การเปรียบเทียบประสิทธิภาพโมเดลการตัดสินใจเลือกกลุ่มวิชาของนักศึกษา สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

Efficiency Comparison of Decision Making Models on Major Selection of Information Technology Students, Faculty of Science, Buriram Rajabhat University

กมลรัตน์ สมใจ[1*]
Kamolrat Somchai[1*]

บทคัดย่อ

งานวิจัยนี้ได้จัดทำขึ้น โดยมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพโมเดลการตัดสินใจเลือกกลุ่มวิชาของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยใช้เหมืองข้อมูล 3 เทคนิควิธี คือ Decision Tree , Naïve Bayes และ Neural Network และใช้ข้อมูลนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ ระหว่างปีการศึกษา 2555-2560 จำนวนทั้งสิ้น 407 ระเบียนประกอบด้วย 16 คุณลักษณะมาสร้างตัวแบบด้วยโปรแกรม RapidMiner Studio 9 รวมทั้งแบ่งข้อมูลทดลองด้วยวิธี 10-Fold Cross Validation เพื่อวัดประสิทธิภาพเปรียบเทียบจากการวัดค่า Accuracy, Recall, Precision และ F-measure

ผลการเปรียบเทียบประสิทธิภาพพบว่า Decision Tree ให้ค่าความถูกต้องดีที่สุด จำแนกเป็นกลุ่มวิชาคือกลุ่มวิชาเทคโนโลยีสารสนเทศ มีค่าความถูกต้อง ร้อยละ 94.17 กลุ่มวิชาเทคโนโลยีคอมพิวเตอร์ มีค่าความถูกต้อง ร้อยละ 88.72 และกลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษามีค่าความถูกต้อง ร้อยละ 92.87 ซึ่งสามารถสรุปได้ว่าตัวแบบ Decision Tree มีประสิทธิภาพดีที่สุด

คำสำคัญ: ประสิทธิภาพโมเดล, การตัดสินใจเลือก, กลุ่มวิชาของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ

Abstract

The objective of this research was to compare the efficiency of decision making models on major selection of Information Technology students of Faculty of Science in Buriram Rajabhat University by using 3 data mining techniques: Decision Tree , Naïve Bayes, and Neural Network. The data derived from the data record with 16 characteristics of 407 students from 2012-2017 were used in developing the model prototype by using RapidMiner Studio 9. Some data were also used in the tryout using 10-Fold Cross Validation to find the efficiency by comparing from Accuracy, Recall, Precision and F-measure.

The findings from the comparison indicated that Decision Tree offered the most accuracy value. When classified by each major, it was found that Information Technology major had the accuracy value of 94.17% while Computer Technology major had 88.72% and Computer Management for Education major had 92.87%. It can be concluded that Decision Tree model had the most efficiency.

Keywords: model efficiency, decision making, majors of Information Technology students

[1]อาจารย์สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์
E-mail: Kamolratsomchai08@gmail.com

บทนำ

หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏบุรีรัมย์ ได้มีการจัดการเรียนการสอนแบ่งเป็นกลุ่มวิชาได้แก่ กลุ่มวิชาเทคโนโลยีสารสนเทศ กลุ่มวิชาเทคโนโลยีคอมพิวเตอร์ และกลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษา โดยการเลือกกลุ่มวิชาดังกล่าวผู้เรียนต้องเลือกเรียนตั้งแต่แรกเริ่มสมัครคัดเลือกเข้าเรียน และจากสัมภาษณ์ผู้เรียนพบว่าส่วนใหญ่ผู้เรียนจะเลือกเรียนตามเพื่อน หรือตามคำแนะนำของผู้ปกครอง มากกว่าจะพิจาณาที่ความสามารถของตนเอง จึงส่งผลให้เกิดปัญหาแก่ผู้เรียนต้องย้ายสาขาวิชาเรียน สอบเข้าเรียนในสาขาวิชาใหม่ รวมถึงการพ้นสภาพทางการเรียน เป็นต้น และเมื่อมีการปรับปรุงหลักสูตรของสาขาวิชาเทคโนโลยี ในปี พ.ศ. 2561 ที่มีการกำหนดให้ชั้นปีที่ 1-2 เป็นการเรียนพื้นฐานที่เปิดโอกาสให้ผู้เรียนได้ศึกษารายวิชาต่าง ๆ ที่จะช่วยให้ผู้เรียนได้ทราบความสามารถเฉพาะด้านของตนเอง และในชั้นปีที่ 3 ผู้เรียนสามารถเลือกเรียนในกลุ่มวิชาต่าง ๆ ได้ ดังนั้นเพื่อไม่ให้ผู้เรียนเกิดปัญหาในการเลือกกลุ่มวิชาวิธีการหนึ่งที่จะช่วยสนับสนุนการตัดสินใจเลือกกลุ่มวิชาของผู้เรียนตามความสามารถของผู้เรียนได้ก็คือ การนำเทคนิคการทำเหมืองข้อมูลมาจำแนกและการเลือกกลุ่มวิชาของผู้เรียน ผ่านการวิเคราะห์ข้อมูลจากคลังข้อมูลนั่นเอง

