



**เทคนิคเหมืองข้อมูลสำหรับพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจ
สำหรับทำนายทักษะด้านโปรแกรมประยุกต์ของนักศึกษาระดับปริญญาตรีในยุคดิจิทัล**
**Data mining techniques of model for decision support system
in predicting the results of the application usage examinations of
undergraduate students in digital age**

ชุตินารัฐ อุตมะสิริเสณี^{1*} และ พรสิริ ชาตปรีชา²

¹ กลุ่มวิชาวิทยาศาสตร์และคณิตศาสตร์ สำนักศึกษาทั่วไป สถาบันการจัดการปัญญาภิวัฒน์ นนทบุรี 11120

² สาขาเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์และเทคโนโลยี สถาบันการจัดการปัญญาภิวัฒน์ นนทบุรี 11120

Chutitanrat Uttamasiriseni^{1*} and Pornsiri Chatpreecha²

¹ Department of Sciences and Mathematics, The Office of General Education,
Panyapiwat Institute of Management, Nonthaburi 11120

² Division of Information Technology, Faculty of Engineering and Technology,
Panyapiwat Institute of Management, Nonthaburi 11120

Received: 2 May 2019/ Revised: 3 December 2019/ Accepted: 20 December 2019

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อสร้างตัวแบบสำหรับสร้างตัวแบบโมเดลการทำนาย (model) ผลทักษะวิชาด้านโปรแกรมประยุกต์ของนักศึกษามหาวิทยาลัยเอกชนแห่งหนึ่ง โดยใช้ข้อมูลกลุ่มตัวอย่างนักศึกษาชั้นปีที่ 1 สาขาการจัดการธุรกิจการค้าสมัยใหม่ ที่ลงเรียนวิชาการใช้โปรแกรมประยุกต์ในองค์กรสมัยใหม่ ในภาคการศึกษาที่ 1 ปีการศึกษา 2561 ข้อมูลที่ได้จากฐานข้อมูลนั้น จะถูกนำมาใช้วิเคราะห์สำหรับทำนายความสามารถในการเตรียมพร้อมเข้าสอบวัดความเชี่ยวชาญในทักษะด้านโปรแกรมประยุกต์ โดยมีรายละเอียดข้อมูลที่น่าสนใจนำมาใช้กำหนดในการทำวิจัยทั้งหมด 16 คุณลักษณะ จำนวนระเบียบที่นำมาใช้วิเคราะห์ข้อมูล ทั้งหมด 554 ระเบียบ มาสร้างตัวแบบการทำนาย โดยใช้เทคนิคการสร้างโมเดลแบบต้นไม้การตัดสินใจ (Decision Tree), Naïve Bay และ Generalized Linear Model (GLM) ในอันที่จะนำมาพัฒนาเป็นระบบสนับสนุนการตัดสินใจด้วยโปรแกรมประยุกต์บนเว็บไซต์โดยใช้ AJAX ร่วมกับ PHP ซึ่งเป็นทางเลือกหนึ่งที่จะสามารถช่วยแนะแนวทางให้กับนักศึกษาได้ทราบข้อมูลในการปรับตัวและจัดการกับข้อมูลทางด้านการเรียนของตนเองให้เหมาะสมมากยิ่งขึ้น

คำสำคัญ: เว็บแอปพลิเคชัน การทำนาย ตัวแบบโมเดล เหมืองข้อมูล ระบบสนับสนุนการตัดสินใจ ต้นไม้การตัดสินใจ
นาอ์เบย์ ตัวแบบเชิงเส้นวางนัยทั่วไป



Abstract

The system used to predict the result of Microsoft Office usability and create a model for prediction of the students at a private university by using the sample of the students the first year who are registered for Applications Usage for Modern Organization classes in semester 1/2019. Each sample has sixteen attributes, the total data sets are 554 which applied in three different data forecasting techniques are Decision Tree, Naïve Bay and Generalized Linear Model (GLM). Then we study course to find the way for develop a decision support system for planning of bachelor degree student. This system prototype was developed by AJAX with PHP. Therefore, we proposed a guideline to develop the system that can be used to manage academic advising data which appropriated for each student.

Keywords: Application, Forecast, Model, Data mining, Decision support systems, Decision Tree, Naïve Bay, Generalized Linear Model (GLM)

