

การวิเคราะห์รายวิชาเอกบังคับที่ส่งผลต่อภาวะการมีงานทำของบัณฑิต ด้วยเทคนิคต้นไม้ตัดสินใจ

Analysis of The Major Compulsory Subjects that Affected to The Employment Status of Graduates by a Decision Tree Technique

รณชัย ชื่นธวัช¹ กฤษณา วัฒนศักดิ์²

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอแนวทางการวิเคราะห์รายวิชาเอกบังคับที่ส่งผลต่อภาวะการมีงานทำของบัณฑิตด้วยเทคนิคต้นไม้ตัดสินใจ ซึ่งเป็นวิธีการด้านเหมืองข้อมูล เพื่อใช้จำแนกภาวะการมีงานทำ จากระดับผลการเรียนในรายวิชาเอกบังคับ ซึ่งกลุ่มตัวอย่างที่นำมาศึกษา คือ บัณฑิตจำนวน 489 ราย ที่จบการศึกษาในหลักสูตรวิทยาศาสตรบัณฑิต ปีการศึกษา 2549-2552 ในสาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา การดำเนินงานวิจัยนี้ ได้นำข้อมูลภาวะการมีงานทำของบัณฑิตที่เป็นกลุ่มตัวอย่างมาผสมผสานกับข้อมูลผลการเรียนรายวิชาเอกบังคับผ่านค่ารหัสนักศึกษาของบัณฑิต จากนั้น แบ่งข้อมูลตามกลุ่มรายวิชา เป็น 4 กลุ่ม แล้วนำข้อมูลแต่ละกลุ่มไปสร้างต้นไม้ตัดสินใจที่ดีที่สุดจากวิธีการประเมินแบบตรวจสอบสลับ 10 ส่วน (10 folds cross validation) และประเมินประสิทธิภาพของต้นไม้ตัดสินใจ ด้วยการวัดค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าการระลึกได้ (Recall)

ผลการวิจัย พบว่า ต้นแบบต้นไม้ตัดสินใจที่ดีที่สุดจากวิธีการประเมินแบบตรวจสอบสลับ 10 ส่วน ให้ค่าความถูกต้อง 78.55% ให้ค่าความแม่นยำ 79.55% และให้ค่าการระลึกได้ 75.72% เมื่อนำต้นแบบไปสร้างกฎความสัมพันธ์ สามารถระบุได้ว่า รายวิชาโครงงานคอมพิวเตอร์ (410714) เป็นรายวิชาที่ส่งผลต่อภาวะการมีงานทำมากที่สุด ดังนั้น ผลลัพธ์ที่ได้จากต้นแบบต้นไม้ตัดสินใจ สามารถนำไปใช้ประเมินความสอดคล้องระหว่างกลุ่มรายวิชาเอกบังคับของหลักสูตรและภาวะการมีงานทำของบัณฑิต และเป็นแนวทางในการพัฒนาหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ของกลุ่มตัวอย่างในงานวิจัยนี้ เพื่อให้อัตราการมีงานทำเพิ่มสูงขึ้นได้

คำสำคัญ : ภาวะการมีงานทำ รายวิชาเอกบังคับ เทคนิคต้นไม้ตัดสินใจ

¹ ผู้ช่วยศาสตราจารย์ สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา

² รองศาสตราจารย์(พิเศษ) สาขาวิชาการแพทย์แผนจีน คณะการแพทย์แผนจีน วิทยาลัยนครราชสีมา

Abstract

This research aims to present the way to analyze a group of major compulsory subjects that affected the employment status of graduates by a decision tree technique that is a data mining approach. A decision tree is used to classify the employment status of graduates from a grade of major compulsory subjects. A study example set is 489 graduated students that graduated from the science curriculum of academic year 2006 to 2009 in a field of computer science, faculty of science and technology Nakhon Ratchasima Rajabhat University. The data were used in this study came from the employment status of graduates data and integrated with a grade report data of the student examples through an identification code. After that, we divide data by a group of subjects and it can be divided into 4 groups. The data in each group is used to generate an optimized decision tree model with 10 folds cross validation and the evolution of our decision tree made through the accuracy, precision and recall.

The research result found that the optimized model from 10 folds cross validation gave an accuracy equal to 78.55%, a precision equal to 79.55% and a recall equal to 75.72%. The model generated association rules and indicated that the subject named “Computer Project (410714)” has most affected to the employment status of graduates. Therefore, we can conclude that this optimized decision tree model can be used to evaluate the consistency between a group of major compulsory subjects and the employment status of graduates. In addition, the model may be used to guideline a science curriculum development in the field of computer science for increasing a rate of the employment status of graduates.

Keywords : Employment status of graduates, Major compulsory subjects, Decision tree technique

บทนำ

การจัดการศึกษาในปัจจุบัน ได้ให้ความสำคัญกับรายวิชาที่มีความสอดคล้องกับความต้องการของตลาดแรงงาน ซึ่งภาวะการมีงานทำของบัณฑิต ที่ระบุอัตราการมีงานทำของบัณฑิต ถือเป็นดัชนีชี้วัดความสำเร็จของการพัฒนาหลักสูตรการเรียนการสอนในระดับอุดมศึกษาทุกหลักสูตร และมีผลให้หลักสูตรที่มีอัตราการมีงานทำในระดับสูง ได้รับความสนใจจากผู้เรียนอย่างมาก โดยมีปัจจัยจำนวน

มาก ที่ส่งผลต่ออัตราการมีงานทำ โดยงานวิจัยนี้ สนใจศึกษาปัจจัยด้านระดับผลการเรียนในรายวิชา เอกบังคับของบัณฑิตขณะที่เป็นนักศึกษาในหลักสูตรวิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ปีพ.ศ. 2549-2552 ของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏ นครราชสีมา ซึ่งระดับผลการเรียนในรายวิชาเอกบังคับ ที่หมายถึง “วิชาแกน วิชาเฉพาะด้าน วิชาพื้นฐานวิชาชีพและวิชาชีพ ที่มุ่งหมายให้ผู้เรียนมีความรู้ ความเข้าใจ และปฏิบัติงานได้” (ประกาศกระทรวงศึกษาธิการ เรื่องเกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548, 2548) น่าจะเป็นปัจจัยหนึ่ง ที่ส่งผลต่อภาวะการมีงานทำของบัณฑิต โดยรายวิชาเอกบังคับตามรายงานหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ปีพ.ศ. 2549 หมายถึง รายวิชาเฉพาะด้านคอมพิวเตอร์ ที่นักศึกษาในสาขาวิชาวิทยาการคอมพิวเตอร์ ต้องเรียนทุกรายวิชา มีจำนวนทั้งสิ้น 15 รายวิชา รวม 45 หน่วยกิต และรหัสวิชาเอกบังคับ จะขึ้นต้นด้วย “410” ส่วนตัวเลขลำดับที่ 4 ของรหัสวิชา คือ การแบ่งกลุ่มย่อยของรายวิชาเอกบังคับ ถ้าเป็นเลข 1 หมายถึง กลุ่มรายวิชาเอกบังคับด้านโครงสร้างพื้นฐานของระบบคอมพิวเตอร์ มีจำนวน 4 รายวิชา รวม 12 หน่วยกิต ถ้าเป็นเลข 2 หมายถึง กลุ่มรายวิชาเอกบังคับด้านการวิเคราะห์ระบบคอมพิวเตอร์ มีจำนวน 3 รายวิชา รวม 9 หน่วยกิต ถ้าเป็นเลข 3 หมายถึง กลุ่มรายวิชาเอกบังคับด้านข้อมูลและงานประยุกต์ มีจำนวน 3 รายวิชา รวม 9 หน่วยกิต ถ้าเป็นเลข 7 หมายถึง กลุ่มรายวิชาเอกบังคับด้านเทคโนโลยีและวิธีการทางซอฟต์แวร์ มีจำนวน 5 รายวิชา รวม 15 หน่วยกิต