การทำเหมืองข้อมูลเพื่อการจำแนกและเลือกกลุ่มวิชา เป็นการนำเอาเทคนิคการวิเคราะห์ข้อมูลจากมุมมองที่แตกต่างกัน และสามารถสรุปเป็นข้อมูลที่มีประโยชน์ กระบวนการทำเหมืองข้อมูลใช้หลายหลักการเช่น เทคนิคการเรียนรู้ของเครื่อง (Machine Learning), สถิติ, เทคนิคการสร้างภาพ (Visualization Techniques) เพื่อค้นพบและนำเสนอความรู้ในรูปแบบที่เข้าใจได้ง่าย (J. Han and M. Kamber, 2006 อ้างอิง จิราภา เลาหะวรนันท์. 2561) โดยอีกความหมายหนึ่งคือกระบวนการที่กระทำกับข้อมูลจำนวนมาก เพื่อสกัดสารสนเทศ รวมถึงรูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลขนาดใหญ่นั้น

ด้วยเหตุนี้ผู้วิจัยได้เล็งเห็นความสำคัญของการใช้เทคนิคการทำเหมืองข้อมูลเพื่อช่วยการตัดสินใจเลือกกลุ่มวิชาของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ ซึ่งจะมีส่วนช่วยสนับสนุนการตัดสินใจเลือกลุ่มวิชาของนักศึกษาตามความเหมาะสมกับความสามารถ เพื่อให้ผู้เรียนเป็นบัณฑิตที่มีคุณภาพ ตรงตามวัตถุของการเรียนการสอนในหลักสูตรเทคโนโลยีสารสนเทศต่อไป

วัตถุประสงค์

เพื่อเปรียบเทียบประสิทธิภาพโมเดลการตัดสินใจเลือกกลุ่มวิชาของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1. เหมืองข้อมูล (Data Mining) การทำเหมืองข้อมูล คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหา

รูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักการทางคณิตศาสตร์ สถิติ การรู้จำ การเรียนรู้ทางเครื่องจักร เพื่อนำความรู้ที่ได้นั้นมาใช้ ในการแก้ปัญหา วางแผน หรือดำเนินกลยุทธ์ขององค์กรให้ประสบความสำเร็จสูงสุด ซึ่งการทำเหมืองข้อมูลสามารถแบ่งประเภทอัลกอริทึมในการทำเหมืองข้อมูลออกเป็น 2
ประเภท คือ

1.1 การสร้างแบบจำลองในการบรรยาย (Descriptive Modeling, Unsupervised Modeling) คือ เป็นการหาความสัมพันธ์ต่างๆ (Association) หรือการจัดกลุ่ม (Clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย

1.2 การสร้างแบบจำลองในการทำนาย (Predictive Modeling, Supervised Modeling) คือ ทุก ๆ ข้อมูลจะมีคุณสมบัติหนึ่งเรียกว่าฉลาก (Label) ซึ่งค่าของคุณสมบัตินี้ จะเป็นค่าที่ใช้ในการทำนายผลของข้อมูล อัลกอริทึมประเภทนี้จะมุ่งเน้นในการแบ่งแยกข้อมูลออกเป็นกลุ่มตามค่าคุณสมบัติของฉลาก ซึ่งถ้าค่าคุณสมบัติของฉลากมีค่าไม่ต่อเนื่องจะเรียกกระบวนการที่ใช้แบ่งแยกว่า การแยกแยะ (Classification) ถ้าค่าคุณสมบัติของฉลากมีค่าต่อเนื่องจะเรียกกระบวนการที่ใช้แบ่งแยกว่า การถดถอย (Regression)