บทนำ

ปัจจุบันเทคโนโลยีในยุคดิจิทัลโดยเฉพาะด้านคอมพิวเตอร์มีการเปลี่ยนแปลงอย่างรวดเร็ว และเป็นส่วนสำคัญที่เข้ามามีบทบาทสำคัญในชีวิตประจำวัน อีกทั้งปัจจุบันวงการด้านการศึกษาได้ให้ความสนใจต่อการปรับเปลี่ยนพฤติกรรมทางด้านการเรียนการสอน และเล็งเห็นถึงความสำคัญของการนำเทคโนโลยีและมาตรฐานประกาศนียบัตรทางด้านความเชี่ยวชาญทางด้านคอมพิวเตอร์ และเพื่อให้สอดคล้องกับแผนพัฒนาการศึกษาของชาติ ตามยุทธศาสตร์ที่ 2 ของกรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสาร พ.ศ. 2554-2563 [1] สำหรับการพัฒนาการเรียนการสอนเพื่อเพิ่มยกระดับคุณภาพและเพิ่มประสิทธิภาพ โดยเฉพาะอย่างยิ่งสถานศึกษาและการพัฒนาศักยภาพและความเชี่ยวชาญให้กับนักศึกษา เพื่อให้ตรงกับความต้องการของบริษัท อีกทั้งกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา (TQF) ได้เห็นถึงความสำคัญของการพัฒนาศักยภาพโดยมีการกำหนดมาตรฐานผลการเรียนรู้ของผู้เรียนไว้ 5 ด้าน ได้แก่ (1) คุณธรรม จริยธรรม (2) ความรู้ (3) ทักษะทางปัญญา (4) ทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ และ (5) ทักษะการวิเคราะห์เชิงตัวเลข การสื่อสาร และการใช้เทคโนโลยีสารสนเทศ [2]

สถาบันการศึกษาหลายแห่งจึงได้มีการพัฒนาหลักสูตรการเรียนการสอนในระดับปริญญาตรีที่เกี่ยวกับด้านเทคโนโลยีสารสนเทศยุคดิจิทัล เพื่อให้สอดคล้องกับกรอบมาตรฐาน TQF ข้อ (5) โดยหลักสูตรวิชาต่าง ๆ ที่เปิดสอนนั้นได้มีการเรียนการสอนเพื่อให้ นักศึกษาได้สามารถเรียนรู้เข้าใจ และใช้เป็น อาชีพ เช่น โปรแกรมสำนักงานอัตโนมัติ ประกอบด้วยโปรแกรมประมวลผลคำ (Word) โปรแกรมตารางการคำนวณ (Excel) โปรแกรมการนำเสนอ (PowerPoint) และระบบสารสนเทศต่างๆ เพื่อเตรียมความพร้อมให้แก่ นักศึกษา ในการเพิ่มโอกาสการเข้ารับการจ้างงานมากขึ้น และทางสถานประกอบการส่วนใหญ่จะนำผลของทักษะด้านโปรแกรมหรือทักษะด้านเทคโนโลยีสารสนเทศมาเป็นองค์ประกอบหนึ่งเพื่อสัมภาษณ์พิจารณาคัดเลือกรับเข้าทำงานด้วย [3, 4]

ทักษะที่สำคัญของคนทำงานรุ่นใหม่ คือ ทักษะด้านคอมพิวเตอร์ซึ่งเป็นทักษะพื้นฐานที่บริษัทต่าง ๆ ระบุคุณสมบัติในการรับเข้าทำงาน ได้แก่ โปรแกรมพื้นฐานที่ต้องใช้ภายในสำนักงาน โปรแกรมประยุกต์เกี่ยวกับวิชาไมโครซอฟท์ออฟฟิศ Microsoft Office (MS Office) เช่น Word Excel PowerPoint [3] ซึ่งในปัจจุบันมีการเปิดโอกาสให้คนไทยได้มีการสอบวัดทักษะเพื่อการได้มาซึ่งใบประกาศนียบัตร Microsoft Office Specialist Certificate (MOS) เป็นใบ

ประกาศนียบัตร เพื่อยืนยันความสามารถในการใช้โปรแกรม Microsoft Office อย่างเป็นมาตรฐานสากลที่ทั่วโลกให้การยอมรับ เพราะถือว่าเป็นผู้ที่มีความรู้ความสามารถในการใช้งานโปรแกรมได้อย่างแท้จริง แต่ทั้งนี้พบว่ายังมีการดำเนินการสอบจริงน้อย เนื่องจากมีค่าใช้จ่ายค่อนข้างสูง ดังนั้นจึงส่งผลให้นักศึกษามีโอกาสเข้าสอบน้อย

ในงานวิจัยนี้ได้นำเทคนิคการทำเหมืองข้อมูล (data mining) [5, 6] มาประยุกต์ใช้กับข้อมูลของนักศึกษา เพื่อเป็นข้อมูลในการนำมาพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจสำหรับทำนายทักษะวิชา Microsoft Office ซึ่งเป็นด้านโปรแกรมประยุกต์ของนักศึกษาระดับปริญญาตรี ซึ่งฐานข้อมูลนักศึกษามีหลายประเภทและมีขนาดใหญ่มาก การวิเคราะห์และหาความเกี่ยวข้องกันของข้อมูล จึงควรต้องมีการดำเนินการอย่างเป็นระบบและด้วยเทคนิควิธีที่ทันสมัย โดยความรู้ที่ได้จากการวิเคราะห์และการค้นหาคำตอบจะถูกเก็บไว้ในรูปของฐานความรู้ ดังนั้นการทำเหมืองข้อมูลจึงเป็นวิธีการหนึ่งที่สามารถนำมาใช้สร้างแบบตัวแบบจำลอง (model) โดยใช้วิธีการจำแนกประเภท (classification) ของนักศึกษาที่มีผลการเรียน Microsoft Office ดีและไม่ดี พร้อมทั้งวิเคราะห์คุณสมบัติ (attribute) ของข้อมูลที่มีผลกับการเรียนในรายวิชาด้านคอมพิวเตอร์ เมื่อสามารถหาตัวแบบที่มีความน่าเชื่อถือได้สูง และสามารถนำข้อมูลของนักศึกษาที่เพิ่งเข้ามาใหม่มาทำการทดสอบโดยผ่านตัวแบบนี้ สามารถทำนายได้ว่านักศึกษาแต่ละคนสามารถมีผลคะแนนได้ดีในระดับใด

