Journal of Information Science Research and Practice

Research Articles

การวิเคราะห์เปรียบเทียบโมเดล ARIMA และ LSTM สำหรับการพยากรณ์ราคาบิตคอยน์: การประยุกต์ใช้เครื่องมือเรียนรู้ของเครื่อง

Comparative Analysis of ARIMA and LSTM Models for Forecasting Bitcoin Prices: A Machine Learning Approach

ธีรภพ แสงศรี [1], เทวา พรหมนุชานนท์ [1*] และ รุจิพันธุ์ โกษารัตน์ [2]

Theeraphop Saengsri [1], Tewa Promnuchanont [1*] and Rujipan Kosarat [2]

[1] หลักสูตรระบบสารสนเทศทางธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ประเทศไทย; Bachelor of Business Administration Program in Business Information System, Faculty of Business Administration and Liberal Arts, Rajamangala University of Technology, Thailand

[2] หลักสูตรวิศวกรรมอิเล็กทรอนิกส์และระบบควบคุมอัตโนมัติ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ประเทศไทย; Bachelor of Engineering Program in Electronics Engineering and Automatic Control Systems, Faculty of Engineering, Rajamangala University of Technology, Thailand

*****Corresponding author email: tewa@rmutl.ac.th

บทคัดย่อ

วัตถุประสงค์: วิเคราะห์และเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ เพื่อระบุโมเดลที่มีความแม่นยำสูงสุดสำหรับการใช้งานในสภาวะตลาดที่แตกต่างกัน และ ประเมินความสามารถของโมเดล ARIMA และ LSTM ในการตอบสนองต่อการเปลี่ยนแปลงของตลาดและความผันผวนของราคาสกุลเงินดิจิทัลเพื่อพัฒนากลยุทธ์การพยากรณ์ราคาที่เหมาะสม

วิธีการศึกษา: ผู้วิจัยได้ดำเนินการเปรียบเทียบประสิทธิภาพของสองโมเดลในการพยากรณ์ราคาบิตคอยน์ ได้แก่ โมเดล ARIMA และ LSTM โดยใช้ชุดข้อมูลราคาบิตคอยน์ ตั้งแต่ปี ค.ศ. 2020 ถึง 2024 จากเว็บไซต์ Yahoo Finance โดยข้อมูลได้มีการแบ่งออกเป็น 2 คือ ชุดฝึกและชุดทดสอบ โดยแบ่งสัดส่วนออกเป็น 80:20 (ชุดฝึก 80% และชุดทดสอบ 20%) และมีการประเมินผลลัพธ์ของการพยากรณ์โดยใช้เมตริกสำคัญ ได้แก่ RMSE และ MAE เพื่อวัดความแม่นยำและประสิทธิภาพของแต่ละโมเดล การวิเคราะห์นี้ช่วยให้สามารถทำความเข้าใจถึงจุดแข็งและจุดอ่อนของแต่ละโมเดลในการจัดการกับความผันผวนและความไม่แน่นอนของตลาดสกุลเงินดิจิทัล

ข้อค้นพบ: จากการศึกษาเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ ข้อค้นพบหลักคือ โมเดล LSTM มีประสิทธิภาพสูงกว่าโมเดล ARIMA อย่างชัดเจนในเงื่อนไขของความแม่นยำ โดยโมเดล LSTM สามารถจับความผันผวนของราคาและปรับตัวตามการเปลี่ยนแปลงของตลาดได้ดีกว่า ซึ่งสะท้อนให้เห็นผ่านค่า RMSE และ MAE ที่ต่ำกว่าเมื่อเทียบกับโมเดล ARIMA นอกจากนี้ ผลการศึกษา

Received: May 9, 2024 Revised: June 17, 2024 Accepted: July 17, 2024 https://doi.org/10.14456/jiskku.2024.22

ยังชี้ให้เห็นว่าโมเดล LSTM มีความเหมาะสมมากขึ้นในการจัดการกับชุดข้อมูลที่มีความซับซ้อนและความผันผวนสูง เช่น ข้อมูลราคาของสกุลเงินดิจิทัล

การประยุกต์ใช้จากการศึกษานี้: จากการศึกษาการเปรียบเทียบโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ ผลการวิจัยมีการประยุกต์ใช้ที่สำคัญหลายประการในแวดวงการเงินและเทคโนโลยี การค้นพบว่าโมเดล LSTM มีประสิทธิภาพสูงกว่าในการจับความผันผวนของตลาด สามารถนำไปพัฒนาเป็นเครื่องมือที่ช่วยในการตัดสินใจลงทุนในตลาดสกุลเงินดิจิทัลได้ นอกจากนี้ผลการวิจัยนี้ยังเป็นประโยชน์สำหรับนักพัฒนาที่ทำงานในการพัฒนาแพลตฟอร์มการเทรดอัตโนมัติ การทำนายที่ดีขึ้นสามารถนำไปสู่การจัดการความเสี่ยงที่ดีขึ้นและโอกาสในการทำกำไรที่เพิ่มขึ้นสำหรับนักลงทุนในตลาดสกุลเงินดิจิทัล

คำสำคัญ: การพยากรณ์ราคาบิตคอยน์ การวิเคราะห์ชุดข้อมูลเชิงเวลา โมเดล LSTM โมเดล ARIMA ความแม่นยำในการพยากรณ์

Abstract

Purpose: The purpose of this study is to evaluate the effectiveness of ARIMA and LSTM models in forecasting Bitcoin prices. This research aims to identify which model provides the highest accuracy and is best suited for application under varying market conditions. Additionally, it seeks to determine the model's ability to adapt to and respond to market dynamics and the volatile nature of cryptocurrency prices.

Methodology: The researchers conducted a comparative analysis of the performance of two models for predicting Bitcoin prices: the ARIMA and LSTM models. The study utilized a dataset of Bitcoin prices from 2020 to 2024, sourced from Yahoo Finance. The data was divided into two sets: a training set and a testing set, with an 80:20 split (80% for training and 20% for testing). The prediction outcomes were evaluated using key metrics, including RMSE and MAE, to measure the accuracy and efficiency of each model. This analysis provided insights into the strengths and weaknesses of each model in handling the volatility and uncertainty of the cryptocurrency market.

Findings: The LSTM model outperforms the ARIMA model in terms of accuracy. It is better equipped to capture the price volatility and adapt to market changes, as reflected by lower RMSE and MAE, compared to the ARIMA model. Furthermore, the study indicates that the LSTM model is more suitable for handling complex and volatile datasets, such as cryptocurrency price data, demonstrating superior predictive reliability and performance.

