การพยากรณ์ราคาหลักทรัพย์กลุ่มสินค้าอุปโภคบริโภค ในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้เทคนิคการเรียนรู้ของเครื่อง

Forecasting Stock Price of Consumer products group in the Stock Exchange of Thailand Using Machine Learning

พัฒนพงศ์ กุมสิน[1] อุทิศ พงศ์จิรวัฒนา[2] และอนุพงศ์ สุขประเสริฐ[3*]

[1,2,3*]คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

Pattanapong Kumsin[1] Utis bhongchirawattana[2] and Anupong Sukprasert[3*]

[1,2,3*]Mahasarakham Business School, Mahasarakham University

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบและเปรียบเทียบประสิทธิภาพการพยากรณ์ราคาหลักทรัพย์กลุ่มสินค้าอุปโภคบริโภค ในตลาดหลักทรัพย์แห่งประเทศไทย ด้วยเทคนิคการเรียนรู้ของเครื่อง โดยใช้ข้อมูลเกี่ยวกับปัจจัยที่ส่งผลต่อราคาหลักทรัพย์กลุ่มสินค้าอุปโภคบริโภค ซึ่งได้มาจากหลายแหล่งข้อมูล ประกอบด้วย ธนาคารแห่งประเทศไทย ตลาดหลักทรัพย์แห่งประเทศไทย ฐานข้อมูลดัชนีเศรษฐกิจการค้า เว็บไซต์ Investing.com และข้อมูลเปิดเผยของกระทรวงพาณิชย์ ระหว่างวันที่ 1 มกราคม 2561 ถึง 1 กันยายน 2566 ที่ถูกจัดเก็บในรูปแบบรายเดือน เพื่อนำมาสร้างแบบจำลองในการพยากรณ์ ตามกระบวนการมาตรฐานในการทำเหมืองข้อมูล (CRISP-DM) โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ทั้ง 4 เทคนิค ได้แก่ เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor) เทคนิคโครงข่ายประสาทเทียม (Neural Network) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคการถดถอยเชิงเส้น (Linear Regression) ผลการวิจัยพบว่า เทคนิคเทคนิคเพื่อนบ้านใกล้ที่สุด เป็นตัวแบบที่ให้ค่าประสิทธิภาพการพยากรณ์แม่นยำมากที่สุด โดยมีค่าเฉลี่ยผิดพลาดกำลังสองเท่ากับ 25.79 ค่ารากที่สองของค่าเฉลี่ยผิดพลาดกำลังสองเท่ากับ 5.08 รากที่สองของค่าผิดพลาดสัมพัทธ์เท่ากับ 0.34 และค่าสัมประสิทธิ์การตัดสินใจพหุคูณมีค่าเท่ากับ 89.80% ซึ่งผลการวิเคราะห์ข้อมูลในครั้งนี้สามารถนำไปใช้เพื่อเป็นแนวทางในการตัดสินใจเพื่อการลงทุนและการวิเคราะห์ตลาดหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค ได้อย่างมีประสิทธิภาพ ช่วยลดอัตราความเสี่ยงของการลงทุนที่เกิดขึ้นได้ในอนาคต

คำสำคัญ : การพยากรณ์; ราคาหลักทรัพย์; สินค้าอุปโภคบริโภค; โครงข่ายประสาทเทียม; เพื่อนบ้านใกล้ที่สุด; ซัพพอร์ตเวกเตอร์แมชชีน; การวิเคราะห์ถดถอยเชิงเส้น;

[1]นิสิตปริญญาโท หลักสูตรบัญชีมหาบัณฑิต (Graduate Student, Master of Accountancy Program)

[2,3*] อาจารย์ประจำคณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม E-mail: anupong.s@acc.msu.ac.th

Abstract

The purpose of this research aims to develop a model and compared the performance of the model for predicting stock prices of consumer products group in the Stock Exchange of Thailand using machine learning. Additionally, the study aims to investigate how well stock prices in this industry forecast using various prediction models. Monthly data was collected from several sources between January 1, 2018 to September 1, 2023 these sources included the Bank of Thailand, the Stock Exchange of Thailand, economic trade index databases, the Investing.com website, and publicly available data from the Ministry of Commerce. The data was analyzed using a Cross-Industry Standard Process for Data Mining (CRISP-DM). Four machine learning techniques K-Nearest Neighbor, Neural Network, Support Vector Machine, and Linear Regression are used in this study. The result of the research shows that K-Nearest Neighbor is the most accurate forecasting model, according to the study's findings. The values of Mean Squared Error (MSE) is 25.79, Root Mean Squared Error (RMSE) is 5.08, Root Relative Squared Error (RRSE) is 0.34, and R-squared are 89.80%. In the consumer products and consumer goods sector, this finding might help guide investment choices and stock market research, thereby reducing future risks.

Keyword: Forecasting; Stock Prices; Consumer Products; Neural Network; K-nearest Neighbors; Support Vector Machine; Linear Regression;

บทนำ

การลงทุนในตลาดหลักทรัพย์ผู้ลงทุนสามารถทำได้ในหลายวิธี ซึ่งรวมถึงการลงทุนในหุ้น, กองทุนรวม, อสังหาริมทรัพย์, สินทรัพย์ดิจิทัลและพันธบัตรหนี้ เมื่อผู้ลงทุนตัดสินใจลงทุนในตลาดหลักทรัพย์ ควรพิจารณาแนวโน้มผลตอบแทนในอนาคต รวมทั้งแนวโน้มภาวะเศรษฐกิจ อุตสาหกรรม ซึ่งเป็นปัจจัยพื้นฐานเพื่อลดความเสี่ยงจากการลงทุน จากการทบทวนวรรณกรรมที่ผ่านมาปัจจัยที่สำคัญที่อาจส่งผลกระทบต่อดัชนีราคาหุ้นของอุตสาหกรรมสินค้าอุปโภคบริโภค ได้แก่ อัตราเงินเฟ้อ ดัชนีราคาผู้บริโภค อัตราดอกเบี้ยนโยบาย อัตราแลกเปลี่ยน ดัชนีตลาดหลักทรัพย์ ดัชนีความเชื่อมั่นผู้บริโภค ราคาน้ำมันดิบ และดัชนีอุตสาหกรรมดาวโจนส์ (นวพล ยิ้มแย้ม, 2564), (สุพาพร ยันศรีสิริชัย, 2561) ดังนั้นผู้ลงทุนควรมีการวางแผน ศึกษาปัจจัยที่ส่งผลต่อดัชนีหลักทรัพย์อุตสาหกรรมสินค้าอุปโภคบริโภค อย่างรอบคอบและมีความพร้อมที่รับมือความเสี่ยงที่เกิดขึ้น หรือนำเทคโนโลยีที่สามารถวิเคราะห์ข้อมูลที่มีความซับซ้อนเพื่อช่วยในการตัดสินใจในการลงทุน

การทำนายลำดับเวลาโดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เป็นวิธีหนึ่งที่ใช้สำหรับการวิเคราะห์ปัจจัยที่เกี่ยวข้องกับการเปลี่ยนแปลงราคาหลักทรัพย์ในอดีต โดยใช้ระบบคอมพิวเตอร์ที่เรียนรู้จากชุดข้อมูลที่มีอยู่ เพื่อนำไปทำนายผลลัพธ์หรือแนวโน้มราคาหลักทรัพย์ในอนาคต การวิเคราะห์ข้อมูลเพื่อช่วยในการตัดสินใจหรือวางแผนที่แม่นยำขึ้น จากการศึกษาพบว่าข้อมูลราคาหลักทรัพย์มักเป็นข้อมูลการแจกแจงที่ไม่ปกติ ในอดีตเคยใช้เทคนิคอนุกรมเวลา เช่น ถั่วเฉลี่ยเคลื่อนที่ (Moving Average), การปรับเรียบเอกซ์โพเนนเชียล (Exponential Smoothing) และ Autoregressive Integrated Moving Average (ARIMA) และพบว่าการทำนายราคาหลักทรัพย์ด้วยเทคนิคทางสถิติเหล่านี้อาจมีความแม่นยำต่ำ (กาญจน์ภิวรรณ จงศิริวิโรจ, 2564) ปัจจุบันมีเทคนิคใหม่ที่นำเข้ามาใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาเพื่อทำนายข้อมูลในอดีตตามช่วงเวลา ซึ่งทำให้นักวิจัยหลายท่านจึงนำเทคนิคการเรียนรู้ของเครื่องมาใช้ในการพยากรณ์ข้อมูลลำดับเวลากันอย่างแพร่หลายในปัจจุบัน เช่น วิรากานต์ กิตติบวรกุล, ศรายุทธ นนท์ศิริ แลพิชิตชัย คำอินทร์ (2565) ได้ศึกษา การเปรียบเทียบ

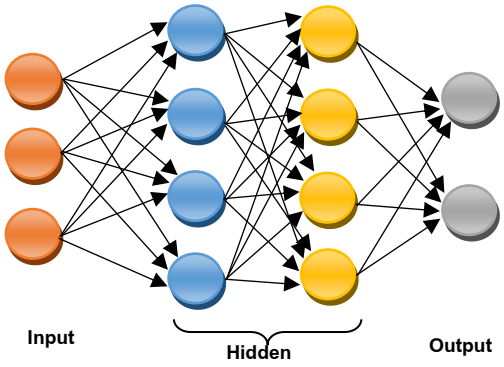
ประสิทธิภาพของเครื่องสำหรับการบำรุงรักษาเชิงคาดการณ์สำหรับเครื่องยนต์อากาศยาน ผลการศึกษาพบว่า เทคนิคโครงข่ายประสาทเทียมให้ผลลัพธ์ที่ดีที่สุดด้วยค่าสัมประสิทธิ์การกำหนดเท่ากับ 0.7578 และค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ยเท่ากับ 0.0016 นัทธมน พลายอินทร์, ชนาพันธุ์ ชนาเนตร และวิกานดา ผาพันธ์ (2563) ได้ศึกษา การพยากรณ์ค่า Scale ของแผ่น PCB ชนิด Multilayer โดยใช้การเรียนรู้ของเครื่อง ผลการศึกษาพบว่า เทคนิคซัพพอร์ตเวกเตอร์รีเกรสชันสามารถช่วยเพิ่มความแม่นยำในการกำหนดค่า Scale และลดเวลาในกระบวนการผลิต PCB Multilayer ด้วยค่า MSE และ MAPE ที่ต่ำสุด เพื่อลดความผิดพลาดในการกำหนดค่า Scale อย่างมีประสิทธิภาพ นอกจากนั้น อมรเทพ พึ่งศรี, สมพร ปั่นโภชา และบำรุง พ่วงเกิด (2560) ที่ได้ศึกษา ประสิทธิภาพการพยากรณ์ดัชนีราคา SET50 Index โดยใช้แบบจำลองโครงข่ายประสาทเทียม ผลการศึกษาพบว่า แบบจำลองโครงข่ายประสาทเทียมสามารถพยากรณ์ได้มีประสิทธิภาพสูงเนื่องจากเป็นแบบจำลองแบบไม่ใช่พารามิเตอร์ (Non-Parametric) โดยใช้หลักการเรียนรู้ของแบบจำลองในการสร้างความสัมพันธ์ของค่าน้ำหนัก และโครงสร้างภายในโครงข่าย

จากข้อมูลที่กล่าวมาข้างต้น ผู้วิจัยมองเห็นถึงความสำคัญของการลงทุนในอุตสาหกรรมสินค้าอุปโภคบริโภคและปัจจัยที่ส่งผลกระทบต่อราคาหลักทรัพย์ในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค ในหลายด้าน อีกทั้งการปรับขึ้นของอัตราดอกเบี้ย ทำให้เกิดความผันผวนในตลาดหลักทรัพย์อย่างต่อเนื่อง ผู้วิจัยจึงตั้งวัตถุประสงค์เพื่อศึกษาการพยากรณ์ราคาหุ้นในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภคด้วยเครื่องมือต่าง ๆ เพื่อหาการเปลี่ยนแปลงราคาหุ้นในอดีตและคาดการณ์ราคาหุ้นในอนาคตผ่านแบบจำลองการพยากรณ์ที่ประกอบด้วยเทคนิคโครงข่ายประสาทเทียม (Neural Network) เทคนิคซับพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor):KNN และเทคนิคการวิเคราะห์ถดถอยเชิงเส้น (Linear Regression) เพื่อใช้เป็นข้อมูลกับนักลงทุนที่สนใจ

