

บทที่ 2

แผนดำเนินการขอบเขตและวิธีการทำวิจัย

2.1 แผนดำเนินการวิจัย

2.1.1 การศึกษาและวิเคราะห์คุณลักษณะของระบบ

ศึกษาถึงทฤษฎีที่เกี่ยวข้องกับการพัฒนาและคุณลักษณะที่พึงมีต่อระบบเพื่อให้ระบบมีประสิทธิภาพอย่างครบถ้วน

2.1.2 การศึกษาทฤษฎีที่นำมาใช้ในระบบ

ศึกษาทฤษฎีที่นำมาใช้ในระบบแบ่งแยกออกเป็นสองส่วน คือ

- **ทฤษฎีการพัฒนาระบบและฐานข้อมูล**

ศึกษาการพัฒนาระบบ โดยการนำ วงจรชีวิตการพัฒนาซอฟต์แวร์ ร่วมด้วยการสร้างแบบจำลองต้นแบบอย่างรวดเร็ว มาใช้ในการพัฒนาระบบเพื่อให้ระบบมีประสิทธิภาพและตรงตามวัตถุประสงค์ อีกทั้งยังมีการจัดการกับเอกสาร ไฟล์ และอื่นๆที่เกี่ยวข้องกับการพัฒนาระบบ รวมไปถึงการพัฒนาโครงสร้างข้อมูลเพื่อการนำเสนอข้อมูลที่ให้ผู้ใช้งานเข้าใจได้ง่าย

- **ทฤษฎีการวิเคราะห์ข้อมูลและการสร้างต้นไม้ตัดสินใจ (Decision Tree)**

ศึกษาทฤษฎีของการวิเคราะห์ข้อมูลในรูปแบบต่างๆเพื่อใช้ในการวิเคราะห์ข้อมูลในระบบขั้นต้นรวมถึงการศึกษาทฤษฎีการสร้างต้นไม้ตัดสินใจ (Decision Tree)