ผู้วิจัยจึงได้เล็งเห็นความสำคัญและเลือกวิชาการใช้โปรแกรมประยุกต์ในด้านการใช้โปรแกรม MS Office (Word, Excel, PowerPoint) ซึ่งอยู่ในรายวิชาการใช้โปรแกรมประยุกต์ในองค์กรสมัยใหม่ของมหาวิทยาลัยเอกชนแห่งหนึ่ง เพื่อนำมาสร้างตัวแบบและพัฒนาระบบต้นแบบของระบบ โดยใช้โปรแกรม rapid miner ซึ่งเป็น open source software ด้าน data mining ซึ่งเหมาะกับการวิเคราะห์ข้อมูลที่มีขนาดใหญ่และจำนวนมาก ในการค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูล เหมาะสำหรับการนำมาใช้เพื่อการบริหารและการ

ตัดสินใจขององค์กร เพื่อใช้ทำนายวัดทักษะการใช้โปรแกรม MS Office ของนักศึกษา และจากการศึกษางานวิจัย [7-9] ได้วิจัยการเปรียบเทียบเทคนิคการคัดเลือกที่เหมาะสมและอัลกอริทึมเพื่อจำแนกพฤติกรรมกรรมการทำความผิดของนักเรียนระดับอาชีวศึกษา โดยใช้เทคนิค 1. เทคนิค Naive Bay 2. เทคนิค Bayesian Belief Network 3. เทคนิค Decision Tree แบบ C4.5 และ 4. เทคนิค RIPPER ซึ่ง C4.5 ให้ค่าความถูกต้องสูงที่สุดและสามารถลดคุณลักษณะที่ต้องนำมาใช้ในการจำแนกได้แม่นยำ [10, 11]

โดยระบบนี้จะสามารถจัดเก็บผลการเรียนและนำเสนอผลการทำนายผลการเรียนโปรแกรม Microsoft Office ให้กับนักศึกษาได้ทราบในรูปแบบออนไลน์ สามารถดูผ่านเว็บไซต์ สมาร์ทโฟน (smart phone) หรือแท็บเล็ต (tablet) ได้อย่างสะดวกและรวดเร็ว สามารถเป็นโมเดลต้นแบบให้กับสถาบันการศึกษาอื่น ๆ ได้ การได้ทราบถึงทักษะของนักศึกษาที่เป็นจุดเด่นหรือต้องปรับปรุงพัฒนาด้านใดนั้น นักศึกษาสามารถนำไปใช้เป็นหลักฐานเพื่อพิจารณาสำหรับการเพิ่มโอกาสในการรับเข้าทำงานในยุคดิจิทัลได้อีกด้วย ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์คือ เพื่อศึกษาและวิเคราะห์เชิงเปรียบเทียบเทคนิคเหมืองข้อมูลที่เหมาะสมสำหรับพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจสำหรับทำนายทักษะด้านโปรแกรมประยุกต์ของนักศึกษาระดับปริญญาตรีในยุคดิจิทัล