การทบทวนวรรณกรรม

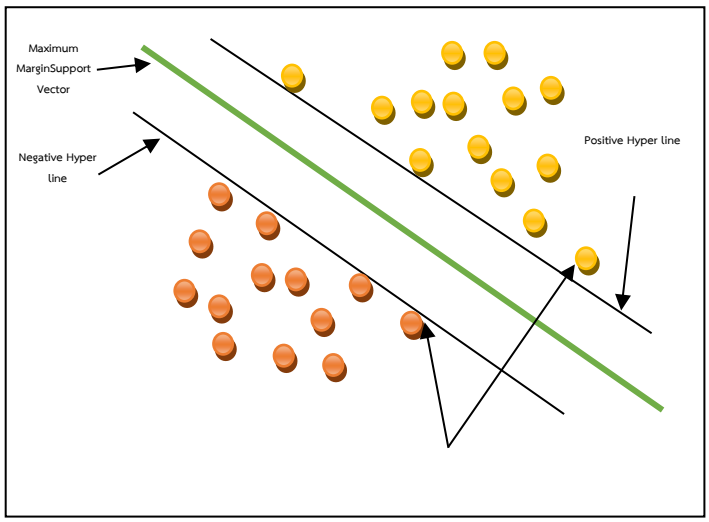
การทำเหมืองข้อมูล (Data mining) เป็นเทคนิคการวิเคราะห์ข้อมูลที่มีจำนวนมากเพื่อหาแนวโน้มหรือความสัมพันธ์ที่ซ่อนภายใน โดยใช้เทคนิคทางสถิติ เชื่อมโยงฐานข้อมูลเพื่อหาความน่าจะเป็นที่จะเกิดขึ้นเพื่อเกิดองค์ความรู้ที่สามารถนำไปพัฒนาหรือเพื่อประกอบการตัดสินใจที่สำคัญในด้านที่เกี่ยวข้อง เทคนิคการพยากรณ์ (Prediction) ถือเป็นเทคนิคหนึ่งในการทำเหมืองข้อมูลโดยใช้การเรียนรู้แบบมีผู้สอน (Supervised Learning) (อนุพงศ์ สุขประเสริฐ, 2564)

เทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network) เป็นตัวแบบทางคณิตศาสตร์ที่จำลองโครงสร้างทางระบบประสาทของมนุษย์เพื่อใช้ในการประมวลผลของข้อมูล ซึ่งจะประกอบไปด้วยโหนดข้อมูลที่เชื่อมต่อกันโดยมีค่าน้ำหนักที่แตกต่างกัน คล้ายกับการส่งสัญญาณของเซลล์ประสาทในระบบประสาท โครงข่ายประสาทเทียมจะประกอบไปด้วยชั้นโหนดที่แบ่งเป็นชั้นอินพุต (Input layer), ชั้นซ่อน (Hidden layers) และชั้นเอาต์พุต (Output Layer) โดยการนำเข้าข้อมูลผ่านโหนดในชั้นอินพุต (Input Layer) ประมวลผลผ่านชั้นซ่อน (Hidden Layer) ก่อนที่จะไปยังชั้นเอาต์พุต (Output Layer) เพื่อให้ผลลัพธ์ที่คาดหวัง ดังภาพที่ 1 (ผุสดี บุญรอด และกรวัฒน์ พลเยี่ยม, 2560)



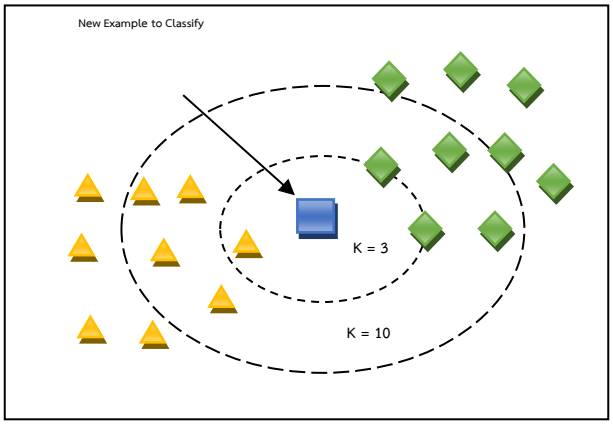
ภาพที่ 1 เทคนิคโครงข่ายประสามเทียม

เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เป็นอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning) โดยใช้ข้อมูลทางสถิติ และเป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) ที่ใช้สำหรับการวิเคราะห์การจำแนกประเภท (Classification) และการวิเคราะห์การถดถอย (Regression) โดยอาศัยหลักการสร้างเส้นแบ่ง (hyperplane) ในการแยกข้อมูลเป็นสองกลุ่มหรือมากกว่า เส้นแบ่งจะต้องอยู่ห่างจากข้อมูลมากที่สุด เพื่อให้สามารถจำแนกข้อมูลได้อย่างถูกต้องและแม่นยำ ดังภาพที่ 2 (Suthaharan. S, 2016)



ภาพที่ 2 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน

เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor) : K-NN เป็นตัวแบบที่ใช้ในการวิเคราะห์ข้อมูลโดยอาศัยการเรียนรู้ของเครื่อง (Machine Learning) และเป็นการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) ใช้ในการจำแนกข้อมูล (Classification) และหาความสัมพันธ์ระหว่างข้อมูล (Regression) โดยวิเคราะห์จากข้อมูลที่ใกล้ที่สุดกับข้อมูลที่ต้องการจำแนก หลักการทำงานของ k-NN คือหาข้อมูลที่ใกล้ที่สุดจากข้อมูลที่ต้องการวิเคราะห์ โดยอาศัยการคำนวณระยะห่างของข้อมูลที่ต้องการจำแนกกับข้อมูลที่อยู่ในชุดข้อมูลเพื่อกำหนดจุดข้อมูลที่ต้องการจำแนกจะอยู่ในกลุ่มหรือคลาสใด ๆ โดยการพิจารณาค่า k (จำนวนข้อมูลใกล้ที่สุดที่จะใช้ในการตัดสินใจ) และนับจำนวนข้อมูลในแต่ละกลุ่มที่ใกล้ที่สุด จากนั้นกำหนดให้จุดข้อมูลใหม่นั้นเป็นของกลุ่มที่มีจำนวนข้อมูลมากที่สุดใน k ข้อมูลที่ใกล้ที่สุด ดังภาพที่ 3 (Kramer. O, 2013)



ภาพที่ 3 เทคนิคเพื่อนบ้านใกล้ที่สุด

เทคนิคการวิเคราะห์ถดถอยเชิงเส้น (Linear Regression) เป็นเทคนิคในการวิเคราะห์ข้อมูลทางสถิติและการเรียนรู้ของเครื่อง (machine learning) ใช้หาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม โดยอาศัยสมการเชิงเส้นในการประมาณค่า โดยทำการปรับพารามิเตอร์ของสมการเส้นตรงเพื่อให้ค่าที่ทำนายมีความเหมาะสมที่สุดกับข้อมูลที่มี โดยมีรูปแบบสมการ (1) ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon \quad (1)$$

โดยที่

Y = ตัวแปรตาม

X_n = ตัวแปรอิสระ

β_n = พารามิเตอร์หรือค่าคงที่ของสมการเส้นตรง

ε = ค่าคงที่ของข้อผิดพลาดทางสถิติ

งานวิจัยที่เกี่ยวข้อง

กรอภิชิต เหง้าพันธ์ และสมพร ปั่นโภชา (2565) ได้ศึกษา การพยากรณ์ความผันผวนของตลาดหลักทรัพย์โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน โดยมีวัตถุประสงค์เพื่อจะศึกษาความสามารถในการพยากรณ์ความผันผวนของดัชนีราคาตลาดหลักทรัพย์แห่งประเทศไทย ผลการวิจัยพบว่า แบบจำลองซัพพอร์ตเวกเตอร์แมชชีนมีค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสองน้อยกว่าแบบจำลอง GARCH ซึ่งแสดงถึงความสามารถในการพยากรณ์ที่ดีกว่า