2.1.3 การวางแผนการพัฒนาระบบ

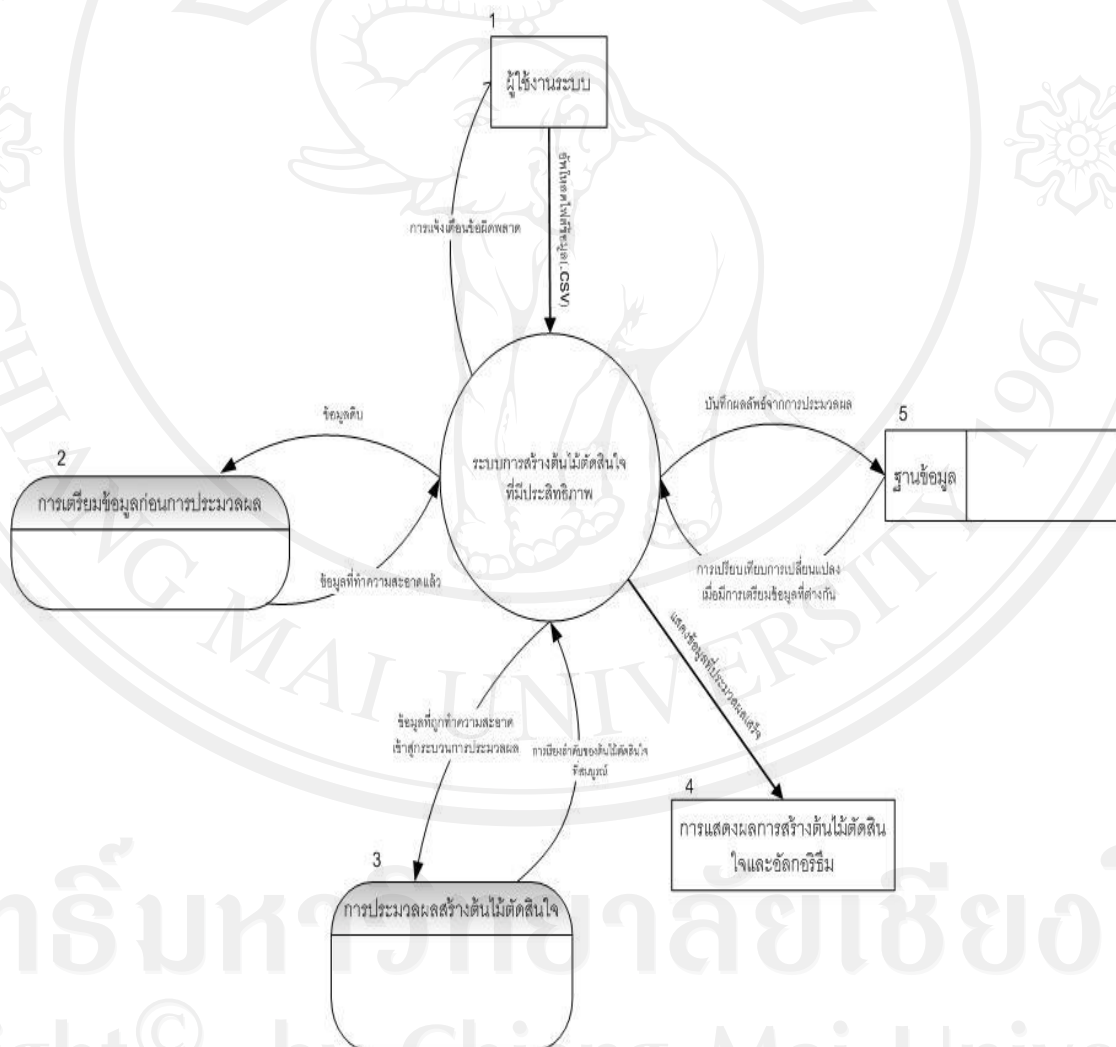
- **การวางแผนการดำเนินงาน**

วางแผนการดำเนินงานของการศึกษาและการพัฒนาระบบช่วยวิเคราะห์และสร้างเงื่อนไขของกลุ่มข้อมูลแบบหลากหลายในรูปแบบต้นไม้ตัดสินใจให้อยู่ในขอบเขตของเวลาที่กำหนดไว้รวมไป

ถึงการจัดการขั้นตอนการพัฒนาระบบให้เป็นไปตามทฤษฎีการสร้างแบบจำลองต้นแบบอย่างรวดเร็ว ภายใต้แนวทางของวงจรชีวิตการพัฒนาซอฟต์แวร์