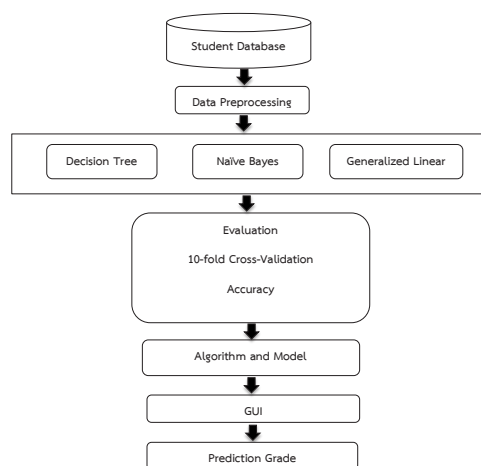
วิธีดำเนินการวิจัย

งานวิจัยนี้ใช้ข้อมูลของนักศึกษาจึงจะต้องเตรียมข้อมูล student database ให้เหมาะสมและให้อยู่ในรูปแบบที่สามารถนำไปใช้กับโปรแกรม rapid miner [12] จากนั้นนำเข้าสู่ขั้นตอน data pre-processing จัดเตรียมและนำข้อมูลมาวิเคราะห์ นำข้อมูลที่ไม่มีค่า ข้อมูลที่ขาดหาย และข้อมูลที่ไม่มีสมบูรณ์ออก เพื่อนำเข้าสู่ขั้นตอนการทำเหมืองข้อมูลเพื่อสร้าง ตัวแบบ โดยใช้เทคนิคเหมืองข้อมูลเพื่อสร้างโมเดลตัวแบบ คือ ต้นไม้การตัดสินใจ (Decision Tree), Naive Bay และ Generalized Linear Model (GLN) หลังจากนั้นทำการแบ่งกลุ่มข้อมูลเพื่อทดสอบ (cross-



validation) ประกอบด้วยข้อมูลส่วน training set และ testing set โดยใช้วิธี 10-fold cross-validation โดยแบ่งข้อมูลเป็น 10 ส่วนเท่า ๆ กัน แล้วทำการทดสอบ โดยเปลี่ยนข้อมูลชุดทดสอบตั้งแต่ส่วนที่ 1 เป็นชุดทดสอบ และส่วนที่ 2-9 เป็นชุดเทรนนิ่งไปเรื่อย ๆ จนกระทั่งถึงชุดทดสอบที่ส่วนที่ 10 เป็นชุดทดสอบ และส่วนที่ 1-9 เป็นชุดเทรนนิ่ง โดยจะนำเอาค่าความแม่นยำของการพยากรณ์ที่ได้นั้นมาหาค่าเฉลี่ย [13, 14] และใช้กฎความสัมพันธ์ (association rule)

โดยพิจารณาความสัมพันธ์ระหว่างปัจจัยที่เกี่ยวข้องในส่วน ของข้อมูลด้านการเรียนและประวัติส่วน of นักศึกษา จากนั้น นำผลจากการทดลองมาวิเคราะห์และสรุปผลเพื่อพัฒนาเป็น ตัวแบบของระบบ เพื่อให้ให้นักศึกษาได้วางแผนการเรียนและ ได้รู้ตนเอง รวมทั้งเตรียมตัวต่อการเรียนต่อไป [3] โดย ขั้นตอนกระบวนการสร้างตัวแบบจำลองการทำวิจัย แสดง ดังภาพที่ 1



ภาพที่ 1 ขั้นตอนการออกแบบการวิจัยและการพัฒนาระบบ

เมื่อกำหนดขั้นตอนการดำเนินการวิจัยแล้ว ในงาน วิจัยนี้ผู้วิจัยได้ใช้ฐานข้อมูลของนักศึกษาชั้นปีที่ 1 ที่ลงทะเบียน วิชาการใช้โปรแกรมประยุกต์ ซึ่งเรียนเกี่ยวกับการใช้โปรแกรม Microsoft Office (Word, Excel, PowerPoint) ในภาค การศึกษาที่ 1 ปีการศึกษา 2561 มหาวิทยาลัยเอกชนแห่งหนึ่ง

ซึ่งมีจำนวนนักศึกษา 554 คน ที่ลงทะเบียนเรียนวิชาดังกล่าว โดยนำมาใช้สำหรับการทดลอง โดยข้อมูลที่นำมาใช้จะเป็น ข้อมูลผลการเรียนและเกรดในรายวิชาของนักศึกษา ข้อมูล ประวัติส่วนตัวของนักศึกษาทั้งหมด 16 คุณลักษณะ แสดง ดังตารางที่ 1



ตารางที่ 1 คุณลักษณะของข้อมูล

ลำดับ	คุณลักษณะ	คำอธิบาย
1	Name DL	ภูมิลำเนา
2	Member_Family	จำนวนสมาชิกในครอบครัว
3	Monthly_expenses_Stu	ค่าใช้จ่ายต่อเดือนของนักศึกษา
4	Place_Status	สถานภาพความเป็นอยู่ของที่อยู่อาศัย
5	Live Status	สถานภาพความเป็นอยู่กับใคร
6	Family_income	รายได้ของครอบครัว
7	Educational_background	วุฒิก่อนเข้าศึกษาปริญญาตรี
8	GPA OLD	เกรดก่อนเข้าศึกษาปริญญาตรี
9	GPA Current	เกรดเฉลี่ยปัจจุบัน High คือ ผลการเรียนอยู่ในระดับสูง (ระดับเกรดเฉลี่ย 3.0-4.00) Medium คือ ผลการเรียนอยู่ในระดับปานกลาง (ระดับเกรดเฉลี่ย 2.01-3.00) Low คือ ผลการเรียนอยู่ในระดับต่ำ (ระดับเกรดเฉลี่ย 0.00-2.00)
10	GPA Match	เกรดด้านคณิตศาสตร์
11	GPA Eng	เกรดภาษาอังกฤษ
12	Hardware_Status	ความพร้อมของอุปกรณ์ที่มี
13	Ms_Word_Test	ผลการทดสอบปฏิบัติไมโครซอฟต์เวิร์ด
14	Ms_Excel_Test	ผลการทดสอบปฏิบัติไมโครซอฟต์เอ็กเซล
15	Ms_PPT_Test	ผลการทดสอบปฏิบัติไมโครซอฟต์พาวเวอร์พอยต์
16	Grade_MS	ผลสอบ Microsoft Office (Word ,Excel, PPT) Excellent คือ A Good คือ B+/B Below คือ C/C+ Poor คือ D/D+/F



