้อมูล (Format data) การเปลี่ยนรูปแบบของข้อมูลโดยใช้การอ้างอิงจากโครงสร้างเดิมการแก้ไขข้อมูลจะไม่เปลี่ยนความหมายแต่จำเป็นต้องเปลี่ยนเพื่อให้มีความเหมาะสมกับเครื่องมือที่ใช้สร้างตัวแบบ

3.1.4 การสร้างตัวแบบ (Modeling)

3.1.4.1 เลือกเทคนิคการสร้างตัวแบบ (Select Modeling Technique) เลือกเทคนิคการสร้างตัวแบบที่ทำได้จริง ถ้าหากมีการใช้เทคนิคหลากหลาย ให้แยกแต่ละเทคนิคออกจากกันโดยเด็ดขาดเพื่อป้องกันความสับสน เทคนิคและเครื่องมือไม่สามารถใช้ได้กับทุกๆ งาน มีเพียงบางเทคนิคเท่านั้นที่เหมาะสม ด้วยข้อจำกัดต่างๆ ทำให้มีเพียงเครื่องมือหรือเทคนิคชนิดเดียวเท่านั้นที่ใช้แก้ปัญหาและอาจจะไม่ใช่เทคนิคที่ดีที่สุดในการแก้ปัญหา

3.1.4.2 การสร้างแบบการทดสอบ (Generate Test Design) ก่อนที่จะสร้างตัวแบบ จำเป็นต้องทดสอบคุณภาพและความถูกต้องของตัวแบบ เช่น การทำเหมืองข้อมูลประเภทมีการเรียนรู้ แบบการจัดหมวดหมู่ จะใช้อัตราความผิดพลาดเป็นตัววัดคุณภาพของตัวแบบ ดังนั้นการออกแบบการทดสอบ ควรจะมีการแยกกลุ่มข้อมูลออก 2 กลุ่ม คือ Training set และ Test set ตัวแบบที่สร้างจาก Training set จะมีคุณภาพใกล้เคียงกับ Test set ยกเว้นกรณีปัญหาตัวแบบที่เข้ากับข้อมูลสำหรับเรียนรู้เท่านั้น (Over-fitting)

3.1.4.3 การสร้างตัวแบบ (Build model) ดำเนินการสร้างตัวแบบจากกลุ่มข้อมูล ตัวแบบที่ได้อาจจะมากกว่า 1 ก็ได้

3.1.4.4 การประเมินตัวแบบ (Assess model) ตัวแบบควรจะได้รับ การประเมินเพื่อให้แน่ใจว่าผ่านมาตรฐานการตัดสินใจ เป็นเทคนิคการประเมินที่ได้จากขั้นตอนการสร้างตัวแบบ

3.1.5 การประเมินผล (Evaluation)

3.1.5.1 การประเมินผล (Evaluate results) การประเมินในขั้นตอนนี้ก่อนหน้านี้อยู่เกี่ยวกับปัจจัยต่างๆ เช่น ความถูกต้อง และลักษณะทั่วไปของตัวแบบ ในขั้นตอนนี้จะประเมินว่าตัวแบบที่ได้ตรงกับเป้าหมายทางธุรกิจและการสนับสนุนการตัดสินใจหรือไม่

3.1.5.2 กระบวนการตรวจทานและแก้ไข (Review process) ในขั้นตอนนี้เป็นการนำผลลัพธ์จากขั้นตอนการประเมินผลมาปรับปรุงแก้ไขหรือ ทบทวนกระบวนการเพื่อตรวจสอบว่ามีปัจจัยสำคัญใดที่ถูกมองข้าม

3.1.5.3 การกำหนดขั้นตอนถัดไป (Determine next steps) หลังจากผ่านการตรวจสอบโครงการต้องตัดสินใจว่าเมื่อโครงการเสร็จสิ้นแล้วจะมีการทบทวน ปรับปรุงโครงการหรือติดตั้งโครงการใหม่

3.1.6 การนำความรู้ไปใช้ (Deployment)

3.1.6.1 การวางแผนการปรับใช้ตัวแบบ (Plan deployment) ขั้นตอนนี้เป็นการนำผลการประเมินและกลยุทธ์ในการตัดสินใจไปใช้งานจริงในธุรกิจ การขยายผล การนำตัวแบบไปเป็นต้นแบบหรือประยุกต์กับงานอื่นๆที่มีลักษณะใกล้เคียง

3.1.6.2 แผนการติดตามและการบำรุงรักษา (Plan monitoring and maintenance) การติดตามและการบำรุงรักษาเป็นประเด็นที่มีความสำคัญ ถ้าผลของการทำเหมืองข้อมูลเป็นส่วนหนึ่งของการทำงาน การเตรียมแผนการซ่อมบำรุงจะช่วยหลีกเลี่ยงการใช้งานข้อมูลจากข้อมูลเหมืองที่ไม่ถูกต้องได้ ในการติดตามผลลัพธ์จากเหมืองข้อมูล โครงการจะต้องมีรายละเอียดของกระบวนการติดตาม

3.1.6.3 จัดทำรายงานขั้นสุดท้าย (Produce final report) เมื่อสิ้นสุดโครงการ ผู้นำโครงการและทีมงานจะต้องจัดทำรายงานขั้นสุดท้าย ซึ่งขึ้นอยู่กับการใช้แผน รายงานนี้อาจนำเสนอผลสรุปของโครงการ หรือนำเสนอผลของการทำเหมืองข้อมูล

3.1.6.4 ทบทวนโครงการ (Review project) ประเมินสิ่งที่ถูกต้อง สิ่งที่ผิด สิ่งที่ดีและสิ่งที่ต้องมีการปรับปรุง

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 ฮาร์ดแวร์

1.1 Shotscope system

1.2 Die casting machine

1.3 Rotor model: Quest 2036-04

3.2.2 ซอฟต์แวร์

1. Microsoft Excel 2007
2. Shotscope software Version 4.1
3. WEKA 3.5.8 Data mining Software

การนำเครื่องมือที่ใช้ในการวิจัยเริ่มจากการเลือกชิ้นงานที่จะทำการศึกษาโดยในการศึกษานี้เลือกชิ้นงานโรเตอร์รุ่น Quest 2036-04 ซึ่งเป็นรุ่นที่มีการผลิตสูง เมื่อกำหนดชิ้นงานได้แล้วขั้นตอนต่อไปคือเลือกเครื่องจักรที่จะใช้ในการศึกษาหลังจากนั้นต้องมีการวางแผนจัดเก็บข้อมูลโดยการนี้จะใช้ข้อมูลที่จัดเก็บอยู่ใน Shotscope Server เมื่อกำหนดขอบเขตเครื่องมือทางด้านฮาร์ดแวร์ได้แล้วขั้นตอนต่อไปคือเลือกโปรแกรมหรือซอฟต์แวร์ที่ใช้ในการศึกษาซึ่งมีขั้นตอนดังนี้

1. การดึงข้อมูลจาก Shotscope Server ใช้โปรแกรม Shotscope Version 4.1
2. การจัดการข้อมูลใช้โปรแกรม Microsoft Excel 2007
3. การสร้างตัวแบบและวิเคราะห์ข้อมูลใช้โปรแกรม WEKA Version 3.5.8