บทความวิจัย

การจำแนกผลการชำระเงิน ด้วยเทคนิคการเรียนรู้ของเครื่องจักร กรณีศึกษา บริษัทให้เช่าพื้นที่จัดเก็บสินค้าแห่งหนึ่ง

ชุษณา เทียนทอง * และ อารียา สุขเอี่ยม
ภาควิชาการบริหารอุตสาหกรรมการผลิตและการบริการคณะพัฒนาธุรกิจและอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

*ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 6976 9912 อีเมล: choosana.t@bid.kmutnb.ac.th
DOI: 10.14416/j.bid.2023.04.003
รับเมื่อ 29 ธันวาคม 2565 แก้ไขเมื่อ 30 มกราคม 2566 ตอบรับเมื่อ 2 มีนาคม 2566 เผยแพร่ออนไลน์ 25 เมษายน 2566
© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

งานวิจัยนี้เป็นการศึกษาการจำแนกผลการชำระเงิน ด้วยเทคนิคการเรียนรู้ของเครื่องจักรโดยใช้โปรแกรม Weka เพื่อการวางแผนการจัดการลูกหนี้ในอนาคต และเพื่อสภาพคล่องทางการเงินที่ดีขึ้นของธุรกิจ โดยมีวัตถุประสงค์เพื่อ 1) ศึกษาระบบการออกใบแจ้งหนี้และติดตามชำระหนี้ 2) สร้างแบบจำลองการจำแนกผลการชำระเงิน โดยใช้เทคนิคการเรียนรู้ของเครื่องจักร 3) เปรียบเทียบค่าความถูกต้องของแต่ละอัลกอริทึมในการจำแนกผลการชำระเงินที่เกินกำหนดและผลการชำระเงินที่ไม่เกินกำหนด โดยใช้เทคนิคการจำแนกประเภท (Classification) และ 4) พยากรณ์ผลการชำระเงินและวัดประสิทธิภาพของแบบจำลอง จากการศึกษาการสร้างแบบจำลองการจำแนกผลการชำระเงินจากอัลกอริทึม 3 แบบ ได้แก่ วิธีข่ายงานเบย์ วิธีการถดถอยโลจิสติกส์ และวิธีการเคลื่อนลงตามความชัน พบว่าแบบจำลองจากวิธีการจำแนกข้อมูลแบบข่ายงานเบย์เป็นแบบจำลองที่มีประสิทธิภาพดีที่สุด ซึ่งให้ค่าความถูกต้องของการจำแนกผลการชำระเงินเท่ากับร้อยละ 82.35 จึงได้ใช้เป็นตัวแบบสำหรับการจำแนกผลการชำระเงินในการศึกษานี้ นอกจากนี้การศึกษาการจำแนกผลการชำระเงินที่เกินกำหนดและผลการชำระเงินที่ไม่เกินกำหนด สอดคล้องกับสมมติฐานที่ตั้งไว้ คือ การจำแนกผลการชำระเงินโดยแบบจำลองจากวิธีการจำแนกข้อมูลที่ถูกเลือก จะมีค่าความคลาดเคลื่อนจากผลลัพธ์จริงไม่เกินร้อยละ 15.00 ซึ่งผลการศึกษามีค่าความคลาดเคลื่อนสัมพัทธ์ (Relative Error) เท่ากับร้อยละ 7.69 และร้อยละ 4.55 ตามลำดับ

คำสำคัญ: การเรียนรู้ของเครื่องจักร เหมืองข้อมูล แบบจำลองการพยากรณ์ การเรียนรู้แบบมีผู้สอน

การอ้างอิงบทความ: ชุษณา เทียนทอง และ อารียา สุขเอี่ยม. (2566). การจำแนกผลการชำระเงิน ด้วยเทคนิคการเรียนรู้ของเครื่องจักร กรณีศึกษา บริษัทให้เช่าพื้นที่จัดเก็บสินค้าแห่งหนึ่ง. *วารสารพัฒนาธุรกิจและอุตสาหกรรม*, 3(1), 27 – 40.

วารสารพัฒนาธุรกิจและอุตสาหกรรม ปีที่ 3, ฉบับที่ 1 มกราคม – เมษายน 2566
Journal of Business and Industrial Development. Vol.3, No.1, Jan. – Apr. 2023

Research Article

The Classification of Payment Result by Machine Learning: Case study of A Storage Space Rental Company

Choosana Tiantong* and Areeya Sukiam

Department of Manufacturing and Service Industry Manufacturing, Faculty of Business and Industrial Development, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.

* Corresponding Author, Tel. 08 6976 9912, E–mail: choosana.t@bid.kmutnb.ac.th
DOI: 10.14416/j.bid.2023.04.003
Received 29 December 2022; Revised 30 January 2023; Accepted 2 March 2023; Published online: 25 April 2023
© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

This research is a study of classification of payment results with machine learning techniques using Weka program to plan future receivable management and for better financial liquidity of the business. The objectives are to 1) study the system of invoicing and debt collection, 2) create models for the payment result classification, 3) compare the accuracy of each algorithm in classifying between due and overdue payments and 4) forecast payments and measure the performance of forecast model. According to the study on three algorithms for creation of classification model, namely, Bayesian network, logistics regression and gradient descent, it is found that the model from the Bayesian network classification yields the best performance with the payment result classification accuracy of 82.35%. Therefore, the model is chosen as the classification model in this study. The results of the classification between due and overdue payments are in accordance to the set hypothesis, that is the error of the payment result classification from the model obtained from the chosen classification method is approximately 15% or less, as compared to the actual results. The relative error of the results is 7.69% and 4.55%, respectively.

Keywords: Machine Learning, Data Mining, Forecasting Model, Supervised Learning

Please cite this article: Tiantong, C., & Sukiam, A. (2023). The Classification of Payment Result by Machine Learning: Case study of A Storage Space Rental Company *Journal of Business and Industrial Development*, 3(1), 27 – 40.

1.บทนำ

ในการประกอบธุรกิจ นอกจากความสามารถของฝ่ายขายในการขายสินค้าหรือบริการที่ทำให้บริษัทมีรายได้และผลกำไรแล้ว การบริหารการส่งใบแจ้งหนี้และการติดตามยอดคงค้างก็มีความสำคัญที่จะทำให้รายรับเข้าบริษัทได้รวดเร็วและตรงเวลา ซึ่งจะพบว่าปัญหาการชำระเงินที่เกินกำหนดของลูกค้านั้น อาจทำให้บริษัทขาดสภาพคล่องในการดำเนินธุรกิจได้ ดังนั้นฝ่ายที่เกี่ยวข้องจึงจำเป็นต้องแสวงหาวิธีควบคุมและในการติดตามลูกหนี้ได้อย่างเหมาะสมกับระยะเวลา ผู้วิจัยได้ตระหนักถึงความสำคัญในเรื่องนี้จึงทำการศึกษาโดยนำข้อมูลการรับชำระเงินจากบริษัทกรณีศึกษา มาทำการวิเคราะห์หาความสัมพันธ์และตัวแปรที่มีผลต่อการทำให้เกิดการชำระเงินเกินวันที่กำหนดในใบแจ้งหนี้ โดยข้อมูลที่นำมาใช้ทั้ง 14 เดือน จะแบ่งเป็นอัตราส่วน 14 : 1 : 1 ดังนี้ 1) ข้อมูลชุดเรียนรู้ (Training Data Set) 2) ข้อมูลชุดทดสอบ (Testing Data Set) และ 3) ข้อมูลชุดการพยากรณ์ (Forecasting Data Set) เพื่อสร้างแบบจำลองในการจำแนกผลการชำระเงินโดยอาศัยเทคนิคการเรียนรู้ของเครื่องจักรและเหมืองข้อมูล ซึ่งใช้โปรแกรมสำเร็จรูป Weka เป็นเครื่องมือในการสร้างแบบจำลองและวัดค่าประสิทธิภาพของแบบจำลองนั้น ๆ เพื่อนำค่าที่ได้มาเปรียบเทียบ และคัดเลือกอัลกอริทึมที่ดีที่สุดสำหรับใช้ในการพยากรณ์ข้อมูลผลการชำระเงินจากการรับชำระค่าบริการ ซึ่งจะสามารถนำแบบจำลองที่สร้างขึ้นจากอัลกอริทึมที่คัดเลือกไปประยุกต์ใช้กับการสนับสนุนการปรับเปลี่ยนนโยบายวิธีการเรียกเก็บค่าบริการ หรือการวางแผนการจัดการลูกหนี้ที่มีแนวโน้มว่าจะชำระเงินเกินกำหนดได้อย่างทันท่วงที โดยใช้เทคนิคการเรียนรู้ของเครื่องจักร (Machine Learning) จากข้อมูลและเครื่องมือทางสถิติเพื่อทำการทำนายผลลัพธ์การพยากรณ์ตามที่ต้องการ