จากตารางที่ 1 คือ ตารางแสดงคุณลักษณะของข้อมูลที่น่ามาใช้สำหรับการสร้างโมเดลการพยากรณ์โดยใช้วิธีการเลือกคุณลักษณะแบบกรอง (filter) จะเป็นกระบวนการประเมินประสิทธิภาพของคุณลักษณะของข้อมูลแต่ละตัวว่ามีความเหมาะสมกับการวิเคราะห์ข้อมูลมากน้อยเพียงใด โดยมีการจัดลำดับความสำคัญตามคุณลักษณะ (ranking) ระดับความสำคัญสูงสุด สำหรับการจัดเก็บข้อมูลพบว่าข้อมูลทั้งหมดที่ได้มีบางส่วนที่ไม่เกี่ยวข้อง เช่น รหัสนักศึกษา นอกจากนี้ยังมีข้อมูลของนักศึกษาและข้อมูลของรายวิชาจำนวนมากที่มีข้อมูลที่ไม่สมบูรณ์ มีค่าที่หายไป จึงคัดเลือกข้อมูลเพื่อความสะดวกในการสร้างแบบจำลอง [5]

เมื่อได้ข้อมูลตามความต้องการแล้วจึงนำข้อมูลมาทำการทดลองด้วยโปรแกรม rapid miner โดยใช้อัลกอริทึมต้นไม้ตัดสินใจ Naïve Bay และ Generalized Linear Model (GLM) รวมไปถึงกฎความสัมพันธ์ เพื่อหาความสัมพันธ์ของรายวิชาในการวิเคราะห์และทำนาย โดยนำมาสร้างแบบจำลองและทำการวัดผลด้วยค่าเปอร์เซ็นต์ของการทำนายถูกต้อง (percentage correct) ค่าที่สามารถสืบค้น

คำตอบสูงสุด (precision) ค่าที่ได้จากการตรวจพบข้อมูล (recall) โดยเลือกอัลกอริทึมที่เหมาะสมกับการข้อมูลและมีความแม่นยำในการทำนายผลมากที่สุด

ผลการวิจัย

สำหรับผลการทดลองได้ใช้จำนวน data set ทั้งหมด 554 คน จากนักศึกษาที่เรียนในรายวิชา ซึ่งผู้วิจัยได้ศึกษาและทำการเปรียบเทียบวิธีทำเหมืองข้อมูล โดยการเปรียบเทียบอัลกอริทึม Decision Tree อัลกอริทึม Naive Bay และอัลกอริทึม Generalized Linear Model เพื่อหาวิธีการที่เหมาะสมกับการทำนายโมเดลที่ดีที่สุดมาใช้เป็นโมเดลตัวแบบเพื่อการพัฒนาระบบ โดยใช้ตัววัดประสิทธิภาพโมเดลของการจำแนกข้อมูลจากการคำนวณ การประเมินผลลัพธ์การทำนาย (confusion matrix) จากค่า precision, recall, accuracy, F-measure สำหรับการทดลองจะแบ่งระดับคะแนนออกเป็นทั้งหมด 4 class คือ excellent, good, poor และ below ผลการทำนายโมเดลตัวแบบแต่ละอัลกอริทึม แสดงตารางที่ 2-4

ตารางที่ 2 ผลการทำนายโมเดลโดยใช้อัลกอริทึม Decision Tree

Class	Excellent	Good	Poor	Below	Class Precision
Pred. Excellent	364	49	20	4	83.30%
Pred. Good	27	24	0	0	47.06%
Pred. Poor	20	1	43	0	67.19%
Pred. Below	2	0	0	0	0.00%
Recall	88.14%	32.43%	68.25%	0.00%	

จากตารางที่ 2 ผลการทำนายจากการใช้อัลกอริทึม Decision Tree การทำนายผลโมเดลแต่ละ class พบว่า ความแม่นยำ (precision) ที่เกิดขึ้นในแต่ละ class คือ 83.30% (excellent), 47.06% (good), 67.19% (poor) และ 0.00% การทำนายผล (below) ส่วนความอ่อนไหวของข้อมูล (recall) คือ 88.14% (excellent), 32.43% (good), 68.25% (poor) และ 0.00% (below) ความแม่นยำของ

