

ระบบแนะนำการเลือกศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์
มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยใช้เทคนิคต้นไม้ตัดสินใจ

The Recommendation System for Further Study Selection on Undergraduate
Level at Faculty of Education, Buriram Rajabhat University using Decision Tree
Technique

วรรณพ ม่วงชาติ* สุภารัตน์ อุไรพันธ์ และวิไลรัตน์ ยาทองไชย
สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์
iamwannaphaa@gmail.com*, Vawadalady@gmail.com, wilairat.yt@bru.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองและพัฒนาาระบบแนะนำการเลือกศึกษาต่อของนักเรียนที่จะศึกษาต่อในระดับปริญญาตรีคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยรวบรวมข้อมูลผู้สมัครเข้าศึกษาระหว่างปี 2556 – 2561 จำนวน 1,521 ข้อมูล มาจำแนกประเภท (Classification) ด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และใช้อัลกอริทึม J48 เพื่อสร้างแบบจำลองในการแนะนำการเลือกศึกษาต่อของนักเรียน ทำการทดสอบประสิทธิภาพของแบบจำลองด้วยวิธีการทดสอบแบบไขว้ทับ แบบ 10 ส่วน (10-fold cross validation) โดยใช้โปรแกรม WEKA จากนั้นนำแบบจำลองมาใช้ในการพัฒนาาระบบแนะนำการเลือกศึกษาต่อของนักเรียนที่จะศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ บนเว็บแอปพลิเคชัน (Web Application)

ผลการศึกษาพบว่า แบบจำลองในการแนะนำการเลือกศึกษาต่อของนักเรียนมีปัจจัยที่สำคัญ 7 ปัจจัย คือ เพศ วุฒิที่รับเข้า GPAXก่อนเข้า ความสามารถพิเศษ รายได้ผู้ปกครอง อาชีพผู้ปกครอง และสถานะการสมัคร โดยมีค่าความถูกต้องสูงสุดร้อยละ 69.56 ซึ่งจากการใช้งานระบบแนะนำการเลือกศึกษาต่อของนักเรียนและอาจารย์พบว่ามีความพึงพอใจในระดับมาก ($\bar{x} = 3.68$, S.D. = 0.58)

คำสำคัญ: ระบบแนะนำการเลือกศึกษาต่อ, ระดับปริญญาตรี, เทคนิคต้นไม้ตัดสินใจ, เว็บแอปพลิเคชัน

ABSTRACT

The purposes of this research were to create the models and develop the recommendation system for further study selection on undergraduate level at Faculty of Education, Buriram Rajabhat University. To collected 1,521 applicant data from 2013 – 2018 for classify using the Decision Tree technique and used the J48 algorithm to create recommendation models for further study selection of students. To evaluated the models with the 10- fold cross validation method using the WEKA program. The models were incorporated in the recommendation system for further study selection on undergraduate level at Faculty of Education, Buriram Rajabhat University which is a Web application.

The research findings showed that the recommendation models for further study selection of students have 7 major factors which are gender, education, GPAX, talent, parent income, parent occupation, and applicant status. The model evaluation is achieved 69.56 % of accuracy. The used of the

recommendation system for further study selection of students and teachers found that in a high level of satisfaction ($\bar{x} = 3.68$, S.D. = 0.58).

Keyword: Recommendation system, Further study selection, Decision tree technique, Web application

บทนำ

คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ เดิมเรียกว่า “หมวดวิชาการศึกษา” ซึ่งเกิดขึ้นพร้อมกันกับวันเปิดทำการสอนของวิทยาลัยครูบุรีรัมย์ ตั้งแต่ปีการศึกษา 2515 เป็นต้นมา โดยนักศึกษารุ่นแรกเป็นนักศึกษาระดับ ป.กศ. (ประกาศนียบัตรวิชาการศึกษา) ปัจจุบันคณะครุศาสตร์มีการเปิดการเรียนการสอนทั้งหมด 12 สาขาวิชา ได้แก่ สาขาวิชาเทคโนโลยีและคอมพิวเตอร์เพื่อการศึกษา สาขาวิชาภาษาไทย สาขาวิชาอังกฤษ สาขาวิชาการศึกษาปฐมวัย สาขาวิชาสังคมศึกษา สาขาคณิตศาสตร์ สาขาวิชาฟิสิกส์ สาขาวิชาวิทยาศาสตร์ทั่วไป สาขาวิชาพลศึกษา สาขาวิชาดนตรีศึกษา สาขาศิลปศึกษา และสาขาวิชานาฏศิลป์ โดยมีนักศึกษาตั้งแต่ชั้นปีที่ 1 - 5 ทั้งหมด 3,300 คน (คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์, 2561) จากข้อมูลการรับสมัครนักศึกษาในแต่ละปีการศึกษาจะมีนักเรียนจำนวนมากให้ความสนใจในการสมัครเรียนคณะครุศาสตร์ ที่สามารถรับเข้าเรียนได้ในจำนวนจำกัด ซึ่งมีอัตราการแข่งขันที่สูง นอกจากนี้การเรียนแต่ละสาขาวิชาจะมีความยากง่ายที่แตกต่างกันออกไป รวมถึงคุณลักษณะที่เหมาะสมและความถนัดของผู้เรียนก็มีส่วนสำคัญที่ส่งผลต่อการเรียน และความ เป็นครูในอนาคต ส่งผลให้เกิดปัญหาคือมีนักเรียนที่มีคุณลักษณะที่เหมาะสมแต่มีเกรดเฉลี่ยที่ไม่ตรงตามเกณฑ์ที่กำหนดไม่ได้รับ การคัดเลือก หรือเมื่อได้เข้าเรียนแล้วเกิดปัญหาในการเรียนทั้งในด้านความถนัด ความรู้พื้นฐานก่อนเรียน ทำให้เกิดผล กระทบต่าง ๆ เช่น การย้ายสาขาวิชา การออกกลางคัน เป็นต้น ดังนั้นการคัดเลือกนักศึกษาในแต่ละปีการศึกษาควรเลือกรับ นักศึกษาที่มีคุณลักษณะที่เหมาะสมและตรงตามความถนัดของแต่ละบุคคลให้มากที่สุด การทำเหมืองข้อมูล (Data Mining) คือกระบวนการที่กระทำกับข้อมูล เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์ ในการทำเหมืองข้อมูลนั้นมีหลายรูปแบบได้แก่ กฎความสัมพันธ์ การ จำแนกประเภทข้อมูล การแบ่งกลุ่มข้อมูล และการสร้างโมโนภาพ (Automated, 2559) ซึ่งการจำแนกประเภทข้อมูล (Data Classification) เป็นกระบวนการที่ใช้ในการหาแบบของชุดข้อมูลที่มีความใกล้เคียงกัน หรือเหมือนกันมากที่สุด เพื่อใช้ในการ ทำนายชุดข้อมูลใหม่ว่าอยู่ในประเภทใดของชุดข้อมูลที่ได้จากการเรียนรู้จากชุดข้อมูลที่มีอยู่แล้ว (Training Data)