วิธีการด้านเหมืองข้อมูล เป็นวิธีการวิเคราะห์ข้อมูลแบบหนึ่ง ที่มีความแตกต่างจากการวิเคราะห์ข้อมูลเชิงสถิติ ซึ่งการวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา ที่อธิบายลักษณะของกลุ่มข้อมูลในเชิงตัวเลขด้วยการวัดค่าเฉลี่ย ร้อยละ การแจกแจงความถี่ การวัดการกระจายของข้อมูล หรือการวิเคราะห์ข้อมูลด้วยสถิติเชิงอนุมาน ที่วิเคราะห์ข้อมูลกลุ่มตัวอย่าง เพื่อเป็นตัวแทนของประชากร ด้วยการประมาณค่า และการทดสอบสมมติฐาน ไม่สามารถค้นหาความสัมพันธ์ระหว่างข้อมูลที่ซ่อนอยู่ภายในกลุ่มข้อมูลได้ แต่การวิเคราะห์ข้อมูลด้วยวิธีการด้านเหมืองข้อมูล จะสามารถค้นหาความสัมพันธ์ที่ซ่อนอยู่ได้ ด้วยการค้นหารูปแบบ และกฎความสัมพันธ์ ที่แฝงอยู่ในข้อมูล สิ่งที่ได้รับจากการทำเหมืองข้อมูล คือ ความรู้ที่สามารถนำไปใช้ให้เกิดประโยชน์ได้ เช่น การสร้างกฎความสัมพันธ์ของสินค้าที่ลูกค้าเลือกซื้อ การวิเคราะห์รูปแบบการตัดสินใจทางการแพทย์ หรือการจัดกลุ่มลักษณะการใช้บริการโทรศัพท์เคลื่อนที่ของลูกค้า เป็นต้น สำหรับประเภทของการทำเหมืองข้อมูล สามารถแบ่งได้เป็น 2 ประเภท ได้แก่ 1) การเรียนรู้แบบมีผู้ฝึกสอน (Supervised learning) เป็นการสร้างตัวแบบในการทำนาย (Predictive model) เพื่อจำแนกข้อมูล (Classification) โดยนำข้อมูลในอดีต มาใช้สร้างตัวแบบทำนายคุณลักษณะที่สนใจในอนาคต ซึ่งเทคนิคที่จัดอยู่ในประเภทนี้คือ เทคนิคต้นไม้ตัดสินใจ (Decision tree) หรือเทคนิค Naïve Bayes เป็นต้น 2) การเรียนรู้แบบไม่มีผู้ฝึกสอน (Unsupervised learning) เป็นการสร้างตัวแบบในการบรรยาย (Descriptive model) เพื่อสร้างกฎความสัมพันธ์หรือจัดกลุ่มข้อมูล (Clustering) โดยนำข้อมูลในอดีต มาใช้สร้างตัวแบบ

จากนั้น เมื่อมีข้อมูลใหม่เข้ามา จะสามารถทราบได้ว่า มีความสัมพันธ์กับข้อมูลใด หรือควรจัดอยู่ในกลุ่มใด ซึ่งเทคนิคที่จัดอยู่ในประเภทนี้ คือ เทคนิคการจัดกลุ่มแบบ K-mean หรือการสร้างกฎความสัมพันธ์ด้วยเทคนิค APRIORI เป็นต้น

ปัจจุบันได้มีนักวิชาการหลายท่าน นำวิธีการด้านเหมืองข้อมูลไปใช้ เพื่อค้นหารูปแบบ หรือกฎความสัมพันธ์ที่แฝงอยู่ในข้อมูล ในศาสตร์ต่างๆ หลายด้าน ไม่ว่าจะเป็น ด้านกฎหมาย ด้านการเงิน หรือด้านการศึกษา ซึ่งจากการทบทวนวรรณกรรม สามารถนำตัวอย่างงานวิจัย ที่มีการนำวิธีการด้านเหมืองข้อมูล มาใช้ในการค้นหารูปแบบหรือกฎความสัมพันธ์ ได้ดังนี้

ชัชชัย แก้วตา และ อัจฉรา มหาวิวัฒน์ (2553) ศึกษาการวินิจฉัยคดี โดยประยุกต์ใช้เทคนิคต้นไม้ตัดสินใจ (Decision tree techniques) เพื่อจำแนกความผิด ซึ่งถูกอธิบายด้วยชุดของคุณลักษณะ (attributes) ออกเป็นมาตราต่างๆ ที่เหมาะสมกับคดีความนั้น และเปรียบเทียบความถูกต้องของการจำแนกข้อมูลด้วยขั้นตอนวิธี ID3 กับขั้นตอนวิธี C4.5 ผลการทดลอง พบว่า ต้นไม้ตัดสินใจที่สร้างจากขั้นตอนวิธี C4.5 สามารถจำแนกความผิดได้ถูกต้องมากกว่าขั้นตอนวิธี ID3 และสามารถนำไปประกอบการพิจารณาวินิจฉัยคดี

ทิพย์ธิดา วงศ์พิพันธ์ (2556) ศึกษาข้อมูลประวัติลูกค้าที่ผ่านการพิจารณาให้สินเชื่อของบริษัท กรุงไทยคาร์เรนท์ แอนด์ ลิสซิ่ง จำกัด (มหาชน) โดยมีวัตถุประสงค์งานวิจัย เพื่อนำความรู้ที่ได้จากการทำเหมืองข้อมูล มาใช้เป็นแนวทางสนับสนุนการตัดสินใจการอนุมัติสินเชื่อของบริษัทได้อย่างมีประสิทธิภาพมากขึ้น โดยใช้เทคนิคการจำแนกกลุ่ม เพื่อหากฎที่ใช้ในการจำแนกลูกค้ากลุ่มดี และลูกค้ากลุ่มไม่ดี ซึ่งได้ทดลองจำแนกกลุ่มข้อมูลด้วยเทคนิคการจำแนกกลุ่ม 3 เทคนิค ได้แก่ เทคนิค Decision tree: C4.5 เทคนิค Decision rule: Part และเทคนิค Naïve Bayes ผลการทดลอง พบว่า เทคนิค Decision tree: C4.5 ให้ผลลัพธ์ที่ดีที่สุด เนื่องจากมีความถูกต้องมากที่สุด และสามารถแบ่งกลุ่มได้ตามเงื่อนไขที่ชัดเจนมากที่สุด จากนั้น นำกฎที่ได้จากเทคนิคดังกล่าว มาเป็นแนวทางและเงื่อนไขในการนำมาเป็นตัวอย่างที่สำคัญ สำหรับการเขียนโปรแกรมเป็นระบบการตัดสินใจอนุมัติสินเชื่อออนไลน์ได้อย่างมีประสิทธิภาพ

