

การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลไวน์ด้วยเทคนิคเหมืองข้อมูล

จิรายุ สิริชัย¹, ไกรุ่ง เสงพระพรหม¹ และสุพจน์ เสงพระพรหม¹

¹สาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

*kairung2011.heng@gmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในการทำนายหาชนิดไวน์ โดยใช้เทคนิคเหมืองข้อมูล 2 วิธี คือ วิธีต้นไม้ตัดสินใจ และวิธีแบบเบย์ นำมาเปรียบเทียบประสิทธิภาพของโมเดลการจำแนกหาตัวแบบที่เหมาะสมเพื่อใช้ทำนายหาชนิดของไวน์ โดยใช้ข้อมูล wine ของเว็บไซต์ UCI วิเคราะห์ข้อมูลบนพื้นฐานของวิธี 10-Fold Cross Validation โดยใช้โปรแกรม R studio ในการสร้างแบบจำลอง ผลการศึกษาพบว่าวิธีต้นไม้ตัดสินใจมีประสิทธิภาพสูงสุด ให้ค่าความถูกต้องเท่ากับ 98.21% รองลงมาคือวิธีแบบเบย์ที่ให้ค่าความถูกต้องเท่ากับ 97.54% จากผลการเปรียบเทียบประสิทธิภาพครั้งนี้ สามารถนำวิธีต้นไม้ตัดสินใจไปใช้ในการทำนายหาชนิดของไวน์

คำสำคัญ: เหมืองข้อมูล วิธีต้นไม้ตัดสินใจ วิธีแบบเบย์ การจำแนกข้อมูล ประสิทธิภาพของโมเดล

A Comparison of Wine Quality Classification Performance by using Data Mining Techniques

Jirayu Sitichai¹, Kairung Hengprapohm¹ and Supojn Hengprapohm¹

¹Data Science, Faculty of Science and Technology, Nakhon Pathom Rajabhat University

Kairung2011.heng@gmail.com

Abstract

The purpose of this research is to compare the efficacy of a model for predicting wine using two data mining techniques, namely the decision tree method and the Bayesian method. The efficiency of the proper classification model for wine prediction was compared using the wine data of the UCI website. Data were analyzed based on the 10-fold cross-validation method using R studio for modeling.

The results showed that the decision tree method was most effective. The accuracy was 98.21%, followed by the Bayesian method with an accuracy of 97.54% from the results of this comparison.

Keywords: data mining, decision tree, naïve bay, classification, accuracy

1. บทนำ

ไวน์ เป็นเครื่องดื่มแอลกอฮอล์ที่ผลิตมาจากองุ่น ต้นองุ่นเป็นพืชพันธุ์ไม้เลื้อยที่สังกัดอยู่ในตระกูล วิตีส (Vitis) ซึ่งถือเป็นพืชไร่ที่ใช้เพาะปลูกที่เก่าแก่ที่สุดในโลก มนุษย์รู้จักการนำผลองุ่นมาผลิตเป็นไวน์ยุคต้นๆ ของประวัติศาสตร์ และยุคต้นๆ ที่มนุษย์เริ่มมีอารยธรรม ไวน์เริ่มมีบทบาทในกิจกรรมต่างๆ ของมนุษย์อย่างกว้างขวาง ทั้งในการฉลองพิธีกรรมทางศาสนาและในเทศกาลรื่นเริงต่างๆ โดยถูกกล่าวอ้างว่านำไวน์เข้ามาเพื่อเสริมสร้างความมั่นใจ เพื่อชำระบาปและเพื่อส่งเสริมสุขภาพ ไวน์หรือเหล้าองุ่นซึ่งผลิตจากผลองุ่นหรือผลไม้มาบดคั้นจนได้น้ำ จากนั้นจึงนำมาหมักให้เกิดปฏิกิริยาและแปรสภาพเป็นแอลกอฮอล์ อีกทั้งยังเป็นเครื่องดื่มที่ได้รับความนิยมและมีการดื่มกันมาเป็นระยะเวลาอันยาวนานนับร้อยปีแล้ว ไวน์นิยมดื่มควบคู่กับอาหารโดยเฉพาะประเทศในแถบยุโรป ซึ่งปัจจุบันคนไทยนิยมดื่มไวน์มากขึ้น ทั้งนี้เรื่องราวพฤติกรรมการดื่มไวน์ของแต่ละบุคคลมีความแตกต่างกัน ซึ่งเกี่ยวเนื่องจากวัฒนธรรมแต่ละเชื้อชาติ อาหารการกิน การดื่มเพื่อให้เข้ากับอาหาร ภูมิอากาศ และจากการเข้าสังคม หรือเป็นความชอบในแต่ละบุคคล แต่อย่างไรก็ตามหากดื่มไวน์โดยไม่ศึกษาคุณลักษณะของไวน์แต่ละชนิดที่ดื่มก็คงไม่แตกต่างเหมือนการดื่มเครื่องดื่มอื่นๆ โดยที่ไวน์จะมีชนิดรสชาติและประเภทแตกต่างกันออกไปเพื่อที่จะตอบสนองต่อวัตถุประสงค์ในการดื่มที่แตกต่างกัน ซึ่งในกรณีนี้จะแตกต่างจากเครื่องดื่มอื่นๆ เช่น น้ำอัดลมหรือน้ำผลไม้ทั่วไป