จากปัญหาดังกล่าว คณะผู้วิจัยมีแนวคิดการนำเหมืองข้อมูล (Data Mining) มาช่วยในการหากฎการจำแนกข้อมูล ของนักเรียนในการเลือกเข้าศึกษาต่อระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยใช้เทคนิคต้นไม้ตัดสินใจ เพื่อให้ได้สารสนเทศที่สามารถนำมาแนะนำการเลือกเรียนของนักเรียนที่สนใจสมัครเข้าศึกษาต่อในระดับปริญญาตรีได้

1. วัตถุประสงค์การวิจัย

1. เพื่อสร้างแบบจำลองในการแนะนำการเลือกศึกษาต่อของนักเรียนที่จะศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยใช้เทคนิคต้นไม้ตัดสินใจ

2. เพื่อพัฒนาระบบแนะนำการเลือกศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ โดยใช้เทคนิคต้นไม้ตัดสินใจ

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

การทำเหมืองข้อมูลเปรียบเสมือนวิวัฒนาการในการตีความหมายข้อมูล จากการจัดเก็บข้อมูลอย่างง่ายมาสู่การ จัดเก็บในรูปแบบข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการค้นหาสารสนเทศที่ซ่อนอยู่ในข้อมูล โอม ศรีนิล (2556) ได้ ให้ความหมายของการทำเหมืองข้อมูลว่าเป็นการค้นหาแบบ และความสัมพันธ์ในชุดข้อมูลขนาดใหญ่โดยทำงานในลักษณะ กึ่งอัตโนมัติที่อาศัยความสามารถในการคำนวณของคอมพิวเตอร์และความรู้เกี่ยวกับธุรกิจในการค้นหาแบบที่เป็นไปได้จาก

ข้อมูลขนาดใหญ่ นอกจากนี้สุรพงษ์ เอื้อวัฒนามงคล (2557) ยังได้ให้ความหมายของการทำเหมืองข้อมูลว่า เป็นกระบวนการวิเคราะห์ข้อมูลที่มีขั้นตอนเพื่อให้ได้มาซึ่งตัวแบบ (Pattern) ที่แสดงความสัมพันธ์ระหว่างข้อมูล โดยผลลัพธ์สามารถนำไปใช้ในการตัดสินใจ รวมถึงสายชล สันสมบูรณ์ทอง (2558) ได้กล่าวว่าการทำเหมืองข้อมูลเป็นกระบวนการทำงานที่สกัดข้อมูลจากฐานข้อมูลที่มีขนาดใหญ่เพื่อให้ได้สารสนเทศที่มีประโยชน์สามารถนำไปช่วยการตัดสินใจในการดำเนินงานต่าง ๆ โดยการทำให้เหมืองข้อมูลเป็นกระบวนการที่สำคัญในการค้นหาความรู้จากฐานข้อมูลขนาดใหญ่ (KDD) ซึ่งสามารถนำมาคาดการณ์อนาคตได้ โดยการจำแนกประเภทข้อมูลเป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดจากกลุ่มข้อมูลสอน โดยมีแอทริบิวต์แบ่ง (Classifying Attribute) เป็นตัวบ่งชี้คลาสของข้อมูล สามารถนำไปประยุกต์ใช้ในหลายด้าน เช่น การจัดกลุ่มลูกค้าทางการตลาด การวิเคราะห์ทางการแพทย์ เป็นต้น

ต้นไม้ตัดสินใจ (Decision Tree)