วสันต์ ศิลปะ, สมพร ปั่นโภชา และบำรุง พ่วงเกิด (2561) ได้ศึกษา การทำนายราคาหลักทรัพย์โดยใช้โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับโดยใช้ตัวแบบอนุกรมเวลาฟังก์ชันแบบไม่เชิงเส้น ผลการวิจัยพบว่า ตัวแบบอนุกรมเวลาฟังก์ชันแบบไม่เชิงเส้น มีความสามารถในการทำนายราคาหุ้นโดยโครงข่ายที่เหมาะสมคือมีสองชั้นซ่อน โดยความคลาดเคลื่อนกำลังสองเฉลี่ย ในชุดข้อมูลฝึกโครงข่ายประสาทเทียมอยู่ที่ 0.0195 และชุดข้อมูล

ทดสอบโครงข่ายประสาทเทียมอยู่ที่ 0.0087 สามารถประยุกต์ใช้ตัวแบบอนุกรมเวลาฟังก์ชันไม่เชิงเส้นในการทำนายราคาหุ้นโดยใช้ปัจจัยพื้นฐานได้ในสถานการณ์จริง

ยศสยา แสงหิรัญ และสมชาย เล็กเจริญ (2561) ได้ศึกษา ปัจจัยที่มีผลต่อการผันแปรของหุ้มกลุ่มเทคโนโลยีโดยใช้เทคนิคการเรียนรู้ของเครื่อง
ผลการศึกษาพบว่า ความแปรปรวนของตัวแปรหุ้นกลุ่มเทคโนโลยีที่มีผลต่อการเปลี่ยนแปลงของราคาปิดของหุ้นมากที่สุดคือ มูลค่าตามบัญชีต่อหุ้น รองลงมาคือ ผลตอบแทนต่อสินทรัพย์ และสุดท้าย
อัตราผลตอบแทนผู้ถือหุ้น เมื่อทดสอบแบบจำลองพบว่าการแยกประเภทแบบการหาเพื่อนบ้านใกล้ที่สุด มีประสิทธิภาพที่ดีที่สุด ซึ่งมีความแม่นยำอยู่ที่ร้อยละ 100

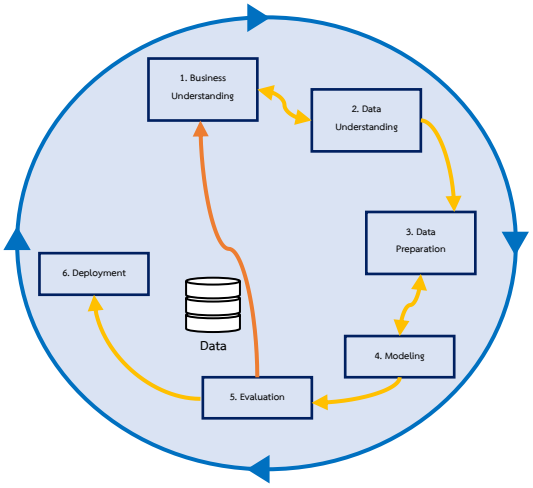
สุพาพร ยันศรีสิริชัย (2561) ได้ศึกษาปัจจัยที่มีผลกระทบต่อราคาหลักทรัพย์ของบริษัทในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค ผลการศึกษาพบว่า ปัจจัยทางเศรษฐกิจที่มีผลต่อราคาหลักทรัพย์ของบริษัทในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภคในทางบวก ได้แก่ อัตราเงินเฟ้อ ส่วนปัจจัยที่ส่งผลในทางลบต่อราคาหลักทรัพย์ของบริษัทในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค ได้แก่ อัตราแลกเปลี่ยน ราคาน้ำมันดิบและผลิตภัณฑ์มวลรวมของประเทศที่ระดับนัยสำคัญทางสถิติ 0.05

นวพล ยิ้มแย้ม (2564) ได้ศึกษา ปัจจัยที่มีผลกระทบต่อดัชนีราคาหลักทรัพย์หมวดกลุ่มสินค้าอุปโภคบริโภคในตลาดหลักทรัพย์แห่งประเทศไทย ผลการศึกษาพบว่า ดัชนีมูลค่าค้าปลีกสามารถอธิบายความผันผวนของดัชนีราคาตลาดหลักทรัพย์ กลุ่มอุตสาหกรรมสินค้าอุปโภค บริโภคได้ร้อยละ 25.04

ผุสดี บุญรอด และกรวัฒน์ พลเยี่ยม (2560) ได้ศึกษาแบบจำลองการพยากรณ์ราคามันสำปะหลัง โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น ผลการศึกษาพบว่า โครงข่ายประสาทเทียมแบบหลายชั้นมีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 3.96 K-NN มีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 7.11 และการวิเคราะห์การถดถอยเชิงเส้นมีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 11.10 แบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นมีประสิทธิภาพดีที่สุด ซึ่งสามารถนำไปประยุกต์ใช้และต่อยอดในการพยากรณ์ราคามันสำปะหลังได้ดีกว่าวิธีอื่น ๆ

วิธีดำเนินการ

กระบวนการในการวิจัยได้ประยุกต์จากกรอบแนวคิด กระบวนการมาตรฐานที่ใช้สำหรับการทำเหมืองข้อมูล (Cross-industry standard process for data mining: CRISP-DM) ประกอบไปด้วย 6 ขั้นตอน ดังนี้ (อนุพงศ์ สุขประเสริฐ, 2564)



ภาพที่ 4 กรอบแนวคิดกระบวนการวิจัยตามรูปแบบ CRISP – DM

ที่มา: ศิรินันท์ มศรีภูมิ และคณะ (2565)

1. ศึกษาข้อมูลผลกระทบที่มีต่อราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค (Business Understanding)