2. เทคนิคเหมืองข้อมูล ในงานวิจัยนี้เลือกวิธีการสร้างแบบจำลองในการทำนาย (Predictive Modeling, Supervised Modeling) โดยใช้เทคนิคการจำแนกข้อมูล (Classification) เป็นกระบวนการจัดแบ่งข้อมูลตามลักษณะของวัตถุประสงค์นั้นๆ โดยใช้ 3 เทคนิควิธี ได้แก่ วิธีต้นไม้การตัดสินใจ (Decision Tree) วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayesian Learning) วิธีโครงข่ายประสาท (Neural Network) โดยมีรายละเอียด ดังนี้

2.1 วิธีต้นไม้การตัดสินใจ (Decision Tree)เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่กำหนดไว้ล่วงหน้า และพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วยรูปแบบของ Tree โครงสร้างประกอบด้วย Root Node, Child และ LeafNode

2.2 วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayes) คือการทำเหมืองข้อมูลที่ถูกสร้างขึ้นโดยหลัก ความน่าจะเป็น ซึ่งจะใช้วิเคราะห์หาความน่าจะเป็น ของสิ่งที่ยังไม่เคยเกิดขึ้นด้วยการคาดเดาจากสิ่งที่เคยเกิดขึ้นมาก่อนโดยใช้ทฤษฎีของ Bayes ในการ แก้ปัญหา

2.3 วิธีโครงข่ายประสาท (Neural Network) คือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณในลักษณะความสัมพันธ์เชิงเชื่อมโยงระหว่างส่วนหรือแบบ connections เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถ ในการเรียนรู้การจดจำแบบรูป และการทำนายอนาคต เช่นเดียวกับความสามารถที่มีในสมองมนุษย์

3. การทดสอบประสิทธิภาพของการจำแนกประเภทข้อมูล การทดสอบประสิทธิภาพของการจำแนกประเภทข้อมูลที่นิยมได้แก่

3.1 ค่าความแม่นยำ (precision) คือค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์

3.2 ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกกี่ตัว เป็นการวัดความถูกต้องของโมเดล

3.3 ค่าความถ่วงดุล (F-measure) คือค่าเฉลี่ย ของค่าความแม่นยำและค่าความระลึก

3.4 ค่าความถูกต้อง (Accuracy) คือจำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส

4. การแบ่งข้อมูลเพื่อทดสอบประสิทธิภาพของโมเดล มี 3 วิธีการใหญ่ ๆ ดังนี้

4.1 วิธี Self- Consistency Test เป็นวิธีการที่ง่ายที่สุดนั่นคือ ข้อมูลที่ใช้ในการสร้างโมเดล และข้อมูลที่ใช้ในการทดสอบโมเดลเป็นข้อมูลชุดเดียวกัน

4.2 วิธี Split Test เป็นการแบ่งข้อมูลด้วยการสุ่มออกเป็น 2 ส่วน เช่น 70% ต่อ 30% หรือ 80% ต่อ 20% โดยข้อมูลส่วนที่หนึ่ง (70% หรือ 80%) ใช้ในการสร้างโมเดล และข้อมูลส่วนที่สอง (30% หรือ 20%) ใช้ในการทดสอบประสิทธิภาพของโมเดล

4.3 วิธี Cross-Validation Test วิธีนี้เป็นวิธีที่นิยมในการทำงานวิจัย โดยจะทำการแบ่งข้อมูลออกเป็น 5 ส่วน โดยที่แต่ละส่วนจะมีจำนวนเท่ากัน หรือ 10-fold cross-validation คือการแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่ากัน หลังจากนั้นข้อมูลหนึ่งส่วนที่ใช้เป็นตัวทดสอบประสิทธิภาพของโมเดล ทำวนไปเช่นนี้จนครบจำนวนที่แบ่งไว้