- การวิเคราะห์และออกแบบพฤติกรรมการใช้งานของระบบ

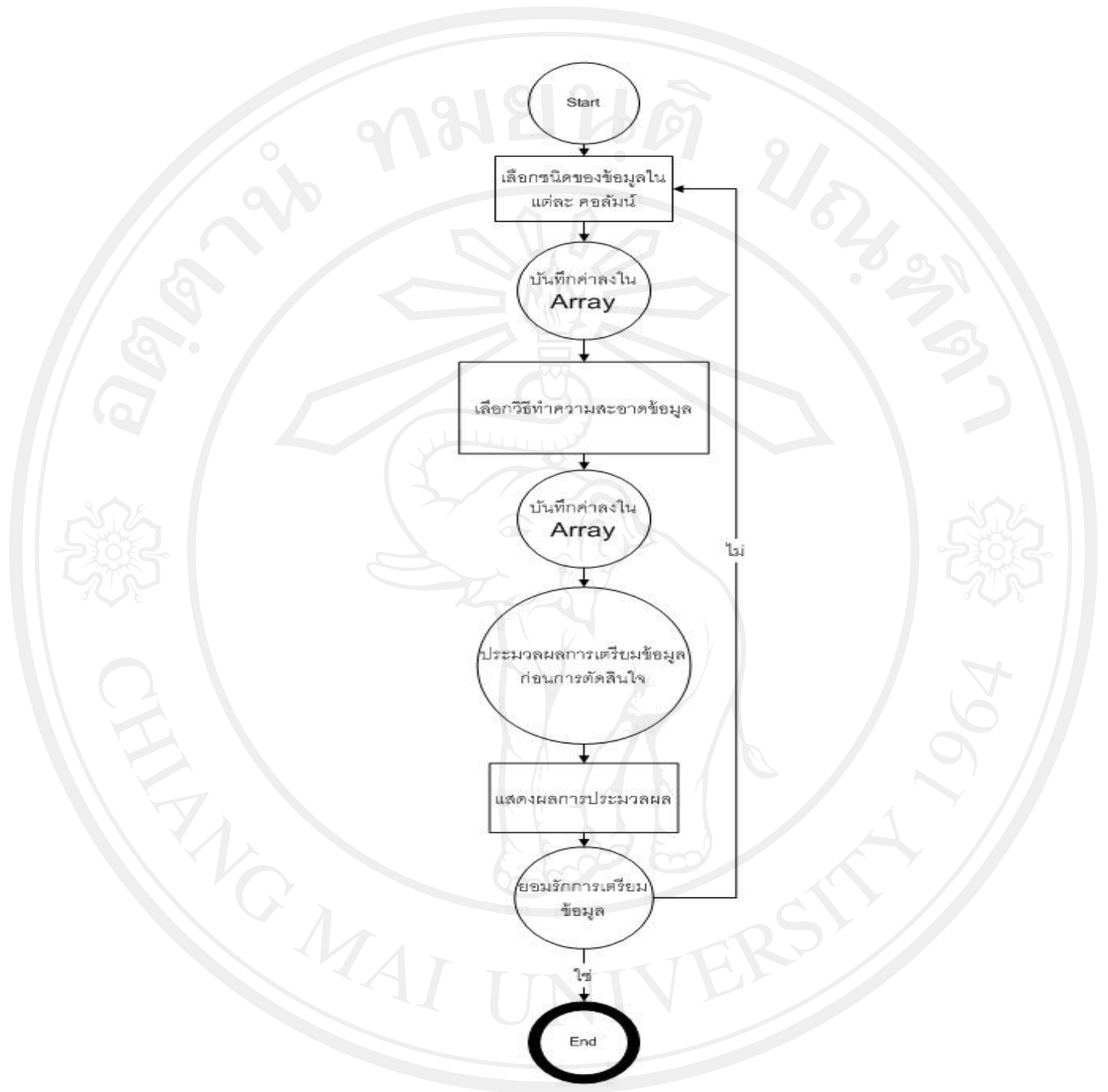
วิเคราะห์และออกแบบพฤติกรรมการใช้งานของระบบ โดยใช้แผนภาพยูเอ็มแอล (Unified Modeling Language : UML) เพื่อแสดงภาพรวมของการใช้งานรวมถึงขั้นตอนการดำเนินการของระบบที่มีความสัมพันธ์กับผู้ใช้



รูปที่ 2.1 แสดงถึงภาพรวมของระบบ (Context Diagram) [13]

จากรูปที่ 2.1 ที่ได้แสดงภาพรวมของระบบการสร้างต้นไม้ตัดสินใจที่มีประสิทธิภาพที่ประกอบด้วย