1.1 วัตถุประสงค์

1.1.1 เพื่อศึกษาระบบการออกใบแจ้งหนี้และติดตามชำระหนี้

1.1.2 เพื่อสร้างแบบจำลองการจำแนกผลการชำระเงิน โดยใช้เทคนิคการเรียนรู้ของเครื่องจักร

1.1.3 เพื่อเปรียบเทียบค่าความถูกต้องของแต่ละอัลกอริทึมในการจำแนกผลการชำระเงินที่เกินกำหนดและผลการชำระเงินที่ไม่เกินกำหนด โดยใช้เทคนิคการจำแนกประเภท (Classification)

1.1.4 เพื่อพยากรณ์ผลการชำระเงินและวัดประสิทธิภาพของแบบจำลอง

1.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

1.2.1 การเรียนรู้ของเครื่องจักร (Machine Learning)

การเรียนรู้ของเครื่องจักร (Machine Learning) คือ ระบบที่สามารถเรียนรู้ได้จากตัวอย่างด้วยตนเอง ปราศจากการป้อนคำสั่งของโปรแกรมเมอร์ สามารถเรียนรู้ได้จากข้อมูลอย่างเดียวเพื่อที่จะผลิตผลลัพธ์ที่แม่นยำออกมาได้ ซึ่งวัตถุประสงค์หลักของการเรียนรู้ของเครื่องจักร คือ การเรียนรู้ (Learning) และการอนุมาน (Inference) โดยจะเรียนรู้ผ่านการค้นพบรูปแบบหรือแบบแผนซ้ำ ๆ เดิม ๆ การค้นพบต้องนั้นต้องมาจากการมีอยู่ข้อมูล (Data) โดยมีคุณสมบัติต่าง ๆ ที่ถูกใช้เพื่อแก้ปัญหาถูกเรียกว่า Feature Vector หลังจากการเรียนรู้และพัฒนาแล้วจะทำให้เกิดสิ่งที่ค้นพบที่เรียกว่าแบบจำลอง (Model) ดังนั้นขั้นของการเรียนรู้ (Learning Stage) ใช้เพื่ออธิบายข้อมูลและรวบรวมให้ข้อมูลเหล่านั้นกลายเป็นแบบจำลอง

1.2.2 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

การเรียนรู้แบบมีผู้สอนเป็นรูปแบบการเรียนรู้ประเภทหนึ่งของการเรียนรู้ของเครื่องจักร (Machine Learning) ซึ่งอัลกอริทึมจำเป็นต้องใช้ข้อมูลส่วนหนึ่งสำหรับสอน (Training Data) และข้อมูลอีกส่วนหนึ่งซึ่งรับกลับมาเพื่อปรับปรุง (Feedback) จากมนุษย์ เพื่อเรียนรู้ความสัมพันธ์ระหว่างข้อมูลที่ถูกป้อนเข้ามาสู่ข้อมูลที่ออกไป เราสามารถใช้ Supervised Learning เมื่อผลลัพธ์ของข้อมูลเป็นสิ่งที่รู้อยู่แล้วและสามารถใช้ในการทำนายข้อมูลใหม่ได้

1.2.3 เหมืองข้อมูล (Data Mining)

เหมืองข้อมูล (Data Mining) คือ การนำข้อมูลจำนวนมากจากฐานข้อมูลหรืออื่น ๆ มาใช้เพื่อวิเคราะห์ ประมวลผล หาความสัมพันธ์ที่ในข้อมูลชุดนั้น ๆ โดยการสร้างแบบจำลองและความสัมพันธ์ทางสถิติ ซึ่งสามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูลได้ การทำเหมืองข้อมูลโดยวิธี Cross-Industry Standard Process for Data Mining หรือ CRISP-DM ประกอบด้วย 6 ขั้นตอน ดังนี้

1) การทำความรู้จักและเข้าใจในธุรกิจ (Business Understanding) ขั้นตอนการเริ่มต้นการจัดทำเหมืองข้อมูล ต้องเริ่มศึกษาปัญหาและสภาพแวดล้อมในการทำงาน และเอกสารต่าง ๆ

2) การทำความเข้าใจข้อมูล (Data Understanding) การรวบรวมข้อมูลที่เกี่ยวข้อง ทำความเข้าใจข้อมูล และคัดข้อมูลที่ไม่เกี่ยวข้องออก จะทำให้กระบวนการเรียนรู้ของเครื่องจักรได้เรียนรู้ผลลัพธ์และพฤติกรรมของข้อมูล

3) การจัดเตรียมข้อมูล (Data Preparation) เมื่อทำความเข้าใจกับข้อมูลและเก็บรวบรวมข้อมูลแล้ว ต้องมีการตรวจสอบและแก้ไขข้อมูลทั้งหมดให้มีความถูกต้อง เพื่อทำการแปลงข้อมูล (Raw Data) ให้กลายเป็นข้อมูลที่สมบูรณ์ จึงสามารถนำไปใช้ในการวิเคราะห์ต่อไปได้

4) การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูล (Modeling) การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูล โดยใช้รูปแบบการเรียนรู้ที่เหมาะสมกับชุดข้อมูล คัดเลือกคุณลักษณะเพื่อสร้างแบบจำลองและเปรียบเทียบค่าประสิทธิภาพของแบบจำลองที่นำมาวิเคราะห์ จากนั้นคัดเลือกแบบจำลองที่เหมาะสมที่สุด เพื่อนำมาใช้เป็นต้นแบบการพยากรณ์ในการจำแนกข้อมูล

5) การประเมินผลลัพธ์ (Evaluation) การประเมินผลลัพธ์ คือ การนำข้อมูลชุดทดสอบ (Testing Data Set) มาทดสอบเพื่อเปรียบเทียบค่าประสิทธิภาพของแบบจำลอง

6) การนำเอาข้อมูลไปใช้งาน (Deployment) การนำแบบจำลองที่สร้างมาพยากรณ์ผลลัพธ์ของข้อมูล จากนั้นเปรียบเทียบผลลัพธ์ระหว่างค่าจริงและค่าพยากรณ์ โดยคำนวณค่าความคลาดเคลื่อนเพื่อชี้วัดประสิทธิภาพของการพยากรณ์นั้น เพื่อจะทำให้ทราบว่าแบบจำลองมีประสิทธิภาพและสามารถนำไปใช้งานต่อไปได้

1.2.4 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า ทิพย์ธิดา [1] ได้ทดสอบข้อมูล โดยใช้เทคนิคการจำแนกกลุ่มเพื่อหากฎที่ใช้ในการจำแนกลูกค้ากลุ่มดีและลูกค้ากลุ่มไม่ดี ซึ่งได้ทดลองจำแนกกลุ่มข้อมูลด้วยเทคนิค การจำแนกกลุ่ม 3 เทคนิค ซึ่งให้ผลลัพธ์เป็นกฎ ได้แก่ โมเดลการจำแนกกลุ่มแบบ Decision Tree : C4.5 หรือเรียกว่าโมเดล Classifies trees J48 โมเดลการจำแนกกลุ่มแบบ Decision Rule : Part หรือเรียกว่า โมเดล Classifier PART และโมเดลการจำแนกกลุ่มแบบ Bayes : Naive Bayes หรือเรียกว่า โมเดล Classifier Naive Bayes ผลการทดลองพบว่า การจำกลุ่มข้อมูลด้วยเทคนิค