เป็นเทคนิคในการจำแนกประเภทข้อมูลที่ใช้ตรวจสอบข้อมูลและสร้างต้นไม้เพื่อการพยากรณ์ ถูกคิดค้นโดย Dr.Dome และ Dr.JA โครงสร้างของต้นไม้ตัดสินใจมีลักษณะแตกแขนงไปตามเงื่อนไขหรือเส้นทางของกิ่งไม้ ซึ่งใช้กฎในรูปแบบ “ถ้า (เงื่อนไข) แล้ว (ผลลัพธ์)” (If-then Rule) มาประกอบการสร้างโครงสร้างต้นไม้ตัดสินใจ ประกอบด้วย

1. โหนดภายใน (Internal Node) คือโหนดที่แสดงถึงคุณลักษณะ (Feature) ที่นำมาใช้ในการแบ่งกลุ่มของข้อมูล ซึ่งมีโหนดราก (Root Node) อยู่บนสุดของโครงสร้าง ซึ่งเป็นโหนดที่มีอิทธิพลต่อการจำแนกกลุ่มมากที่สุด
2. กิ่ง (Branch) เป็นตัวเชื่อมระหว่างโหนดที่ใช้เป็นเงื่อนไขหรือทางเลือกของการกระทำ ซึ่งมาจากผลลัพธ์แต่ละตัวของทุกตัวทำนาย (Predictor) หรือคุณลักษณะ
3. โหนดใบ (Leaf Node) เป็นโหนดที่แสดงผลลัพธ์ของเงื่อนไขหรือการกระทำตามเงื่อนไขที่เกิดขึ้น

การทำเหมืองข้อมูลถูกนำมาใช้ในงานวิจัยด้านการศึกษาหลากหลายงาน อาทิ การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่เปิดหลักสูตรการเรียนตามแขนงวิชาเพื่อเป็นเครื่องมือช่วยตัดสินใจเลือกแขนงวิชา โดยใช้ข้อมูลผลการเรียนและผลการวัดความสามารถด้านต่าง ๆ ที่เกี่ยวข้องมาสร้างแบบจำลองโดยเปรียบเทียบเทคนิคการทำเหมืองข้อมูล 5 เทคนิค และพยากรณ์ผ่านเทคนิค “Ensemble” พบว่าการพยากรณ์มีความแม่นยำ 72.92% โดยการพัฒนากระบวนการนี้ทำให้นักศึกษาทราบถึงแขนงวิชาที่เหมาะสมกับตนเองมากที่สุดถึงน้อยที่สุดพร้อมระดับความน่าจะเป็น ซึ่งส่งผลให้นักศึกษาตัดสินใจเลือกแขนงวิชาที่เหมาะสมกับตนเองได้ และมีโอกาสประสบความสำเร็จในการเรียน (จิราภา เลหาหะวันนท์, รชต ลัมสุทธิวันภูมิ และบัณฑิต ฐานะโสภณ, 2561) รวมถึงธีรพงษ์ สังข์ศรี (2557) ได้ทำการวิเคราะห์พฤติกรรมการเลือกสมัครสาขาวิชาเรียนและการเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่ โดยนำแนวความคิดการทำเหมืองข้อมูลมาช่วยในการวิเคราะห์ข้อมูล ผู้สมัครเข้าศึกษาต่อเพื่อช่วยการวางแผนการรับสมัครในอนาคต แบ่งการทำงานออกเป็น 2 ส่วนคือ 1) การวิเคราะห์พฤติกรรมการเลือกสมัครสาขาวิชาเรียน ด้วยเทคนิคการจัดกลุ่มข้อมูล (Clustering) แบบ Simple K-means สามารถแบ่งข้อมูลพฤติกรรมของผู้สมัครได้ 4 กลุ่ม และใช้การหาความสัมพันธ์ของข้อมูล (Association Rules) ด้วยเทคนิคอะพริอริ (Apriori) เพื่อหาความสัมพันธ์ของข้อมูลในแต่ละกลุ่มพฤติกรรมผู้สมัคร มีค่าความเชื่อมั่น 0.9 และ 2) การเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่โดยตัวแบบพยากรณ์ถูกสร้างด้วยเทคนิคต้นไม้ช่วยตัดสินใจ (Decision Tree) มีค่าความถูกต้อง 93.76% เทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network) แบบเพอร์เซ็ปตรอนหลายชั้น (Multi-layer Perceptron) มีค่าความถูกต้อง 93.60% เช่นเดียวกับชั้นทอง ประชุมชาติ และพิมรินทร์ คีรินทร์ (2555) ทำการวิจัยเพื่อวิเคราะห์พฤติกรรมการเลือกสมัครสาขาวิชาเรียนของนักศึกษาใหม่ของมหาวิทยาลัยราชภัฏพิบูลสงคราม โดยใช้เทคนิคการจัดกลุ่ม เพื่อจัดกลุ่มของผู้สมัคร และสร้างตัวแบบจำลองการหาความสัมพันธ์ ด้วยขั้นตอนวิธีอะพริอริ เพื่อหาความสัมพันธ์ของพฤติกรรมการเลือกสาขาวิชาเรียนของผู้สมัคร ซึ่งผลการวิจัยที่ได้นำไปใช้สนับสนุนการวางแผนการรับนักศึกษาและการประชาสัมพันธ์ในการรับนักศึกษาใหม่ได้อย่างเหมาะสม อาทิทยาพร โรจรัตน์ (2556) ได้เสนอรูปแบบการ