บุญมา เฟ่งชวน (2548) ศึกษาข้อมูลภาวะการมีงานทำของบัณฑิต โดยนำมาวิเคราะห์ด้วยเทคนิคเหมืองข้อมูล คือ การค้นหากฎความสัมพันธ์ ซึ่งมีวัตถุประสงค์เพื่อศึกษา และพัฒนาตัวแบบเพื่อใช้ในการทำนายแนวโน้มเลือกอาชีพแรกหลังสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรี และนำข้อมูลนักศึกษารายบุคคลของสถาบันอุดมศึกษาของรัฐ ปีการศึกษา 2548 มาวิเคราะห์กับตัวแบบที่สร้างได้ และนำเสนอในรูปแบบตารางและกราฟเพื่อนำไปสนับสนุนการตัดสินใจในด้านการผลิตบัณฑิต ผลการวิจัย พบว่า ตัวแบบที่พัฒนาขึ้น มีค่าเฉลี่ยร้อยละความถูกต้องเป็น 65.39 ถ้าแบ่งสัดส่วนข้อมูล Training set เป็นร้อยละ 60 และค่าเฉลี่ยร้อยละความถูกต้องเป็น 74.72 ถ้าแบ่งสัดส่วนข้อมูล Training set เป็นร้อยละ 70 และค่าเฉลี่ยร้อยละความถูกต้องเป็น 81.89 ถ้าแบ่งสัดส่วนข้อมูล

Training set เป็นร้อยละ 80 ดังนั้น สามารถนำตัวแบบมาหาแนวโน้มการเลือกอาชีพแรกหลังสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรีได้

ไพฑูรย์ จันทร์เรือง (2550) ศึกษาการพัฒนาาระบบสนับสนุนการตัดสินใจเลือกสาขาการเรียนของนักศึกษาระดับปริญญาตรีโดยใช้เทคนิคต้นไม้ตัดสินใจ ซึ่งจากการทดลองพบว่าในการสร้างตัวแบบสำหรับพัฒนาระบบสนับสนุนการตัดสินใจเลือกสาขาการเรียนของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิคต้นไม้ตัดสินใจนั้น ควรแยกสร้างตัวแบบสำหรับแต่ละสาขาการเรียน เนื่องจากคุณสมบัติของผู้เรียนแต่ละสาขา มีความแตกต่างกัน เพื่อให้ได้ตัวแบบที่สามารถทำนายแนวโน้มของผลการเรียนที่เหมาะสมสำหรับแต่ละสาขา แต่เนื่องจากคะแนนเฉลี่ยของนักศึกษาที่นำมาพัฒนาตัวแบบนั้น ส่วนใหญ่จะมีเกณฑ์คะแนนเกาะกลุ่มกันอยู่ในช่วงกลางของข้อมูล(2.00 – 3.00) ทำให้ผลการตัดสินใจส่วนใหญ่จะโน้มเอียงไปในเกณฑ์พอใช้ (ช่วงคะแนน 2.00 – 2.49) และปานกลาง (ช่วงคะแนน 2.50 – 2.99)

เรวดี ศักดิ์ดุจรธรรม (2553) ศึกษาการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยการประยุกต์ใช้เทคนิคคตาต้นไม้ในการสร้างฐานความรู้ เพื่อนำไปใช้ประกอบการตัดสินใจให้คำปรึกษาทางการเรียนของอาจารย์ที่ปรึกษา ซึ่งเป็นการทำนายเพื่อหาความถนัดของนักศึกษาว่าควรเรียนในสาขาวิชาใด โดยนำข้อมูลส่วนตัวและข้อมูลผลการลงทะเบียนของนักศึกษาจากคณะบริหารธุรกิจ ทั้ง 4 สาขา คือ สาขาวิชาคอมพิวเตอร์ธุรกิจ สาขาวิชาการตลาดสาขาวิชาการจัดการ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว มาทำการสร้างโมเดล 2 โมเดล คือ โมเดลสำหรับวิเคราะห์หาพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา และโมเดลที่ใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา ผลการทดลอง พบว่า โมเดลวิเคราะห์หาพฤติกรรมทางการเรียน บอกได้ว่า ผลการเรียนกลุ่มวิชาแกนการเงิน มีผลต่อผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาคอมพิวเตอร์ธุรกิจ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยวมากที่สุด และผลการเรียนกลุ่มวิชาแกนธุรกิจ มีผลต่อผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาการตลาดและการจัดการมากที่สุด ส่วนโมเดลทำนายผลการเรียนของสาขาวิชาคอมพิวเตอร์ มีความถูกต้องคิดเป็น 73.49% โมเดลทำนายผลการเรียนของสาขาวิชาการตลาด มีความถูกต้องคิดเป็น 83.58% โมเดลทำนายผลการเรียนของสาขาวิชาการจัดการ มีความถูกต้องคิดเป็น 78.12% และโมเดลทำนายผลการเรียนของสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยวมีความถูกต้องคิดเป็น 86.67%

จากที่กล่าวมา การค้นหาความสัมพันธ์ระหว่างระดับผลการเรียนในรายวิชาเอกบังคับของบัณฑิตขณะเป็นนักศึกษาและภาวะการมีงานทำของบัณฑิต มีความสำคัญในการปรับปรุงรายวิชาเอกบังคับในหลักสูตรที่ช่วยให้ภาวะการมีงานทำเพิ่มขึ้นได้ แต่การศึกษาและวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนาหรือสถิติเชิงอนุมาน ไม่สามารถค้นหารูปแบบหรือกฎความสัมพันธ์ที่แฝงอยู่ได้ ดังนั้นงานวิจัยนี้ จึงนำเสนอแนวทางการวิเคราะห์ระดับผลการเรียนในรายวิชาเอกบังคับที่ส่งผลต่อภาวะการมีงานทำของบัณฑิตด้วยวิธีการด้านเหมืองข้อมูล โดยใช้เทคนิคต้นไม้ตัดสินใจ เพื่อจำแนกภาวะการมีงานทำ จากระดับผลการเรียนในรายวิชาบังคับ ซึ่งการเลือกเทคนิคต้นไม้ตัดสินใจมาใช้ใน

การดำเนินการวิจัย เนื่องจากเป็นเทคนิคที่นิยมใช้และให้ความถูกต้องในระดับสูง ดังตัวอย่างงานวิจัยที่ได้กล่าวไว้ข้างต้น

วัตถุประสงค์การวิจัย

1. เพื่อนำเสนอแนวทางการวิเคราะห์ความสัมพันธ์ระหว่างระดับผลการเรียนในรายวิชาเอก บัณฑิตและภาวะการปฏิบัติงานทำของบัณฑิต โดยใช้วิธีการด้านเหมืองข้อมูล ด้วยเทคนิคต้นไม้ตัดสินใจ
2. เพื่อสร้างตัวแบบต้นไม้ตัดสินใจที่ดีที่สุดด้วยการตรวจสอบแบบสลับ 10 ส่วน สำหรับใช้สนับสนุนการตัดสินใจในการปรับปรุงกลุ่มรายวิชาเอกบัณฑิตที่ส่งผลต่อภาวะการปฏิบัติงานทำของบัณฑิต

ขอบเขตของปัญหา

การวิจัยนี้มีขอบเขตการศึกษาระดับผลการเรียนในรายวิชาเอกบัณฑิตและภาวะการปฏิบัติงานทำของบัณฑิตที่ผ่านการศึกษาในหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ปีพ.ศ. 2549-2552 ของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา จำนวน 489 คน ด้วยเทคนิคต้นไม้ตัดสินใจ ซึ่งเป็นวิธีการด้านเหมืองข้อมูล และประเมินประสิทธิภาพของตัวแบบ โดยใช้วิธีตรวจสอบแบบสลับ 10 ส่วน ด้วยการวัดค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าการระลึกได้ (Recall)