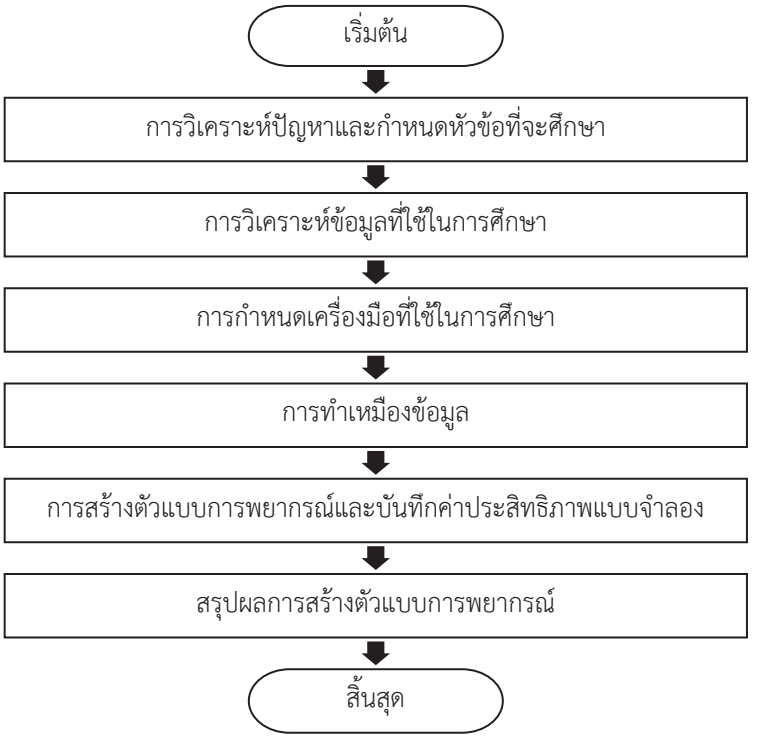
J48 ให้ผลลัพธ์ดีที่สุด เนื่องจากมีความถูกต้องมากที่สุดและสามารถแบ่งกลุ่มได้ตามเงื่อนไขที่ชัดเจนมากที่สุด สุทธิศานต์ [2] ศึกษาจุดเด่นของสถิติแบบเบส์ ซึ่งกล่าวว่า สถิติแบบเบส์มีจุดเด่นด้านการวิเคราะห์เมื่อเปรียบเทียบกับสถิติแบบความถี่หลายประการ ได้แก่ 1) สถิติแบบเบส์ใช้ตัวอย่างวิเคราะห์ (Sample) ที่ค่อนข้างน้อย 2) มีความยืดหยุ่นเกี่ยวกับการแจกแจงของตัวอย่าง 3) มีความยืดหยุ่นสูงในการวิเคราะห์โมเดลที่มีโครงสร้างลำดับชั้นที่ซับซ้อนมีข้อมูลขาดหายหรือมีพารามิเตอร์จำนวนมาก วิฐรา [3] ได้ทำการศึกษาวิธีวิเคราะห์การถดถอยเชิงเส้นสำหรับข้อมูลที่มีมิติสูง การวิเคราะห์การถดถอยเชิงเส้นเป็นเครื่องมือที่ไม่สลับซับซ้อน แต่มีประโยชน์อย่างมากในการวิเคราะห์ มีวัตถุประสงค์เพื่อนำเสนอเครื่องมือเพื่อเป็นทางเลือกในการจัดการกับข้อมูลที่มีมิติสูงซึ่งรวมถึงวิธี Penalized Regression และวิธีเบส์ ตลอดจนข้อควรคำนึงถึงในการเลือกใช้เครื่องมือสำหรับการวิเคราะห์การถดถอยสำหรับข้อมูลที่มีมิติสูง สุพัตรา และ ธิปไตย [4] ทำการศึกษาข่ายงานเบย์และการประยุกต์โดยกล่าวว่า ข่ายงานเบย์ เป็นแบบจำลองกราฟิกที่ใช้อธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไขระหว่างตัวแปร ซึ่งโครงสร้างของแบบจำลองกราฟิกนี้ จะประกอบไปด้วยเซตของโหนดและเส้นเชื่อมโดยที่โหนดจะแทนตัวแปรสุ่ม และเส้นเชื่อมจะแสดงถึงการขึ้นต่อกันของตัวแปรสุ่ม ข่ายงานเบย์สามารถอนุมานค่าโดยไม่จำเป็นต้องอาศัยทฤษฎีความน่าจะเป็นเชิงความถี่ที่ต้องอาศัยการทดลองเสมอ และด้วยเหตุนี้จึงทำให้วิธีข่ายงานเบย์ได้รับความนิยมกันอย่างแพร่หลาย ในศาสตร์หรืองานด้านต่าง ๆ ในปัจจุบัน สายชล [5] เปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี ได้แก่ วิธีเพื่อนบ้านใกล้สุด k ตัว โดยใช้อัลกอริทึมชนิด IBk วิธีต้นไม้ การตัดสินใจโดยใช้อัลกอริทึมชนิด 48 วิธีโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมชนิดเพอร์เซปตรอนแบบหลายชั้น วิธีซัพพอร์ตเวกเตอร์แมชชีนโดยใช้อัลกอริทึม SMO ชนิดโพลิโนเมียลเคอร์เนล วิธีฐานกฎโดยใช้อัลกอริทึมชนิด ตารางการตัดสินใจ วิธีการถดถอยลอจิสติกทวิภาค และวิธีนาอีฟเบส์ โดยวิธีการจำแนกกลุ่มที่มีประสิทธิภาพมากที่สุดจะให้ ค่าความแม่น ความเที่ยง ความระลึก และค่าวัดเอฟที่มีค่ามากที่สุด แต่จะให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อย อรอนงค์ [6] ได้ศึกษากระบวนการตั้งแต่การสั่งซื้อไปจนถึงการชำระเงิน (Order-to-Cash) ของบริษัทกรณีศึกษา และการสร้างแบบจำลองผ่านเทคนิคการเรียนรู้ของเครื่องจักร (Machine Learning) เพื่อสร้างแบบจำลองในโปรแกรมเวก้า (WEKA) และเปรียบเทียบค่าความถูกต้อง (Accuracy) ในการจำแนกใบแจ้งหนี้เกินกำหนดของขั้นตอนวิธี (Algorithm) ได้แก่ ขั้นตอน วิธีข่ายงานเบย์ (Bayesian Network : Bayes Net) ขั้นตอนวิธีการเคลื่อนลงตามความชันแบบสุ่ม (Stochastic Gradient Descent : SGD) และขั้นตอนวิธีต้นไม้ตัดสินใจแบบเจ 48 (048 Decision Tree) เริ่มจากการทำความเข้าใจกระบวนการและข้อมูล แล้วเก็บรวบรวมข้อมูลใบแจ้งหนี้ทั้งหมด 766 ระเบียน (Record) ในช่วงเวลาทั้งหมด 12 เดือน แบ่งเป็น ชุดข้อมูลสอน (Training Dataset) ชุดข้อมูลตรวจสอบ (Validation Dataset) และชุดข้อมูลทดสอบ (Test Dataset) โดยการศึกษานี้ใช้ อัตราส่วน 8 : 2 : 2 เพื่อตรวจสอบและประเมินผลแบบจำลองที่ได้จากขั้นตอนวิธีที่เหมาะสมที่สุด สุธีธิดา [7] ศึกษาเครื่องมือการพยากรณ์ราคาด้วยการเรียนรู้ของเครื่องจักร ด้วยโปรแกรม Weka 3.8 เพื่อสร้างแบบจำลองการพยากรณ์ราคาโดยเทคนิคการเรียนรู้ของเครื่องจักรและเพื่อสร้างแบบจำลองที่มีประสิทธิภาพในการพยากรณ์ราคาผักและผลไม้ และนำไปประยุกต์ใช้ประกอบการพิจารณายอดคำสั่งซื้อและช่วยให้เกิดการจัดการข้อมูลราคาอย่างเป็นระบบ และรับมือการเปลี่ยนแปลงราคาในอนาคต จากการศึกษาด้วยการเปรียบเทียบความคลาดเคลื่อนของข้อมูลก่อนและหลังปรับค่าพารามิเตอร์ด้วยอัลกอริทึม 3 แบบประกอบไปด้วย ด้วยอัลกอริทึม Gaussian Processes, Linear Regression, SMOreg ด้วยโปรแกรม Weka 3.8

ผลการศึกษาที่ได้ อัลกอริทึม SMOreg ก่อนปรับค่าพารามิเตอร์ ได้ให้ค่าความคลาดเคลื่อนเฉลี่ยรวมน้อยที่สุดในการนำไปสร้างเป็นรูปแบบจำลองเพื่อพยากรณ์ราคาในอนาคต และวัดประสิทธิภาพด้วยค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยจากค่าจริงเปรียบเทียบกับค่าพยากรณ์