Applications of this study: This comparative study of ARIMA and LSTM models has significant practical applications in the finance and technology sectors. The findings that the LSTM model is superior in capturing market volatility can be leveraged to develop tools that aid investment decisions in the cryptocurrency market. Additionally, this research is beneficial for developers working on automated trading platforms, as improved predictions can lead to better risk management and increased profit opportunities for investors in the cryptocurrency market.
Keywords: Bitcoin Price Forecasting, Time Series Analysis, LSTM Model, ARIMA Model, Forecasting Accuracy

1. บทนำ

การพยากรณ์ราคาบิตคอยน์ ได้กลายเป็นประเด็นสำคัญในด้านการเงินและเทคโนโลยีสารสนเทศ เนื่องจาก บิตคอยน์เป็นสกุลเงินดิจิทัลชั้นนำที่มีมูลค่าตลาดสูงและมีอิทธิพลต่อตลาดการเงินโลก การทำนายราคาบิตคอยน์ อย่างแม่นยำมีความสำคัญต่อนักลงทุน ผู้ค้า และผู้กำหนดนโยบาย เนื่องจากช่วยให้พวกเขาตัดสินใจได้อย่างมีข้อมูลและจัดการความเสี่ยงได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม ความผันผวนสูงและลักษณะที่ไม่แน่นอนของตลาดสกุลเงินดิจิทัล ทำให้การพยากรณ์ราคาบิตคอยน์ เป็นงานที่ท้าทายอย่างมาก (Prabpala, Tuamsuk & Chansanam, 2023)

บิตคอยน์ ซึ่งเปิดตัวในปี ค.ศ. 2009 ได้กลายเป็นสกุลเงินดิจิทัลชั้นนำที่ไม่เพียงแต่มีอิทธิพลต่อตลาดการเงินเท่านั้น แต่ยังรวมถึงกรอบการกำกับดูแลและภูมิทัศน์ของเศรษฐกิจโลก ลักษณะการกระจายอำนาจของบิตคอยน์ และเทคโนโลยีบล็อคเชนที่อยู่เบื้องหลัง ได้มีการนำเสนอความท้าทายและโอกาสที่ไม่เหมือนใครในการพยากรณ์ราคาไม่เหมือนกับตลาดการเงินแบบดั้งเดิม ตลาดสกุลเงินดิจิทัลถูกขับเคลื่อนโดยปัจจัยหลากหลายรวมถึงการเปลี่ยนแปลงทางเทคโนโลยี เหตุการณ์ทางภูมิรัฐศาสตร์ และการเปลี่ยนแปลงของอารมณ์นักลงทุน ซึ่งทั้งหมดนี้มีส่วนทำให้ราคามีความผันผวน (Latif, Selvam, Kapse, Sharma & Mahajan, 2023)

งานวิจัยก่อนหน้านี้ได้สำรวจวิธีการต่าง ๆ ในการพยากรณ์ราคาบิตคอยน์ รวมถึงโมเดลทางสถิติ เช่น ARIMA และ GARCH (Bhongchirawattana, Sukprasert, Paengsab & Saraphap, 2023; Boonmana & Kulvanich, 2017) และโมเดลการเรียนรู้ของเครื่อง เช่น Support Vector Machines (SVM), Random Forests และ Long Short-Term Memory (LSTM) networks (Rizwan, Narejo & Javed, 2019) โมเดล ARIMA ได้รับความนิยมเนื่องจากเรียบง่ายและสามารถตีความได้ แต่ประสิทธิภาพอาจลดลงเมื่อต้องจัดการกับข้อมูลไม่เชิงเส้นและมีความผันผวนสูง (Junhan & Thongkam, 2022) ในทางกลับกัน โมเดล LSTM ซึ่งเป็นสถาปัตยกรรมเครือข่ายประสาทเทียมแบบ Recurrent Neural Network (RNN) ได้แสดงให้เห็นถึงศักยภาพในการดักจับความสัมพันธ์ระยะยาวและจัดการกับข้อมูลที่ไม่เป็นเชิงเส้น ทำให้เป็นตัวเลือกที่น่าสนใจสำหรับการพยากรณ์ราคาบิตคอยน์ (Bukhari et al., 2020)

อย่างไรก็ตาม ยังคงมีความต้องการการวิจัยเพิ่มเติมเพื่อเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM โดยตรงในการพยากรณ์ราคาบิตคอยน์ การศึกษาดังกล่าวจะช่วยให้เข้าใจถึงจุดแข็งและจุดอ่อนของแต่ละโมเดล และระบุว่าโมเดลใดเหมาะสมที่สุดสำหรับการใช้งานในสถานการณ์ต่าง ๆ นอกจากนี้ การวิจัยก่อนหน้านี้ยังไม่ได้ตรวจสอบอย่างละเอียดถึงผลกระทบของช่วงเวลาและความถี่ของข้อมูลที่แตกต่างกันต่อประสิทธิภาพการทำนายของโมเดล ARIMA และ LSTM (Yamak, Yujian & Gadosey, 2019)

เพื่อจัดการกับช่องว่างเหล่านี้ การศึกษานี้จึงมุ่งเน้นไปที่การเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ โดยใช้ข้อมูลราคาย้อนหลังเป็นเวลา 4 ปี โดยจะทำการทดสอบโมเดลด้วยชุดข้อมูลย้อนหลังและวิเคราะห์ผลลัพธ์เพื่อระบุว่าโมเดลใดมีประสิทธิภาพสูงสุดในการทำนายราคาบิตคอยน์ ในสภาวะตลาดที่แตกต่างกัน ผลการวิจัยนี้จะมีประโยชน์อย่างมากในการช่วยให้ผู้ลงทุนและผู้ใช้งานทั่วไปเข้าใจถึงความแม่นยำและความเชื่อถือได้ของโมเดลที่ใช้ในการพยากรณ์ราคาบิตคอยน์ และอาจช่วยในการวิเคราะห์แนวโน้มของตลาดในอนาคตได้

2. วัตถุประสงค์

1) วิเคราะห์และเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ เพื่อระบุโมเดลที่มีความแม่นยำสูงสุดสำหรับการใช้งานในสภาวะตลาดที่แตกต่างกัน

2) ประเมินความสามารถของโมเดล ARIMA และ LSTM ในการตอบสนองต่อการเปลี่ยนแปลงของตลาดและความผันผวนของราคาสกุลเงินดิจิทัลเพื่อพัฒนากลยุทธ์การพยากรณ์ราคาที่เหมาะสม

3. วิธีการศึกษา

เพื่อทำการวิเคราะห์และเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ การศึกษานี้จะใช้วิธีการดังต่อไปนี้