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยแบบผสมผสาน ด้วยการนำวิธีการด้านเหมืองข้อมูล ซึ่งเป็นวิธีการด้านวิทยาการคอมพิวเตอร์มาใช้ในการวิเคราะห์รายวิชาเอกบัณฑิตที่ส่งผลเชิงบวกต่อภาวะการปฏิบัติงานทำของบัณฑิต เพื่อเป็นแนวทางสนับสนุนการตัดสินใจในการบริหารการศึกษาให้เกิดประสิทธิภาพสอดคล้องกับพันธกิจของมหาวิทยาลัย เรื่องการผลิตบัณฑิตที่มีคุณภาพ และเป็นที่ต้องการของผู้ใช้บัณฑิต โดยวิธีการด้านเหมืองข้อมูล เป็นวิธีการหนึ่งที่มีความแตกต่างจากการวิเคราะห์ด้วยสถิติเชิงพรรณนาหรือสถิติเชิงอนุมาน ดังที่ได้กล่าวมาข้างต้น ซึ่งเทคนิคหนึ่งที่ใช้ในการทำเหมืองข้อมูล คือ เทคนิคต้นไม้ตัดสินใจ (Decision tree) โดยเป็นเทคนิคที่ใช้ในการจำแนกประเภท (Class) ของข้อมูล ซึ่งใช้วิธีการประมาณฟังก์ชันที่มีค่าไม่ต่อเนื่อง (discrete-value function) ด้วยแผนผังต้นไม้ แล้วนำมาอธิบายในรูปแบบของกฎ “ถ้า-แล้ว” (if-then rule) เพื่อให้สามารถอ่านและเข้าใจการตัดสินใจของต้นไม้ได้ ซึ่งมีการนำเสนอขั้นตอนวิธีการสร้างต้นไม้ตัดสินใจโดย John Ross Quinlan (1986) โดยใช้การวัดค่าความไม่บริสุทธิ์ของข้อมูล เรียกว่า เอนโทรปี (Entropy) ที่นิยามได้ดังสมการที่ 1 (Quinlan, 1986) จากนั้น นำไปคำนวณค่าคาดหวังความรู้ (Information gain) ที่นิยามได้ดังสมการที่ 2 (Quinlan, 1986) แล้วเลือกคุณลักษณะ (Attribute) ที่มีค่าคาดหวังความรู้สูงที่สุด มาเป็นโหนดราก (Root node) ของต้นไม้ตัดสินใจ แล้วดำเนินการค้นหาโหนดในลำดับถัดไปของต้นไม้ตัดสินใจ

ด้วยการเขียนโปรแกรมในลักษณะแบบเวียนเกิด (Recursive) จนกระทั่ง โหนดใบไม้ (Leaf node) ทุกโหนด คือ ค่าการจำแนกประเภท (Class) และตัวอย่างของต้นไม้ตัดสินใจ แสดงได้ดังรูปที่ 1

$$E(s) = - \sum_{j=1}^n P_s(j) \log_2 P_s(j) \tag{1}$$

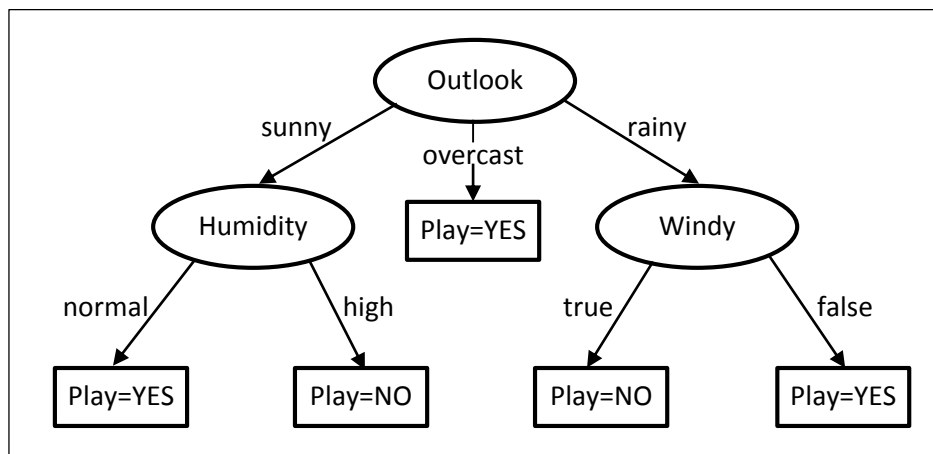
$$\text{Information Gain (class, Attribute)} = \text{Entropy (class)} - \text{Entropy (Attribute)} \tag{2}$$

เมื่อ

E(s) คือ เอนโทรปีของ s

s คือ ตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปรตามหลาย ๆ กรณี

P_s(j) คือ อัตราส่วนของกรณีใน S ที่ตัวแปรตามหรือผลลัพธ์มีค่า j



รูปที่ 1 ตัวอย่างของต้นไม้ตัดสินใจในการออกไปเล่นกีฬาจากสภาพอากาศ

การวิจัยนี้ ใช้วิธีการทดสอบความถูกต้องของตัวแบบต้นไม้ตัดสินใจ ด้วยวิธีตรวจสอบแบบ สลับ 10 ส่วน (10 folds cross validation) เนื่องจากเป็นวิธีการตรวจสอบที่มีความน่าเชื่อถือสูง เพราะข้อมูลทุกตัวจะถูกนำไปใช้ฝึกสอนตัวแบบ จากนั้น จะถูกสลับไปใช้ทดสอบตัวแบบด้วย โดยการ วัดค่าความถูกต้องของตัวแบบ จะเปรียบเทียบกับชุดข้อมูลทดสอบ ด้วยการสร้างตารางประเมินผล การทำนาย เรียกว่า ตาราง Confision matrix ดังรูปที่ 2 (Fawcett, 2006) จากนั้น นำไปคำนวณค่า ความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ดังสมการที่ 3-5 (Fawcett, 2006) ตามลำดับ

$$\text{Accuracy} = (TP+TN)/(TP+FP+FN+TN) \tag{3}$$

$$\text{Precision} = TP/(TP+FP) \tag{4}$$

$$\text{Recall} = TP/(TP+FN) \tag{5}$$

เมื่อ

TP = จำนวนผลการทำนายว่าเป็นคลาสบวก (Positive class/Yes) และข้อมูลทดสอบเป็นคลาสบวก (Positive class/Yes) ผลการทำนายเป็นจริง (True)

FP = จำนวนผลการทำนายว่าเป็นคลาสบวก (Positive class/Yes) แต่ข้อมูลทดสอบเป็นคลาสลบ (Negative class/No) ผลการทำนายเป็นเท็จ (False)

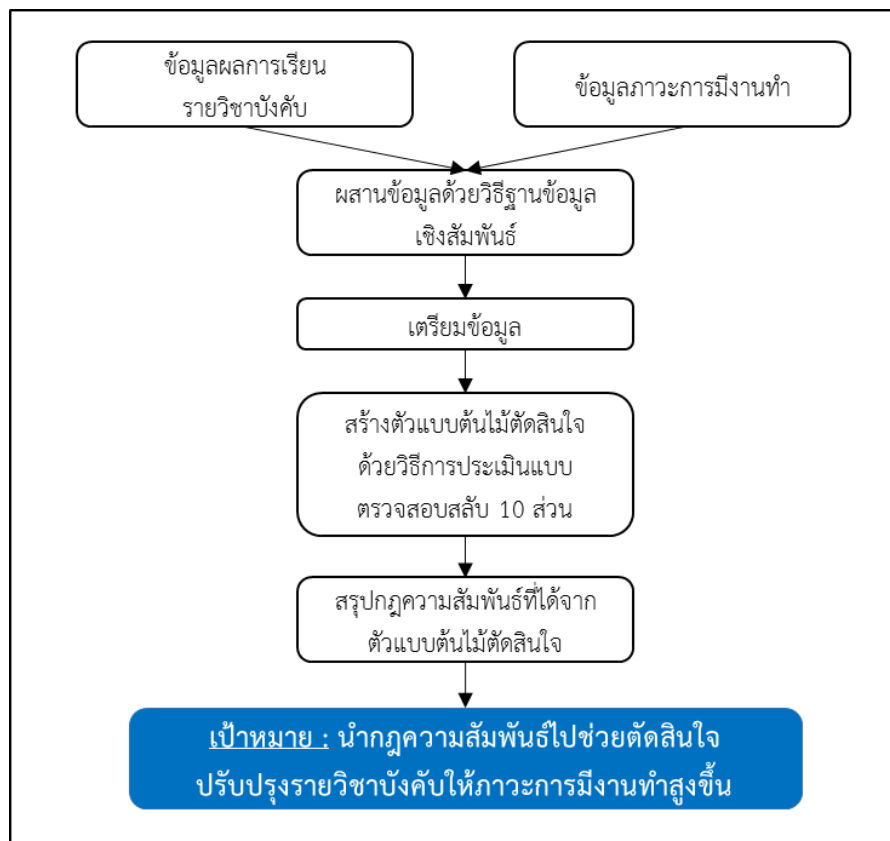
FN = จำนวนผลการทำนายว่าเป็นคลาสลบ (Negative class/No) แต่ข้อมูลทดสอบเป็นคลาสบวก (Positive class/Yes) ผลการทำนายเป็นเท็จ (False)