โมเดล คือ 77.80% (accuracy) จากผลดังกล่าวสามารถวิเคราะห์การทำนายผลของการพยากรณ์ทำให้เห็นว่าความสามารถการทำนายผลของอัลกอริทึมของ Decision Tree 97.43 จากค่า precision แสดงว่าสามารถทำนายผลได้ดี ส่วนการวัดความถูกต้องของโมเดล คือ 94.19 จากค่า recall

**ตารางที่ 3** ผลการทำนายโมเดลโดยใช้อัลกอริทึม Naïve Bay

Class	True Excellent	True Good	True Poor	True Below	Class Precision
Pred. Excellent	369	49	21	4	83.30%
Pred. Good	18	21	1	0	52.50%
Pred. Poor	15	2	40	0	70.18%
Pred. Below	11	2	1	0	0.00%
Class Recall	89.35%	28.38%	63.49%	0.00%	

จากตารางที่ 3 ผลการทำนาย จากการใช้อัลกอริทึม Naïve Bay การทำนายผลโมเดลแต่ละ class พบว่า ความแม่นยำ (precision) ที่เกิดขึ้นในแต่ละ class คือ 83.30% (excellent), 52.50% (good), 70.18% (poor) และ 0.00% (below) ส่วนความอ่อนไหวของข้อมูล (recall) คือ 89.35% (excellent), 28.38% (good), 63.49% (poor) และ 0.00% (below) ความแม่นยำของโมเดล คือ 77.65% (accuracy) จากผลดังกล่าวสามารถวิเคราะห์การทำนายผลของการพยากรณ์ทำให้เห็นว่า ความสามารถในการทำนายผลของอัลกอริทึมของ Naïve Bay 98.58 จากค่า precision แสดงว่าสามารถทำนายผลได้ดี ส่วนการวัดความถูกต้องของโมเดล คือ 87.84 จากค่า recall

ตารางที่ 4 ผลการทำนายโมเดลโดยใช้อัลกอริทึม Generalized Linear Model (GLN)

Class	True Excellent	True Good	True Poor	True Below	Class Precision
Pred. Excellent	405	73	30	4	79.10%
Pred. Good	1	0	0	0	0.00%
Pred. Poor	7	1	33	0	80.49%
Pred. Below	0	0	0	0	0.00%
Class Recall	98.06%	0.00%	52.38%	0.00%	

จากตารางที่ 4 สำหรับการใช้อัลกอริทึม Generalized Linear Model การทำนายผลโมเดลแต่ละ class พบว่า ความแม่นยำ (precision) ที่เกิดขึ้นในแต่ละ class คือ 79.10% (excellent), 0.00% (good), 80.49% (poor) และ 0.00% (below) ส่วนความอ่อนไหวของข้อมูล (recall) คือ 98.06% (excellent), 0.00% (good), 52.38% (poor) และ 0.00% (below) ความแม่นยำของโมเดล คือ 79.07% (accuracy) จากผลดังกล่าวเราสามารถวิเคราะห์การทำนายผลของการพยากรณ์ทำให้เห็นว่า ความสามารถในการทำนายผลของอัลกอริทึมของ Generalized Linear Model 98.48 จากค่า precision แสดงว่าสามารถทำนายผลได้ดี ส่วนการวัดความถูกต้องของโมเดล คือ 97.96 จากค่า recall และผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมที่เหมาะสมมาสร้างโมเดลการทำนาย แสดงดังตารางที่ 5



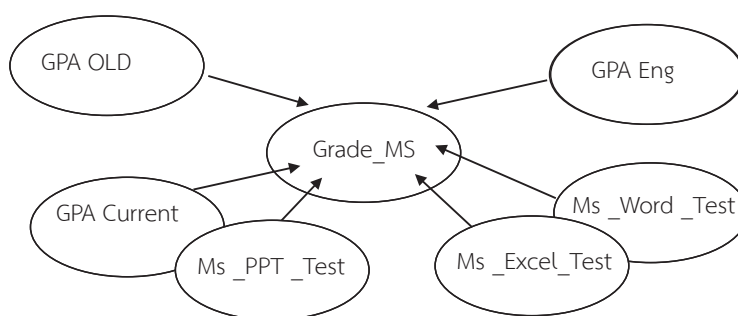
ตารางที่ 5 เปรียบเทียบผลการทดลองของแต่ละอัลกอริทึม

Algorithm	Precision (%)	Recall (%)	Accuracy (%)	F-Measure
Decision Tree	97.43	94.19	92.06	0.93
Naïve Bay	98.58	87.84	86.85	0.93
Generalized Linear Mode (GLM)	98.48	97.96	96.54	0.98