ทำนายผลการศึกษาเพื่อใช้วางแผนการลงทะเบียนเรียนโดยนำวิธีความสัมพัทธ์ ต้นไม้ตัดสินใจ และวิธีการของการเรียนรู้เบย์ (Naive Bayes) มาวิเคราะห์และเปรียบเทียบหาความสัมพัทธ์ที่ก่อให้เกิดการทำนายผลการศึกษานักศึกษา สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยขอนแก่น จากผลการวิเคราะห์ความสัมพัทธ์ทั้ง 3 วิธีการ พบว่ารูปแบบเพื่อการวิเคราะห์การวางแผนการลงทะเบียนเรียนด้วยวิธีความสัมพัทธ์มี 26 กฎ วิธีต้นไม้ตัดสินใจมี 26 กฎ แบ่งเป็นการวิเคราะห์โดยใช้รายวิชาทั้งหมด 13 กฎความสัมพัทธ์มีค่าความถูกต้อง 77.92 % ประเมินโดยใช้ 10-folds Cross Validation และการวิเคราะห์โดยแยกเอาแต่รายวิชาที่เป็นวิชาเลือกจำนวน 13 กฎมีค่าความถูกต้อง 68.56 % โดยใช้การประเมิน 40-folds Cross Validation และวิธีการเรียนรู้เบย์ มีค่าความถูกต้อง 85.77% โดยใช้การประเมิน 10-fold Cross Validation ซึ่งวิธีการความสัมพัทธ์และต้นไม้ตัดสินใจจะเห็นความสัมพันธ์ของรายวิชาเพื่อการวางแผนลงทะเบียนได้ดีกว่าวิธีการเรียนรู้เบย์ รวมถึงสรุปพัฒนาการ ภาค (2556) ทำการวิจัยเพื่อค้นหาปัจจัยที่มีผลต่อการเลือกแผนการเรียน และความสามารถในการศึกษาในแผนการเรียนนั้นได้อย่างประสบความสำเร็จ ด้วยเทคนิคเหมืองข้อมูล ซึ่งข้อมูลที่ใช้ คือ ผลการเรียนในรายวิชา และข้อมูลแบบสอบถามวิธีการตัดสินใจเลือกแผนการเรียนของนักเรียนระดับมัธยมศึกษาตอนปลาย โรงเรียนสตรีศรีสุริโยทัย ปีการศึกษา 2555 ทั้งสิ้น 850 คน ผลการวิจัยพบว่า ตัวแบบการเลือกแผนการเรียนระดับมัธยมศึกษาตอนปลายมีปัจจัยที่มีผลต่อการตัดสินใจเลือกแผนการเรียนของนักเรียนคือ ผลการเรียนวิชาพื้นฐานระดับมัธยมศึกษาตอนต้น คือ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา ศาสนาและวัฒนธรรม และภาษาอังกฤษ รวมทั้งผลการเรียนเฉลี่ยสะสมระดับมัธยมศึกษาตอนต้น และให้ค่าความถูกต้องในการแนะนำแผนการเรียนร้อยละ 79.76

วิธีดำเนินการวิจัย

1. กรอบการศึกษา

การวิจัยนี้มีกรอบในการศึกษา 5 ขั้นตอน คือ การเตรียมข้อมูลก่อนประมวลผล อัลกอริทึมสำหรับการจำแนกประเภท การวัดประสิทธิภาพการจำแนกประเภท การสร้างแบบจำลอง และการพัฒนาและประเมินผลระบบ

- **การเตรียมข้อมูลก่อนประมวลผล (Data Pre-processing)** เป็นขั้นตอนแรกก่อนการทำเหมืองข้อมูลเนื่องจากข้อมูลในฐานข้อมูลที่จะนำมาวิเคราะห์อาจมีข้อมูลไม่สมบูรณ์ (Incomplete Data) ข้อมูลมีค่าผิดพลาด (Error) หรือมีค่าผิดปกติ (Outliers) ตลอดจนข้อมูลไม่สอดคล้อง (Inconsistent Data) ที่ต้องตรวจสอบความถูกต้องในเรื่องของรูปแบบข้อมูล ค่าของข้อมูลก่อนการประมวลผล โดยงานวิจัยนี้ได้ทำการคัดข้อมูลที่เป็นส่วนรบกวน หรือข้อมูลที่ไม่เกี่ยวข้องออกไป (Data Cleaning) ทำการคัดเลือกข้อมูลและแปลงข้อมูลให้เหมาะกับการทำเหมืองข้อมูล (Data Selection and Transformation)

- **อัลกอริทึมสำหรับการจำแนกประเภท (Classifier Algorithm)** ทำการจำแนกประเภทด้วยอัลกอริทึม J48 เพื่อให้ได้ผลลัพธ์ที่มีความยืดหยุ่นสามารถนำไปใช้โดยแปลงเป็นกฎได้ ให้ความแม่นยำสูง และนำไปประยุกต์ใช้กับการค้นหาข้อมูลด้วยภาษาสอบถามเชิงโครงสร้าง (SQL) ได้

- **การวัดประสิทธิภาพการจำแนกประเภท (Evaluation Classifier)** โดยใช้วิธีการตรวจสอบไขว้ (K-fold cross-validation) เพื่อประเมินผลแบบจำลองจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-measure)

- **การสร้างแบบจำลอง (Academic DSS Model)** เมื่อทำการสร้างแบบจำลองทำให้ได้กฎจากอัลกอริทึมการจำแนกประเภทเพื่อนำไปสู่การพัฒนาระบบแนะนำการเลือกศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ที่นำมาแนะนำการเรียนของนักเรียนที่สนใจสมัครเข้าศึกษาต่อ คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