จากการศึกษางานวิจัยเกี่ยวข้องดังกล่าว ผู้วิจัยสรุปได้ว่า เทคนิคเหมืองข้อมูลสามารถนำไปทำนายผลลัพธ์ของข้อมูลในลักษณะของการจำแนกกลุ่มข้อมูลได้ดี โดยสามารถวัดประสิทธิภาพของแบบจำลองที่สร้างขึ้นและนำไปใช้ได้จากค่าความถูกต้อง (Correctly) ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error : MAE) และค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสอง (Root Mean Square Error : RMSE)

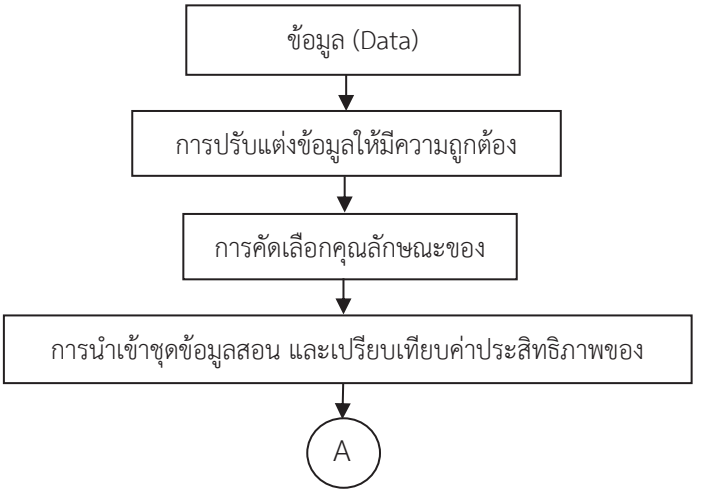
2. วิธีการวิจัย

จากการรวบรวมข้อมูลพื้นฐานต่าง ๆ ของบริษัทประกอบธุรกิจการให้บริการพื้นที่เก็บของและสัมภาระ ซึ่งมีรูปแบบธุรกิจเป็นแบบ B2B และ B2C และศึกษางานวิจัยที่เกี่ยวข้อง จึงนำมาสู่การสร้างกระบวนการวิจัย ดังแสดงในรูปที่ 1 เพื่อดำเนินการวางแผนกระบวนการในแต่ละขั้นตอน ไม่ว่าจะเป็นการวิเคราะห์ปัญหาและข้อมูลที่ใช้การกำหนดเครื่องมือในการวิจัย การทำเหมืองข้อมูล การสร้างตัวแบบการพยากรณ์ ตลอดจนการบันทึกค่าประสิทธิภาพแบบจำลอง