ศึกษาปัจจัยที่ส่งผลต่อราคาของหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภคเพื่อนำมาสร้างตัวแบบและเปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์ราคาหุ้นด้วยเทคนิคการเรียนรู้ของเครื่อง จากการศึกษาพบว่า ปัจจัยที่ส่งผลกระทบต่อราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค มีจำนวน 8 ปัจจัย ได้แก่ 1) อัตราเงินเฟ้อพื้นฐาน (Inflation rate) 2) ดัชนีราคาผู้บริโภค (Consumer Price Index) 3) อัตราดอกเบี้ยนโยบาย (Policy Interest rate) 4) อัตราแลกเปลี่ยน (Exchange rate) 5) ดัชนีราคาตลาดหลักทรัพย์ 6) ดัชนีความเชื่อมั่นผู้บริโภค (Consumer Confidence Index) 7) ราคาน้ำมันดิบ (Crude Oil Price) และ 8) ดัชนีหลักทรัพย์อุตสาหกรรมดาวโจนส์ (Dow Jones Index) (นวพล ยิ้มแย้ม, 2566), (สุพาพร ยันศรีสิริชัย, 2561)

2. การทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding)

ขั้นตอนการรวบรวมข้อมูล การทำความเข้าใจข้อมูล ผลกระทบที่มีต่อราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค และพิจารณาตรวจสอบความถูกต้องของข้อมูลเพื่อให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้

แหล่งข้อมูลการวิจัยในครั้งนี้ได้จาก ธนาคารแห่งประเทศไทย ตลาดหลักทรัพย์แห่งประเทศไทย ฐานข้อมูลดัชนีเศรษฐกิจการค้า
เว็ปไซต์ Investing.com และข้อมูลเปิดเผยของกระทรวงพาณิชย์ ระหว่างวันที่ 1 มกราคม 2561 ถึงวันที่ 30 กันยายน 2566 จำนวน 69 แถว ดังตารางที่ 1

ตารางที่ 1 ข้อมูลที่ใช้ในงานวิจัย

ลำดับที่	ชื่อแอตทริบิวท์	ชนิดข้อมูล	ตัวอย่าง
1	วัน เดือน ปี	Date	1/1/2563
2	อัตราเงินเฟ้อ	real	0.45
3	ดัชนีราคาผู้บริโภค	real	99.81
4	อัตราดอกเบี้ยนโยบาย	real	0.50

ลำดับที่	ชื่อแอตทริบิวท์	ชนิดข้อมูล	ตัวอย่าง
5	อัตราแลกเปลี่ยน	real	32.89
6	ดัชนีตลาดหลักทรัพย์	real	1521.92
7	ดัชนีความเชื่อมั่นผู้บริโภค	real	36.7
8	ราคาน้ำมันดิบ	real	72.9
9	ดัชนีอุตสาหกรรมดาวโจนส์	real	34935.5
10	ดัชนีหลักทรัพย์สินค้าอุปโภคบริโภค	real	99.15

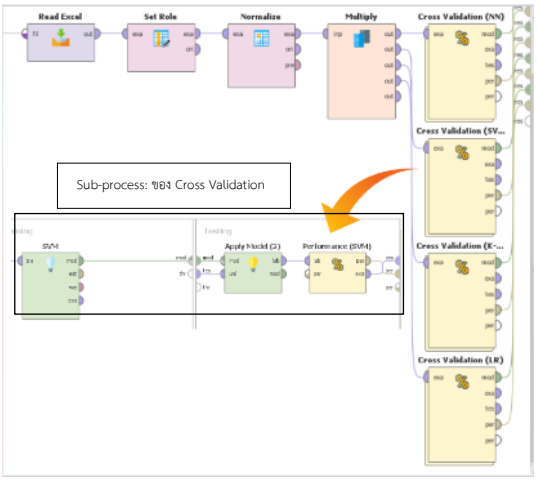
3. การเตรียมข้อมูล (Data Preparation)

3.3.1 การคัดเลือกข้อมูล (Data Selection) เป็นการคัดเลือกข้อมูลที่เหมาะสมเพื่อใช้สำหรับการวิเคราะห์ข้อมูล โดยข้อมูลที่จำเป็นสำหรับการวิเคราะห์ในครั้งนี้มีจำนวน 8 แอตทริบิวต์ ได้แก่ 1) อัตราเงินเฟ้อ 2) ดัชนีราคาผู้บริโภค 3) อัตราดอกเบี้ยนโยบาย 4) อัตราแลกเปลี่ยน 5) ดัชนีตลาดหลักทรัพย์ 6) ดัชนีความเชื่อมั่นผู้บริโภค 7) ราคาน้ำมันดิบ และ 8) ดัชนีอุตสาหกรรมดาวโจนส์

3.3.2 การกำหนดหน้าที่ให้กับตัวแปร เป็นการกำหนดหน้าที่ให้กับแต่ละแอตทริบิวต์เพื่อนำไปใช้ในการวิเคราะห์ โดยกำหนดแอตทริบิวต์ วัน/เดือน/ปี (Date) ให้ทำหน้าที่เป็นไอดี (ID) และกำหนดแอตทริบิวต์ดัชนีหลักทรัพย์อุตสาหกรรมสินค้าอุปโภคบริโภค ให้ทำหน้าที่เป็นตัวแปรตาม (Label) ที่ใช้ในการพยากรณ์ และแอตทริบิวต์อื่น ๆ ให้ทำหน้าที่เป็นตัวแปรอิสระ

3.3.3 การแปลงค่าข้อมูล (Data Transformation) เป็นการแปลงข้อมูลด้วยวิธีนอร์มอลไลซ์ปรับการกระจายค่าของข้อมูล และค่าเบี่ยงเบนมาตรฐาน (Z-Transformation) โดยทำการปรับค่าข้อมูลจำนวน 8 แอตทริบิวต์ ได้แก่ 1) อัตราเงินเฟ้อ 2) ดัชนีราคาผู้บริโภค 3) อัตราดอกเบี้ยนโยบาย 4) อัตราแลกเปลี่ยน 5) ดัชนีตลาดหลักทรัพย์ 6) ดัชนีความเชื่อมั่นผู้บริโภค 7) ราคาน้ำมันดิบ และ 8) ดัชนีอุตสาหกรรมดาวโจนส์ เพื่อปรับปรุงคุณภาพข้อมูล ความเข้ากันได้ และการใช้งาน เพื่อให้มั่นใจว่าการแสดงข้อมูลมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น

4. การสร้างแบบจำลอง (Modeling) เป็นขั้นตอนในการวิเคราะห์ข้อมูลโดยเลือกใช้เทคนิคแบบจำลองการพยากรณ์เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด โดยออกแบบและสร้างแบบจำลองจำนวน 4 แบบจำลอง ประกอบด้วย เทคนิคโครงข่ายประสาทเทียม เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคเพื่อนบ้านใกล้ที่สุด และเทคนิคการวิเคราะห์ถดถอยเชิงเส้น โดยใช้ข้อมูลทางสถิติและสมการทางคณิตศาสตร์ในการคำนวณ และใช้โปรแกรม RapidMiner Studio 10.0 ในการประมวลผล เนื่องจากเป็นโปรแกรมที่ใช้งานง่ายโดยไม่จำเป็นจะต้องมีความรู้เกี่ยวกับการเขียนโปรแกรมก็สามารถวิเคราะห์ข้อมูลได้ ดังภาพที่ 5 (อนุพงศ์ สุขประเสริฐ, 2564)