- **การพัฒนาและประเมินผลระบบ (Development and Evaluation System)** โดยเลือกใช้ภาษา PHP HTML JAVASCRIPT และระบบจัดการฐานข้อมูล MySQL ในรูปแบบเว็บแอปพลิเคชัน (Web Application)

2. การรวบรวมข้อมูลเพื่อการประมวลผล

คลังข้อมูลของมหาวิทยาลัยราชภัฏบุรีรัมย์จะเก็บฐานข้อมูลที่เกี่ยวข้องกับการทำงานของมหาวิทยาลัย เช่น ฐานข้อมูลการศึกษา ฐานข้อมูลการเงิน ฐานข้อมูลการประกันคุณภาพ เป็นต้น เพื่อนำมาใช้ประโยชน์ในการบริหารงานของมหาวิทยาลัย โดยข้อมูลที่น่าสนใจในการวิจัยครั้งนี้คือ ข้อมูลผู้สมัครสอบเข้าศึกษาต่อมหาวิทยาลัยราชภัฏบุรีรัมย์ ระหว่างปีการศึกษา 2556 – 2561 ใน 2 ด้านคือ ข้อมูลส่วนบุคคล เช่น รหัสอ้างอิงผู้สมัคร ประเภทการสมัครสอบ เพศ จุดที่รับเข้า เกรดเฉลี่ย สถานะผู้สมัคร วันเกิด รหัสจังหวัดที่เกิด รหัสเชื้อชาติ ความสามารถพิเศษ สถานะบิดา ข้อมูลรายได้บิดา เป็นต้น และข้อมูลด้านการศึกษา เช่น จุดที่จบ ชื่อ ชื่ออังกฤษ ชื่อย่อ ชื่อย่ออังกฤษ ISEC ระดับการศึกษา และสายการศึกษา เป็นต้น จากนั้นนำข้อมูลจากหลายตารางมารวมกันทำให้ได้ข้อมูล จำนวน 20,987 ระเบียบ 152 แอตทริบิวต์

3. การคัดเลือกและแปลงข้อมูล

-ขั้นตอนนี้จะได้ชุดข้อมูลที่สามารถนำมาใช้ในการสร้างแบบจำลองเพื่อแนะนำการเลือกสาขาวิชาในคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ จำนวน 7 แอตทริบิวต์ 1,521 ระเบียบรายละเอียดของแต่ละปัจจัยดังแสดงในตารางที่ 1

ตารางที่ 1 รายละเอียดปัจจัย

ตัวแปร	คำอธิบาย	ค่าที่เป็นไปได้
APPLICANTSEX	เพศ	F = ชาย , M = หญิง
ENTRYDEGREE	วุฒิการศึกษาที่รับเข้า	Ac = ม.6 (ศิลป์-คำนวณ) ; AL = ม.6 (ศิลป์-ภาษา) HVC = ปวส ; M6 = ม.6 ; ME = ม.6 (คณิต-อังกฤษ) SM = ม.6 (วิทย์-คณิต) ; VC = ปวช.
ENTRYGPAX	เกรดเฉลี่ยจากโรงเรียนมัธยม	W = 2.00 – 2.49 ; M = 2.50 – 2.99 G = 3.00 – 3.49 ; B = 3.50 – 4.00
TALENT	ความสามารถพิเศษ	01 = คอมพิวเตอร์ ; 02 = ศิลปะ ; 03 = ภาษา 04 = กีฬา ; 05 = ดนตรี
PARENTREVENUE	รายได้ผู้ปกครอง	1 =< 150,000 ต่อปี ; 2 = 150,000-300,000 ต่อปี 3 => 300,000 ต่อปี
PARENTOCCUP	อาชีพผู้ปกครอง	1 = รับราชการ; 2 = รัฐวิสาหกิจ; 3 = หน่วยงานเอกชน; 4 = ค้าขาย/ธุรกิจส่วนตัว; 5 = เกษตรกร; 6 = อาชีพอิสระ/รับจ้าง; 7 = ไม่มีเงินได้; 9 = พนักงานราชการ/ลูกจ้างหน่วยงานราชการ
SELECTIONSTATUS	สถานะการสมัคร	Th = ภาษาไทย; Eng = ภาษาอังกฤษ; Da = นาฏศิลป์; Pe = ละครศึกษา; Sa = สังคมศึกษา ; Gs = วิทยาศาสตร์ทั่วไป; Mu = ดนตรีศึกษา ; Ph = ฟิสิกส์ ; Art = ศิลปะศึกษา; Ma = คณิตศาสตร์ ; Ec = การศึกษาปฐมวัย; Etc = เทคโนโลยีและ คอมพิวเตอร์ฯ ; No = ไม่ผ่าน

4. การวิเคราะห์ข้อมูล

การวิจัยนี้ใช้เทคนิคต้นไม้ตัดสินใจในการจำแนกประเภทของข้อมูลเพื่อสร้างแบบจำลอง (Model) ดังนี้

4.1 การสร้างและทดสอบแบบจำลอง งานวิจัยนี้ใช้โปรแกรม WEKA ในการสร้างตัวแบบการวิเคราะห์ปัจจัยและทดสอบตัวแบบด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree Technique) โดยใช้วิธีการจำแนกประเภท เลือกใช้ J48 คือ อัลกอริทึม C4.5 เพื่อใช้ในการเรียนรู้ (Learning) และสร้างแบบจำลองต้นไม้เพื่อการวิเคราะห์ปัจจัย โดยตัวแบบที่ได้อยู่ในรูปของกฎ