3.1 การเตรียมข้อมูล

ข้อมูลราคาปิดรายวันของบิตคอยน์ ที่ใช้ในการวิจัยนี้ได้รับการรวบรวมจากแหล่งข้อมูล Yahoo Finance (https://finance.yahoo.com/quote/BTC-USD/history) ในรูปแบบ CSV ซึ่งประกอบด้วยราคาปิดของบิตคอยน์ ตั้งแต่วันที่ 25 เมษายน ค.ศ. 2020 ถึงวันที่ 23 เมษายน ค.ศ. 2024 เป็นระยะเวลาทั้งสิ้น 4 ปี โดยมีจำนวนชุดข้อมูลทั้งสิ้น 1,461 ชุดข้อมูล การรวบรวมข้อมูลในช่วงเวลานี้ ช่วยให้สามารถวิเคราะห์และทดสอบประสิทธิภาพของโมเดลการพยากรณ์ภายใต้เงื่อนไขตลาดที่แตกต่างกันในระยะเวลาที่ยาวนาน รวมถึงการตอบสนองต่อเหตุการณ์ทางเศรษฐกิจและตลาดที่สำคัญต่าง ๆ นอกจากนี้ ข้อมูลที่ได้จาก Yahoo Finance ยังได้รับการตรวจสอบอย่างละเอียดเพื่อความเสถียรและความสมบูรณ์ และไม่พบ Missing value ก่อนนำมาใช้ในการวิจัย

ตารางที่ 1 ตารางแสดงข้อมูล วันที่ (Date)/ราคาเปิด (Open)/ราคาสูงสุด (High)/ราคาต่ำสุด (Low)/ราคาปิด (Close)/ราคาปิดที่ปรับแล้ว (Adj close)/ปริมาณการซื้อขาย (Volume) (USD) ของบิตคอยน์ ตั้งแต่ 25 เมษายน ค.ศ. 2020 ถึงวันที่ 23 เมษายน ค.ศ. 2024

Date	Open	High	Low	Close	Adj Close	Volume (USD)
2020-04-25	7,550.48	7,641.36	7,521.67	7,569.94	7,569.94	32,941,541,447.00
2020-04-26	7,570.14	7,700.59	7,561.41	7,679.87	7,679.87	33,070,154,491.00
2020-04-27	7,679.42	7,795.60	7,679.42	7,795.60	7,795.60	36,162,144,725.00
2020-04-28	7,796.97	7,814.53	7,730.81	7,807.06	7,807.06	33,187,959,921.00
2020-04-29	7,806.71	8,871.75	7,786.05	8,801.04	8,801.04	60,201,052,203.00
...						
2024-04-19	63,510.75	65,481.60	59,651.39	63,843.57	63,843.57	49,920,425,401.00
2024-04-20	63,851.10	65,442.46	63,172.40	64,994.44	64,994.44	23,097,485,495.00
2024-04-21	64,992.82	65,723.24	64,277.72	64,926.64	64,926.64	20,506,644,853.00
2024-04-22	64,935.63	67,233.96	64,548.18	66,837.68	66,837.68	28,282,686,673.00
2024-04-23	66,839.89	67,199.24	65,864.87	66,407.27	66,407.27	24,310,975,583.00

ในการวิจัยนี้แบ่งข้อมูลออกเป็นสองส่วน คือ ชุดข้อมูลฝึกและชุดข้อมูลทดสอบ เป็นขั้นตอนที่ช่วยให้สามารถประเมินประสิทธิภาพของแต่ละโมเดลได้อย่างเป็นกลาง โดยมีวิธีการแบ่งข้อมูลดังนี้

1) ชุดข้อมูลฝึก (Training set) ชุดข้อมูลนี้คิดเป็น 80% ของข้อมูลทั้งหมดที่มีอยู่ และจะถูกใช้ในการสร้างและฝึกโมเดล ARIMA และ LSTM ชุดข้อมูลฝึกมีหน้าที่ในการปรับโมเดลและปรับพารามิเตอร์ต่าง ๆ เพื่อให้โมเดลสามารถเรียนรู้และปรับตัวเข้ากับลักษณะเฉพาะของข้อมูลราคาบิตคอยน์

2) ชุดข้อมูลทดสอบ (Testing set) ชุดข้อมูลนี้คิดเป็น 20% ของข้อมูลทั้งหมด มีหน้าที่เพื่อใช้ในการทดสอบประสิทธิภาพของโมเดลที่ได้รับการฝึกแล้ว การทดสอบโมเดลด้วยข้อมูลชุดนี้ทำให้สามารถประเมินความสามารถของโมเดลในการทำนายและประยุกต์ใช้กับข้อมูลใหม่ที่อาจเกิดขึ้นจริงได้

3.2 การพัฒนาโมเดล

การพัฒนาโมเดลสำหรับการวิจัยนี้รวมถึงการสร้างและฝึกโมเดล ARIMA และ LSTM สำหรับการพยากรณ์ราคาบิตคอยน์ โดยใช้ข้อมูลที่ได้จากชุดข้อมูลฝึก ขั้นตอนแรกในการวิจัยนี้คือการโหลดข้อมูลราคาบิตคอยน์ จากไฟล์ CSV ที่ได้เตรียมไว้ โดยใช้ไลบรารี Pandas ใน Python เพื่อโหลดข้อมูลและตั้งค่าวันที่เป็นดัชนีของ DataFrame ขั้นตอนนี้เป็นส่วนหนึ่งของกระบวนการเตรียมข้อมูลเพื่อให้พร้อมสำหรับการวิเคราะห์และการทำนายต่อไป จากการวิเคราะห์ข้อมูลเบื้องต้น พบว่าข้อมูลราคาปิด (Close) ของบิตคอยน์ มีความผันผวนสูง และมีแนวโน้มที่ไม่เป็น Stationary ดังนั้น จึงเลือกใช้ราคาปิด (Close) เป็นตัวแปรตาม (y) ในการสร้างโมเดล ARIMA และ LSTM โดยใช้ดัชนีเวลาเป็นตัวแปรต้น (x) เพื่อให้โมเดลสามารถเรียนรู้รูปแบบและความสัมพันธ์ของราคาบิตคอยน์ ในช่วงเวลาที่ต่างกันได้