ภาพที่ 5 การเตรียมข้อมูลและสร้างแบบจำลองการพยากรณ์ โดยใช้โปรแกรม RapidMiner Studio

5. การประเมินผล (Evaluation)

เป็นการประเมินประสิทธิภาพของแบบจำลองการพยากรณ์ได้ดำเนินการวัด ค่าเฉลี่ยความผิดพลาดกำลังสอง รากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง รากที่สองของค่าผิดพลาดสัมพัทธ์ และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ โดยวิธี Cross Validation test โดยแบ่งชุดข้อมูล เป็น 10 ส่วน เพื่อให้ข้อมูลกระจายเท่า ๆ กัน เพื่อสร้างแบบจำลองการพยากรณ์

ค่าเฉลี่ยผิดพลาดกำลังสอง (Mean Squared Error: MSE) คือ ค่าสถิติที่ใช้วัดความแม่นยำของโมเดลการพยากรณ์ โดยคำนวณจากผลต่างระหว่างค่าจริงกับค่าพยากรณ์ที่ได้จากโมเดล แล้วยกกำลังสอง แล้วหาค่าเฉลี่ยของผลลัพธ์ที่ได้คำนวณได้จากสมการที่ (2)

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \quad (2)$$

โดยที่

n คือ จำนวนตัวอย่างข้อมูล

y_t คือ ค่าจริงของข้อมูลตัวอย่างที่ i

$\hat{Y}_t$ คือ ค่าที่โมเดลทำนายสำหรับข้อมูลตัวอย่างที่ t

รากที่สองของค่าเฉลี่ยผิดพลาดกำลังสอง (Root Mean Squared Error: RMSE) คือ ค่าที่บอกถึงความแม่นยำในการพยากรณ์ โดยคำนวณจากผลต่างระหว่างค่าจริงกับค่าพยากรณ์ที่ยกกำลังสอง แล้วหาค่าเฉลี่ย แล้วจึงหารากที่สองของค่าเฉลี่ยนั้น คำนวณได้จากสมการที่ (3)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2} \quad (3)$$

โดยที่

n คือ จำนวนตัวอย่างข้อมูล

y_t คือ ค่าจริงของข้อมูลตัวอย่างที่ i

$\hat{Y}_t$ คือ ค่าที่โมเดลทำนายสำหรับข้อมูลตัวอย่างที่ t

รากที่สองของค่าผิดพลาดสัมพัทธ์ (Root Relative Squared Error: RRSE) คือ ค่าที่บ่งบอกถึงระดับของความแม่นยำในการพยากรณ์ โดยคำนวณจากผลต่างระหว่างค่าจริงกับค่าพยากรณ์ที่ยกกำลังสอง แล้วหารด้วยค่าจริง แล้วหารากที่สองของค่าที่ได้ คำนวณได้จากสมการที่ (4)

$$RRSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (4)$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

$\hat{Y}_t$ คือ ค่าที่ได้จากการพยากรณ์ที่ เวลา t ใด

$\bar{Y}_t$ คือ ค่าเฉลี่ยของค่าจริงที่เวลา t ใด

ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple. Determination: R^2) คือ ค่าที่บ่งบอกถึงปริมาณของความสัมพันธ์ระหว่างตัวแปรสองตัว โดยคำนวณจากค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) ยกกำลังสองคำนวณได้จากสมการที่ (5)

$$R^2 = 1-\left(\frac{\sum(Y-\hat{Y})^2}{\sum(Y-\bar{Y})^2}\right) \qquad (5)$$

โดยที่

n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

$\hat{Y}_t$ คือ ค่าที่ได้จากการพยากรณ์ที่ เวลา t ใด ๆ

$\bar{Y}_t$ คือ ค่าเฉลี่ยของค่าจริงที่เวลา t ใด ๆ

6. การนำไปใช้งาน (Deployment)

เผยแพร่ผลการวิเคราะห์หรือนำผลลัพธ์ที่ได้ไปใช้งาน เพื่อให้นักลงทุนที่สนใจนำไปใช้ประโยชน์และเป็นข้อมูลการพยากรณ์ราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค เพื่อลดความเสี่ยงและมีข้อมูลที่เพียงพอต่อการตัดสินใจ ลงทุน

ผลการศึกษา

ผู้วิจัยได้ทำการสร้างแบบจำลองในการพยากรณ์ราคาหลักทรัพย์ของอุตสาหกรรมสินค้าอุปโภคบริโภคด้วยเทคนิคการเรียนรู้ของเครื่อง และได้ทดสอบประสิทธิภาพในการพยากรณ์ของตัวแบบพยากรณ์ราคาหลักทรัพย์ ซึ่งได้นำผลลัพธ์ของประสิทธิภาพมาเปรียบเทียบ แสดงดังตารางที่ 2

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพการประมาณค่าข้อมูล

Regression Technique	Regression Performance			
	Mean Square Error	Root Mean Square Error	Root Relative Square Error	R-square
Neural Network	27.21	5.22	0.36	88.10%
Linear Regression	77.44	8.80	0.62	77.40%
k-Nearest Neighbor**	25.79	5.08	0.34	89.80%
Support Vector Machine	211.38	14.54	0.78	64.80%

** คือเทคนิคที่มีความเหมาะสมในการสร้างตัวแบบการพยากรณ์ราคาหลักทรัพย์

จากผลการวิเคราะห์ในตารางที่ 2 พบว่าเทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) มีประสิทธิภาพในการพยากรณ์ราคาหลักทรัพย์สูงสุด โดยมีค่าเฉลี่ยผิดพลาดกำลังสอง (MSE) ต่ำสุด เท่ากับ 25.79 รองลงมาคือเทคนิคโครงข่ายประสาทเทียม (Neural Network) 27.21 และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) 211.38 เมื่อพิจารณารากที่สองของค่าเฉลี่ยผิดพลาดกำลังสอง (RMSE) พบว่าเทคนิคเพื่อนบ้านใกล้ที่สุดยังคงมีประสิทธิภาพสูงสุด 5.08 รองลงมาคือเทคนิคโครงข่ายประสาทเทียม 5.22 และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 4.54 สำหรับค่ารากที่สองของค่าเฉลี่ยสัมพัทธ์ (MRE) เทคนิคเพื่อนบ้านใกล้ที่สุดมีค่าต่ำสุด 0.34 รองลงมาคือเทคนิคโครงข่ายประสาทเทียม 0.36 และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 0.78 สำหรับค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R-squared) เทคนิคเพื่อนบ้านใกล้ที่สุดมีค่าสูงสุด 89.80% รองลงมาคือเทคนิคโครงข่ายประสาทเทียม 88.10% และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 64.80% จากผลลัพธ์ข้างต้น สรุปได้ว่าเทคนิคเพื่อนบ้านใกล้ที่สุดมีประสิทธิภาพในการพยากรณ์ราคาหลักทรัพย์สูงสุด