ปัจจุบันกลุ่มผู้บริโภคไวน์ขยายเพิ่มมากขึ้น ทำให้มีหลายประเทศได้ผลิตและจำหน่ายไวน์กันมากขึ้น โดยเฉพาะประเทศไทย ประกอบกับมีรายงานการวิจัยทางการแพทย์จากต่างประเทศ ว่าการดื่มไวน์ในปริมาณที่เหมาะสมจะช่วยเรื่องระบบสูดโลหิตของร่างกาย ทั้งยังช่วยลดอัตราการเป็นโรคหลอดเลือดและหัวใจ ส่งผลให้มีการบริโภคเพิ่มมากขึ้น

การนำเทคโนโลยีเหมืองข้อมูล (Data Mining) มาช่วยในการทำนายเพื่อการวางแผนให้ตรงกับความต้องการใช้งาน เป็นการนำความรู้จากข้อมูลที่มีอยู่มาใช้ในการทำนายข้อมูลใหม่ที่จะเกิดขึ้นในอนาคต ซึ่งการทำเหมืองข้อมูลคือการกระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำมาประยุกต์ใช้งานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

ในการวิจัยครั้งนี้ ได้นำข้อมูลของไวน์มาวิเคราะห์เพื่อพัฒนาหาโมเดลที่สามารถทำนายหาค่าความแม่นยำที่จะได้ไวน์ต่างสายพันธุ์แต่ปลูกในภูมิภาคเดียวกัน

2. วัตถุประสงค์การวิจัย

- 2.1 เพื่อศึกษาวิธีการจำแนกข้อมูลด้วยเทคนิคเหมืองข้อมูล
- 2.2 เพื่อเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจและเทคนิคนาอิวเบย์

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 ทฤษฎีที่เกี่ยวข้อง

3.1.1 เหมืองข้อมูล

เหมืองข้อมูล (Data Mining) คือการค้นหาหรือการสกัดความรู้จากฐานข้อมูลขนาดใหญ่ กระบวนการที่กระทำกับข้อมูลจำนวนมาก เพื่อค้นหารูปแบบแนวทางและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักการทางคณิตศาสตร์ สถิติ เพื่อนำความรู้ที่ได้นั้นมาใช้ในการแก้ปัญหา วางแผน หรือการดำเนินกลยุทธ์ขององค์กรให้ประสบความสำเร็จสูงสุด ซึ่งการวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูลสามารถแบ่งได้เป็น 2 ประเภทหลักๆ คือ

1. เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) จะเน้นที่การพิจารณาข้อมูลเป็นหลัก เช่น พิจารณาว่าข้อมูลมีความสัมพันธ์กันในลักษณะใดบ้าง เทคนิคในประเภทนี้จะแบ่งย่อยได้อีกคือ เทคนิคการค้นหากฎความสัมพันธ์ (Association Rule) และการแบ่งกลุ่มข้อมูล (Clustering)

2. เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) เน้นการเรียนรู้จากข้อมูลที่มีอยู่ในอดีตเพื่อนำมาสร้างโมเดลสำหรับทำนายหรือคาดการณ์สิ่งที่จะเกิดขึ้นในอนาคต สามารถแบ่งย่อยได้อีกคือ การจำแนกประเภทข้อมูล (Classification) และการประมาณค่าข้อมูล (Regression) ซึ่งทั้งสองเทคนิคจะมีความแตกต่างกันที่คำตอบที่ต้องการทำนายซึ่งการจำแนกประเภทข้อมูล (Classification) จะทำนายข้อมูลที่มีค่าเป็น Nominal เช่น เพศชาย เพศหญิง หรือค่าที่ไม่ใช่ตัวเลขนั่นเอง ส่วนการประมาณค่าข้อมูล (Regression) จะใช้กับข้อมูลคำตอบที่เป็นตัวเลขเท่านั้น