ผลจากการทดลองในตารางที่ 5 แสดงให้เห็นว่า อัลกอริทึม Naïve bay และ GLM มีประสิทธิภาพสูงใกล้เคียงกันสามารถทำนายผลข้อมูลได้ใกล้เคียงกันคือ สามารถสืบค้นคำตอบสูงสุด (precision) 98.58% และ 98.48% ส่วน Decision Tree คือ 97.43 ส่วนค่าตรวจพบข้อมูล (recall) คือ 87.84%, 97.96% และ 94.19% ค่าความแม่นยำของโมเดลโดยรวม ทุกคลาส (accuracy) 92.06%, 96.54% และ 92.06 และค่าที่แสดงความสัมพันธ์ระหว่างค่าที่สามารถสืบค้นคำตอบสูงสุดกับค่าที่ได้จากการตรวจพบข้อมูล (F-measure) สำหรับ GLM 0.98, Naïve bay 0.93 และ Decision Tree 0.93

ผลจากการวิเคราะห์ดังกล่าวนี้ผู้วิจัยจึงนำแบบจำลองที่สร้างได้นั้นมาประยุกต์ใช้ส่วนของการทำนายผลโดยนำ อัลกอริทึม Decision Tree มาใช้ในการพัฒนาระบบต้นแบบสำหรับพัฒนาเป็นระบบสนับสนุนการตัดสินใจ

สำหรับการวางแผนการเรียน เพื่อทำนายผลการสอบวัดทักษะวิชาด้านโปรแกรมประยุกต์บนเว็บไซต์ เพราะเหมาะสมกับประเภทของข้อมูลและคุณลักษณะของข้อมูลที่กำหนดเอาไว้ รวมถึงความสอดคล้องกันของกฎความสัมพันธ์ข้อมูล และผลการทำนายใกล้เคียงกับผลที่เกิดขึ้นจริง จึงเหมาะที่จะนำมาใช้เป็นต้นแบบในการพัฒนาระบบ โดยผลการทำนายตัวแปรที่มีผลต่อเกรดวิชาโปรแกรม Microsoft Office โดยการเซตค่าความแม่นยำในการทำนายตัวแปรที่มีผลต่อเกรดวิชาโปรแกรม Microsoft Office คือ เกรดก่อนเข้าศึกษาปริญญาตรี เกรดเฉลี่ยปัจจุบัน เกรดภาษาอังกฤษ ผลการทดสอบปฏิบัติไมโครซอฟต์เวิร์ด ผลการทดสอบปฏิบัติไมโครซอฟต์เอ็กเซล และผลการทดสอบปฏิบัติไมโครซอฟต์พาวเวอร์พอยต์ แสดงตัวแบบทำนายเกรดวิชาโปรแกรมประยุกต์ (ภาพที่ 2)



ภาพที่ 2 แบบจำลองทำนายเกรดวิชาโปรแกรมประยุกต์

ระบบที่พัฒนา

ผลจากการวิเคราะห์ดังกล่าวนี้จึงนำแบบจำลองที่สร้างได้นั้นมาทำการออกแบบและพัฒนาเป็นระบบเพื่อ

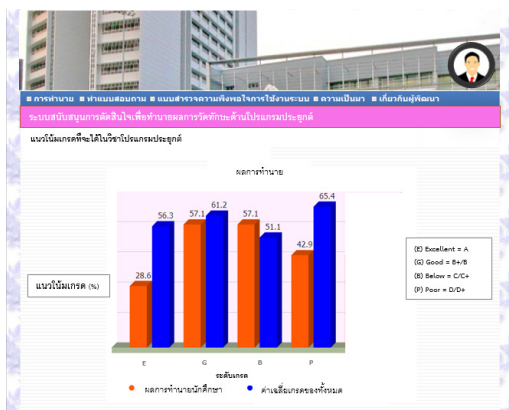
เป็นต้นแบบสำหรับพัฒนาระบบการทำนายผลการสอบวัดทักษะโปรแกรมประยุกต์เกี่ยวกับ Microsoft Office ดังภาพที่ 3-4



ภาพที่ 3 แสดงหน้าจอหลักต้นแบบของโปรแกรม

ขั้นตอนต่อไปแสดงต้นแบบเว็บไซต์สำหรับการทำนายผลวัดทักษะโปรแกรมประยุกต์ ซึ่งเกี่ยวกับวิชา Microsoft Office เพื่อเป็นแนวทางให้กับนักศึกษาได้ทราบ

ว่าตนเองว่าตนเองจะได้เกรดอยู่ในระดับใด เพื่อจะได้ปรับตัวและพัฒนาตนเองเพิ่มขึ้นในอนาคต แสดงหน้าจอต้นแบบผลการทำนายดังภาพที่ 4



ภาพที่ 4 หน้าจอแสดงผลการทำนาย

อภิปรายและสรุปผลการวิจัย

แบบจำลองที่สร้างขึ้นได้ใช้วิธีทดสอบข้อมูลของนักศึกษาโดยวิธี 10 fold Cross-Validation เป็นวิธีที่นักวิจัยนิยมใช้กันเป็นส่วนใหญ่ โดยในงานวิจัยนี้ผู้วิจัยได้เปรียบเทียบกับอัลกอริทึม Decision Tree อัลกอริทึม Naïve bay และอัลกอริทึม GLM เพื่อหาวิธีการที่เหมาะสมกับการทำนายโมเดลที่ดีที่สุดมาใช้ในการพัฒนาระบบ โดยวิเคราะห์ด้วยโปรแกรม rapid miner ซึ่งเป็นโปรแกรมทางด้าน data mining โดยมีวิจัยเลือกใช้อัลกอริทึม Decision Tree เพราะ