TN = จำนวนผลการทำนายว่าเป็นคลาสลบ (Negative class/No) และข้อมูลทดสอบเป็นคลาสลบ (Negative class/No) ผลการทำนายเป็นจริง (True)

		True class	
		Positive (Yes)	Negative (No)
Hypothesized class	Yes	True Positives (TP)	False Positives (FP)
	No	False Negatives (FN)	True Negatives (TN)

รูปที่ 2 Confusion matrix

การวิจัยนี้มีวัตถุประสงค์เพื่อ วิเคราะห์ระดับผลการเรียนรายวิชาบังคับของบัณฑิตหลักสูตรวิทยาศาสตร์บัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา ที่ส่งผลกระทบต่อภาวะการมีงานทำของบัณฑิต ซึ่งใช้ชุดข้อมูล 2 ชุด คือ 1) ข้อมูลระดับผลการเรียนในรายวิชาบังคับขณะที่บัณฑิตมีสถานะเป็นนักศึกษา และ 2) ข้อมูลภาวะการมีงานทำของบัณฑิต โดยนำข้อมูลทั้ง 2 ชุด มาผสมกันด้วยวิธีการของฐานข้อมูลเชิงสัมพันธ์ผ่านคาร์ทสัณศึกษา จากนั้น ดำเนินการวิเคราะห์ข้อมูลด้วยวิธีด้านเหมืองข้อมูล โดยใช้เทคนิคต้นไม้ตัดสินใจ เพื่อจำแนกภาวะการมีงานทำ จากระดับผลการเรียนในรายวิชาบังคับ ภายใต้กรอบแนวคิดงานวิจัย ดังรูปที่ 3



รูปที่ 3 กรอบแนวคิดงานวิจัย

จากกรอบแนวคิดงานวิจัย อธิบายวิธีการดำเนินงานวิจัยตามกรอบแนวคิด ได้ดังนี้

1. ขอข้อมูลภาวะการมีงานทำและข้อมูลระดับผลการเรียนในรายวิชาบังคับของบัณฑิตที่ผ่านการเรียนหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ ปีพ.ศ. 2549 จากสำนักคอมพิวเตอร์ มหาวิทยาลัยราชภัฏนครราชสีมา ซึ่งเป็นหน่วยงานที่เป็นผู้ดูแลฐานข้อมูลที่จัดเก็บข้อมูลดังกล่าวไว้

2. ผสานข้อมูลทั้ง 2 ชุด ด้วยวิธีด้านฐานข้อมูลเชิงสัมพันธ์ โดยใช้โปรแกรม Microsoft access เป็นเครื่องมือ และใช้รหัสนักศึกษาเป็นตัวเชื่อมความสัมพันธ์ ซึ่งข้อมูลภาวะการมีงานทำ สนใจเฉพาะค่าสถานะการมีงานทำ เพื่อนำมาใช้กำหนดเป็น Class attribute ในขั้นตอนการฝึกสอนตัวแบบต้นไม้ตัดสินใจ ดังรูปที่ 5

3. ขั้นตอนเตรียมข้อมูล เพื่อทำความสะอาดข้อมูล โดยจัดกลุ่มข้อมูลระดับผลการเรียนในรายวิชาเอกบังคับ ตามค่าตัวเลขลำดับที่ 4 ของรหัสวิชา ได้เป็น 4 กลุ่ม ดังที่ได้กล่าวไว้ในบทนำ ก่อนนำไปฝึกสอนตัวแบบต้นไม้ตัดสินใจ แยกตามกลุ่มวิชาเอกบังคับ

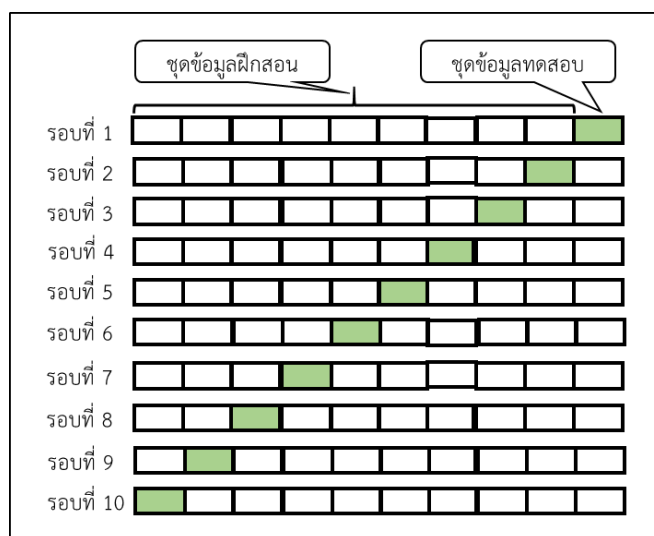
4. สร้างตัวแบบต้นไม้ตัดสินใจ แยกตามกลุ่มวิชาเอกบังคับ ด้วยวิธีการตรวจสอบสลับ 10 ส่วน ซึ่งจะมีการสร้างตัวแบบต้นไม้ตัดสินใจ จำนวน 10 รอบ แต่ละรอบ จะมีการแบ่งข้อมูลออกเป็น

10 ส่วน แต่ละส่วน มีจำนวนข้อมูลเท่ากัน โดยใช้ข้อมูล 9 ส่วน เป็นชุดข้อมูลฝึกสอนสำหรับสร้างตัวแบบ แล้วใช้ข้อมูล 1 ส่วน เป็นชุดข้อมูลทดสอบสำหรับประเมินความถูกต้องของตัวแบบ ซึ่งข้อมูลทดสอบจะถูกสลับไปเรื่อย ๆ ในแต่ละรอบ สุดท้าย เลือกตัวแบบที่มีความถูกต้องมากที่สุดจากทั้ง 10 รอบ มาสรุปเป็นกฎความสัมพันธ์ และวิธีการประเมินตัวแบบ ด้วยการตรวจสอบแบบสลับ 10 ส่วน แสดงได้ดังรูปที่ 6

5. สรุปกฎความสัมพันธ์ที่ได้จากตัวแบบต้นไม้ตัดสินใจที่มีความถูกต้องมากที่สุด ของวิธีการตรวจสอบแบบสลับ 10 ส่วน