ในการวิจัยครั้งนี้ ได้ใช้เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) แบบการจำแนกประเภทข้อมูล (Classification) โดยใช้อัลกอริทึมเพื่อใช้ทำนายแนวโน้มให้แก่ข้อมูลที่ใช้ทำนาย

3.1.2 วิธีต้นไม้ตัดสินใจ (Decision Tree)

เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่กำหนดไว้ล่วงหน้า และพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วยรูปแบบของ Tree

3.1.3 วิธีแบบเบย์ (Naïve Bayes)

คือการทำเหมืองข้อมูลที่ถูกสร้างขึ้นโดยหลักความน่าจะเป็น ซึ่งจะใช้การวิเคราะห์หาความน่าจะเป็นของสิ่งที่ยังไม่เคยเกิดขึ้นด้วยการคาดเดาจากสิ่งที่เคยเกิดขึ้นมาก่อน โดยใช้ทฤษฎีของ Bayes ในการแก้ปัญหา

4. งานวิจัยที่เกี่ยวข้อง

ปีสุดา ดาวเรือง และคณะ (2561) การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลเพื่อทำนายแขนงวิชาเรียนของนักศึกษา เป็นงานวิจัยเกี่ยวกับปัญหาในศึกษาออกกลางคันในเวลาเรียน โดยใช้เทคนิคเหมืองข้อมูล 3 วิธีคือ ต้นไม้ตัดสินใจ วิธีแบบเบย์ และวิธีฐานกฎ นำมาเปรียบเทียบประสิทธิภาพของโมเดลการจำแนกหาตัวแบบที่เหมาะสมเพื่อใช้ทำนายแขนงวิชาเรียนที่เหมาะสมกับนักศึกษา ผลการเปรียบเทียบประสิทธิภาพ พบว่าการจำแนกประเภทแบบวิธีฐานกฎให้ค่าความถูกต้องที่สุด (78.66) รองลงมาคือวิธีต้นไม้ตัดสินใจ (78.26) และวิธีแบบเบย์ (78.12)

ทิพย์ธิดา วงศ์พิพันธ์ (2556) การใช้เหมืองข้อมูลช่วยในการตัดสินใจการใช้สินเชื่อ เป็นงานวิจัยเกี่ยวกับให้บริการรถเช่าระยะยาว สัญญาเช่าระยะเวลา 3-5 ปี และระยะสั้น สัญญาเช่าระยะเวลาเป็นรายวัน รายสัปดาห์ และรายเดือน โดยใช้เทคนิคเหมืองข้อมูล 3 วิธีคือ ต้นไม้ตัดสินใจ วิธีแบบเบย์ เพื่อการวิเคราะห์ข้อมูลลูกค้าที่มา

เช่ารถ เพื่อช่วยการตัดสินใจพิจารณาข้อมูลสินเชื่อของลูกค้าให้มีประสิทธิภาพมากยิ่งขึ้น และลดความเสี่ยงด้านหนี้สูญในการรออนุมัติให้เช่ารถยนต์ต่อลูกค้าที่ไม่เหมาะสม ผลการเปรียบเทียบประสิทธิภาพ พบว่าการจำแนกวิธีต้นไม้ตัดสินใจให้ผลลัพธ์ที่ดีที่สุด มีค่าความถูกต้อง (90.47)

กมลลาสน อุดมล้ำเลิศ และอานนท รุ่งสว่าง (2553) การคัดแยกหมวดหมู่เว็บเพจภาษาไทยอัตโนมัติ เป็นงานวิจัยเกี่ยวกับระบบแบ่งเว็บเพจออกเป็นหมวดหมู่ตามเนื้อหาที่ปรากฏ หรือที่เรียกว่า “เว็บไดเรกทอรี” (web directory) จากการทดลองกับข้อมูลทดสอบ 2 ชุด ซึ่งคัดเลือกมาจากเว็บไดเรกทอรีภาษาไทย 5 หมวดหมู่หลัก ทำให้เห็นว่าการลดคุณลักษณะโดยวิธีการคัดเลือกคุณลักษณะที่เหมาะสมจะไม่ส่งผลต่อความถูกต้องในการจำแนกหมวดหมู่ และการลดคุณลักษณะ ประกอบกับการใช้ตัวคัดแยก Naïve Bayes จะให้ค่าประสิทธิภาพผลดีที่สุด โดยมีค่าความถูกต้อง (83.4)