การจำแนกประเภทจากการเรียนรู้ด้วยชุดการเรียนรู้ แล้วนำไปทดสอบด้วยชุดทดสอบ โดยใช้วิธีการตรวจสอบไขว้ โดยวิธีการตรวจสอบไขว้เป็นการแบ่งข้อมูลแบบสุ่มออกเป็นจำนวนกลุ่มตามตัวเลข คือ ถ้าใส่ตัวเลข ค่า N เข้าไปจะมีการแบ่งข้อมูลออกเป็นส่วนย่อยจำนวน N และจะเก็บข้อมูล 1 ส่วนย่อยเพื่อไว้ทดสอบ ส่วนข้อมูลที่เหลือนำมาสร้างแบบจำลอง และจะทำงานไปจนกว่าข้อมูลถูกใช้ทดสอบจนหมดทุกส่วนการวิจัยครั้งนี้ได้กำหนดค่า k ในวิธีการตรวจสอบไขว้ คือ 10, 100

4.2 การวัดค่าประสิทธิภาพของตัวแบบการพยากรณ์ การวัดค่าประสิทธิภาพของตัวแบบการวิเคราะห์ปัจจัย จากการนำตัวแบบที่ได้จากชุดการเรียนรู้มาทดสอบด้วยชุดข้อมูลทดสอบ โดยการวัดค่าที่ได้ คือ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-measure)

5. การพัฒนาและการประเมินผลระบบ

สถาปัตยกรรมของระบบถูกสร้างขึ้นเพื่อใช้งานผ่านระบบเครือข่ายผ่านเว็บเบราว์เซอร์ (Web Browser) ซึ่งใช้ระบบจัดการฐานข้อมูล (Database Management System) คือ MySQL โดยใช้ภาษาสคริปต์เป็นพีเอชพี (PHP Hypertext Preprocessor) สำหรับทำงานบนเว็บเซิร์ฟเวอร์ (Web Server) ซึ่งมีการทำงานในรูปแบบ Client-Side Script ตามผู้ใช้งานระบบแต่ละระดับ และทำการศึกษาความพึงพอใจจากอาจารย์ และนักเรียน จากการสุ่มแบบเจาะจง จำนวน 30 คน เพื่อเป็นกลุ่มตัวอย่าง โดยการตอบแบบสอบถามความพึงพอใจเกี่ยวกับความสามารถในการใช้งานระบบ ที่มีลักษณะเป็นแบบเลือกตอบมาตราส่วนประมาณค่า (Likert Scale) 5 ระดับตามเกณฑ์ของลิเคิร์ท (ปรีชา บุณรอด, 2541)