Field:	Course	StudentID	ClassJob	S410101	S410102	S410103
Table:	AllClassJob	AllClassJob	AllClassJob	HaveJobCom	HaveJobCom	HaveJobCom
Sort:						
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:						
or:						

รูปที่ 5 การผสมข้อมูลภาวะการมีงานทำและข้อมูลระดับผลการเรียนเข้าด้วยกัน



รูปที่ 6 การแบ่งข้อมูลเพื่อสร้างต้นไม้ตัดสินใจด้วยวิธีการตรวจสอบแบบสลับ 10 ส่วน

ผลการวิจัย

เมื่อดำเนินการตามกรอบแนวคิดงานวิจัย หลังจากนำข้อมูลภาวะการปฏิบัติงานทำและข้อมูลผลการเรียน มาผสานเข้าด้วยกัน จากนั้นแบ่งข้อมูลเป็น 4 ชุด ตามตัวเลขลำดับที่ 4 ของรหัสวิชา และแต่ละกลุ่มมีความหมาย ดังนี้

1. กลุ่มที่ 1 คือกลุ่มที่มีตัวเลขลำดับที่ 4 ของรหัสวิชา เป็นเลข 1 ประกอบด้วย รหัสวิชา 410101, 410102, 410103 และ 410105

2. กลุ่มที่ 2 คือกลุ่มที่มีตัวเลขลำดับที่ 4 ของรหัสวิชา เป็นเลข 2 ประกอบด้วย รหัสวิชา 410201, 410202 และ 410203

3. กลุ่มที่ 3 คือกลุ่มที่มีตัวเลขลำดับที่ 4 ของรหัสวิชา เป็นเลข 3 ประกอบด้วย รหัสวิชา 410301, 410303 และ 410304

4. กลุ่มที่ 4 คือกลุ่มที่มีตัวเลขลำดับที่ 4 ของรหัสวิชา เป็นเลข 7 ประกอบด้วย รหัสวิชา 410701, 410702, 410703, 410713 และ 410714

เมื่อแบ่งข้อมูลตามกลุ่มวิชา จะนำข้อมูลแต่ละกลุ่มมาทำความสะอาดข้อมูล โดยกำจัดค่าสูญหาย (Missing value) ด้วยวิธีการลบแถวข้อมูลที่เกิดค่าสูญหาย หลังจากกำจัดค่าสูญหาย ข้อมูลแต่ละกลุ่ม สรุปได้ดังรูปที่ 7-10.

Id	StudentID	Polynomial	0	Least	5440207336 (1)	Most	4930207101 (1)	Values	4930207101 (1), 4930207202 (1), ...[399 more]
Label	ClassJob	Polynomial	0	Least	no (159)	Most	yes (242)	Values	yes (242), no (159)
✓	S410101	Polynomial	0	Least	B+ (23)	Most	C (112)	Values	C (112), D+ (79), ...[5 more]
✓	S410102	Polynomial	0	Least	A (22)	Most	C (104)	Values	C (104), C+ (93), ...[5 more]
✓	S410103	Polynomial	0	Least	A (16)	Most	C (120)	Values	C (120), D+ (80), ...[5 more]
✓	S410105	Polynomial	0	Least	D (13)	Most	C (109)	Values	C (109), C+ (90), ...[5 more]

รูปที่ 7 สรุปข้อมูลกลุ่มที่ 1 หลังจากกำจัดค่าสูญหาย

Id	StudentID	Polynomial	0	Least	5440207336 (1)	Most	4930207101 (1)	Values	4930207101 (1), 4930207202 (1), ...[399 more]
Label	ClassJob	Polynomial	0	Least	no (159)	Most	yes (242)	Values	yes (242), no (159)
✓	S410201	Polynomial	0	Least	D (15)	Most	C+ (107)	Values	C+ (107), B (94), ...[5 more]
✓	S410202	Polynomial	0	Least	A (2)	Most	D (178)	Values	D (178), D+ (114), ...[5 more]
✓	S410203	Polynomial	0	Least	A (14)	Most	C (109)	Values	C (109), C+ (92), ...[5 more]

รูปที่ 8 สรุปข้อมูลกลุ่มที่ 2 หลังจากกำจัดค่าสูญหาย

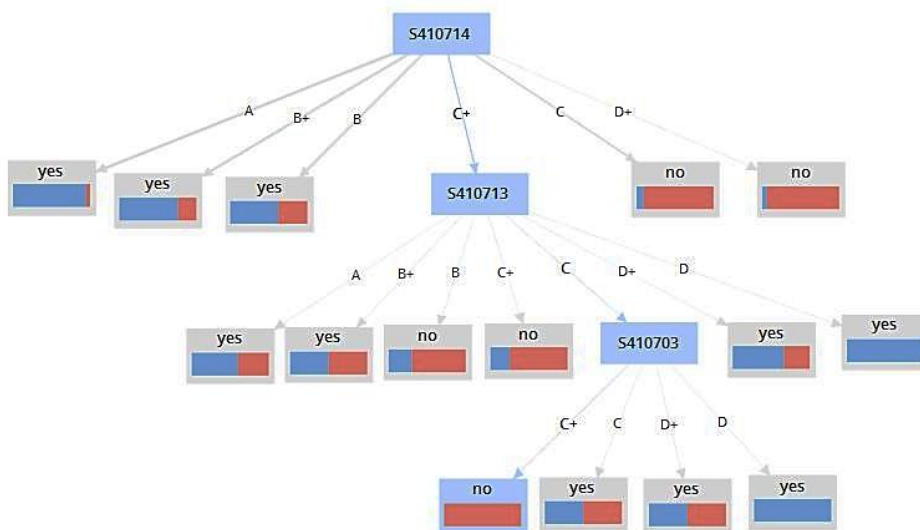
StudentID	Polynomial	0	Least 5440207336 (1)	Most 4930207101 (1)	Values 4930207101 (1), 4930207202 (1), ...[399 more]
ClassJob	Polynomial	0	Least no (159)	Most yes (242)	Values yes (242), no (159)
S410301	Polynomial	0	Least B+ (14)	Most C (102)	Values C (102), D+ (102), ...[5 more]
S410303	Polynomial	0	Least D (27)	Most C (85)	Values C (85), C+ (80), ...[5 more]
S410304	Polynomial	0	Least A (41)	Most C (76)	Values C (76), D (68), ...[5 more]

รูปที่ 9 สรุปข้อมูลกลุ่มที่ 3 หลังจากกำจัดค่าสูญหาย

StudentID	Polynomial	0	Least 5440207336 (1)	Most 4930207101 (1)	Values 4930207101 (1), 4930207202 (1), ...[399 more]
ClassJob	Polynomial	0	Least no (159)	Most yes (242)	Values yes (242), no (159)
S410701	Polynomial	0	Least B+ (13)	Most C (95)	Values C (95), D+ (88), ...[5 more]
S410702	Polynomial	0	Least B+ (14)	Most C (103)	Values C (103), D+ (96), ...[5 more]
S410703	Polynomial	0	Least A (11)	Most C (100)	Values C (100), D+ (93), ...[5 more]
S410713	Polynomial	0	Least D (31)	Most C (92)	Values C (92), B (81), ...[5 more]
S410714	Polynomial	0	Least D+ (16)	Most A (100)	Values A (100), B+ (88), ...[4 more]

รูปที่ 10 สรุปข้อมูลกลุ่มที่ 4 หลังจากกำจัดค่าสูญหาย

หลังจากกำจัดค่าสูญหายของแต่ละกลุ่มข้อมูล จึงนำแต่ละกลุ่มข้อมูลไปสร้างต้นไม้ตัดสินใจ ด้วยวิธีการประเมินแบบตรวจสอบสลับ 10 ส่วน ผลการทดลอง พบว่า มีเฉพาะข้อมูลกลุ่มที่ 4 เท่านั้นที่สามารถสร้างต้นไม้ตัดสินใจได้ ดังรูปที่ 11