สรุปและอภิปรายผล

จากการศึกษานี้มีวัตถุประสงค์เพื่อสร้างตัวแบบในการพยากรณ์ราคาหลักทรัพย์ในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค และเพื่อเปรียบเทียบประสิทธิภาพการพยากรณ์ราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค ของแบบจำลองการพยากรณ์ โดยใช้ข้อมูลที่ได้ทำการเก็บรวบรวมจาก ธนาคารแห่งประเทศไทย ตลาดหลักทรัพย์แห่งประเทศไทย ฐานข้อมูลดัชนีเศรษฐกิจการค้า เว็ปไซต์ Investing.com และข้อมูลเปิดเผยของกระทรวงพานิชย์ ระหว่างวันที่ 1 มกราคม 2561 ถึงวันที่ 30 กันยายน 2566 จำนวน 69 แถว ซึ่งปัจจัยที่ส่งผลต่อราคาหลักทรัพย์มีทั้งหมด 8 ตัวแปร ได้แก่ อัตราเงินเฟ้อ ดัชนีราคาผู้บริโภค อัตราดอกเบี้ยนโยบาย อัตราแลกเปลี่ยน ดัชนีตลาดหลักทรัพย์ ดัชนีความเชื่อมั่นผู้บริโภค ราคาน้ำมันดิบ และดัชนีอุตสาหกรรมดาวโจนส์ (นวพล ยิ้มแย้ม, 2566), (สุพาพร ยันศรีสิริชัย, 2561) โดยนำมาวิเคราะห์ข้อมูลตามกระบวนการของมาตรฐานการทำเหมืองข้อมูลเพื่อพัฒนาแบบจำลองการพยากรณ์ราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค โดยได้พัฒนาขึ้นจาก 4 แบบจำลอง ได้แก่ เทคนิคโครงข่ายประสาทเทียม เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคเพื่อนบ้านใกล้ที่สุด และเทคนิคการวิเคราะห์ถดถอยเชิงเส้น จากผลการวิจัยได้พบว่า แบบจำลองการพยากรณ์โดยใช้เทคนิคเพื่อนบ้านใกล้ที่สุด มีประสิทธิภาพการพยากรณ์ที่แม่นยำที่สุด โดยมีค่าผิดพลาดกำลังสองเฉลี่ย เท่ากับ 25.79 รากที่สองของค่าผิดพลาดกำลังสองเฉลี่ย เท่ากับ 5.08 รากที่สองของค่าสัมพัทธ์ เท่ากับ 0.34 และค่ากำลังสองของค่าสหสัมพัทธ์ เท่ากับ 89.80% ซึ่งแสดงให้เห็นว่า มีความเหมาะสมต่อการนำ

แบบจำลองดังกล่าวไปพยากรณ์ราคาหลักทรัพย์กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภคได้ดีกว่า ซึ่งสอดคล้องกับงานวิจัยของยศสยา แสงหิรัญ และสมชาย เล็กเจริญ (2561) ที่ได้ศึกษา ปัจจัยที่มีผลต่อการผันแปรของหุ้นกลุ่มเทคโนโลยีโดยใช้เทคนิคการทำเหมืองข้อมูล ผลการวิจัยพบว่า เมื่อทดสอบแบบจำลองพบว่าการแยกประเภทแบบการหาเพื่อนบ้านใกล้ที่สุด มีประสิทธิภาพที่ดีที่สุด ซึ่งมีความแม่นยำอยู่ที่ร้อยละ 100 เทคนิคโครงข่ายประสาทเทียมที่มีค่า R^2 เท่ากับ 88.10% และการวิเคราะห์ถดถอยเชิงเส้น ที่มีค่า R^2 ท่ากับ 81.10% ตามลำดับ อีกทั้งยังสอดคล้องกับ Samruddhi and Kumar (2020) ที่ได้ศึกษาการคาดการณ์ราคารถยนต์มือสองโดยใช้โมเดลเพื่อนบ้านใกล้ที่สุด ผลโดยศึกษาพบว่า โมเดลเพื่อนบ้านใกล้ที่สุดให้ค่าความแม่นยำที่ 85% โดยได้แบ่งกลุ่ม K-Fold เป็น 5 และ 10 ส่วน ในการพัฒนาแบบจำลองซึ่งแสดงให้เห็นว่าโมเดลมีความเหมาะสมที่สุด นอกจากนั้นยังสอดคล้องกับ Kück and Freitag (2021) ที่ได้ศึกษา การพยากรณ์ความต้องการของลูกค้าสำหรับการวางแผนการผลิตโดยใช้โมเดลเพื่อนบ้านใกล้ที่สุด ผลการศึกษาพบว่า โมเดลเพื่อนบ้านใกล้ที่สุด เป็นทางเลือกที่น่าสนใจสำหรับการพยากรณ์ความต้องการของลูกค้าในภาคอุตสาหกรรม เนื่องจากมีความแม่นยำสูง และเหมาะกับชุดข้อมูลเวลาระยะสั้น

ข้อเสนอแนะ

เพื่อให้ได้ข้อมูลการวิเคราะห์ที่ครอบคลุมมากยิ่งขึ้นอาจเปรียบเทียบโดยใช้แบบจำลองการพยากรณ์อื่นเพิ่มเติม เช่น เทคนิคต้นไม้ตัดสินใจ เทคนิคการเรียนรู้เชิงลึก รวมถึงการประยุกต์ใช้เทคนิคแบบผสมผสาน (Hybrid Technique) หรือการเพิ่มชุดข้อมูลให้มีจำนวนมากขึ้น เพื่อพัฒนาประสิทธิภาพแบบจำลอง และปรับค่าพารามิเตอร์อื่น ๆ ให้เหมาะสม เพื่อให้เกิดประโยชน์สูงสุด

บรรณานุกรม

กันต์สุดา อินทรสุวรรณ์ (2564). ปัจจัยทางเศรษฐกิจที่มีผลกระทบต่อดัชนีราคาหลักทรัพย์หมวดขนส่งและโลจิสติกส์. วารสารสังคมศาสตร์ปัญญาพัฒน์, 3(2) 15-24.

กาญจน์ภิวรรณ จงศิริวิโรจ (2564). การเปรียบเทียบวิธีในการพยากรณ์ราคาหุ้นด้วยแบบจำลองอารีม่า, โครงข่ายประสาทเทียม และตัวแบบผสม.Chulalongkorn University Theses and Dissertations (Chula ETD). 5592.