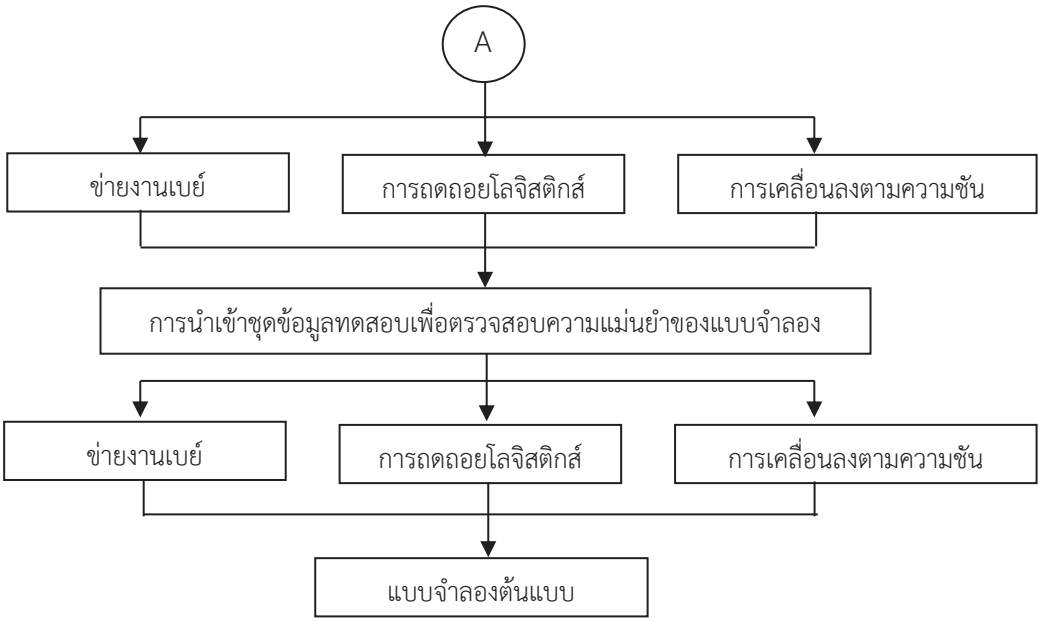


รูปที่ 1 แผนผังแสดงขั้นตอนการดำเนินการศึกษา

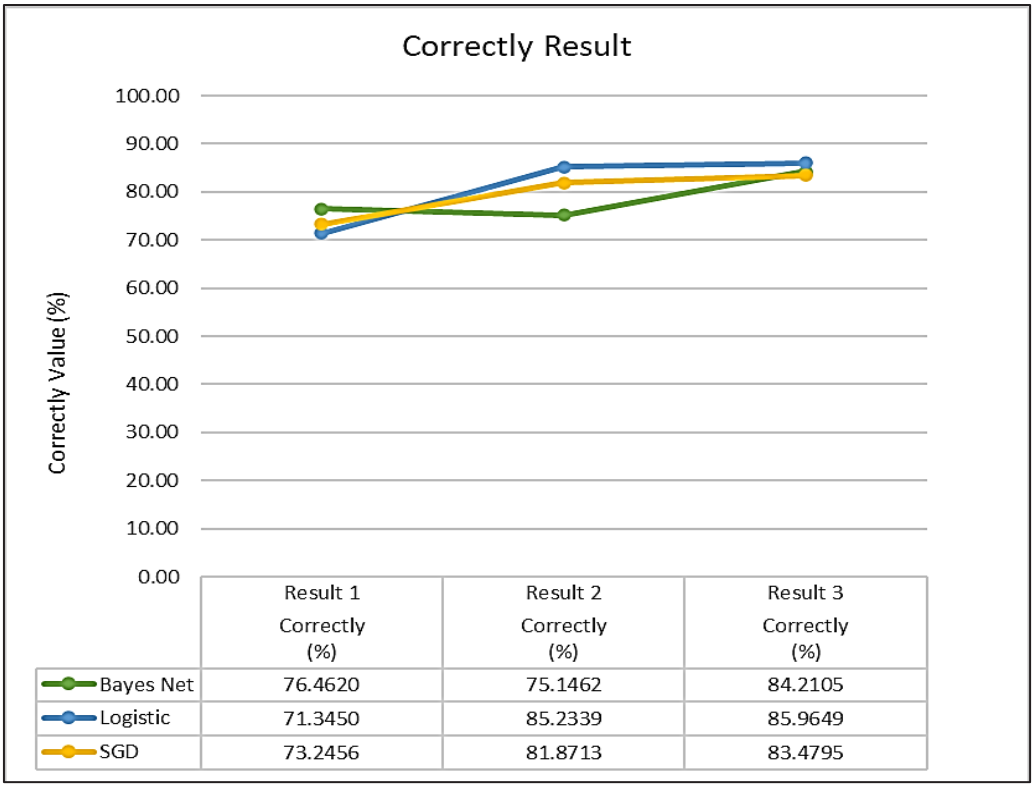
จากรูปที่ 1 สามารถอธิบายกระบวนการได้ดังนี้ (1) การวิเคราะห์ปัญหาและกำหนดหัวข้อที่จะศึกษาวิจัย การศึกษากิจกรรมต่าง ๆ ที่เกิดขึ้น การสังเกตปัญหาและภาพรวมของงาน โดยใช้เครื่องมือ 5W 1H เป็นองค์ประกอบในการวิเคราะห์ปัญหา (2) การวิเคราะห์ข้อมูลที่ใช้ในการศึกษาวิจัย โดยใช้ข้อมูลจากรายงานการรับชำระเงินในแต่ละเดือน รวมถึงตัวแปรต่าง ๆ ที่เป็นปัจจัยที่ทำให้ลูกค้าชำระเงินเกินกำหนดเวลา เพื่อนำตัวแปรเหล่านั้นเข้าสู่ขั้นตอนการคัดเลือกคุณสมบัติของข้อมูล และทำการคัดเลือกคุณลักษณะ (Attribute) เพื่อให้ได้ชุดข้อมูลที่มีความสัมพันธ์ในกระบวนการเรียนรู้ของเครื่องจักร (3) การกำหนดเครื่องมือที่ใช้ในการศึกษา ผู้วิจัยได้เลือกใช้โปรแกรม Weka เวอร์ชั่น 3.8.4 มาสนับสนุนการทำเหมืองข้อมูลและการสร้างแบบจำลองการจำแนกประเภทข้อมูล โดยโปรแกรมจะมีอัลกอริทึมให้เลือกใช้จำนวนมาก ซึ่งอัลกอริทึมที่ผู้วิจัยเลือกใช้ในการจำแนกประเภทข้อมูลจำนวน 3 อัลกอริทึม ได้แก่ ข่ายงานเบย์ การถดถอยโลจิสติกส์ และการเคลื่อนลงตามความชัน จากอัลกอริทึมทั้ง 3 ประเภท ได้จากการคัดเลือก และพิจารณาจากลักษณะพื้นฐานทางเทคนิคที่แตกต่างกัน รวมถึงความนิยมในการนำมาใช้ในการวิจัยต่าง ๆ (4) การทำเหมืองข้อมูล (Data Mining) โดยใช้วิธี Cross-Industry Standard Process for Data Mining หรือ CRISP-DM ซึ่งประกอบด้วย 6 ขั้นตอน ได้แก่ การทำความรู้จักและเข้าใจในธุรกิจ การทำความเข้าใจข้อมูล การจัดเตรียมข้อมูลโดยแบ่งชุดข้อมูล ตามเทคนิคการเรียนรู้ของเครื่องจักร การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูล การประเมินผลลัพธ์ และการนำเอาข้อมูลไปใช้งาน (5) การสร้างตัวแบบการพยากรณ์จากชุดข้อมูลสอน แสดงดังรูปที่ 2 โดยเริ่มจากการคัดเลือกคุณลักษณะโดยใช้ฟังก์ชัน Select Attribute ในโปรแกรม Weka ใช้วิธีการคัดเลือกคุณลักษณะ Information Gain Attribute Eval จากฟังก์ชัน Select Attribute จึงได้คุณลักษณะที่ส่งผลให้ประสิทธิภาพของแบบจำลองมีค่าสูงสุด จากนั้นจึงทดสอบการปรับค่าพารามิเตอร์ที่มีผลต่อการเรียนรู้ของอัลกอริทึม เพื่อหาจุดที่ทำให้การเรียนรู้จากข้อมูลสอนนี้มีประสิทธิภาพมากที่สุด (6) การสรุปผลการวัดค่าประสิทธิภาพของแบบจำลองที่ใช้ในการศึกษา พบว่าทุกวิธีของการจำแนกข้อมูลที่ใช้ มีค่าความถูกต้องเพิ่มขึ้น ค่า MAE และ RMSE ลดลง ซึ่งแสดงให้เห็นว่าแบบจำลองจาก 3 วิธีการจำแนกประเภทข้อมูล มีค่าความคลาดเคลื่อนต่ำลง และมีประสิทธิภาพเพิ่มขึ้น จากทุกขั้นที่ทำการปรับปรุง ดังรูปที่ 3 ที่เป็นกราฟแสดงค่าความถูกต้องของแบบจำลองที่เพิ่มขึ้นและแสดงผลลัพธ์ดังตารางที่ 1



รูปที่ 2 แผนผังแสดงกระบวนการสร้างแบบจำลองสำหรับการพยากรณ์



รูปที่ 2 แผนผังแสดงกระบวนการสร้างแบบจำลองสำหรับการพยากรณ์ (ต่อ)



	Result 1 Correctly (%)	Result 2 Correctly (%)	Result 3 Correctly (%)
Bayes Net	76.4620	75.1462	84.2105
Logistic	71.3450	85.2339	85.9649
SGD	73.2456	81.8713	83.4795

รูปที่ 3 ผลลัพธ์ค่าความถูกต้อง (Correctly) หลังจากการคัดเลือกคุณลักษณะและปรับปรุงค่าพารามิเตอร์

ตารางที่ 1 สรุปผลการเปรียบเทียบค่าประสิทธิภาพของแบบจำลองในแต่ละขั้นจากข้อมูลชุดเรียนรู้

Classifier Method	Result 1st*			Result 2nd*			Result 3rd*		
	Correctly (%)	MAE	RMSE	Correctly (%)	MAE	RMSE	Correctly (%)	MAE	RMSE
Bayes Net	76.4620	0.2803	0.4125	75.1462	0.3025	0.4007	84.2105	0.2054	0.3367
Logistic	71.3450	0.2928	0.5158	85.2339	0.1561	0.3621	85.9649	0.1494	0.3505
SGD	73.2456	0.2675	0.5172	81.8713	0.1813	0.4258	83.4795	0.1652	0.4065

หมายเหตุ* :

Result 1st หมายถึง ผลการวัดค่าประสิทธิภาพของแบบจำลองหลังทำการการคัดเลือกคุณลักษณะโดยฟังก์ชัน Select Attribute

Result 2nd หมายถึง ผลการวัดค่าประสิทธิภาพของแบบจำลองหลังจากการคัดเลือกคุณลักษณะเพิ่มเติมโดยผู้ปฏิบัติงานและใช้งานข้อมูล

Result 3rd หมายถึง ผลการวัดค่าประสิทธิภาพของแบบจำลองหลังปรับค่าพารามิเตอร์ของอัลกอริทึม

3. ผลการวิจัย

3.1 ผลการสร้างแบบจำลอง

จากการประมวลผลการเรียนรู้ด้วยของเครื่องจักรโดยใช้โปรแกรม Weka ผู้วิจัยได้นำข้อมูลเข้าสู่โปรแกรมเพื่อทำการประมวลผล และวัดค่าประสิทธิภาพของแบบจำลองจากข้อมูลชุดเรียนรู้จำนวนข้อมูล 12 เดือน โดยทดสอบด้วยอัลกอริทึม 3 ชนิด ได้แก่ 1) ข่ายงานเบย์ (Bayesian Network Classifiers) 2) การถดถอยโลจิสติกส์ (Logistic Regression) และ 3) การเคลื่อนลงตามความชัน (Stochastic Gradient Descent) จากนั้นทำการปรับปรุง ด้วยการกำจัดตัวแปรต้นที่ไม่จำเป็นออก เพื่อลดมิติของข้อมูลจากเทคนิคการคัดเลือกคุณลักษณะด้วยวิธี Information Gain Attribute Evaluation แล้ววัดค่าประสิทธิภาพของทั้ง 3 อัลกอริทึม หลังจากปรับปรุงข้อมูลครั้งที่ 1 แล้วจึงทำการปรับปรุงข้อมูลครั้งที่ 2 โดยการวิเคราะห์คุณลักษณะที่เกี่ยวข้องเพิ่มเติมให้ข้อมูลประกอบด้วยคุณลักษณะที่สมบูรณ์มากยิ่งขึ้น และจากการปรับปรุงครั้งที่ 3 ซึ่งเป็นการปรับค่าพารามิเตอร์ของแบบจำลอง เพื่อทำการทดสอบประสิทธิภาพของแบบจำลองด้วยข้อมูลชุดเรียนรู้ในขั้นสุดท้าย เพื่อให้แบบจำลองมีประสิทธิภาพยิ่งขึ้น

3.2 ผลการวิเคราะห์ข้อมูลชุดทดสอบและเปรียบเทียบค่าประสิทธิภาพของแบบจำลอง

ผู้วิจัยนำข้อมูลชุดทดสอบ (Testing Data Set) จำนวนข้อมูล 119 ระเบียน ไปทำการทดสอบแบบจำลองทั้ง 3 อัลกอริทึม ด้วยการประมวลผลจากการเรียนรู้ของเครื่องจักรโดยโปรแกรม Weka เพื่อเปรียบเทียบค่าเมทริกซ์วัดประสิทธิภาพของแบบจำลองจากวิธีการจำแนกประเภทข้อมูลทั้ง 3 วิธี ซึ่งให้ผลการทดสอบดังตารางที่ 2