รัชพล กลัดชื่น และ จริญญา แสนราช (2561) ได้เปรียบเทียบประสิทธิภาพของอัลกอริทึมในการทำนายและคุณลักษณะที่มีต่อผลสัมฤทธิ์ทางการเรียนของนักศึกษาระดับอาชีวศึกษา โดยทำการศึกษาข้อมูลนักศึกษาระดับประกาศนียบัตรวิชาชีพ จำนวน 5,100 ระเบียบงาน ตั้งแต่ปีการศึกษา 2550 - 2559 9 สาขาวิชา 27 คุณลักษณะโดยใช้เทคนิคการจำแนกข้อมูล 3 เทคนิค ได้แก่ Decision Tree : J4 8graft, Naïve Bayes และ Rule Induction พบว่าการใช้เทคนิค Decision Tree : J48graft ด้วยการเลือกคุณลักษณะแบบ Forward Selection และ การเลือกคุณลักษณะทั้งหมด มีค่าความถูกต้องเท่ากับ 83.08% และ 81.71% ตามลำดับ และทดสอบด้วยวิธี T-Test พบว่าการทดสอบทั้งสองแบบมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากผลการเปรียบเทียบประสิทธิภาพในครั้งนี้ สามารถนำเทคนิค Decision Tree : J48graft ไปใช้ในการทำนายผลสัมฤทธิ์ทางการเรียนและเป็นแนวทางในการสอนเสริมหรือแนะแนวให้กับนักศึกษาต่อไป

สุรวีธร ศรีเปารยะ และสายชล สนิสมบูรณ์ทอง (2560) ได้ทำการเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง ซึ่งเป็นงานวิจัยเกี่ยวกับการเป็นโรคไตเรื้อรัง งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการจำแนกกลุ่ม โดยเลือกใช้วิธีความใกล้เคียงกันมากที่สุด วิธีต้นไม้ตัดสินใจ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีฐานกฎ และวิธีแบบเบย์ ผลที่ได้คือวิธีการจำแนกกลุ่มที่มีประสิทธิภาพการจำแนกที่ดีที่สุดคือ วิธีต้นไม้ตัดสินใจ ซึ่งให้ค่าความถูกต้อง คือ 100% และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0059

5. วิธีดำเนินการวิจัย

CRISP-DM เป็นกระบวนการในการวิเคราะห์ที่ใช้กันอย่างแพร่หลายซึ่งประกอบด้วย 6 ขั้นตอนได้แก่

1. Business Understanding เป็นขั้นตอนการทำความเข้าใจโจทย์โดยการวิเคราะห์ข้อมูล

ไวน์ (Wine) เป็นเครื่องดื่มแอลกอฮอล์ทำจากองุ่นหรือผลไม้อื่นหมัก สมดุลเคมีธรรมชาติขององุ่นทำให้มันหมักโดยไม่ต้องเพิ่มน้ำตาล กรด เอ็นไซม์ น้ำหรือสารอาหารอื่น ยีสต์บริโภคน้ำตาลในองุ่นแล้วเปลี่ยนเป็นเอทานอลและคาร์บอนไดออกไซด์ พันธุ์ขององุ่นและสายพันธุ์ของยีสต์ที่ต่างกันทำให้ได้ไวน์คนละแบบ แบบที่รู้จักกันดีเกิดจากอันตรกิริยาที่ซับซ้อนยิ่งระหว่างการเจริญทางชีวเคมีของผลไม้ ปฏิกริยาที่เกี่ยวข้องในการหมัก แหล่งที่ปลูก (terrior) และการระบุแหล่ง (appellation)

2. Data Understanding เป็นขั้นตอนการดูว่ามีข้อมูลอะไรบ้างที่สามารถนำมาใช้ในการวิเคราะห์ของโปรเจกต์นี้ได้บ้างจากการศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้องกับงานวิจัย และการวิเคราะห์ข้อมูล โดยวิเคราะห์จากข้อมูลที่เลือกมา ดังนี้

- 1) Alcohol
- 2) Malic acid
- 3) Ash
- 4) Alcalinity of ash
- 5) Magnesium
- 6) Total phenols
- 7) Flavanoids
- 8) Nonflavanoid phenols
- 9) Proanthocyanins
- 10) Color intensity
- 11) Hue
- 12) OD280/OD315 of diluted wines
- 13) Proline