กรอภิชิต เหง้าพันธ์ และสมพร ปั่นโภชา (2565). การพยากรณ์ความผันผวนของตลาดหลักทรัพย์โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน. รายงานสืบเนื่องจากการประชุม การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 17 ปีการศึกษา 2565 (น.10-20). บัณฑิตวิทยาลัย มหาวิทยาลัยหอการค้าไทย.

ตลาดหลักทรัพย์แห่งประเทศไทย (2566). สืบค้นจาก https://www.set.or.th/th/market/index/set/consump.

ธนาคารแห่งประเทศไทย (2566). สืบค้นจาก https://www.bot.or.th/th/ourroles/monetary-policy/mpc-publication/policy-interest-rate.html

นัทธมน พลายอินทร์, ชนาพันธุ์ ชนาเนตร และวิกานดา ผาพันธ์ (2563). การพยากรณ์ค่า Scale ของแผ่น PCB

ชนิด Multilayer โดยใช้การเรียนรู้ของเครื่อง. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม, 39(6), 692-701.

นรินทร์ จิวิตัน (2564). การเปรียบเทียบประสิทธิภาพในการทำนายผลความไม่สมดุลของข้อมูลการเสียชีวิตจากอุบัติเหตุบนโครงข่ายถนน โดยใช้เทคนิคการทำเหมืองข้อมูล. วารสารเทคโนโลยีสารสนเทศ, 18(1), 75-84.

นวพล ยิ้มแย้ม (2566). ปัจจัยที่มีผลกระทบต่อดัชนีราคาหลักทรัพย์หมวดกลุ่มสินค้าอุปโภคบริโภคในตลาดหลักทรัพย์แห่งประเทศไทย.วารสารวิชาการนอร์ทเทิร์น, 10(1), 18-25.

ผุสดี บุญรอด และกรวัฒน์ พลเยี่ยม (2560). แบบจำลองการพยากรณ์ราคามันสำปะหลังโดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น. วารสารวิทยาศาสตร์และเทคโนโลยี, 25(3), 533-543.

เมธาวี พิมพ์จันทร์ (2563). ปัจจัยที่มีอิทธิพลต่อดัชนีกลุ่มอุตสาหกรรมในตลาดหลักทรัพย์แห่งประเทศไทยในสถานการณ์การแพร่ระบาดของโรค COVID-19.(วิทยานิพนธ์ปริญญามหาบัณฑิต มหาวิทยาลัยธุรกิจบัณฑิตย์).

ยศสยา แสงหิรัญ และสมชาย เล็กเจริญ (2561). ปัจจัยที่มีผลต่อการผันแปรของหุ้นกลุ่มเทคโนโลยีโดยใช้เทคนิคการทำเหมืองข้อมูล. รายงานสืบเนื่องจากการประชุม การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 13 ปีการศึกษา 2561 (น.2582-2590). บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต.

วสันต์ ศิลปะ สมพร ปั่นโภชา และบำรุง พ่วงเกิด (2561). การทำนายราคาหลักทรัพย์โดยใช้ โครงข่ายประสาทแบบแพร่ย้อนกลับโดยใช้ตัวแบบอนุกรมเวลาฟังก์ชันแบบไม่เชิงเส้น. รายงานสืบเนื่องจากการประชุม การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 12 ปีการศึกษา 2560 (น.1508-1518), บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต.

วิรากานต์ กิตติบวรกุล, ศรายุทธ นนท์ศิริ แลพิชิตชัย คำอินทร์ (2565). การเปรียบเทียบประสิทธิภาพของเครื่องสำหรับการบำรุงรักษาเชิงคาดการณ์สำหรับเครื่องยนต์อากาศยาน. วารสารวิชาการสมาคมสถาบันอุดมศึกษาเอกชนแห่งประเทศไทย (สสอท.), 11(1), 1-14.

ศิรินันท์ มศรีภูมิ, อนุพงศ์ สุขประเสริฐ และธัญญธร ศรีวิเชียร (2565). ประสิทธิภาพของเทคนิคเหมืองข้อมูลสำหรับพยากรณ์ความสาเร็จทางการศึกษา. โครงการนำเสนอผลงานทางวิชาการระดับชาติ ครั้งที่ 10 (น.114-128). การก้าวไปข้างหน้าของธุรกิจ เศรษฐศาสตร์และการสื่อสารเพื่อสังคมพลวัตและนวัตกรรมแห่งสมัย คณะบริหารธุรกิจ เศรษฐศาสตร์และการสื่อสาร มหาวิทยาลัยนเรศวร.

สุพาพร ยันศรีสิริชัย (2561). ปัจจัยที่มีผลกระทบต่อราคาหลักทรัพย์ของบริษัทในกลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค. รายงานสืบเนื่องจากการประชุม การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 13 ปีการศึกษา 2561 (น.1490-1499). บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต.

สุภาภรณ์ ทิมสำราญ (2564). การพยากรณ์ประสิทธิภาพของหุ้น ในตลาดหลักทรัพย์แห่งประเทศไทย. วารสารรามคำแหง ฉบับบัณฑิตศึกษา, 4(3) 75-87.

อนุพงศ์ สุขประเสริฐ (2564). คู่มือการทำเหมืองข้อมูลด้วย โปรแกรม RAPIDMINER STUDIO (พิมพ์ครั้งที่ 4). มหาสารคาม: คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม.

อมรเทพ พึ่งศรี, สมพร ปั่นโภชา และบำรุง พ่วงเกิด (2560). ประสิทธิภาพการพยากรณ์ดัชนีราคา SET50 Index โดยใช้แบบจำลองโครงข่ายประสาทเทียม. การประชุมนำเสนอผลงานระดับบัณฑิตศึกษา ครั้งที่ 12 มหาวิทยาลัยรังสิต, (น.1519-1528).

K.Samruddhi and Dr R.Ashok Kumar (2020). Used Car Price Prediction using K-Nearest Neighbor Based Model. International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), 4(2), 629-632.

Kramer, O. (2013). K-Nearest Neighbors. In: Dimensionality Reduction with Unsupervised Nearest Neighbors. Intelligent Systems Reference Library, 51.

Kück, Mirko & Freitag Michael (2021). Forecasting of customer demands for production planning by local k-nearest neighbor models, International Journal of Production Economics, Elsevier, 231(C).

Mierswa, I., & Klinkenberg, R. (2018). RapidMiner Studio (10.0) [Data science, machine learning, predictive analytics]. Retrieved from https://rapidminer.com/.

Suthaharan, S. (2016). Support Vector Machine. In: Machine Learning Models and Algorithms for Big Data Classification. Integrated Series in Information Systems, 36, 207-235.