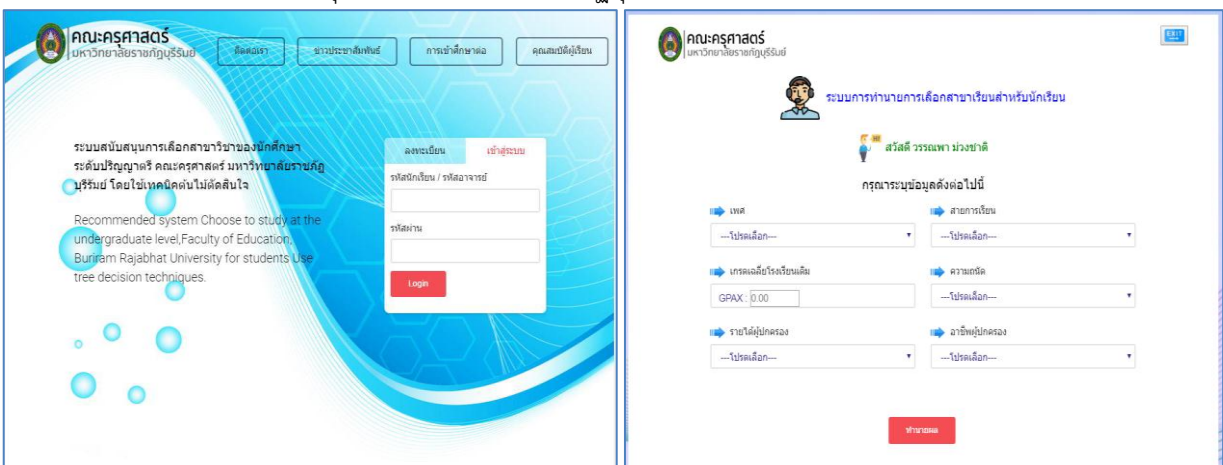
ผลการวิจัยและอภิปรายผล

1. ผลการสร้างแบบจำลอง

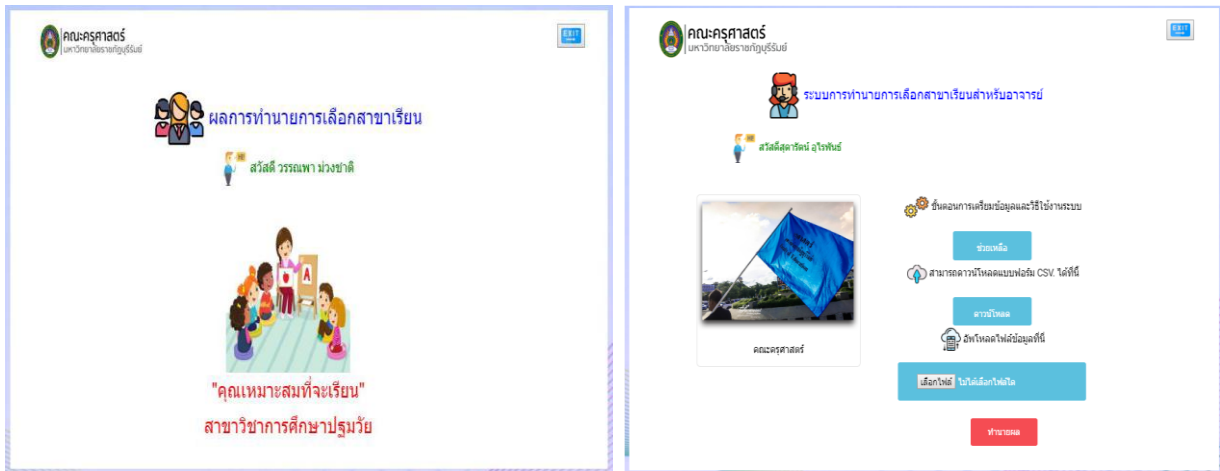
แบบจำลองเพื่อแนะนำการเลือกสาขาวิชาในคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ พบว่า รูปแบบเพื่อการแนะนำการเลือกสาขาวิชาวิธีต้นไม้ตัดสินใจมีกฎ จำนวน 23 กฎ ประเมินโดยใช้ 10-folds Cross Validation มีค่าความถูกต้อง 69.89% ค่าความแม่นยำ 59.50% ค่าความระลึก 69.60% และค่าความถ่วงดุล 63.60% ด้วยเทคนิคต้นไม้ตัดสินใจเป็นเทคนิคในวิธีการแบบ Vote Ensemble จะสามารถพยากรณ์ได้ดีและมีความน่าเชื่อถือเมื่อมีความแม่นยำอย่างต่ำ 50 % (จิราภา เลหาหะวรรณท์ และคณะ, 2561) ซึ่งจะสอดคล้องกับงานวิจัยของสุพัฒน์กุล ภัคโชค (2556)

2. ผลการพัฒนาและประเมินผลระบบ

การพัฒนาและประเมินผลระบบแนะนำการเลือกสาขาวิชาในคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ประกอบด้วยกระบวนการทำงานที่สำคัญ คือ 1) ส่วนการใช้งานสำหรับนักเรียน เพื่อนำปัจจัยสู่ระบบในการทำนายผลการแนะนำการเลือกสาขาวิชา 2) ส่วนการใช้งานสำหรับอาจารย์แนะแนว ที่สามารถดาวน์โหลดไฟล์ข้อมูลของนักเรียน เพื่อนำมาทำนายผลการแนะนำการเลือกสาขาวิชาในการสมัครสอบคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ หน้าจอการทำงานดังภาพที่ 1-2



ภาพที่ 1 การนำเข้าปัจจัยเพื่อการแนะนำการเลือกสาขาวิชา



ภาพที่ 2 การทำนายการเลือกสาขาเรียน และหน้าจอการทำงานสำหรับอาจารย์

ในการประเมินผลการทำงานของระบบโดยการศึกษาความพึงพอใจของการใช้งานระบบจากอาจารย์ และนักเรียน ด้วยแบบสอบถามความพึงพอใจเกี่ยวกับความสามารถในการใช้งานระบบ (System Usability) มีผลการศึกษาดังตารางที่ 2

ตารางที่ 2 ผลการศึกษาความพึงพอใจโดยรวมของระบบ

รายการ	\bar{x}	S.D.	ระดับความคิดเห็น
1. ระบบง่ายต่อการใช้งาน	3.80	0.45	มาก
2. ผลลัพธ์ที่ได้จากการทำนายมีความถูกต้อง น่าเชื่อถือ	3.60	0.55	มาก
3. ความชัดเจนของผลการทำนาย	3.80	0.84	มาก
4. ผลการทำนายที่สามารถใช้ประกอบการวางแผนการศึกษาได้	3.40	0.55	ปานกลาง
5. ความเร็วในการประมวลผลของระบบ	3.80	0.45	มาก
6. ความชัดเจนของข้อความที่แสดงบนจอ	4.20	0.45	มาก
7. ความเหมาะสมของการใช้สีของตัวอักษรและรูปภาพ	3.40	0.55	ปานกลาง
8. การใช้ถ้อยคำบนจอภาพสามารถสื่อความหมายให้เข้าใจได้ง่าย	3.40	0.55	ปานกลาง
9. การกำหนดรหัสผู้ใส่ และรหัสผ่านในการตรวจสอบผู้เข้าใช้ระบบ	3.60	0.55	มาก
10. ผลลัพธ์ที่ได้จากการทำนายตรงกับความต้องการของผู้ใช้	3.80	0.84	มาก
โดยรวม	3.68	0.58	มาก