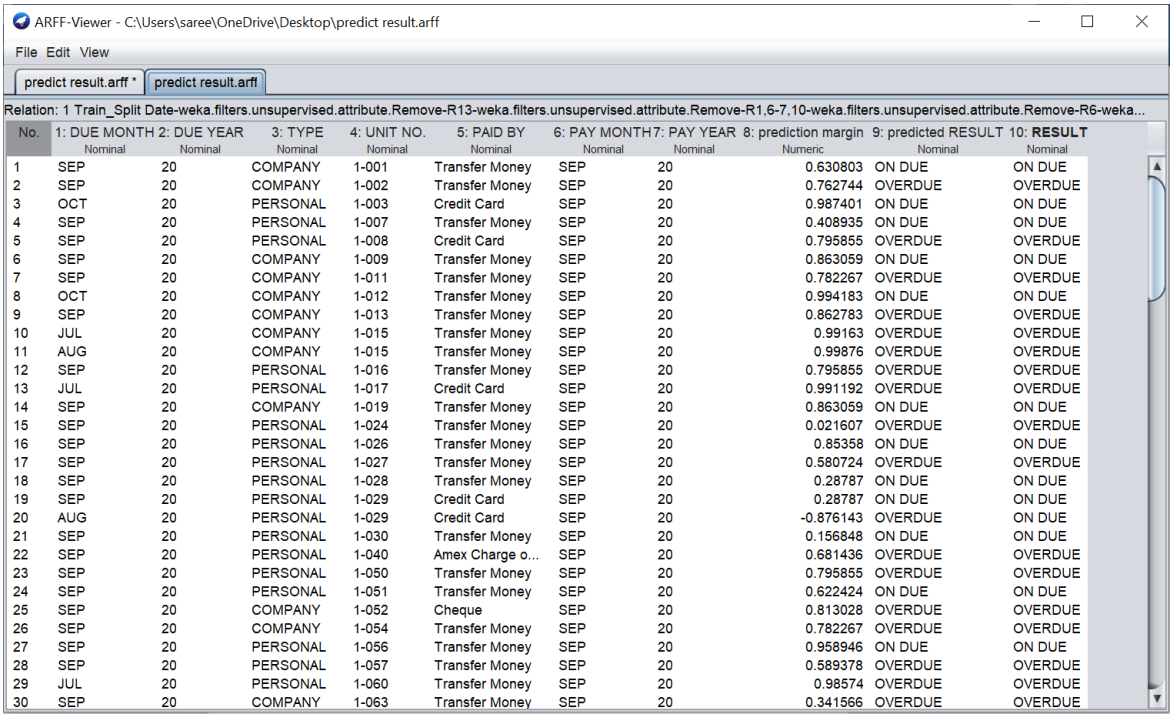
ตารางที่ 2 ตารางแสดงผลลัพธ์ของค่าเมทริกซ์วัดประสิทธิภาพจากการทดสอบแบบจำลองผ่านโปรแกรม Weka โดยใช้ข้อมูลชุดทดสอบ (Testing Data Set)

Classify Method	Result of Efficiency Metrix from Testing Data Set		
	Correctly (%)	MAE	RMSE
Bayes Net	82.3529	0.2562	0.3689
Logistic	36.1345	0.6387	0.7992
SGD	36.1345	0.6387	0.7992

จากการทดสอบทำให้ทราบว่า ผลลัพธ์ของแบบจำลองที่สร้างขึ้นโดยวิธีการจำแนกข้อมูลแบบข่ายงานเบย์ (Bayesian Network Classifiers) เป็นแบบจำลองที่สามารถค้นหาความสัมพันธ์ของข้อมูลชุดนี้ได้ดี และมีประสิทธิภาพที่สุด ซึ่งให้ค่าความถูกต้องของการจำแนกประเภทข้อมูลผลการชำระเงินเท่ากับร้อยละ 82.3529 ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) เท่ากับ 0.2562 และค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสอง (RMSE) เท่ากับ 0.3689 โดยผู้วิจัยจะนำแบบจำลองจากวิธีการจำแนกข้อมูลแบบข่ายงานเบย์ไปใช้เพื่อพยากรณ์ข้อมูลชุดการพยากรณ์ในขั้นตอนต่อไป

3.3 ผลการพยากรณ์ผลลัพธ์จากแบบจำลองที่ถูกคัดเลือก

ผู้วิจัยได้ใช้ข้อมูลเพิ่มอีกจำนวน 1 ชุด ระยะเวลาข้อมูล 1 เดือน จำนวนข้อมูล 122 ระเบียน เพื่อทำการพยากรณ์ผลลัพธ์ด้วยวิธีการจำแนกประเภทด้วยวิธีข่ายงานเบย์ (Bayesian Network Classifiers) โดยใช้ฟังก์ชัน Visualize Classifier Error ในโปรแกรม Weka สำหรับการแสดงผลลัพธ์การพยากรณ์ข้อมูลอย่างเต็มรูปแบบ ดังจะเห็นได้จากการแสดงผลลัพธ์ของข้อมูลการชำระเงินทั้งหมด 122 ระเบียน แบบจำลองจากวิธีการจำแนกข้อมูลแบบข่ายงานเบย์สามารถทำนายผลลัพธ์ได้ถูกต้อง 114 ระเบียน คิดเป็นร้อยละ 93.4426 และทำนายผลลัพธ์ไม่ถูกต้อง 8 ระเบียน คิดเป็นร้อยละ 6.5574 จากการแสดงผลการพยากรณ์จากฟังก์ชัน ARFF Viewer ของโปรแกรม Weka ดังรูปที่ 4 สามารถสรุปได้ว่า แบบจำลองจากการจำแนกข้อมูลแบบข่ายงานเบย์สามารถจำแนกกลุ่มผลการชำระเงินที่เกินกำหนด (OVERDUE) ได้ถูกต้อง 72 ระเบียน จากผลการชำระเงินที่เกินกำหนดทั้งหมด 78 ระเบียน และจำแนกกลุ่มผลการชำระเงินที่ไม่เกินกำหนด (ON DUE) ได้ถูกต้อง 42 ระเบียน



ARFF-Viewer - C:\Users\saree\OneDrive\Desktop\predict result.arff

File Edit View

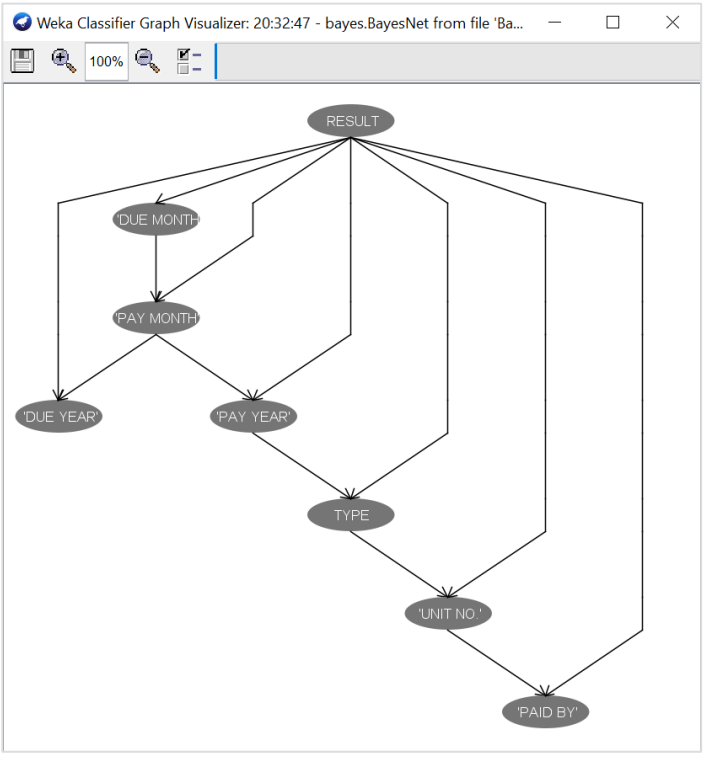
predict result.arff * | predict result.arff

Relation: 1 Train_Split Date-weka.filters.unsupervised.attribute.Remove-R13-weka.filters.unsupervised.attribute.Remove-R1,6-7,10-weka.filters.unsupervised.attribute.Remove-R6-weka...