รูปที่ 11 ต้นไม้ตัดสินใจของข้อมูลกลุ่มที่ 4

เมื่อได้ต้นไม้ตัดสินใจของข้อมูลกลุ่มที่ 4 จึงนำมาประเมินประสิทธิภาพของตัวแบบต้นไม้ตัดสินใจ ด้วยการวัดค่า Accuracy ค่า Precision และค่า Recall ได้ผลดังรูปที่ 12 ซึ่งค่าการประเมินประสิทธิภาพของตัวแบบฯ อยู่ในเกณฑ์ที่ยอมรับได้ จึงนำไปสร้างเป็นกฎความสัมพันธ์ เพื่อนำไปช่วยในการตัดสินใจปรับปรุงรายวิชาเอกบังคับ ที่ส่งผลกระทบต่อภาวะการมีงานทำ ได้ดังรูปที่ 13

PerformanceVector	
PerformanceVector:	
accuracy: 78.55% +/- 3.93% (mikro: 78.55%)	
ConfusionMatrix:	
True:	yes no
yes:	216 60
no:	26 99
weighted_mean_recall: 75.72% +/- 5.24% (mikro: 75.76%), weights: 1, 1	
ConfusionMatrix:	
True:	yes no
yes:	216 60
no:	26 99
weighted_mean_precision: 79.55% +/- 4.41% (mikro: 78.73%), weights: 1, 1	
ConfusionMatrix:	
True:	yes no
yes:	216 60
no:	26 99

รูปที่ 12 ค่าการประเมินประสิทธิภาพของตัวแบบต้นไม้ตัดสินใจ

Tree
S410714 = A: yes {yes=95, no=5}
S410714 = B: yes {yes=53, no=32}
S410714 = B+: yes {yes=67, no=21}
S410714 = C: no {yes=6, no=58}
S410714 = C+
S410713 = A: yes {yes=3, no=2}
S410713 = B: no {yes=3, no=7}
S410713 = B+: yes {yes=1, no=1}
S410713 = C
S410703 = C: yes {yes=1, no=1}
S410703 = C+: no {yes=0, no=3}
S410703 = D: yes {yes=2, no=0}
S410703 = D+: yes {yes=1, no=1}
S410713 = C+: no {yes=4, no=12}
S410713 = D: yes {yes=3, no=0}
S410713 = D+: yes {yes=2, no=1}
S410714 = D+: no {yes=1, no=15}

รูปที่ 13 กฎความสัมพันธ์ที่ได้จากต้นไม้ตัดสินใจของข้อมูลกลุ่มที่ 4

อภิปรายผลการวิจัย

การวิจัยนี้ มีประเด็นสำคัญ คือ การวิเคราะห์รายวิชาเอกบังคับในหลักสูตรวิทยาศาสตร์บัณฑิต ปีพ.ศ. 2549 สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา ที่ส่งผลกระทบต่อภาวะการมีงานทำ ผ่านการทำเหมืองข้อมูล โดยใช้เทคนิคต้นไม้ตัดสินใจและนำมาอภิปรายผลตามวัตถุประสงค์ของการวิจัยที่ตั้งไว้ 2 ข้อ ได้ดังนี้

1. สามารถสร้างแนวทางการวิเคราะห์ความสัมพันธ์ระหว่างระดับผลการเรียนในรายวิชาเอกบังคับและภาวะการมีงานทำของบัณฑิต ด้วยเทคนิคต้นไม้ตัดสินใจ โดยดำเนินการวิจัยด้วยการนำระดับผลคะแนนในรายวิชาเอกบังคับ มาผสานกับภาวะการมีงานทำของบัณฑิต ด้วยการเชื่อมความสัมพันธ์ผ่านรหัสนักศึกษาตามหลักการของระบบฐานข้อมูลเชิงสัมพันธ์ จากนั้นแบ่งข้อมูลตามกลุ่มรายวิชาเป็น 4 กลุ่ม แล้วทำความสะอาดข้อมูลด้วยการกำจัดค่าสูญหาย จากนั้น จึงนำข้อมูลแต่ละกลุ่ม มาสร้างตัวแบบต้นไม้ตัดสินใจ ผลการวิจัย พบว่า เมื่อดำเนินการวิจัยตามแนวทางดังกล่าว สามารถสร้างต้นไม้ตัดสินใจที่ดีที่สุด จากการประเมินด้วยการตรวจสอบแบบสลับ 10 ส่วน ได้ในข้อมูลกลุ่มที่ 4 ซึ่งเป็นกลุ่มรายวิชาเอกบังคับด้านเทคโนโลยีและวิธีการทางซอฟต์แวร์ โดยมีจำนวน 5 รายวิชา รวม 15 หน่วยกิต ส่วนข้อมูลอีก 3 กลุ่มที่เหลือ ไม่สามารถสร้างต้นไม้ตัดสินใจได้ ซึ่งจากผลการวิจัย วิเคราะห์ได้ว่า ผลการเรียนรายวิชาเอกบังคับ กลุ่มที่ 1 ถึง 3 ไม่มีความสัมพันธ์ต่อภาวะการมีงานทำของบัณฑิตอย่างมีนัยยะ มีเฉพาะผลการเรียนรายวิชาเอกบังคับ ในกลุ่มที่ 4 เท่านั้นที่มีความสัมพันธ์และสามารถนำไปใช้ทำนายภาวะการมีงานทำของบัณฑิตได้อย่างมีนัยยะ เนื่องจากสามารถสร้างต้นไม้ตัดสินใจได้ โดยประเมินประสิทธิภาพของตัวแบบต้นไม้ตัดสินใจด้วยการวัดค่า Accuracy ค่า Precision และค่า Recall ซึ่งพบว่า ค่า Accuracy ที่หมายถึง การทำนายว่ามีการมีงานทำ (คลาสบวก) และไม่มีการมีงานทำ (คลาสลบ) ได้ถูกต้อง เมื่อเทียบกับข้อมูลทั้งหมด วัดค่าได้เท่ากับ 78.55% ส่วนค่า Precision ที่หมายถึง การทำนายว่ามีการมีงานทำ (คลาสบวก) ได้ถูกต้อง เมื่อเทียบกับจำนวนการทำนายว่ามีการมีงานทำทั้งหมด วัดค่าได้เท่ากับ 79.55% และค่า Recall ที่หมายถึง การทำนายว่ามีการมีงานทำได้ถูกต้อง เมื่อเทียบกับจำนวนข้อมูลที่มีการมีงานทำทั้งหมด วัดค่าได้เท่ากับ 75.72% ทั้ง 3 ค่า มีเปอร์เซ็นต์ความถูกต้องอยู่ในช่วง 75-80% ซึ่งเป็นช่วงเกณฑ์ที่ยอมรับได้ ดังนั้น แนวทางการวิเคราะห์ความสัมพันธ์ระหว่างระดับผลการเรียนในรายวิชาเอกบังคับและภาวะการมีงานทำของบัณฑิต โดยใช้วิธีการด้านเหมืองข้อมูล ด้วยเทคนิคต้นไม้ตัดสินใจ ที่งานวิจัยนี้ได้นำเสนอ สามารถนำไปใช้ประโยชน์ได้อย่างมีนัยยะ

2. เมื่อสร้างตัวแบบต้นไม้ตัดสินใจที่ดีที่สุด จากการประเมินด้วยการตรวจสอบแบบสลับ 10 ส่วน สำหรับข้อมูลกลุ่มที่ 4 โดยมีผลการประเมินประสิทธิภาพของต้นไม้ตัดสินใจอยู่ในเกณฑ์ที่ยอมรับได้ จึงนำมาสร้างกฎความสัมพันธ์ เพื่อช่วยตัดสินใจว่า รายวิชาใดส่งผลกระทบต่อภาวะการมีงานทำ ดังรูปที่ 13 จากกฎความสัมพันธ์ที่ได้ อภิปรายผลได้ว่า รายวิชาที่มีรหัสวิชาเป็น 410714 ซึ่งหมายถึง รายวิชา วิศวกรรมคอมพิวเตอร์ ซึ่งเป็นรายวิชาที่นักศึกษาชั้นปีที่ 4 ทุกคน จะต้องลงทะเบียนเรียน เป็นรายวิชา

ที่มีความสัมพันธ์กับภาวะการมีงานทำมากที่สุด ถ้าได้ระดับผลการเรียนในรายวิชา 410714 อยู่ใน ระดับ A, B+ หรือ B จะทำนายได้ว่า บัณฑิตมีแนวโน้มว่าได้งานทำ ขณะเดียวกัน ถ้าได้ระดับผลการ เรียน C หรือ D+ จะทำนายได้ว่า บัณฑิตมีแนวโน้มว่าไม่มีงานทำ ที่ระดับความแม่นยำในช่วง 75- 80% เมื่อพิจารณาคำอธิบายรายวิชาโครงการคอมพิวเตอร์ (410714) จากคู่มือการจัดการเรียนการ สอน สาขาวิชาวิทยาการคอมพิวเตอร์ (วท.บ.) ปีการศึกษา 2549-2552 มหาวิทยาลัยราชภัฏ นครราชสีมา ระบุไว้ว่า “ศึกษาค้นคว้าด้วยตนเอง (เป็นกลุ่มหรือศึกษาเดี่ยว) เพื่อนำไปสู่การนำเสนอ ผลการศึกษาแก้ปัญหาเชิงปฏิบัติการ หรือการทดลองวิเคราะห์และเปรียบเทียบเชิงวิชาการ ใน สาขาวิชาวิทยาการคอมพิวเตอร์ ภายใต้การดูแลของอาจารย์ที่ปรึกษา” สามารถวิเคราะห์ได้ว่า การ กำหนดให้นักศึกษาชั้นปีที่ 4 ได้ใช้ทฤษฎีและทักษะทางด้านคอมพิวเตอร์ที่ได้เรียนมา แล้วนำมา ปฏิบัติได้ ภายใต้การกำกับดูแลของอาจารย์ที่ปรึกษา จะช่วยให้นักศึกษาสามารถนำวิชาความรู้ที่ได้ เรียนมาตลอดหลักสูตร ไปประยุกต์ใช้ในการทำงาน หลังจากจบการศึกษาได้ ซึ่งมีความสมเหตุสม ผล ดังนั้น ถ้าต้องการให้ภาวะการมีงานทำของบัณฑิตหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏ นครราชสีมา เพิ่มสูงขึ้น อาจพิจารณา เพิ่มหน่วยกิตในรายวิชาโครงการคอมพิวเตอร์เพิ่มมากขึ้น เพื่อให้นักศึกษาได้มีเวลาสะสมความรู้และทักษะด้านคอมพิวเตอร์ได้มากขึ้น อันจะส่งผลให้ภาวะการ มีงานทำของนักศึกษาเพิ่มสูงขึ้น และสรุปได้ว่า เทคนิคต้นไม้ตัดสินใจ สามารถนำมาใช้สนับสนุนการ ตัดสินใจในการปรับปรุงกลุ่มรายวิชาเอกบังคับ ที่ช่วยให้ภาวะการมีงานทำของบัณฑิตเพิ่มสูงขึ้นได้

ข้อเสนอแนะ

ข้อเสนอแนะทั่วไป

1. การทำเหมืองข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ แม้ว่าจะสามารถสร้างกฎความสัมพันธ์ให้ สามารถนำไปใช้ได้โดยง่าย แต่บางครั้ง ถ้าข้อมูลมีจำนวนน้อยเกินไป อาจไม่สามารถสร้างต้นไม้ตัดสินใจ ได้ ดังนั้น การนำแนวทางที่งานวิจัยนี้ได้นำเสนอไปใช้ประโยชน์ ควรดำเนินการกับข้อมูลที่มีขนาดใหญ่ จะทำให้สามารถสร้างต้นไม้ตัดสินใจ ที่สามารถทำนายค่าของเป้าหมายที่สนใจได้อย่างมีนัยยะมากขึ้น

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งต่อไป จะได้นำแนวทางที่งานวิจัยนี้ได้นำเสนอ ไปใช้กับข้อมูลที่มีขนาดใหญ่ มากกว่าข้อมูลที่ใช้ในงานวิจัยนี้ เช่น การวิเคราะห์รายวิชาพื้นฐานทั่วไปที่ส่งผลต่อภาวะการมีงานทำ เพราะรายวิชาพื้นฐานทั่วไป เป็นรายวิชาที่นักศึกษาชั้นปีที่ 1 ทุกหลักสูตร ต้องผ่านการลงทะเบียน เรียน ดังนั้น จึงมีจำนวนข้อมูลที่น่ามาใช้ในการวิจัยมากกว่าข้อมูลที่ใช้ในงานวิจัยนี้

2. อาจมีการศึกษาเทคนิคอื่นที่สามารถสร้างกฎความสัมพันธ์ได้เช่นเดียวกับต้นไม้ตัดสินใจ แล้วนำมาเปรียบเทียบความแม่นยำกับเทคนิคต้นไม้ตัดสินใจ จากนั้น เลือกใช้เทคนิคที่มีความแม่นยำ มากที่สุด ไปใช้สร้างกฎความสัมพันธ์ที่สามารถสนับสนุนการตัดสินใจ ในการปรับปรุงรายวิชาเรียน ให้ ภาวะการมีงานทำเพิ่มสูงขึ้นได้

บรรณานุกรม

- ซัดชัย แก้วตา และ อัจฉรา มหาวีรวัฒน์. (2553). การวินิจฉัยคดีด้วยเทคนิคต้นไม้ตัดสินใจ. *Proceeding of NCIT 2010 the National Conference on Information Technology Bangkok Thailand*, 308-313.
- ทิพย์ธิดา วงศ์พิพันธ์. (2556). การใช้เหมืองข้อมูลช่วยในการตัดสินใจการให้สินเชื่อ กรณีศึกษา: บริษัท ไทยคาร์เร็นท์ แอนด์ ลีส จำกัด (มหาชน). งานค้นคว้าอิสระในหลักสูตร วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิต.
- บุญมา เฟ่งชวน. (2548). การใช้เทคนิคเหมืองข้อมูลเพื่อพัฒนาระบบสนับสนุนการตัดสินใจด้านการผลิตบัณฑิตระดับปริญญาตรี. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร.
- ประกาศกระทรวงศึกษาธิการ เรื่องเกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ.2548. (2548, 25 พฤษภาคม). *ราชกิจจานุเบกษา*. เล่ม 122 ตอนพิเศษ 39 ง. หน้า 7-13.
- ไพฑูรย์ จันทร์เรือง. (2550). ระบบสนับสนุนการตัดสินใจเลือกสาขาการเรียนของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิคต้นไม้ตัดสินใจ. สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- เรวดี ศักดิ์อุยธรรม. (2553). การใช้เทคนิคดาต้าไมนิงในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์. รายงานการวิจัย วิทยาลัยราชพฤกษ์.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*. 27: 861–874.
- Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*. 1(1): 81-106.