จากตารางที่ 2 ผลการศึกษาความพึงพอใจเกี่ยวกับความสามารถในการใช้งานระบบ พบว่า โดยภาพรวมผู้ใช้มีระดับความพึงพอใจในระดับมาก ($\bar{x} = 3.68$, S.D. = 0.58) โดยประเด็นที่ผู้ใช้มีความพึงพอใจมากที่สุดคือ ความชัดเจนของข้อความที่แสดงบนจอที่มีความพึงพอใจในระดับมาก ($\bar{x} = 4.20$, S.D. = 0.45) รองลงมาคือ ระบบง่ายต่อการใช้งาน ความชัดเจนของผลการทำนาย ความเร็วในการประมวลผลของระบบ และผลลัพธ์ที่ได้จากการทำนายตรงกับความต้องการของผู้ใช้ ซึ่งแสดงให้เห็นว่าโดยภาพรวมระบบสามารถช่วยในการตัดสินใจเพื่อเลือกสาขาวิชาที่เหมาะสมกับตัวของนักเรียนในการสมัครเข้าศึกษาต่อคณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

ข้อเสนอแนะ

1. ในกระบวนการทำเหมืองข้อมูลขั้นตอนในการเตรียมข้อมูลถือว่ามีความสำคัญเป็นอย่างมาก ประสพการณ์ที่ได้จากการวิจัยนี้ ที่ใช้ข้อมูลที่เก็บไว้ในฐานข้อมูลเพื่อการศึกษาของมหาวิทยาลัยราชภัฏบุรีรัมย์มาทำการวิเคราะห์ที่มีจำนวน

ตารางจำนวนมาก และข้อมูลมีหลากหลายรูปแบบ รวมถึงการมีค่าผิดพลาด (Missing value) เก็บไว้จำนวนมาก ดังนั้นในขั้นตอนการเตรียมข้อมูลจึงต้องวางแผนเป็นอย่างดี ซึ่งต้องใช้เวลาและต้องกระทำอย่างระมัดระวัง

2. การเลือกแอททริบิวต์ (Attributes Selection) ที่จะนำมาวิเคราะห์รูปแบบเพื่อการแนะนำการเลือกสาขาวิชาที่มีความสำคัญอีกขั้นตอนหนึ่ง โดยงานวิจัยนี้พบว่า ความเหมาะสมของแอททริบิวต์สำหรับการจัดประเภทข้อมูล ค่าข้อมูลจะต้องมีค่าที่วนซ้ำ และไม่หลากหลาย

3. ผลจากการวิจัย จะช่วยให้นักเรียน และอาจารย์แนะแนวได้สารสนเทศที่สามารถนำมาแนะแนวการเลือกเรียนของนักเรียนที่สนใจสมัครเข้าศึกษาต่อในระดับปริญญาตรี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ได้ นอกจากนี้ยังช่วยให้นักเรียนตัดสินใจเลือกในสาขาวิชาที่เหมาะสมกับคุณลักษณะของตนได้ซึ่งจะช่วยลดปัญหาในการเรียนของผู้เรียน

เอกสารอ้างอิง

- ขันทอง ประชุมชาติ และพิมรินทร์ คีรินทร์. (2555). *การวิเคราะห์พฤติกรรมกรรมการเลือกสมัครสาขาวิชาเรียนของนักศึกษาใหม่โดยใช้เทคนิคการทำเหมืองข้อมูล*. การประชุมวิชาการระดับชาติ ครั้งที่ 2 สถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏกาแพงเพชร สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพิบูลสงคราม. จาก <http://compcenter.bu.ac.th/news-information/data-mining>.
- คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์. (2561). ประวัติความเป็นมา. [ออนไลน์]. สืบค้นจาก : <http://edu.bru.ac.th/edu.bru.ac.th/ประวัติความเป็นมา/>
- จิราภา เลหาหวนันท์, รชต ลัมสุทธิวันภูมิ และบัณฑิต ฐานะโสภณ. (2561). *การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ*. คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ธีรพงษ์ สังข์ศรี. (2555). *การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียนและการเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่โดยใช้เทคนิคการทำเหมืองข้อมูล*. โปรแกรมวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา.
- ปรีชา บุญรอด. (2541). *มาตรวัดการวิจัยทางสังคมศาสตร์ (ตอนที่ 2)* [ออนไลน์]. สืบค้นจาก : http://www.journal.au.edu/abac_newsletter/1998/june98/scale.html.
- สายชล สีนสมบูรณ์ทอง. (2558). *การทำเหมืองข้อมูล (Data Mining)*. กรุงเทพฯ: จามจุรีโปรดักส์.
- สุพัฒน์กุล ภัคโชค. (2556). *ตัวแบบการเลือกแผนการเรียนระดับมัธยมศึกษาตอนปลายโดยพิจารณาผลการเรียนรายวิชาหลักด้วยเทคนิคเหมืองข้อมูล*. วิทยานิพนธ์ตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิศวกรรมเว็บ คณะเทคโนโลยีสารสนเทศมหาวิทยาลัยธุรกิจบัณฑิต.
- สุรพงศ์ เอื้อวัฒนามงคล. (2557). *การทำเหมืองข้อมูล (Data Mining)*. กรุงเทพฯ: บางกอกบล็อก.
- อาทิตย์พร โรจรัตน์. (2556). *การทำนายผลการศึกษาเพื่อการวางแผนการลงทะเบียนของนักศึกษาโดยใช้การทำเหมืองข้อมูล*. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยขอนแก่น.
- โอม ศรีนิล. (2556). *การออกแบบและพัฒนาคลังข้อมูล*. กรุงเทพฯ: บางกอกบล็อก.
- Automated. (2559). *เหมืองข้อมูล (data mining)*. [ออนไลน์]. สืบค้นเมื่อ 12 มิถุนายน 2562 จาก <http://www.autosoft.in.th/data-warehouse/เหมืองข้อมูล-data-mining/>.