No.	1: DUE MONTH Nominal	2: DUE YEAR Nominal	3: TYPE Nominal	4: UNIT NO. Nominal	5: PAID BY Nominal	6: PAY MONTH Nominal	7: PAY YEAR Nominal	8: prediction margin Numeric	9: predicted RESULT Nominal	10: **RESULT** Nominal
1	SEP	20	COMPANY	1-001	Transfer Money	SEP	20	0.630803	ON DUE	ON DUE
2	SEP	20	COMPANY	1-002	Transfer Money	SEP	20	0.762744	OVERDUE	OVERDUE
3	OCT	20	PERSONAL	1-003	Credit Card	SEP	20	0.987401	ON DUE	ON DUE
4	SEP	20	PERSONAL	1-007	Transfer Money	SEP	20	0.408935	ON DUE	ON DUE
5	SEP	20	PERSONAL	1-008	Credit Card	SEP	20	0.795855	OVERDUE	OVERDUE
6	SEP	20	COMPANY	1-009	Transfer Money	SEP	20	0.863059	ON DUE	ON DUE
7	SEP	20	COMPANY	1-011	Transfer Money	SEP	20	0.782267	OVERDUE	OVERDUE
8	OCT	20	COMPANY	1-012	Transfer Money	SEP	20	0.994183	ON DUE	ON DUE
9	SEP	20	COMPANY	1-013	Transfer Money	SEP	20	0.862783	OVERDUE	OVERDUE
10	JUL	20	COMPANY	1-015	Transfer Money	SEP	20	0.99163	OVERDUE	OVERDUE
11	AUG	20	COMPANY	1-015	Transfer Money	SEP	20	0.99876	OVERDUE	OVERDUE
12	SEP	20	PERSONAL	1-016	Transfer Money	SEP	20	0.795855	OVERDUE	OVERDUE
13	JUL	20	PERSONAL	1-017	Credit Card	SEP	20	0.991192	OVERDUE	OVERDUE
14	SEP	20	COMPANY	1-019	Transfer Money	SEP	20	0.863059	ON DUE	ON DUE
15	SEP	20	PERSONAL	1-024	Transfer Money	SEP	20	0.021607	OVERDUE	OVERDUE
16	SEP	20	PERSONAL	1-026	Transfer Money	SEP	20	0.85358	ON DUE	ON DUE
17	SEP	20	PERSONAL	1-027	Transfer Money	SEP	20	0.580724	OVERDUE	OVERDUE
18	SEP	20	PERSONAL	1-028	Transfer Money	SEP	20	0.28787	ON DUE	ON DUE
19	SEP	20	PERSONAL	1-029	Credit Card	SEP	20	0.28787	ON DUE	ON DUE
20	AUG	20	PERSONAL	1-029	Credit Card	SEP	20	-0.876143	OVERDUE	ON DUE
21	SEP	20	PERSONAL	1-030	Transfer Money	SEP	20	0.156848	ON DUE	ON DUE
22	SEP	20	PERSONAL	1-040	Amex Charge o...	SEP	20	0.681436	OVERDUE	OVERDUE
23	SEP	20	PERSONAL	1-050	Transfer Money	SEP	20	0.795855	OVERDUE	OVERDUE
24	SEP	20	PERSONAL	1-051	Transfer Money	SEP	20	0.622424	ON DUE	ON DUE
25	SEP	20	COMPANY	1-052	Cheque	SEP	20	0.813028	OVERDUE	OVERDUE
26	SEP	20	COMPANY	1-054	Transfer Money	SEP	20	0.782267	OVERDUE	OVERDUE
27	SEP	20	PERSONAL	1-056	Transfer Money	SEP	20	0.958946	ON DUE	ON DUE
28	SEP	20	PERSONAL	1-057	Transfer Money	SEP	20	0.589378	OVERDUE	OVERDUE
29	JUL	20	PERSONAL	1-060	Transfer Money	SEP	20	0.98574	OVERDUE	OVERDUE
30	SEP	20	COMPANY	1-063	Transfer Money	SEP	20	0.341566	OVERDUE	OVERDUE

รูปที่ 4 แสดงผลการพยากรณ์ (บางส่วน) จากฟังก์ชันการแสดงผล ARFF Viewer

จากผลการชำระเงินที่ไม่เกินกำหนด ทั้งหมด 44 ระเบียน ซึ่งความถูกต้องคิดเป็นร้อยละของการจำแนกผลการชำระเงินที่เกินกำหนดและไม่เกินกำหนด เท่ากับร้อยละ 92.31 และร้อยละ 95.45 ตามลำดับ อีกทั้งการใช้ฟังก์ชันการแสดงผลแบบ Graph Visualizer แสดงให้เห็นความสัมพันธ์ของคุณลักษณะต่าง ๆ ในการทำนายผลลัพธ์ของการชำระเงินจากข้อมูลชุดพยากรณ์ ด้วยการจำแนกประเภทข้อมูลโดยวิธีข่ายงานเบย์ ซึ่งตัวแปรแต่ละตัวมีการเชื่อมต่อความสัมพันธ์กันไปในทิศทางใดดังรูปที่ 5



รูปที่ 5 การแสดงการเชื่อมต่อโหนดของความสัมพันธ์จากการพยากรณ์ผลการชำระเงินด้วยแบบจำลองจากวิธีการจำแนกประเภทข้อมูลแบบข่ายงานเบย์

3.4 ผลการทดสอบสมมติฐาน

สมมติฐาน คือ ผลการพยากรณ์การจำแนกผลการชำระเงิน โดยแบบจำลองจากวิธีการจำแนกข้อมูลที่ถูกเลือกจะมีค่าความคลาดเคลื่อนจากผลลัพธ์จริงไม่เกิน ร้อยละ 15.00 จากผลจากการพยากรณ์ผลลัพธ์การจำแนกกลุ่มผลการชำระเงินด้วยข้อมูลชุดการพยากรณ์ (Forecasting Data Set) ซึ่งใช้แบบจำลองจากวิธีการจำแนกข้อมูลแบบข่ายงานเบย์ พบว่าการพยากรณ์ผลลัพธ์ด้วยการจำแนกประเภทข้อมูลผลการชำระเงินมีค่าความถูกต้องเท่ากับร้อยละ 93.4426 สามารถจำแนกกลุ่มผลการชำระเงินที่ชำระเงินเกินกำหนด (Overdue) ได้ถูกต้องร้อยละ 92.31 และจำแนกกลุ่มผลการชำระเงินที่ไม่เกินกำหนด (On Due) ได้ถูกต้องร้อยละ 95.45 จะเห็นได้ว่าการพยากรณ์ผลการชำระเงินที่เกินกำหนด และไม่เกินกำหนดมีค่าความคลาดเคลื่อนสัมพัทธ์ (Relative Error) เท่ากับร้อยละ 7.69 และร้อยละ 4.55 ตามลำดับ

4. อภิปรายผลและสรุป

จากการศึกษา ทำให้ได้ตัวแบบการเรียนรู้และการพยากรณ์ที่มีค่าความถูกต้องสูงในการจำแนกผลการชำระเงินของลูกค้า เพื่อใช้ทำนายข้อมูลการชำระเงินของลูกค้าในอนาคต เมื่อมีการทดสอบด้วยชุดข้อมูลทดสอบแล้วแบบจำลอง

ที่ใช้เป็นตัวแบบการพยากรณ์ที่ได้จากการจำแนกประเภทข้อมูลด้วยวิธีข่ายงานเบย์ [1] เป็นแบบจำลองเพียงหนึ่งเดียวที่ค่าความถูกต้อง และมีค่าความคลาดเคลื่อนน้อยที่สุดตามผลลัพธ์เมทริกซ์วัดประสิทธิภาพ ซึ่งแสดงให้เห็นว่าแบบจำลองที่ได้จากวิธีการจำแนกข้อมูลแบบข่ายงานเบย์ เป็นแบบจำลองที่มีความเหมาะสมที่สุดในการจำแนกผลการชำระเงินในการศึกษาครั้งนี้ โดยสามารถนำแบบจำลองการจำแนกผลการชำระเงินนี้ไปต่อยอด เพื่อสนับสนุนให้มีการปรับเปลี่ยนนโยบายในการวางแผนการจัดการลูกหนี้ที่มีแนวโน้มว่าจะชำระเงินเกินกำหนดได้อย่างทันท่วงที อีกทั้งทำให้เกิดสภาพคล่องทางการเงินที่ดีขึ้นขององค์กร ตลอดจนสามารถวางแผนการหมุนเวียนเงินสดภายในองค์กรได้อย่างเหมาะสม