งานวิจัยที่เกี่ยวข้อง

จิราภา เลาหะวรนันท์ (2558 : บทคัดย่อ) ได้ทำวิจัยเรื่องการใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ โดยใช้ข้อมูลผลการเรียนและผลการวัดความสามารถด้านต่าง ๆ ที่เกี่ยวข้องมาสร้างแบบจำลองพยากรณ์โดยเปรียบเทียบเทคนิคการทำเหมืองข้อมูล 5 เทคนิค และพยากรณ์ผ่านเทคนิค “Ensemble” พบว่าการพยากรณ์มีความแม่นยำ อยู่ที่ 72.92% โดยแขนงวิศวกรรมซอร์ฟแวร์สามารถทำนายได้แม่นยำถึง 86.67%

เสกสรรค์ วิลัยลักษณ์ (2558 : บทคัดย่อ) ได้ทำวิจัยเรื่องการใช้เทคนิคการทำเหมืองข้อมูลเพื่อพยากรณ์ผลการเรียนของนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา โดยมีวัตถุประสงค์เพื่อพัฒนาคลังข้อมูลและสร้างตัวแบบพยากรณ์ผลการเรียนของนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา ผลการวิจัยพบว่า คลังข้อมูลนักเรียนระดับมัธยมศึกษาที่ใช้งานโดย ผู้บริหาร หัวหน้ากลุ่มสาระการ เรียนรู้ และอาจารย์ประจำชั้นมีความพึงพอใจการใช้งานคลังข้อมูลอยู่ในระดับดี และในการทำเหมืองข้อมูล พยากรณ์ พบว่า ชุดข้อมูลแบบไม่จัดกลุ่มนำมาคัดเลือกคุณลักษณะด้วยวิธี Correlation-based Feature Selection (CFS) ร่วมกับเทคนิคโครงข่ายประสาทเทียมแบบมัลติเลเยอร์เพอร์เซ็ปตรอน ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 94.48 และมีค่ารากที่สองของความคลาดเคลื่อนน้อยที่สุดที่ 0.1880 เหมาะสมสำหรับการสร้าง ระบบพยากรณ์ผลการเรียนของนักเรียน

วิธีดำเนินการวิจัย

1. การทำความเข้าใจถึงปัญหา (Business Understanding) โดยผู้วิจัยได้ทำการศึกษาข้อมูลจากการเลือกกลุ่มการเรียน 3 กลุ่มของนักศึกษาสาขาวิชาเทคโนโลยี ได้แก่ ข้อมูลอัตราการพ้นสภาพ อัตราการย้ายสาขาวิชา อัตราการสำเร็จการศึกษา เป็นต้น พบว่าสาเหตุสำคัญของปัญหาเนื่องมาจาก ผู้เรียนไม่ได้เลือกกลุ่มวิชาที่ตรงกับความสามารถของผู้เรียน เลือกตามเพื่อน หรือเลือกตามผู้ปกครองแนะนำ ทำให้ผลการเรียนต่ำกว่าเกณฑ์ที่มหาวิทยาลัยกำหนด จึงส่งผลให้เกิดการลาพักการศึกษา ย้ายสาขา และพ้นสภาพนักศึกษา เป็นต้น ดังนั้นเมื่อสาขาวิชาได้มีการปรับปรุงหลักสูตรโดยกำหนดให้นักศึกษาสามารถเลือกกลุ่มวิชาเรียนได้เมื่อนักศึกษาขึ้นชั้น ปีที่ 3 ผู้วิจัยจึงมีแนวคิดแก้ปัญหาเรื่องผลการเรียนต่ำกว่าเกณฑ์อันเนื่องมาจากการเลือกกลุ่มวิชาที่ได้ตรงกับความสามารถของผู้เรียน ในลักษณะการทำเหมืองข้อมูลเพื่อช่วยวิเคราะห์ให้ผู้เรียนตัดสินใจได้เลือกกลุ่มการเรียนที่ตรงกับความสามารถของตนเอง และสามารถศึกษาในกลุ่มวิชานั้น ๆ ได้อย่างมีประสิทธิภาพ