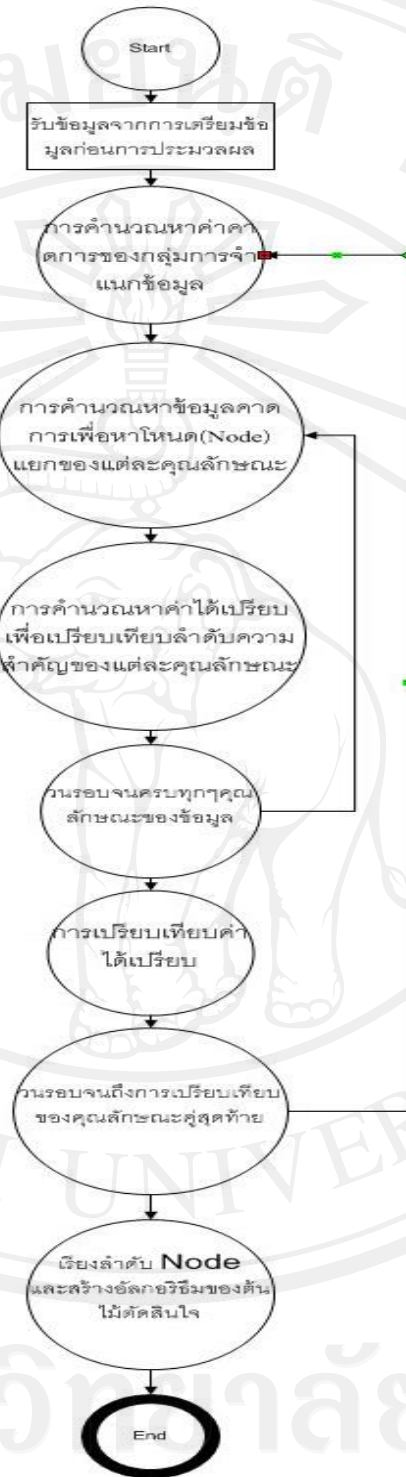
- ผู้ใช้งานระบบ หมายถึง ผู้ใช้งานระบบการสร้างต้นไม้ตัดสินใจที่มีประสิทธิภาพ
- ฐานข้อมูล หมายถึง ส่วนที่เก็บข้อมูลของระบบการสร้างต้นไม้ตัดสินใจที่มีประสิทธิภาพ
 - ส่วนการเตรียมข้อมูลก่อนการประมวลผล หมายถึง ส่วนของการวิเคราะห์ข้อมูลก่อนการประมวลผลเพื่อให้ข้อมูลมีความเหมาะสมก่อนการเข้าสู่กระบวนการประมวลผลสร้างต้นไม้ตัดสินใจ
 - การประมวลผลสร้างต้นไม้ตัดสินใจ หมายถึง ส่วนของการวิเคราะห์ข้อมูลการสร้างต้นไม้ตัดสินใจเพื่อหาการเชื่อมโยงของข้อมูลและมีการสร้างและทดสอบกลุ่มข้อมูลเพื่อหาร้อยละของความถูกต้องของอัลกอริทึมของต้นไม้ตัดสินใจ
 - การแสดงผลการสร้างต้นไม้ตัดสินใจและอัลกอริทึม หมายถึง ส่วนของการแสดงผลของต้นไม้ตัดสินใจที่ได้ประมวลผลเสร็จแล้ว



รูปที่ 2.2 แสดงถึงระบบย่อยการประมวลผลต้นไม้ตัดสินใจ (Data Flow Diagram) [13]

จากรูปที่ 2.2 ที่ได้แสดงถึงระบบย่อยการประมวลผลในการสร้างต้นไม้ตัดสินใจที่มีประสิทธิภาพมีลำดับในการทำงานดังนี้

- การคำนวณหาค่าคาดการณ์ของกลุ่มจำแนกข้อมูล
 การคำนวณหาค่าคาดการณ์ของกลุ่มการจำแนกข้อมูลโดยใช้สูตรการคำนวณจากสมการที่ 10 โดยอ้างอิงจากจำนวนคลาส (class) ของข้อมูลทั้งหมดซึ่งโดยมาตรฐานของชุดข้อมูลจะอยู่ในคุณลักษณะขวาคสุด
- การคำนวณหาข้อมูลคาดการณ์เพื่อหาโหนดแยก
 การคำนวณหาข้อมูลคาดการณ์เพื่อหาโหนดแยกของคุณลักษณะแต่ละโหนดย่อยโดยใช้สมการที่ 10 โดยอ้างอิงจากค่าในแต่ละคุณลักษณะของชุดข้อมูลทั้งหมด (ยกเว้นคลาส)
- การคำนวณหาค่าได้เปรียบเพื่อเปรียบเทียบความสำคัญ
 การคำนวณหาค่าได้เปรียบจะคำนวณโดยใช้ข้อมูลจากขั้นตอนข้างต้นด้วยสมการที่ 11 และเมื่อได้ค่าที่คำนวณแล้วจะมีการหาผลต่างของผลที่ได้จากการคำนวณค่าคาดการณ์ของกลุ่มจำแนกข้อมูลกับสมการที่ 11 โดยค่าดังกล่าวจะถูกนำไปใช้เปรียบเทียบเมื่อมีการวนรอบครบทุกคุณลักษณะ
- การวนรอบคำนวณหาค่าได้เปรียบเพื่อเปรียบเทียบความสำคัญ
 การวนรอบการคำนวณหาค่าได้เปรียบเพื่อเปรียบเทียบความสำคัญเพื่อใช้ในการเปรียบเทียบค่าที่มากที่สุดสำหรับใช้เป็นโหนด (node) ของต้นไม้ตัดสินใจในแต่ละระดับ
- การวนรอบจนถึงการเปรียบเทียบของคุณลักษณะสุดท้าย
 การวนรอบทำซ้ำของการคำนวณเพื่อหาลำดับความสำคัญของโหนดทั้งหมดของต้นไม้ตัดสินใจที่มีความสัมพันธ์ต่อคุณลักษณะของแต่ละคุณลักษณะของชุดข้อมูล
- เรียงลำดับโหนดและสร้างอัลกอริทึมของต้นไม้ตัดสินใจภายใต้ความสัมพันธ์ของคุณลักษณะของชุดข้อมูลนั้นๆ

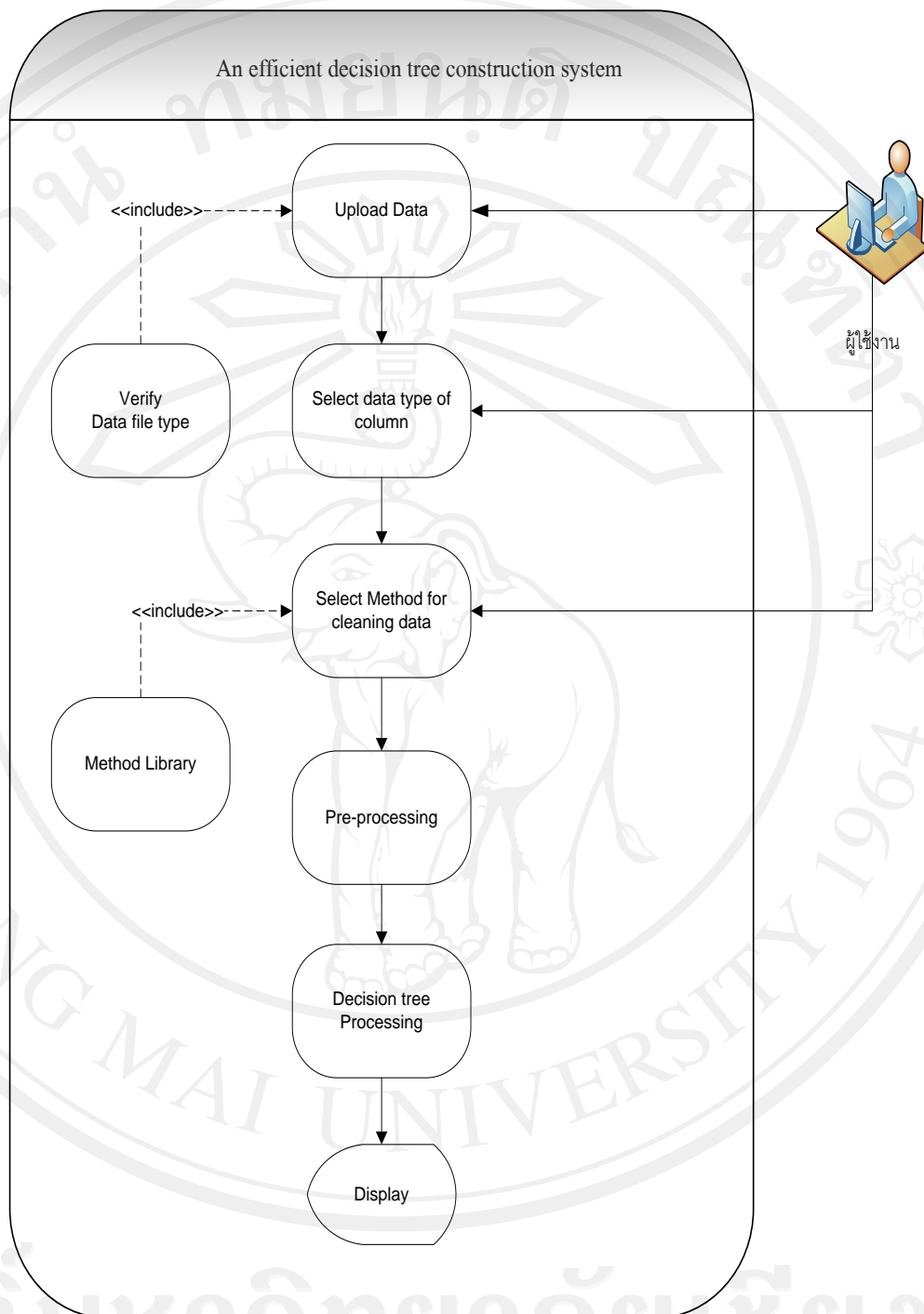


รูปที่ 2.3 แสดงถึงระบบย่อยการเตรียมข้อมูลก่อนการประมวลผล

(Data Flow Diagram) [13]

จากรูปที่ 2.3 ที่ได้แสดงถึงระบบย่อยการเตรียมข้อมูลก่อนการประมวลผลมีลำดับในการทำงานดังนี้

- การเลือกชนิดข้อมูลสำหรับแต่ละคุณลักษณะเลือกชนิดข้อมูลที่สอดคล้องกับค่าของข้อมูลในชุดข้อมูล
- เลือกวิธีการทำความสะอาดข้อมูลเลือกการทำความสะอาดข้อมูลให้เหมาะสมกับข้อมูลที่จะใช้ในกระบวนการสร้างต้นไม้ตัดสินใจ
- แสดงข้อมูลหลังจากการเตรียมข้อมูลแสดงข้อมูลที่ผ่านกระบวนการเตรียมข้อมูล



รูปที่ 2.4 แสดงถึงพฤติกรรมการทำงานของระบบเมื่อมีผู้ใช้งาน

(Use case Diagram) [13]

จากรูปที่ 2.4 ที่ได้แสดงถึงพฤติกรรมการทำงานของระบบเมื่อมีผู้ใช้งานระบบ มีลำดับในการทำงานดังนี้

- อัปโหลดข้อมูล (Upload Data)
 ผู้ใช้งานอัปโหลด (Upload) ไฟล์ (File) ข้อมูลของชุดข้อมูลที่ใช้ในการสร้างต้นไม้มัดตัดสินใจโดยที่ชนิดของไฟล์เป็น .CSV (เป็นชนิดไฟล์ข้อมูลที่เก็บข้อมูลโดยการใช้อุปกรณ์หมายถึงลูกภาคันระหว่างแต่ละคุณลักษณะและมีการเว้นบรรทัดเพื่อลำดับข้อมูลในแต่ละแถว)
- เลือกชนิดของข้อมูลในแต่ละคุณลักษณะ (Select data type of column)
 ผู้ใช้งานจะต้องทำการเลือกชนิดของข้อมูลให้ตรงกับความเป็นจริงของแต่ละคุณลักษณะของชุดข้อมูล
- เลือกชนิดในการทำความสะอาดข้อมูลก่อนการประมวลผล (Select method for cleaning data and preprocessing)
 ผู้ใช้งานจะต้องทำการเลือกวิธีการทำความสะอาดข้อมูลในแต่ละคุณลักษณะก่อนที่จะเข้าสู่กระบวนการสร้างต้นไม้มัดตัดสินใจ
- ระบบประมวลผลการเตรียมข้อมูลก่อนการประมวลผลการสร้างต้นไม้มัดตัดสินใจ (Preprocessing)
 ระบบจะทำการประมวลผลการเตรียมข้อมูลตามที่ผู้ใช้งานได้เลือกสำหรับแต่ละคุณลักษณะ
- ระบบประมวลผลการสร้างต้นไม้มัดตัดสินใจ (Decision tree processing)
 ระบบทำการประมวลผลการสร้างต้นไม้มัดตัดสินใจตามข้อมูลที่ได้ทำการเตรียมในขั้นตอนก่อนหน้า
- แสดงผล (Display)
 ระบบทำการแสดงผลของการสร้างต้นไม้มัดตัดสินใจ โดยมีการแสดงผลตารางข้อมูลที่ถูกนำเข้าสู่ระบบและผ่านการประมวลผลก่อนการตัดสินใจเรียบร้อยแล้ว ชุดอัลกอริทึมของต้นไม้มัดตัดสินใจ 4 ชุดโดยมีการทดสอบโดยการทำ Cross validations และผลการทดสอบเปอร์เซ็นต์ความถูกต้องเมื่อเทียบกับข้อมูลจริง

- **การออกแบบฐานข้อมูลของระบบ**

การออกแบบฐานข้อมูลเพื่อเก็บข้อมูลรวมไปถึงความสัมพันธ์ของข้อมูลแต่ละส่วนและการสืบค้นข้อมูลด้วยภาษา SQL (Standard Query Language) จากฐานข้อมูลให้มีประสิทธิภาพและการผิดพลาดน้อยที่สุด

- **การออกแบบส่วนตอบโต้กับผู้ใช้งาน**

การออกแบบโครงสร้างของข้อมูล (Information Architecture) ส่วนต่อประสานกับผู้ใช้งาน และวิเคราะห์ความเสี่ยงของการใช้งานที่ส่งผลต่อความผิดพลาดของระบบจากผู้ใช้งานเอง (เช่น กรอกข้อมูลผิดชนิด กรอกข้อมูลไม่ตรงกับที่ระบบต้องการ เป็นต้น)

- **การออกแบบระบบ**

การออกแบบคุณลักษณะของระบบ (Feature) ย่อยๆ เมื่อมีการรวมผสาน (Integration) ของคุณลักษณะแล้วสามารถทำงานได้เป็นระบบ

- **การออกแบบกรณีการทดสอบระบบ**

การออกแบบกรณีการทดสอบของระบบเพื่อทดสอบประสิทธิภาพการทำงานของระบบโดยแบ่งเป็นสองกรณีดังนี้

- **กรณีการทดสอบระบบตามขอบเขตที่กำหนดไว้**

การทดสอบที่มีข้อมูลและขั้นตอนการทำงานที่อยู่ภายใต้ขอบเขตของระบบเพื่อนำไปสู่การปรับปรุงแก้ไขให้มีประสิทธิภาพครอบคลุมขอบเขตทั้งหมด

- **ทดสอบระบบนอกเหนือขอบเขตที่กำหนดไว้**

การทดสอบที่มีข้อมูลและขั้นตอนการทำงานที่นอกเหนือขอบเขตของระบบเพื่อระบุแนวโน้มการพัฒนาต่อในรุ่นถัดไป

2.1.4 การพัฒนาระบบ

พัฒนาระบบตามที่ได้ออกแบบภายใต้ระยะเวลาของการวางแผนที่ได้กำหนด

2.1.5 การทดสอบระบบและปรับปรุงระบบ

การทดสอบตามกรณีการทดสอบที่ได้ออกแบบไว้รวมไปถึงการปรับปรุงระบบเมื่อมีข้อผิดพลาดเพื่อให้ระบบมีประสิทธิภาพตามที่วางแผนไว้

2.1.6 การสรุปและประเมินประสิทธิภาพของระบบ

สรุปผลการทดสอบระบบรวมถึงการประเมินประสิทธิภาพของระบบ

2.2 ขอบเขตการทำวิจัย

2.2.1 ด้านระบบ

- ระบบจะสามารถปฏิบัติการบนโปรแกรม AppServ (Microsoft Windows 7)
 - ภาษาที่ใช้ในการพัฒนาระบบ
 - PHP5
 - HTML5
 - CSS3
 - Sql
 - JavaScript