นอกจากนี้ จากการศึกษาการสร้างแบบจำลองต้นแบบสำหรับการพยากรณ์การจำแนกผลการชำระเงินนั้น ผลลัพธ์ที่ได้เป็นไปตามวัตถุประสงค์ที่ตั้งไว้คือ เพื่อสร้างแบบจำลองการจำแนกผลการชำระเงินโดยเปรียบเทียบค่าความถูกต้องของแต่ละอัลกอริทึมในการจำแนกข้อมูลผลการชำระเงินโดยใช้เทคนิคการจำแนกประเภท [6] ซึ่งจากผลการทดสอบแบบจำลองด้วยข้อมูลชุดทดสอบ (Testing Data Set) [7] พบว่าการจำแนกประเภทข้อมูลผลการชำระเงินมีแบบจำลองที่สามารถจำแนกประเภทโดยมีค่าความถูกต้องมากกว่า ร้อยละ 80.00 เพียง 1 วิธีการจำแนกประเภทข้อมูล นั่นคือ วิธีการจำแนกข้อมูลแบบข่ายงานเบย์ ซึ่งมีค่าความถูกต้องเท่ากับ ร้อยละ 82.3529 แสดงให้เห็นว่าสมมติฐานข้อที่ 1 ที่กล่าวว่า แบบจำลองการจำแนกผลการชำระเงินในการศึกษานี้มีค่าความถูกต้องมากกว่าร้อยละ 80.00 ทุกวิธีการจำแนกข้อมูล และสมมติฐานข้อที่ 2 ที่กล่าวว่าแบบจำลองการจำแนกผลการชำระเงินในการศึกษานี้จะมีแบบจำลองที่มีค่าความถูกต้องมากกว่าร้อยละ 90.00 อย่างน้อย 1 วิธี นั้นไม่เป็นจริง ซึ่งสอดคล้องกับงานวิจัยของ สุรชัย และคณะ [8] เกี่ยวกับชุดข้อมูลที่นำมาใช้ซึ่งเป็นข้อมูลทางการเงินว่า ข้อมูลทางการเงินโดยส่วนใหญ่จะเป็นระบบที่ไม่เป็นเชิงเส้นและข้อมูลมักจะไม่มีความนิ่ง มีความยากที่จะเรียนรู้พฤติกรรมของข้อมูล ทำให้ยากต่อการสร้างตัวแบบพยากรณ์ที่จะให้ค่าการพยากรณ์ที่แม่นยำ เช่น ราคาหลักทรัพย์ที่ซื้อขายกันในตลาดหลักทรัพย์ ซึ่งเป็นข้อมูลที่มีความผันผวนสูง และยากต่อการคาดคะเน

4.1 ข้อเสนอแนะจากการศึกษา

ในการศึกษาครั้งนี้ ผลการศึกษาได้ตัวแบบการเรียนรู้และการพยากรณ์ที่มีค่าความถูกต้องสูงในการจำแนกผลการชำระเงินของลูกค้า เพื่อใช้ทำนายข้อมูลการชำระเงินของลูกค้าในอนาคต จึงมีความเป็นไปได้ที่องค์กรสามารถนำไปประยุกต์ใช้เพื่อให้เกิดประโยชน์ในด้านการติดตามหนี้สินจากลูกค้าตามกำหนดระยะเวลาได้ องค์กรควรจัดให้มีการอบรมการใช้โปรแกรมเพื่อประโยชน์ทางด้านการเงินโดยควรพิจารณาข้อจำกัดในการนำไปใช้ด้วย

สำหรับข้อจำกัดในการศึกษาทดลองในครั้งนี้ พบว่า ด้านระยะเวลาการจัดทำโครงงานพิเศษ ส่งผลให้ข้อมูลที่เตรียมนั้นมีจำนวนข้อมูลไม่เพียงพอต่อการสนับสนุนการเรียนรู้ของโปรแกรม Weka ทำให้แบบจำลองยังขาดความแม่นยำในการจำแนกประเภทข้อมูลด้วยอัลกอริทึมที่หลากหลาย ดังนั้นผู้ศึกษาจึงมีความเห็นว่าการเก็บรวบรวมข้อมูลนั้นหากข้อมูลยิ่งมีจำนวนมากจะส่งผลให้การวิเคราะห์ข้อมูลด้วยการเรียนรู้ของเครื่องจักรนั้นมีประสิทธิภาพที่สูงขึ้นและมีความแม่นยำมากขึ้น นอกจากนี้ผลการพิสูจน์สมมติฐานในเรื่องของค่าความถูกต้องของแบบจำลองแม้ว่าจะไม่เป็นไปตามคาดหวังที่ตั้งไว้ อาจเป็นเพราะข้อมูลการรับชำระเงินนั้นมีตัวแปรหลักมาจากความสามารถในการใช้จ่ายของลูกค้าในช่วงสถานการณ์ที่ไม่ปกติขณะเกิดโรคระบาดจากเชื้อไวรัสโคโรน่า (Covid-19) ซึ่งเป็นสถานการณ์วิกฤติที่ไม่คาดคิด จึงส่งผลทำให้การพยากรณ์จากแบบจำลองที่สร้างขึ้นด้วยชุดข้อมูลที่มีความคาบเกี่ยวระหว่างช่วงภาวะเศรษฐกิจที่ผกผันนั้น

มีความคลาดเคลื่อนไปมาก แต่ก็ยังสามารถนำมาเป็นข้อพิสูจน์ได้ว่าการใช้ข้อมูลสอนและชุดข้อมูลทดสอบเป็นข้อมูลในช่วงเวลาที่เศรษฐกิจคงที่ และใช้ชุดข้อมูลที่มีเงื่อนไขครอบคลุมก็จะทำให้องค์กรได้รับประโยชน์อย่างเต็มที่

4.2 ข้อเสนอแนะในการศึกษาครั้งต่อไป

ในการศึกษาเพื่อให้เป็นประโยชน์ในการบริหารธุรกิจ องค์กรควรแสวงหาโปรแกรมที่สามารถเข้ามาช่วยในงานด้านการบริหารธุรกิจ และนำมาทำการศึกษาทดลองเพื่อให้เป็นประโยชน์ต่อการบริหารงานในองค์กร เช่น การปรับปรุงคุณภาพการเก็บข้อมูลทางการตลาด การปรับปรุงกระบวนการอนุมัติการเบิกจ่ายงบประมาณ เป็นต้น

เอกสารอ้างอิง

[1] Wongpipan, T. (2013). *The Use of Data Mining to Assist in Credit Decision Making. A Case Study of Krungthai Car Rent and Lease Public Company Limited.* [Unpublished master's thesis]. Dhurakij Pundit University Bangkok. (in Thai).

[2] Chumvijarn, S. (2015). Bayesian Statistics for the Social Science. *Journal of Research Methodology, 2(28),* 271-277. (in Thai).

[3] Pungpapong, V. (2015). A Brief Review on High-dimensional Linear Regression. *Thai Science and Technology Journal*, *23(2)*, 212–223. (in Thai).

[4] Panyakhun, S., Klaisongkhram, T., & Pongsart, T. (n.d.). *Bayesian Networks and Applications*. Khon Kaen University. (in Thai).

[5] Sinsomboonthong, S. (2018). An Efficiency Comparison in Prediction of Child and Adolescence Game Addition in Bangkok. *Thai Science and Technology Journal*, *3(26),* 405–414. (in Thai).

[6] Kotuangchan, O. (2018). *Classification of overdue invoices. with machine learning techniques*. [Bachelor's thesis]. King Mongkut's University of Technology North Bangkok. (in Thai).

[7] Naphaaroonchai, S. (2019). Price Forecasting by Machine Learning Using Weka Program, Case Study of Conrad Hotel, Bangkok. [Bachelor's thesis]. King Mongkut's University of Technology North Bangkok. (in Thai).

[8] Chancharat, S., Musikapodoke, R., & Khermkhan, J. (2013). Prediction Efficiency with Artificial Neural Network:Case Study of the SET Index. *The Journal of King Mongkut's University of Technology North Bangkok, 23(3*), 706–714. (in Thai).