3.2.1 โมเดล ARIMA

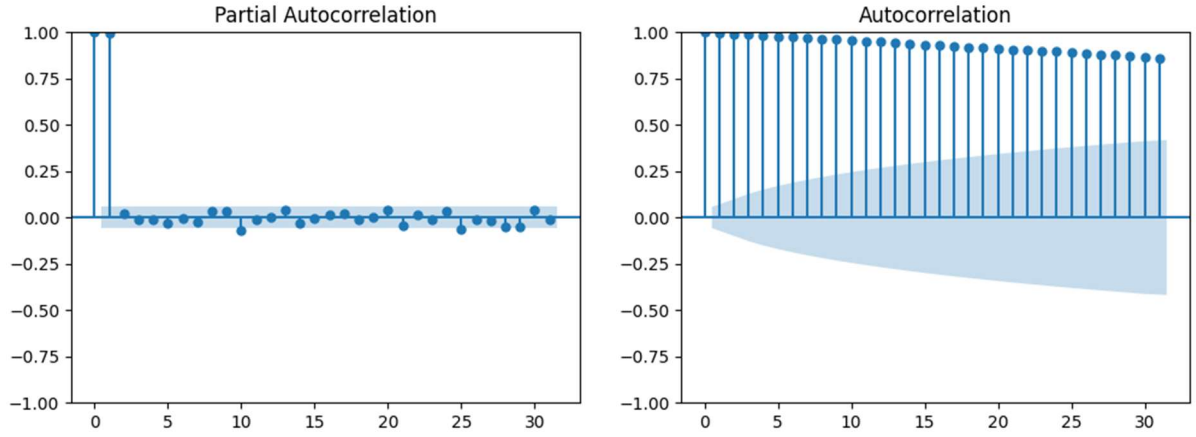
การเลือกพารามิเตอร์ (p, d, q)

การเลือกพารามิเตอร์ d (Differencing)

ทำการการทดสอบ Augmented Dickey-Fuller (ADF) เพื่อทำให้ข้อมูลที่ โดยการทำ differencing หนึ่งครั้ง (d = 1) โดยใช้ฟังก์ชัน .diff() โดยค่า p-value ของการทดสอบเหลือน้อยกว่า 0.05 หลังจากการ differencing ครั้งแรก (p-value = 3.002090194876249e-19)

การเลือกพารามิเตอร์ p (Autoregressive) และ q (Moving Average)

ผู้วิจัยได้ใช้ฟังก์ชัน plot_acf และ plot_pacf จากไลบรารี statsmodels เพื่อวิเคราะห์และตัดสินใจเลือกพารามิเตอร์ p และ q ของโมเดล ได้ผลดังนี้



ภาพที่ 1 กราฟ PACF และ ACF

Partial Autocorrelation Function (PACF) จากกราฟ PACF จะเห็นได้ว่าค่า Autocorrelations มีการลดลงอย่างรวดเร็วหลังจาก lag 1 ซึ่งชี้ให้เห็นว่า p = 1 เหมาะสมสำหรับโมเดล ARIMA

Autocorrelation Function (ACF) จากกราฟ ACF แสดงให้เห็นว่าค่า Autocorrelations นั้นค่อย ๆ ลดลงและเข้าใกล้ 0 หลังจาก lag 1 ซึ่งบ่งชี้ว่า q = 1 เหมาะสมสำหรับส่วน MA (Moving Average) ของโมเดล

โดยสรุป โมเดล ARIMA (1,1,1) ถูกเลือกสำหรับการวิเคราะห์นี้ เนื่องจากการปรับให้ข้อมูลให้คงที่ด้วยการ Differencing หนึ่งครั้ง และลักษณะของกราฟ PACF และ ACF ที่ชี้ให้เห็นความเหมาะสมของการใช้ลักษณะ Auto Regressive และ Moving Average ที่เหมาะสม เนื่องจากไม่ได้มีการแบ่ง Validation set อย่างชัดเจน จึงมีการใช้ข้อมูลชุด Test ทั้งหมดในการประเมินประสิทธิภาพของโมเดลหลังการฝึกเสร็จสิ้

3.2.2 โมเดล LSTM

ในการเตรียมข้อมูลสำหรับการวิจัย ผู้วิจัยได้ปรับขนาดข้อมูลโดยใช้ MinMaxScaler จากไลบรารี Scikit-learn เพื่อนำค่าข้อมูลมาอยู่ในช่วง 0 ถึง 1 ขั้นตอนนี้มีจุดมุ่งหมายเพื่อเสริมประสิทธิภาพใน

การฝึกฝนโมเดลให้ได้ผลลัพธ์ที่ดีขึ้น เนื่องจากไม่ได้มีการแบ่ง Validation set อย่างชัดเจน จึงมีการใช้ validation_split = 0.2 ในขั้นตอนการฝึก ซึ่งหมายความว่า 20% ของข้อมูลในชุด Train จะถูกแบ่งออกมาเป็น Validation set โดยอัตโนมัติ เพื่อใช้ในการประเมินประสิทธิภาพของโมเดลระหว่างการฝึก

```
# สร้างโมเดล LSTM
model = Sequential([
    LSTM(50, return_sequences=True, input_shape=(60, 1)),
    LSTM(50, return_sequences=False),
    Dropout(0.2),
    Dense(1)
])
```

ภาพที่ 2 แสดงการออกแบบโมเดล LSTM

โมเดล LSTM ที่ออกแบบนี้มีโครงสร้างที่ประกอบด้วยชั้น LSTM หลายชั้นเพื่อเรียนรู้จากข้อมูลเวลาซีรีส์ที่มีความซับซ้อนและผันผวนสูงของราคาสกุลเงินดิจิทัล ชั้น LSTM แรกถูกตั้งค่าให้มี Return Sequences เป็น True เพื่อส่งผลลัพธ์ที่เป็นลำดับความยาวเต็มของข้อมูลไปยังชั้นถัดไป ช่วยให้โมเดลสามารถรักษารายละเอียดของลำดับข้อมูลได้ ชั้น LSTM ที่สอง ไม่ได้ส่งต่อลำดับข้อมูลแต่สรุปเป็น Vector ของข้อมูลที่จะใช้ในการทำนาย นอกจากนี้ยังมีชั้น Dropout ที่อัตรา 20% เพื่อลดปัญหาการเกินตัวและช่วยให้โมเดลไม่จำรายละเอียดของชุดข้อมูลฝึกเกินไป ทำให้สามารถทำนายข้อมูลใหม่ได้แม่นยำยิ่งขึ้น ชั้นสุดท้ายคือชั้น Dense ที่ใช้สำหรับทำนายค่าเอาต์พุตสุดท้าย เป็นค่าราคาปิดของบิตคอยน์

```
# ฝึกอบรมโมเดลจากข้อมูลฝึก 80% และบันทึกประวัติการฝึกอบรมลงในตัวแปร history
history = model.fit(trainX, trainY, epochs=20, batch_size=64, validation_split=0.2, verbose=1)
```

ภาพที่ 2 แสดงการฝึกโมเดล LSTM

epochs = 20 คือ การกำหนดให้มีการวนซ้ำทั้งหมด 20 รอบในชุดข้อมูลฝึก เพื่อให้โมเดลมีโอกาสเรียนรู้และปรับเทียบน้ำหนักจากข้อมูลหลาย ๆ รอบ การเลือกจำนวน epochs นั้นสำคัญต่อการทำให้สมดุลระหว่างการเรียนรู้ไม่เพียงพอ (Underfitting) และการเรียนรู้มากเกินไป (Overfitting)

batch_size = 64 คือ กำหนดขนาดของแบทช์ที่ 64 หมายถึง โมเดลจะปรับน้ำหนักหลังจากได้เรียนรู้จาก 64 ตัวอย่างในชุดข้อมูล ขนาดแบทช์ส่งผลต่อความเร็วในการฝึกและการประมวลผลของแต่ละ epoch

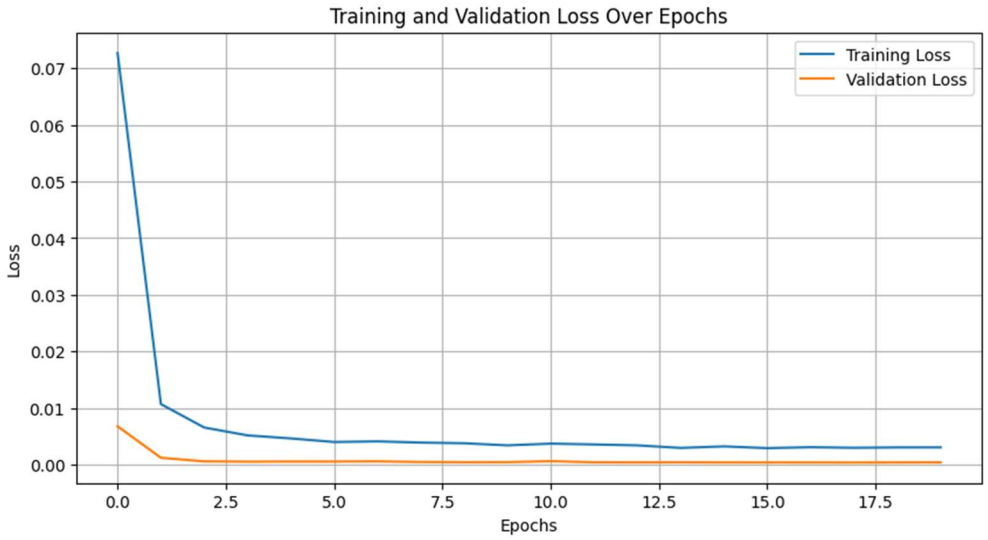
validation_split = 0.2 คือ ชุดข้อมูลทดสอบเพื่อประเมินประสิทธิภาพของโมเดล

verbose = 1 คือ การแสดงผลขั้นตอนการฝึกอบรมบนหน้าจอ ช่วยให้ติดตามความคืบหน้าและดูการเปลี่ยนแปลงของฟังก์ชันสูญเสียในระหว่างการฝึก

จากนั้นได้ทำการบันทึกโมเดล และสเกลเลอร์ที่ใช้ปรับขนาดข้อมูลเพื่อนำไปใช้ในอนาคตได้ โมเดล LSTM ที่ฝึกเสร็จถูกบันทึกเป็นไฟล์โดยใช้รูปแบบไฟล์ .h5 ซึ่งเป็นรูปแบบมาตรฐานสำหรับเก็บโมเดล Keras ที่ไม่เพียงแต่บันทึกโครงสร้างของโมเดลเท่านั้น แต่ยังรวมถึงน้ำหนัก การตั้งค่าการฝึก และ

สถานะของการ optimizer ด้วย การบันทึกโมเดลในรูปแบบนี้ช่วยให้สามารถโหลดโมเดลกลับมาใช้งานหรือทำการทดสอบต่อได้โดยที่ไม่ต้องฝึกใหม่

สเกลเลอร์ ซึ่งใช้ในการปรับขนาดข้อมูลราคาบิตคอยน์ ให้อยู่ในช่วงที่เหมาะสมก่อนการฝึก ถูกบันทึกโดยใช้ไลบรารี joblib ในรูปแบบไฟล์ .pkl การบันทึกสเกลเลอร์นี้เป็นสิ่งจำเป็นเมื่อต้องการใช้โมเดลในการทำนายข้อมูลใหม่ที่ต้องปรับขนาดข้อมูลให้เป็นไปในลักษณะเดียวกับชุดข้อมูลที่ใช้ฝึกโมเดล โดยสามารถรักษาความสอดคล้องและความแม่นยำของโมเดลในการทำนาย



ภาพที่ 3 แสดงประสิทธิภาพของชุดข้อมูลฝึกและค่า loss

3.3 การวัดประสิทธิภาพของโมเดลทั้ง 2

ในการประเมินประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ ผู้วิจัยใช้ค่า RMSE และ MAE เป็นเกณฑ์วัดความแม่นยำ โดยอ้างอิงจากงานวิจัยที่เกี่ยวข้อง (Latif et al., 2023; Hua, 2020) และพิจารณาความผันผวนของราคาบิตคอยน์ ในช่วงเวลาที่ศึกษา โดยพิจารณาว่าค่า RMSE ที่ต่ำกว่า 5,000 และ MAE ที่ต่ำกว่า 4,000 เป็นเกณฑ์ที่ยอมรับได้สำหรับโมเดลที่มีประสิทธิภาพในการพยากรณ์ที่ดี

ในการวัดประสิทธิภาพของโมเดล ARIMA(1,1,1) และ LSTM ที่พัฒนาจากชุดข้อมูลฝึก 80% สำหรับการทำนายราคาบิตคอยน์ จากข้อมูลประวัติราคาปิด ค่า RMSE (Root mean squared error) และ MAE (Mean absolute error) ได้ถูกคำนวณโดยใช้ชุดข้อมูลการทดสอบ 20% เพื่อประเมินความแม่นยำและประสิทธิภาพของโมเดลที่ทำนาย ผู้วิจัยใช้ไลบรารี scikit-learn ฟังก์ชัน mean_squared_error และ mean_absolute_error ใช้ในการคำนวณค่า RMSE และ MAE