2.2.2 ด้านข้อมูล

ข้อมูลนำเข้าจากการอัปโหลด (Upload) โดยใช้ไฟล์นามสกุล .CSV ซึ่งข้อมูลภายในข้อมูลต้องมีจำนวนคอลัมน์ไม่เกิน 70 คอลัมน์และจำนวนบรรทัดไม่เกิน 500 บรรทัด อีกทั้งชนิดของข้อมูลในคอลัมน์เดียวกันต้องเป็นข้อมูลที่มีชนิดเดียวกัน เช่น ตัวอักษร ตัวเลข เป็นต้น ทั้งนี้เพื่อให้ระบบที่ทำงานบน Internet มีประสิทธิภาพดี เมื่อมีการเรียกใช้งานของผู้ใช้งานหลายคนในเวลา

เดียวกัน ข้อมูลภายในจะต้องเป็นในรูปแบบ ข้อความ (Text) ข้อมูลต่อเนื่อง (Continuous) และ ข้อมูลที่ไม่ต่อเนื่อง (Discrete) เท่านั้น ซึ่งความถูกต้องของการประมวลผลนั้นจะขึ้นอยู่กับ การเลือกใช้การเตรียมข้อมูลก่อนประมวลผลที่เหมาะสมจึงจะทำให้ผลที่ออกมามีความผิดพลาดน้อย ส่วนการแสดงผลของระบบจะเป็นการนำเสนอข้อมูลออกมาในรูปแบบรูปร่างต้นไม้มัดตัดสินใจและ ตารางที่วิเคราะห์ข้อมูลทดสอบ อีกทั้งยังแสดงเปอร์เซ็นต์ความผิดพลาดของต้นไม้ตัดสินใจและ อัลกอริทึมเพื่อใช้ในการประยุกต์ใช้ของผู้ใช้งาน

2.3 วิธีการทำวิจัย

2.3.1 วิเคราะห์และออกแบบระบบต้นแบบ

วิเคราะห์และออกแบบระบบต้นแบบโดยมีคุณลักษณะตามวัตถุประสงค์ มีการจัดการรุ่น (Sub Version :SVN) ของระบบต้นแบบและเอกสารที่ประกอบการพัฒนา

2.3.2 สร้างระบบต้นแบบและทดลองการใช้งานของระบบต้นแบบ

การสร้างระบบต้นแบบตามที่ได้ออกแบบไว้และทดลองการใช้งานระบบต้นแบบเพื่อหา ข้อผิดพลาดของระบบและเพิ่มคุณลักษณะที่ส่งผลต่อประสิทธิภาพของระบบตามขอบเขตที่กำหนด

2.3.3 ปรับปรุงระบบต้นแบบ

ปรับปรุงระบบต้นแบบให้มีประสิทธิภาพตามที่ได้ออกแบบไว้ตามแผนการ เพื่อให้ระบบ ต้นแบบมีประสิทธิภาพตรงตามเป้าหมาย

2.3.4 การพัฒนาระบบต้นแบบที่นำไปสู่ระบบที่ใช้งานจริง

พัฒนาระบบต้นแบบให้สามารถใช้งานได้จริงและระบบต้นแบบต้องมีการทำงานตามวัตถุประสงค์ตามแผนที่วางไว้

2.3.5 ทดสอบระบบตามกรณีการทดสอบที่ได้กำหนดไว้

ทดสอบระบบตามกรณีการทดสอบที่ได้กำหนดไว้เพื่อหาข้อผิดพลาดของระบบนำไปสู่การปรับปรุงแก้ไข

2.3.6 การติดตั้งและทดสอบการใช้งานจริง

การติดตั้งระบบ พร้อมทั้งมีการทดสอบการใช้งานภายใต้สถานการณ์และสิ่งแวดล้อมจริงโดยข้อมูลที่มีรูปแบบอยู่ในขอบเขตตามแผนที่วางไว้

2.3.7 ตรวจสอบและปรับปรุงระบบ

ตรวจสอบข้อผิดพลาดและปรับปรุงระบบให้มีประสิทธิภาพ รวมไปถึงการเปลี่ยนแปลงส่วนต่อประสานกับผู้ใช้งานให้เหมาะสม

2.3.8 การประเมินคุณภาพของระบบ

ประเมินคุณภาพของระบบช่วยวิเคราะห์และสร้างเงื่อนไขของกลุ่มข้อมูลแบบหลากหลายในรูปแบบต้นแบบไม่ตัดสินใจ