เหมาะกับคุณลักษณะของข้อมูลที่จะนำมาใช้รวมไปถึงการเชื่อมโยงกฎความสัมพันธ์ และเมื่อนำมาใช้ในการพัฒนาระบบส่วนการทำนายผลสามารถทำนายได้แม่นยำกับผลที่เกิดขึ้นจริง ซึ่งถือว่าเหมาะสมในการนำมาสร้างแบบจำลองช่วยนักศึกษาในการวิเคราะห์ทำนาย และเพื่อแนะแนวให้กับนักศึกษาในการทำนายผลการเรียนในวิชาโปรแกรม Microsoft Office ได้มากยิ่งขึ้นและได้นำเสนอต้นแบบออกมาในรูปแบบเว็บไซต์เพื่อให้ใช้งานได้ง่ายและสะดวก



เอกสารอ้างอิง

1. กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร. แผนแม่บทเทคโนโลยีสารสนเทศและการสื่อสาร (ฉบับที่2) ของประเทศไทย พ.ศ. 2552-2556 [อินเทอร์เน็ต]. 2552 [เข้าถึงเมื่อ 15 มี.ค. 2562]. เข้าถึงได้จาก <http://www.mict.go.th>
2. เพลินพิศ ศิริสมบุญ. คุณลักษณะบัณฑิตตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติของวิทยาลัยเซาธ์อีสท์บางกอก ประจำปีการศึกษา 2556-2557 (รุ่นที่ 15). วารสารวิชาการเซาธ์อีสท์บางกอก 2559;2 (2):55-67.
3. ชูติมา อุตมะบุญย์, ประสงค์ ปราณีตพลกรัง. การพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา. Journal of Information Science and Technology 2553;1(2):39-48.
4. Jobthai com. 10 คุณสมบัติในการทำงานที่ นายจ้างทุกคนต้องการจากผู้สมัครงาน. [อินเทอร์เน็ต]. 2560 [เข้าถึงเมื่อ 30 มี.ค. 2562]. เข้าถึงได้จาก <https://www.jobthaicom/REACH/career-tips.html>
5. สายชล สันสมบุญทอง. การทำเหมืองข้อมูล. (พิมพ์ครั้งที่ 1). กรุงเทพฯ: จามจุรีโปรดักท์; 2558.
6. Hsia T, Shie A, Chen L. Course planning of extension education to meet market demand by using data mining techniques-an example of chinkuo technology university in Taiwan. Expert Syst Appl 2008;34(1):596-602.
7. วันวิสาข์ ชนะประเสริฐ. การประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อแนะนำอาชีพสำหรับนักศึกษาปริญญาตรี คณะโบราณคดี มหาวิทยาลัยศิลปากร. ใน: เอกสารประกอบการประชุมวิชาการบัณฑิตศึกษาระดับชาติและนานาชาติ มหาวิทยาลัยศิลปากร ครั้งที่ 7 วันที่ 20-21 กรกฎาคม 2560. บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร. กรุงเทพฯ; 2560.หน้า 42-50.
8. สุนทรทิพย์ วงศ์พันธ์. การเปรียบเทียบเทคนิคการคัดเลือกคุณลักษณะที่เหมาะสมและอัลกอริทึมเพื่อจำแนกพฤติกรรมการกระทำผิดของนักเรียนระดับอาชีวศึกษา. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์, บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์. กรุงเทพฯ; 2551.
9. Quinlan JR. Machine learning: introduction of decision trees. Boston: Kluwer Academic Publishers; 1986.
10. Villacampa O. Feature selection and classification methods for decision making: a comparative analysis. Ph.D. Dissertation, Nova Southeastern University. USA; 2015.
11. Sakhare NN, Joshi SA. Classification of criminal data using J48-Decision Tree algorithm. IFRSA Int J Data Warehous Mining 2014;4:167-71.
12. มหาวิทยาลัยเชียงใหม่. Data mining with rapid miner (สำหรับผู้เริ่มต้น) [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 30 เม.ย.2562]. เข้าถึงได้จาก:<https://jupiter.ba.cmu.ac.th>
13. จูติมา ช่างชัย. การวิเคราะห์หารูปแบบการเรียนรู้โดยใช้เหมืองข้อมูลของนักศึกษาต่อการจัดทำปริญญานิพนธ์. วารสารบัณฑิตศึกษา มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ในพระบรมราชูปถัมภ์ 2559;10(2):53-62.
14. Sinsomboon S. Data mining. Bangkok: Jamjuree Product; 2015.