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM

Metric	ARIMA	LSTM
RMSE	18,520.98	2,442.73
MAE	13,266.63	1,693.21

จากตารางข้างต้น พบว่าโมเดล LSTM มีประสิทธิภาพที่ดีกว่าโมเดล ARIMA อย่างมีนัยสำคัญในการทำนายราคาบิตคอยน์ โดย LSTM มีค่า RMSE อยู่ที่ 2,442.73 ซึ่งต่ำกว่าเกณฑ์ที่ยอมรับได้ที่ 5,000 และมีค่า MAE ที่ 1,693.21 ซึ่งต่ำกว่าเกณฑ์ที่ 4,000 ในขณะที่ ARIMA มีค่า RMSE ที่ 18,520.98 และ MAE ที่ 13,266.63 ซึ่งสูงกว่าเกณฑ์ที่ยอมรับได้ค่อนข้างมาก ซึ่งสะท้อนให้เห็นว่าโมเดล LSTM มีความแม่นยำและความผิดพลาดในการทำนายที่น้อยกว่า ARIMA ในบริบทของการลงทุน ผลลัพธ์นี้บ่งชี้ว่าโมเดล LSTM มีศักยภาพในการนำไปใช้เป็นเครื่องมือช่วยตัดสินใจในการลงทุนบิตคอยน์ ได้ดีกว่าโมเดล ARIMA เนื่องจากสามารถจับความผันผวนของราคาและปรับตัวเข้ากับการเปลี่ยนแปลงของตลาดได้แม่นยำกว่า อย่างไรก็ตาม นักลงทุนควรใช้โมเดล LSTM ร่วมกับเครื่องมือและข้อมูลอื่น ๆ ในการตัดสินใจลงทุน เนื่องจากโมเดล LSTM เป็นเพียงเครื่องมือหนึ่งในการวิเคราะห์และไม่สามารถรับประกันความแม่นยำได้ 100%

4. ผลการศึกษา

4.1 ประสิทธิภาพของโมเดล

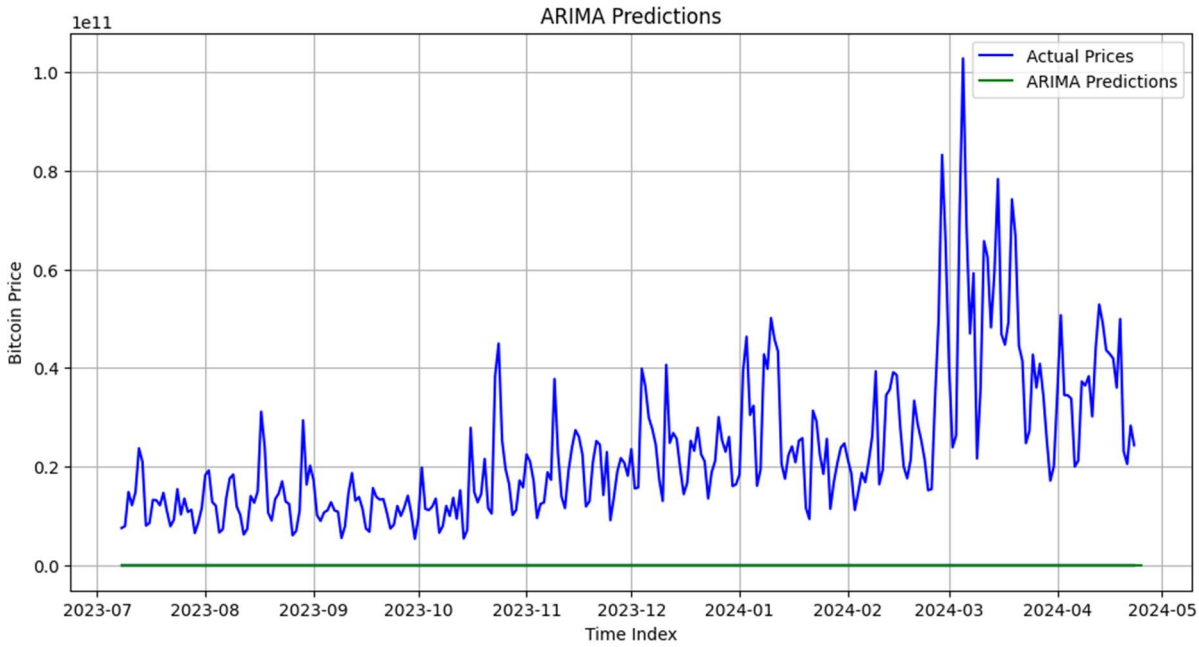
โมเดล LSTM สามารถทำนายราคาบิตคอยน์ ได้ดีกว่าโมเดล ARIMA อย่างชัดเจน ซึ่งสะท้อนผ่านค่า Root Mean Squared Error (RMSE) และ Mean Absolute Error (MAE) ที่ต่ำกว่าค่อนข้างมากเมื่อเทียบกับโมเดล ARIMA โดย LSTM มี RMSE อยู่ที่ 2,442.73 และ MAE ที่ 1,693.21 ในขณะที่ ARIMA มี RMSE ที่ 18,520.98 และ MAE ที่ 13,266.63

4.2 การวิเคราะห์ข้อมูล

โมเดล LSTM มีความสามารถในการจับความสัมพันธ์ระยะยาวและซับซ้อนในข้อมูลราคาบิตคอยน์ ได้ดีกว่าโมเดล ARIMA ด้วยโครงสร้างของเครือข่ายประสาทเทียมที่เหมาะสมกับข้อมูลที่มีความผันผวนและไม่แน่นอนสูง โมเดล ARIMA อาจเหมาะสมกับข้อมูลที่มีความสัมพันธ์เชิงเส้นและรูปแบบฤดูกาล แต่ไม่สามารถจับความสัมพันธ์ซับซ้อนในระดับที่ LSTM ทำได้

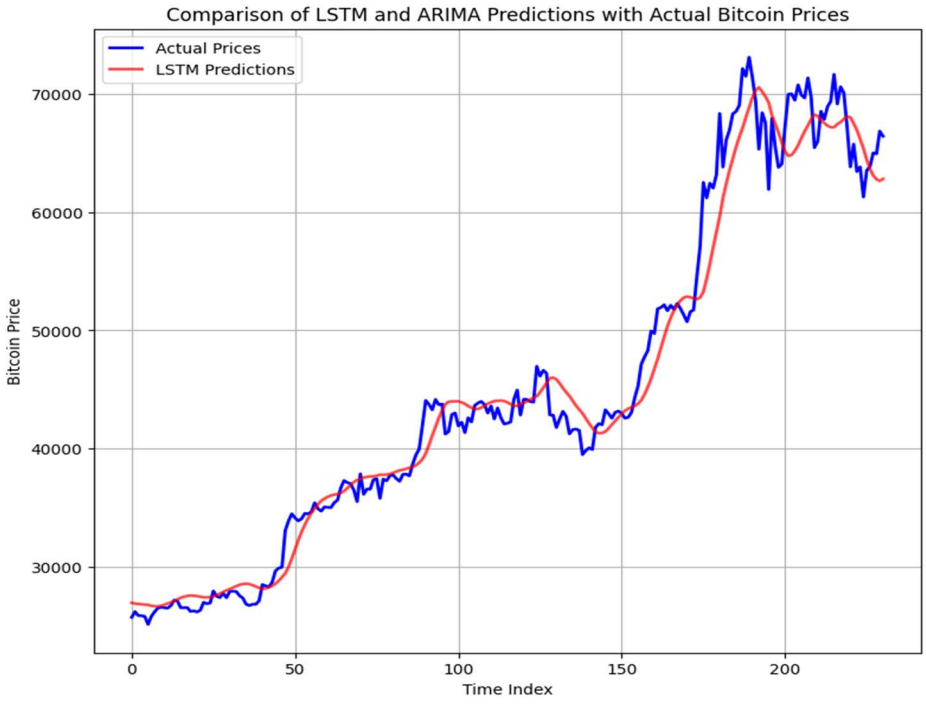
4.3 การใช้งานและความเหมาะสม

การศึกษานี้แสดงให้เห็นว่าโมเดล LSTM เหมาะสำหรับการทำนายราคาในตลาดที่มีความผันผวนสูง เช่นตลาดสกุลเงินดิจิทัล และสามารถให้ข้อมูลที่มีความแม่นยำสูงสำหรับการตัดสินใจของนักลงทุน โมเดล ARIMA ยังคงเป็นเครื่องมือที่มีค่าสำหรับการวิเคราะห์ข้อมูลทางการเงินที่ต้องการรูปแบบที่เรียบง่ายและการตีความที่ตรงไปตรงมา



ภาพที่ 4 กราฟการทำนายของโมเดล ARIMA เทียบกับราคาจริง (Actual Price)

แสดงผลการเปรียบเทียบราคาจริงของบิตคอยน์ (เส้นสีน้ำเงิน) กับราคาที่ทำนายโดยโมเดล ARIMA (เส้นสีเขียว) ในชุดข้อมูลทดสอบ ตั้งแต่เดือนกรกฎาคม ค.ศ. 2020 ถึง 2024 จะเห็นได้ว่าโมเดล ARIMA ไม่สามารถจับความผันผวนของราคาบิตคอยน์ ได้ เนื่องจากเส้นการทำนายมีลักษณะคงที่เกือบตลอดช่วงเวลาที่ทำการทดสอบ ในขณะที่ราคาบิตคอยน์ จริงมีความผันผวนสูง ซึ่งแสดงให้เห็นข้อจำกัดของโมเดล ARIMA ในการทำนายข้อมูลอนุกรมเวลาที่มีความไม่เป็นเชิงเส้นสูง เช่น ราคาบิตคอยน์



ภาพที่ 5 กราฟการทำนายของโมเดล LSTM เทียบกับราคาจริง (Actual Price)

แสดงการทำนายของโมเดล LSTM ที่ติดตามราคาจริง (Actual prices) ของบิตคอยน์ ได้อย่างใกล้ชิดมากกว่าโมเดล ARIMA ซึ่งชี้ให้เห็นถึงความสามารถของโมเดลในการจับความผันผวนและตอบสนองต่อการเปลี่ยนแปลงของตลาดได้เร็วกว่า โดยให้การทำนายที่แม่นยำและสามารถจับแนวโน้มระยะยาวของราคาได้ดีกว่าโมเดล ARIMA สำหรับชุดข้อมูลนี้ แกน X ในกราฟแสดงดัชนีเวลา (Time index) ไม่ใช่วันที่จริง แต่เป็นลำดับของข้อมูลตามเวลาที่เกิดขึ้น ตัวเลขเริ่มต้นที่ 0 และเพิ่มขึ้นไปจนถึงประมาณ 200 ซึ่งหมายถึงจำนวนวันหรือช่วงเวลาข้อมูลตามที่ตั้งในการวิเคราะห์

5. อภิปรายผล

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของโมเดล ARIMA และ LSTM ในการพยากรณ์ราคาบิตคอยน์ โดยใช้ข้อมูลราคาย้อนหลังเป็นเวลา 4 ปี (ค.ศ. 2020 - 2024) ซึ่งประกอบด้วย 1,461 ชุดข้อมูล ผลการศึกษาพบว่าโมเดล LSTM มีประสิทธิภาพที่เหนือกว่าโมเดล ARIMA อย่างมีนัยสำคัญในการพยากรณ์ราคาบิตคอยน์ ในทุกสภาวะตลาด โดยสามารถจับความผันผวนและความไม่แน่นอนของราคาบิตคอยน์ได้ดีกว่า ซึ่งสอดคล้องกับสมมติฐานที่ว่าโมเดล LSTM ที่มีความสามารถในการเรียนรู้รูปแบบที่ซับซ้อนและไม่เป็นเชิงเส้น จะมีประสิทธิภาพดีกว่าโมเดล ARIMA ที่เหมาะกับข้อมูลที่มีความสัมพันธ์เชิงเส้นมากกว่า

ข้อค้นพบนี้สอดคล้องกับงานวิจัยที่ผ่านมาหลายชิ้น เช่น Latif et al. (2023) และ Hua (2020) ที่พบว่าโมเดล LSTM มีประสิทธิภาพที่ดีกว่าในการพยากรณ์ราคาบิตคอยน์ เมื่อเทียบกับโมเดล ARIMA อย่างไรก็ตาม งานวิจัยนี้ได้เพิ่มความน่าเชื่อถือให้กับข้อค้นพบนี้ โดยการใช้ข้อมูลในระยะเวลาที่ยาวนานขึ้น (4 ปี) และจำนวนชุดข้อมูลที่มากขึ้น (1,461 ชุดข้อมูล) ซึ่งช่วยให้มั่นใจได้ว่าผลการวิจัยมีความน่าเชื่อถือและสามารถนำไปประยุกต์ใช้ได้จริง

งานวิจัยนี้และงานวิจัยที่อ้างอิง (Hua, 2020) พบว่ามีความคล้ายคลึงกันในเรื่องของหัวข้อการวิจัย โมเดลที่ใช้ และเมตริกที่ใช้ในการประเมินประสิทธิภาพ คือ RMSE และ MAE แต่มีความแตกต่างกันในเรื่องของระยะเวลาของข้อมูลที่ใช้, สถาปัตยกรรมของโมเดล LSTM ที่ไม่ได้ระบุรายละเอียดไว้ ทำให้ไม่สามารถเปรียบเทียบได้ และงานวิจัยนี้ได้เพิ่มการกำหนดเกณฑ์ในการประเมินประสิทธิภาพของโมเดล โดยพิจารณาค่า RMSE ที่ต่ำกว่า 5,000 และ MAE ที่ต่ำกว่า 4,000 เป็นเกณฑ์ที่ยอมรับได้ ซึ่งเป็นส่วนที่แตกต่างจากงานวิจัยของ Hua (2020) อย่างไรก็ตาม งานวิจัยทั้งสองมีข้อค้นพบที่สอดคล้องกัน คือ โมเดล LSTM มีประสิทธิภาพที่ดีกว่าโมเดล ARIMA ในการพยากรณ์ราคาบิตคอยน์ แต่ในงานวิจัยนี้ได้ให้รายละเอียดเกี่ยวกับประสิทธิภาพของโมเดล LSTM ที่มากกว่า เช่น การแสดงให้เห็นว่าโมเดล LSTM สามารถจับความผันผวนและความไม่แน่นอนของราคาบิตคอยน์ ได้ดีกว่าโมเดล ARIMA

นอกจากนี้ ผลการศึกษาที่แสดงให้เห็นว่าโมเดล LSTM สามารถปรับตัวเข้ากับการเปลี่ยนแปลงของตลาดและความผันผวนของราคาบิตคอยน์ ได้ดีกว่าโมเดล ARIMA ยังสอดคล้องกับแนวคิดของเครือข่ายประสาทเทียม LSTM ที่มีหน่วยความจำระยะยาว (Long short-term memory) ทำให้สามารถจดจำข้อมูลในอดีตและนำมาใช้ในการพยากรณ์ได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม ผลการศึกษาครั้งนี้ไม่ได้ขัดแย้งกับงานวิจัยที่ผ่านมาที่พบว่าโมเดล ARIMA ยังคงเป็นเครื่องมือที่มีประโยชน์ในการวิเคราะห์ข้อมูลทางการเงินที่มีรูปแบบที่เรียบง่ายและการตีความที่ตรงไปตรงมา แต่สำหรับการพยากรณ์ราคาบิตคอยน์ ที่มีความผันผวนสูง โมเดล LSTM อาจเป็นทางเลือกที่ดีกว่า

6. ข้อเสนอแนะ

6.1 ข้อเสนอแนะสำหรับการนำไปใช้ประโยชน์เชิงวิชาการ

1) การพัฒนาโมเดลการพยากรณ์ ผลการวิจัยนี้สามารถเป็นพื้นฐานสำหรับการพัฒนาโมเดลการพยากรณ์ราคาบิตคอยน์ ที่ซับซ้อนและแม่นยำยิ่งขึ้น โดยอาจนำโมเดล LSTM มาปรับปรุงหรือรวมกับเทคนิคอื่น ๆ เช่น การวิเคราะห์ปัจจัยพื้นฐาน หรือการวิเคราะห์ความเชื่อมั่นของตลาด เพื่อเพิ่มความแม่นยำในการพยากรณ์

2) การศึกษาเปรียบเทียบโมเดล ควรมีการศึกษาเปรียบเทียบโมเดล LSTM กับโมเดลการเรียนรู้ของเครื่องอื่น ๆ เช่น Support Vector Machines (SVM) หรือ Random Forests เพื่อให้เห็นภาพรวมของประสิทธิภาพของโมเดลต่าง ๆ ในการพยากรณ์ราคาบิตคอยน์

3) การศึกษาปัจจัยที่มีผลต่อราคาบิตคอยน์ ควรมีการศึกษาเพิ่มเติมเกี่ยวกับปัจจัยต่าง ๆ ที่มีผลต่อราคาบิตคอยน์ เช่น ข่าวสาร กฎระเบียบของรัฐบาล และความเชื่อมั่นของตลาด เพื่อนำมาปรับปรุงโมเดลการพยากรณ์ให้มีความแม่นยำมากยิ่งขึ้น

6.2 ข้อเสนอแนะสำหรับการนำไปใช้ประโยชน์เชิงนโยบาย

1) การกำกับดูแลตลาด ผลการวิจัยนี้สามารถช่วยผู้กำหนดนโยบายในการเข้าใจถึงความผันผวนและความเสี่ยงของตลาด Bitcoin ซึ่งจะเป็นประโยชน์ในการกำหนดนโยบายและกฎระเบียบที่เหมาะสมเพื่อควบคุมและส่งเสริมการเติบโตของตลาดสกุลเงินดิจิทัล

2) การจัดการความเสี่ยง โมเดลการพยากรณ์ราคาบิตคอยน์ ที่มีความแม่นยำสูงสามารถช่วยนักลงทุนและผู้ค้าในการจัดการความเสี่ยงได้อย่างมีประสิทธิภาพมากขึ้น โดยสามารถคาดการณ์แนวโน้มของตลาดและปรับกลยุทธ์การลงทุนได้อย่างเหมาะสม

เอกสารอ้างอิง

Bhongchirawattana, U., Sukprasert, A., Paengsab, N., & Saraphap, M. (2023). Cryptocurrency price forecasting for futures using Arima Model. (In Thai). **Journal of Management Science Udon Thani Rajabhat University, 5**(4), 1-11.

Boonmana, C. & Kulvanich, N. (2017). A comparative prediction accuracy of hybrid time series models. (In Thai). **Thai Science and Technology Journal, 25**(2), 178-184.

Bukhari, A.H., Raja, M.A.Z., Sulaiman, M., Islam, S., Shoaib, M., & Kumam, P. (2020). Fractional Neuro-Sequential ARFIMA-LSTM for Financial Market Forecasting. **IEEE Access, 8**, 71326-71338. https://doi.org/10.1109/ACCESS.2020.2985763

Hua, Y. (2020). Bitcoin price prediction using ARIMA and LSTM. **E3S Web of Conferences, 218**, 01050. https://doi.org/10.1051/e3sconf/202021801050

Junhan, N., & Thongkam, J. (2022). Performance Comparison of ARIMA and Machine Learning Regression Techniques in Time Series Forecasting of Bitcoin Prices. (In Thai). **Journal of Science and Technology, Ubon Ratchathani University, 24**(1), 62-67.

Latif, N., Selvam, J. D., Kapse, M., Sharma, V., & Mahajan, V. (2023). Comparative performance of LSTM and ARIMA for the short-term prediction of Bitcoin prices. **Australasian Accounting, Business and Finance Journal, 17**(1), 256-276. https://doi.org/10.14453/aabfj.v17i1.15

Prabpala, S., Tuamsuk, K., & Chansanam, W. (2023). Digital Assets Price Prediction Using Sentiment Analysis on Crowd Trading Idea. (In Thai). **Journal of Information Science Research and Practice, 41**(3), 73-92. https://doi.org/10.14456/jiskku.2023.21

Rizwan, M., Narejo, S., & Javed, M. (2019). Bitcoin price prediction using Deep Learning Algorithm. **Conference: 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)**. (1-7). https://doi.org/10.1109/MACS48846.2019.9024772

Yamak, P.T., Yujian, L., Gadosey, P.K. A comparison between ARIMA, LSTM, and GRU for time series forecasting. **Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence**. (49–55). https://doi.org/10.1145/3377713.3377722