3. Data Preparation เป็นขั้นตอนของการเตรียมข้อมูลโดยอาจจะเป็นการเชื่อมโยงข้อมูล (join) สำหรับการวิเคราะห์ข้อมูลในงานวิจัยครั้งนี้ด้วยโปรแกรม R studio

4. Modeling เป็นขั้นตอนการวิเคราะห์ข้อมูลโดยใช้เทคนิคต่างๆ ของ Data Mining หรือ Machine Learning

การจำแนกประเภทข้อมูล เป็นขั้นตอนการทดสอบการจำแนกประเภทข้อมูล หลังจากทำการเลือกคุณลักษณะของข้อมูลที่มีความเหมาะสมและสอดคล้องมากที่สุด โดยทำการเปรียบเทียบค่าเฉลี่ยความถูกต้องการจำแนกประเภทข้อมูลของแต่ละกลุ่มการทดสอบ โดยใช้เครื่องมือ R Studio และ Excel ซึ่งกลุ่มตัวอย่าง คือ <https://archive.ics.uci.edu/ml/datasets/Wine>

5. Evaluation เป็นขั้นตอนการวัดผลลัพธ์ที่ได้จากการวิเคราะห์ว่าประสิทธิภาพของโมเดลที่ได้เป็นอย่างไร

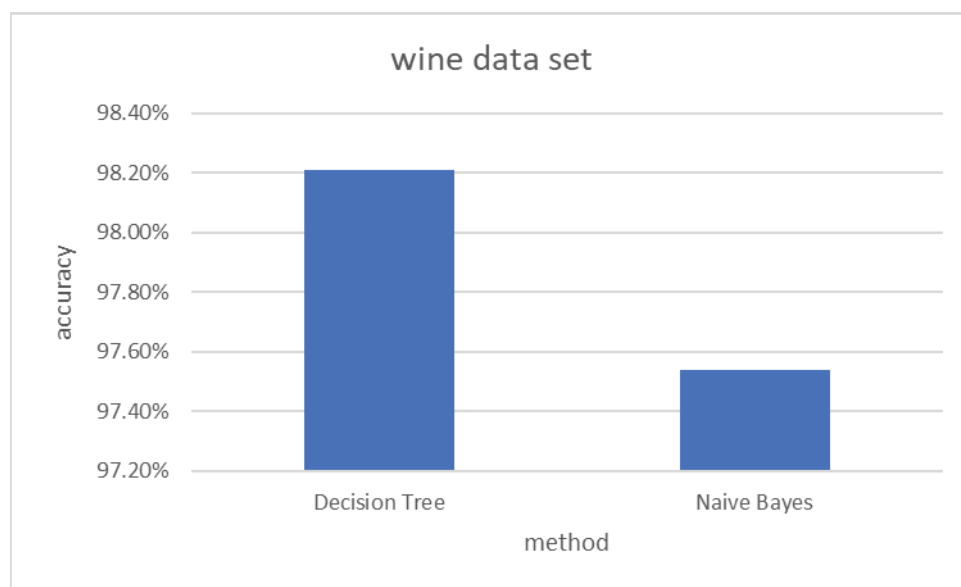
6. Deployment เป็นขั้นตอนการนำผลลัพธ์การวิเคราะห์ที่ได้ไปใช้งานต่อไป เช่น เสนอข้อมูลให้กับฝ่ายการตลาดเพื่อทำการติดต่อต่อไป

6. ผลการวิจัย

ผู้วิจัยได้เปรียบเทียบประสิทธิภาพเทคนิคการเลือกคุณลักษณะที่เหมาะสมสำหรับการจำแนกประเภทข้อมูลผลการเปรียบเทียบการคัดเลือกคุณลักษณะในการคัดเลือกคุณลักษณะด้วยเทคนิค 2 แบบ คือ Decision Tree, Naïve Bayes ดังนี้

ตารางที่ 1 การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดย เทคนิค 2 แบบ คือ Decision Tree, Naïve Bayes จะแสดงในรูปแบบตาราง

Algorithms	Accuracy (%)
Decision Tree	98.21%
Naive Bayes	97.54%



ภาพที่ 1 การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดย เทคนิค 2 แบบ คือ Decision Tree, Naïve Bayes จะแสดงในรูปแบบกราฟแท่ง