2. การทำความเข้าใจข้อมูล (Data Undenstanding) ผู้วิจัยได้ทำการศึกษาข้อมูลของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ตั้งแต่ปีการศึกษา 2555-2560 ซึ่งมีจำนวนข้อมูลทั้งหมด 407 แถว ซึ่งมีจำนวนแอทริบิวต์ ทั้งหมด 16 แอทริบิวต์ ได้แก่ ข้อมูลเพศ , ผลการเรียนในรายวิชาหลัก 14 วิชา

และข้อมูลกลุ่มวิชา 3 กลุ่ม ได้แก่ กลุ่มวิชาเทคโนโลยีสารสนเทศ, กลุ่มวิชาเทคโนโลยีคอมพิวเตอร์,กลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษา โดยมีคลาส ได้แก่ ผลการเรียนที่สำเร็จการศึกษาอยู่ในระดับดี ระดับปานกลาง และระดับต่ำ

3. การเตรียมข้อมูล (Data Preparation) ผู้วิจัยได้ทำการเตรียมข้อมูล เพื่อทำการทดลองโดยทำการคัดเลือกข้อมูลที่จะนำมาวิเคราะห์ทดลองนั้น จะมีความสัมพันธ์ของแต่ละคุณลักษณะซึ่งเป็นตัวแปรอิสระ กับคำตอบที่ต้องการ (Class) ซึ่งเป็นตัวแปรตาม ซึ่งมีรายละเอียดข้อมูลดังตาราง

ตารางที่ 1 แสดงการแทนค่าข้อมูล

ปัจจัย	รายละเอียด	ค่าที่แทน
Sex : เพศ	ชาย	Male
	หญิง	Female
4091109 : ผลการเรียนวิชาคณิตศาสตร์พื้นฐาน 4131401: ผลการเรียนวิชาพื้นฐานคอมพิวเตอร์และเทคโนโลยีสารสนเทศ 4132101: ผลการเรียนวิชาระบบสารสนเทศเพื่อการจัดการ 4132201: ผลการเรียนวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ 4132102: ผลการเรียนวิชาการบริหารและการจัดการฐานข้อมูล 4131101: ผลการเรียนวิชาระบบฐานข้อมูล 4133202: ผลการเรียนวิชาเทคโนโลยีอินเทอร์เน็ต 4131301: ผลการเรียนวิชาหลักการเขียนโปรแกรมคอมพิวเตอร์ 4131302: ผลการเรียนวิชาการเขียนโปรแกรมภาษาระดับสูง 4132301: ผลการเรียนวิชาการเขียนโปรแกรมเชิงวัตถุ 4132402: ผลการเรียนวิชาโครงสร้างข้อมูลและอัลกอริทึม 4132403: ผลการเรียนวิชาการสื่อสารข้อมูลและเครือข่ายคอมพิวเตอร์ 4133401: ผลการเรียนวิชาระบบปฏิบัติการ 4132401: ผลการเรียนวิชาคณิตศาสตร์ดีสครีตและโครงสร้าง	● ระดับสูง ได้แก่ผลการเรียน A, B+, B	Good
	● ระดับกลาง ได้แก่ผลการเรียน C+ และ C	Normal
	● ระดับต่ำ ได้แก่ผลการเรียน D+ และ D	Bad
GPA: ผลการเรียนที่สำเร็จการศึกษา	● ระดับดี ได้แก่ผลการเรียนเฉลี่ยระหว่าง 3.00-4.00	Good
	● ระดับปานกลาง ได้แก่ผลการเรียนเฉลี่ยระหว่าง 2.50-2.99	Normal

	● ระดับต่ำ ได้แก่ผลการเรียนเฉลี่ยระหว่าง 2.00-2.49	Bad
Major: กลุ่มวิชา	กลุ่มวิชาเทคโนโลยีสารสนเทศ	IT
	กลุ่มวิชาเทคโนโลยีคอมพิวเตอร์และเครือข่าย	CT
	กลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษา	CE

4. การสร้างแบบจำลอง (Modeling) งานวิจัยนี้ได้ใช้เทคนิคในการทำเหมืองข้อมูล 3 เทคนิค ได้แก่ วิธีการตัดสินใจแบบต้นไม้ , วิธีการเรียนรู้เบย์อย่างง่าย และวิธีโครงข่ายประสาท โดยโปรแกรม RapidMiner Studio วิเคราะห์และกำหนดรูปแบบการทดลอง

5. การประเมินผล (Evaluation) โดยแบ่งข้อมูลทดลองโดยใช้วิธี 10-Fold Cross Validation คือการแบ่งข้อมูลออกเป็น 10 กลุ่มเท่า ๆ กัน จากนั้นทำการวัดประสิทธิภาพแบบจำลอง ดังต่อไปนี้

5.1 วัดค่าความถูกต้องของแบบจำลองโดยใช้ค่าความถูกต้องของข้อมูล (Accuracy)

5.2 วัดค่าสัดส่วนของข้อมูลที่ตรงตามความต้องการที่ถูกค้นคืนกับข้อมูลที่ต้องการทั้งหมด ค่าความระลึก (Recall)

5.3 วัดค่าสัดส่วนของข้อมูลที่ถูกต้อง และตรงตามความต้องการ ส่วนด้วยข้อมูลที่ถูกค้นคืนทั้งหมดค่าความแม่นยำ (Precision)

5.4 การวัดค่าเฉลี่ยค่าความแม่นยำในการตรวจพบและค่าความครบถ้วนในการตรวจพบเข้าด้วยกันจึงเป็นเหมือนค่าวัดความแม่นยำโดยรวม ค่าความถ่วงดุล (F-measure)

ผลและอภิปรายผลการวิจัย

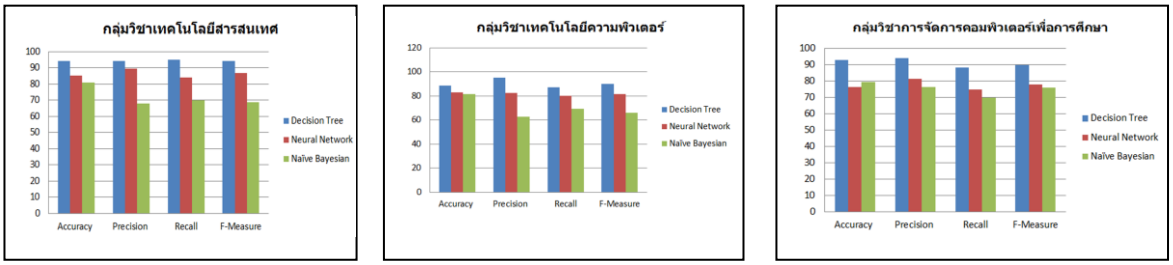
ผลการเปรียบเทียบประสิทธิภาพโมเดลด้วยเทคนิควิธีต้นไม้การตัดสินใจ (Decision Tree) วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayes) และวิธีโครงข่ายประสาท (Neural Network) โดยใช้หลักการแบ่งข้อมูลและข้อมูลเรียนรู้ด้วยวิธีการ 10-fold cross validation และการวิเคราะห์หาประสิทธิภาพด้วยการหาค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) สามารถวิเคราะห์ประสิทธิภาพของโมเดลดังตารางที่ 2 และภาพที่ 1-3

ตารางที่ 2 ผลการวิเคราะห์ประสิทธิภาพของโมเดล

Quality Assurance Group	Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F-Measure
กลุ่มวิชาเทคโนโลยีสารสนเทศ	Decision Tree	94.17	94.33	95.00	94.01
	Neural Network	85.00	89.28	84.17	86.65
	Naïve Bayesian	80.69	68.06	69.72	68.88
กลุ่มวิชาเทคโนโลยีคอมพิวเตอร์	Decision Tree	88.72	95.24	87.16	90.12
	Neural Network	83.08	82.38	80.22	81.28

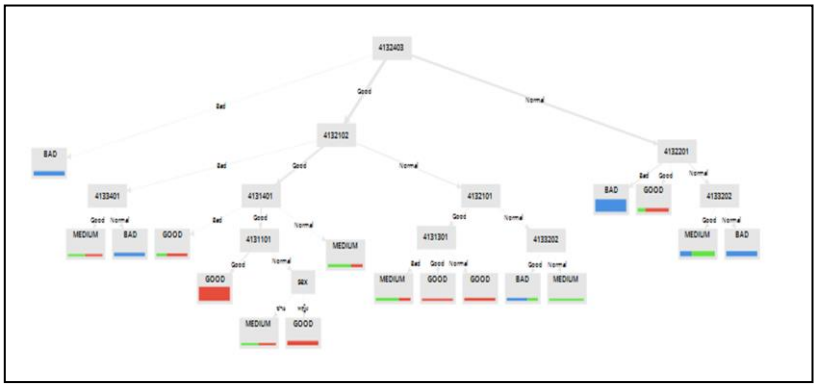
	Naïve Bayesian	81.54	62.59	69.44	65.83
กลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษา	Decision Tree	92.87	94.17	88.21	89.98
	Neural Network	76.58	81.28	74.88	77.95
	Naïve Bayesian	79.39	76.34	70.06	75.94

จากตารางที่ 2 แสดงถึงผลการวิเคราะห์ประสิทธิภาพของโมเดลการทำนายการเลือกกลุ่มวิชาของสาขาวิชาเทคโนโลยีสารสนเทศ ระหว่างวิธีต้นไม้การตัดสินใจ (Decision Tree) วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayes) และวิธีโครงข่ายประสาท (Neural Network) จากการทดลองพบว่า ทุกกลุ่มข้อมูลทดลองวิธีการต้นไม้ตัดสินใจ (Decision Tree) ได้ค่าความถูกต้อง (Accuracy) สูงสุด

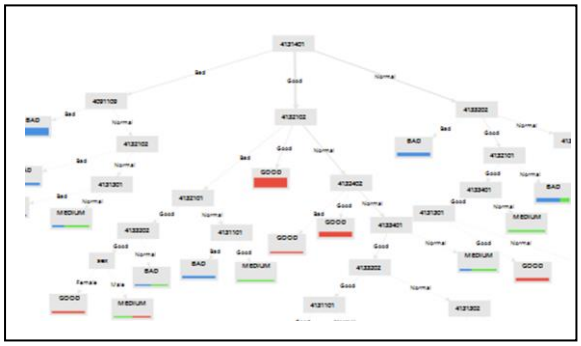


ภาพที่ 1 กราฟเปรียบเทียบค่าความแม่นยำ ค่าความระลึกและค่าความถ่วงดุลของการจำแนกข้อมูล

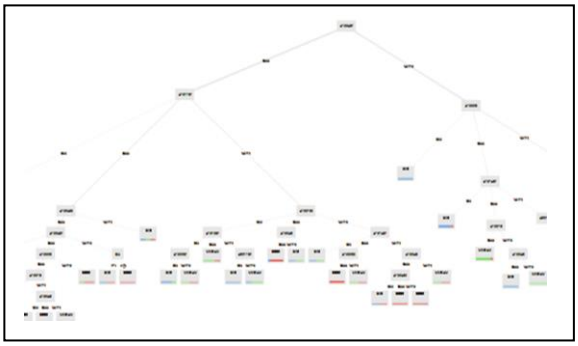
ซึ่งผลการเปรียบเทียบประสิทธิภาพของโมเดลจากงานวิจัยนี้ สามารถนำไปเป็นข้อมูลสำหรับสำหรับสร้างตัวแบบการตัดสินใจเลือกกลุ่มวิชาให้กับนักศึกษาได้ดังแสดงในภาพที่ 2-4



ภาพที่ 2 แสดงต้นไม้ตัดสินใจข้อมูลกลุ่มวิชา เทคโนโลยีสารสนเทศ



ภาพที่ 3 แสดงต้นไม้ตัดสินใจข้อมูลกลุ่มวิชาเทคโนโลยีคอมพิวเตอร์



ภาพที่ 4 แสดงต้นไม้ตัดสินใจข้อมูลกลุ่มวิชาการจัดการคอมพิวเตอร์เพื่อการศึกษา

จากตัวแบบผลการทำนายการเลือกกลุ่มวิชาที่เหมาะสมสำหรับการศึกษาสาขาวิชาเทคโนโลยีสารสนเทศในภาพที่ 1-3 สามารถนำไปสร้างเป็นกฎ (Rule) เพื่อทำนายผลการเลือกกลุ่มวิชาว่าต้องพิจารณาจากผลการเรียนในรายวิชาใดบ้าง จึงจะทำให้ผลการเรียนเฉลี่ยในกลุ่มวิชาที่เลือกอยู่ในระดับดี เช่น ผู้เรียนในกลุ่มวิชาเทคโนโลยีสารสนเทศที่มีผลการเรียนในรายวิชาการสื่อสารข้อมูลและเครือข่ายคอมพิวเตอร์ (4132403) ระดับต่ำ (เกรด D+,D) เมื่อเลือกเรียนในกลุ่มวิชาเทคโนโลยีสารสนเทศ มีโอกาสสำเร็จการศึกษาได้ผลการเรียนเฉลี่ย (GPA) อยู่ในระดับต่ำ (BAD) ผลการเรียนระหว่าง 2.00-2.49 เป็นต้น ซึ่งผลการทำนายดังกล่าวสามารถนำไปใช้เป็นแนวทางช่วยตัดสินใจเลือกกลุ่มวิชาของนักศึกษาสาขาวิชาเทคโนโลยีต่อไปได้

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จได้ด้วยความกรุณาจากสาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ที่ได้ให้ทุนสนับสนุนงานวิจัยในครั้งนี้จนเสร็จสมบูรณ์ ผู้วิจัยขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณผู้บริหารมหาวิทยาลัยราชภัฏบุรีรัมย์ ที่ให้การสนับสนุนและให้คำแนะนำเป็นอย่างดีตลอดมา ขอขอบพระคุณผู้เชี่ยวชาญตรวจสอบความถูกต้องของงานวิจัยพร้อมทั้งให้ความรู้ และคำแนะนำต่าง ๆ ที่มีประโยชน์ต่อการวิจัย

เอกสารอ้างอิง

จิราภา เลาหะวรนันท์ และคณะ. (2558). “การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือก แขนงวิชาสำหรับนักศึกษา คณะเทคโนโลยีสารสนเทศ”. วารสารเทคโนโลยีสารสนเทศลาดกระบัง, 4(2), 39-47.

สุธีรา วงศ์อนันทรัพย์. (2559). “การใช้เทคนิคเหมืองข้อมูลในการประเมินความรู้ และหาความถนัด เพื่อพัฒนาศักยภาพของนักศึกษา”. วารสารสังคมศาสตร์(มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก), 5(1), 12-19.

เสกสรรค์ วิลัยลักษณ์. (2558). “การใช้เทคนิคการทำเหมืองข้อมูลเพื่อพยากรณ์ผลการเรียนของนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและพัฒนาการศึกษา”. วารสารสาขาวิชาวิทยาศาตร์และเทคโนโลยี(มหาวิทยาลัยเกษตรศาสตร์), 2(2), 1-17.

อนันต์ ปินะตะ. (2557). **การใช้เทคนิคเหมืองข้อมูลในการเลือกกลุ่มวิชาที่เหมาะสมสำหรับการศึกษาต่อระดับปริญญาตรี.** วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ. มหาวิทยาลัยมหาสารคาม.

อนุพงศ์ สุขประเสริฐ์. (2562). **คู่มือการทำเหมืองข้อมูลด้วยโปรแกรม RapidMiner Studio**. พิมพ์ครั้งที่ 2. มหาสารคาม : มหาวิทยาลัยมหาสารคาม.

อรนุช พันโท และมนต์ชัย เทียนทอง. (2557). “การเปรียบเทียบประสิทธิภาพการจำแนกรูปแบบการเรียนรู้ VARK ด้วยเทคนิคเหมืองข้อมูล”. วารสารเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏอุบลราชธานี, 4(1), 1-11.

เอกสิทธิ์ พัชรวงศ์ศักดา. (2557). **การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิง เบื้องต้น.** พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : บริษัทเอเชีย ดิจิตอลการพิมพ์ จำกัด.

C. Shearer, “The CRISP-DM model: The new blueprint for data mining,” Journal of Data Warehousing, vol. 5, no.4, pp. 13 – 22, 2000.

M. Brame, Principles of Data Mining, Springer-Verlag London Limited: London, 2007