7. สรุป อภิปรายผลการวิจัย

ผลการวิจัย ครั้งนี้โดยพัฒนาและเปรียบเทียบตัวแบบการจำแนกทั้ง 2 เทคนิค ได้แก่ Decision Tree, Naïve Bayes ซึ่งผลการประเมินประสิทธิภาพตัวแบบ คือ Decision Tree ซึ่งได้ค่าที่สูงที่สุดจากข้อมูล ค่าความถูกต้อง(Accuracy) ได้ 98.21 % จึงสรุปได้ว่า Decision Tree เป็นตัวแบบที่เหมาะสมที่สุดที่จะนำมาใช้จำแนกข้อมูล ผลการทดลองใช้โปรแกรม R Studio พบว่า Decision Tree มีความแม่นยำมากที่สุด

8. ข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยที่ใช้โปรแกรม R Studio เพื่อนำมาประยุกต์ใช้ในการจำแนกข้อมูล โดยที่มีเครื่องมือต่างๆให้เลือกใช้ อาทิเช่น Naïve Bayes, Decision Tree โดยสามารถนำเครื่องมือเหล่านี้มาเปรียบเทียบการจำแนกเพื่อหาค่าความถูกต้องมากที่สุด

9. กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลงได้ด้วยดี เนื่องจากได้รับความกรุณาอย่างสูงจาก ผศ.ดร.ไกรรุ่ง เสงพระพรหม อาจารย์ที่ปรึกษางานวิจัย ที่กรุณาให้คำแนะนำปรึกษา ตลอดจนปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่เป็นอย่างดี ผู้วิจัยตระหนักถึง ความตั้งใจจริงและความทุ่มเทของอาจารย์และขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้ อนึ่ง ผู้วิจัยหวังว่า งานวิจัยฉบับนี้จะมีประโยชน์อยู่ไม่น้อย จึงขอมอบส่วนดี ทั้งหมดนี้ให้แก่เหล่าคุณอาจารย์ที่ได้ประสิทธิประสาทวิชาจนทำให้ผลงานวิจัยเป็น ประโยชน์ต่อผู้ที่เกี่ยวข้องและขอมอบความกตัญญูทเวทิตาคุณ แต่บิดา มารดา และผู้มี พระคุณทุกท่าน สำหรับข้อบกพร่องต่าง ๆ ที่อาจจะเกิดขึ้นนั้น ผู้วิจัยขอน้อมรับผิดเพียง ผู้เดียว และยินดีที่จะรับฟังคำแนะนำจากทุกท่านที่ได้เข้ามาศึกษา เพื่อเป็นประโยชน์ใน การพัฒนางานวิจัยต่อไป

10. เอกสารอ้างอิง

- ปัสุดา ดาวเรือง จรรย์ แสนราช และ อนิราช มิ่งขวัญ.(2561). **การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลเพื่อทำนายแขนงวิชาเรียนของนักศึกษา** ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ วิทยาเขตปราจีนบุรี
- ทิพย์ธิดา วงศ์พิพันธ์ (2556) **การใช้เหมืองข้อมูลช่วยในการตัดสินใจการให้สินเชื่อ** สาขาวิชาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิต
- กมลასน์ อุดมล้ำเลิศ และ อานนท์ รุ่งสว่าง. (2553). **เรื่องเต็มการประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 48: สาขาสถาปัตยกรรมศาสตร์และวิศวกรรมศาสตร์. การคัดแยกหมวดหมู่เว็บเพจภาษาไทยอัตโนมัติด้วยนาอิวเบย์.** (น. 527 หน้า). กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์;สำนักงานคณะกรรมการอุดมศึกษา;กระทรวงศึกษาธิการ;กระทรวงเกษตรและสหกรณ์;กระทรวงวิทยาศาสตร์และเทคโนโลยี;กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม;กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร;สำนักงานคณะกรรมการวิจัยแห่งชาติ;สำนักงานกองทุนสนับสนุนการวิจัย.
- รัชพล กลัดชื่น และ จรรย์ แสนราช. (2561). **การเปรียบเทียบประสิทธิภาพอัลกอริทึมและการคัดเลือกคุณลักษณะที่เหมาะสม เพื่อการทำนายผลสัมฤทธิ์ทางการเรียนของนักศึกษาระดับอาชีวศึกษา.** *Research Journal Rajamangala University of Technology Thanyaburi*, Vol.18, Issue 1, P.1-10.
- สุรวุฒ ศรีเปารยะ และสายชล สนิสมบูรณ์ทอง. (2560). **การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง** ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง