

การเปรียบเทียบประสิทธิภาพของเทคนิคอริมาและเทคนิคถดถอยการเรียนรู้ของเครื่อง ในการพยากรณ์ราคาบิตคอยน์

Performance Comparison of ARIMA and Machine Learning Regression Techniques in Time Series Forecasting of Bitcoin Prices

นิติกร จันทาญ* และ จารี ทองคำ

Nitikorn Junhan* and Jaree Thongkam

กลุ่มสารสนเทศเชิงประยุกต์ ภาควิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

Applied Information Group, Department of Information Technology, Faculty of Informatics,

Maharakham University

*E-mail: happygottji@gmail.com

Received: Sep 15, 2021

Revised: Nov 13, 2021

Accepted: Dec 03, 2021

บทคัดย่อ

ในปัจจุบันบิตคอยน์เป็นสกุลเงินอิเล็กทรอนิกส์ที่ได้รับความนิยมและมีการซื้อขายกันมากที่สุด ทำให้ราคาจึงมีความผันผวนไม่แน่นอนในแต่ละวัน ดังนั้นงานวิจัยเพื่อหาวิธีที่เหมาะสมในการพยากรณ์ราคาบิตคอยน์จึงได้รับความสนใจ เทคนิคอริมา และเทคนิคถดถอยการเรียนรู้ของเครื่องเป็นเทคนิคที่ใช้ในการพยากรณ์สถานการณ์ต่างๆ เช่น พยากรณ์ราคาทองคำ และราคาน้ำมัน โดยอริมาเป็นเทคนิคดั้งเดิม ส่วนเทคนิคถดถอยการเรียนรู้ของเครื่องเป็นเทคนิคใหม่ เช่น Multilayer perceptron, Radial basis function และ Support vector regression งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิคอริมา และเทคนิคถดถอยการเรียนรู้ของเครื่องในการพยากรณ์ราคาบิตคอยน์ โดยนำเทคนิคดังกล่าวทั้งสองเทคนิคมาใช้สร้างแบบจำลองอนุกรมเวลาเพื่อพยากรณ์ราคาบิตคอยน์เป็นรายวัน ข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง ได้แก่ ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณซื้อขาย และมูลค่าตลาดรวม จากเว็บไซต์ www.coinmarketcap.com ตั้งแต่วันที่ 1 มกราคม ค.ศ. 2017 ถึง 31 ธันวาคม ค.ศ. 2019 ในการวัดประสิทธิภาพของแบบจำลอง หลักการเลื่อนหน้าต่างได้ถูกนำมาใช้ในการแบ่งข้อมูลเป็นชุดข้อมูลฝึกฝน และชุดข้อมูลทดสอบ นอกจากนี้แล้วค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ และค่าความคลาดเคลื่อนเฉลี่ยรากที่สองได้ถูกนำมาเป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพของแบบจำลอง จากการทดลองพบว่า เทคนิคอริมา มีประสิทธิภาพสูงสุด โดยมีค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ และค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง สำหรับการพยากรณ์ ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณซื้อขายในแต่ละวัน และมูลค่าตลาดรวมต่ำที่สุด

คำสำคัญ: บิตคอยน์ การพยากรณ์ เทคนิคอริมา เทคนิคถดถอยการเรียนรู้ของเครื่อง

Abstract

At present, Bitcoin is the most popular and traded cryptocurrency, causing its price to fluctuate daily. Therefore, research to find a suitable method for forecasting the Bitcoin price has attracted attention. ARIMA and machine learning regression techniques are techniques used for forecasting various situations such as gold price and oil price forecasting. ARIMA is a classical technique while machine learning regression techniques are new techniques such as Multilayer perceptron, Radial basis function, and Support vector regression. This research aimed to compare the performance of ARIMA and machine learning regression techniques in forecasting Bitcoin prices. Both techniques were used to create time series forecasting models for daily Bitcoin prices. The data used for creating models were the opening prices, the highest prices, the lowest prices, the closing prices, the volumes, and market caps from www.coinmarketcap.com from 1 January 2017 to 31 December 2019. In performance evaluation of the models, sliding window technique was used to separate data into training and testing sets. Furthermore, mean absolute error (MAE) and root mean square error (RMSE) were also used as criteria to compare performances of the models. From the experiments, it was found that ARIMA showed the best performance with the lowest MAE and RMSE for forecasting the opening prices, the highest prices, the lowest prices, the closing prices, the volumes, and market caps.

Keywords: Bitcoin, Forecast, ARIMA, Machine learning regression techniques

1. บทนำ

บิตคอยน์ คือ สกุลเงินอิเล็กทรอนิกส์ ได้ถูกคิดค้นขึ้นในปี ค.ศ. 2008 การซื้อขายบิตคอยน์อาศัยเครือข่ายระหว่างเครื่องคอมพิวเตอร์โดยไม่มีตัวกลางหรือรัฐเข้ามาควบคุม [1], [2] ทำให้ได้รับความนิยมในหมู่นักลงทุนและผู้สนใจ จากการที่ราคาบิตคอยน์มีความไม่คงที่สูง จึงทำให้นักลงทุนต้องการเครื่องมือที่ช่วยในการพยากรณ์ มีนักวิจัยได้สร้างแบบจำลองเพื่อทำการพยากรณ์ราคาบิตคอยน์ โดยข้อมูลราคาบิตคอยน์เป็นข้อมูลอนุกรมเวลา (Time series data) ที่เป็นชุดข้อมูลที่ถูกเก็บรวบรวมตามระยะเวลา เป็นช่วง ๆ อย่างต่อเนื่อง [3] ในการวิเคราะห์ข้อมูลอนุกรมเวลาโดยส่วนใหญ่จะทำการคำนวณด้วยเทคนิคอริมา (Autoregressive Integrated Moving Average; ARIMA) [4]

เทคนิค ARIMA เป็น เทคนิคถดถอยที่นิยมนำมาใช้ในการสร้างแบบจำลองเพื่อใช้ในการพยากรณ์

ข้อมูลอนุกรมเวลาโดยเทคนิคนี้เหมาะสำหรับข้อมูลอนุกรมเวลาแบบไม่คงที่ (Non-stationary time series) ส่วนเทคนิคการเรียนรู้ของเครื่อง (Machine learning) เป็น ปัญญาประดิษฐ์แบบหนึ่งที่มีความสามารถเรียนรู้ได้ด้วยตัวเองทำให้เกิดเป็นแบบจำลองจากข้อมูลเก่าเพื่อใช้ในการพยากรณ์เหตุการณ์ในอนาคตได้อย่างแม่นยำ [5], [6] นักวิจัยหลายท่านจึงได้นำเทคนิคการเรียนรู้ของเครื่องมาใช้ในการพยากรณ์ข้อมูลอนุกรมเวลากันอย่างแพร่หลาย เช่น Cibikdiken และ Karakoyun [4] ได้นำเทคนิค ARIMA และ เทคนิคโครงข่ายประสาทเทียมแบบหลายสำหรับการถดถอย (Multilayer perceptron; MLP) มาสร้างแบบจำลองเพื่อพยากรณ์ราคาบิตคอยน์ล่วงหน้า โดยใช้ข้อมูลตั้งแต่วันที่ 28 เมษายน ค.ศ. 2013 ถึง 29 ตุลาคม ค.ศ. 2017 แล้ววัดประสิทธิภาพด้วยค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (Mean absolute percentage error; MAPE) พบว่า เทคนิค

MLP มีประสิทธิภาพสูงกว่าเทคนิค ARIMA ส่วน Khedmati และคณะ [7] ได้เปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ราคาเปิดตลาด (Open) ราคาปิดตลาด (Close) ราคาสูงสุด (High) และราคาต่ำสุด (Low) ของราคาบิตคอยน์ล่วงหน้า ด้วยเทคนิค ARIMA, เทคนิคโครงข่ายประสาทเทียม (Artificial neural network; ANN), เทคนิคการสุ่มป่าไม้ (Random Forest, RF) และ เทคนิคจำแนกเชิงเส้นแบบถดถอย (Support vector machine for regression; SVR) โดยใช้ค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง (Root mean square error; RMSE) เป็นเกณฑ์ในการวัดประสิทธิภาพ พบว่า เทคนิค SVR ให้ค่า RMSE ต่ำที่สุด นอกจากนี้ Lekla และ Thongkam [3] ได้เปรียบเทียบประสิทธิภาพของแบบจำลองพยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศจำนวน 5 สกุลเงิน (EURUSD, USDJPY, USDCHEF, NZDUSD และ GBPUSD) ที่สร้างจากเทคนิค LR, MLP และ SVR ข้อมูลอนุกรมเวลาจาก 1 กุมภาพันธ์ ค.ศ. 2014 ถึง 30 มกราคม ค.ศ. 2017 ในการแบ่งข้อมูลออกเป็นชุดฝึกสอนและชุดทดสอบด้วยหลักการเลื่อนหน้าต่าง (Sliding window) และใช้ค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ (Mean absolute error; MAE) และ ค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เป็นเกณฑ์ในการแสดงประสิทธิภาพในการพยากรณ์ของแบบจำลอง พบว่า เทคนิค SVR มีค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ และความคลาดเคลื่อนเฉลี่ยรากที่สองต่ำกว่าเทคนิค Linear regression และ MLP

ดังนั้นงานวิจัยนี้ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ราคาเปิดตลาด (Open) ราคาสูงสุด (High) ราคาต่ำสุด (Low) ราคาปิดตลาด (Close) ปริมาณการซื้อขายในแต่ละวัน

(Volume) และมูลค่าตลาดรวมของบิตคอยน์ (Market cap) ของราคาบิตคอยน์ในตลาดหุ้น ที่สร้างเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และ เทคนิค SVR โดยหลักการเลื่อนหน้าต่าง ซึ่งเทคนิค ARIMA เป็นเทคนิคดั้งเดิมที่มีการใช้กันมาอย่างแพร่หลายในการพยากรณ์ข้อมูลอนุกรมเวลา แต่ในหลายงานวิจัยยังไม่ได้มีข้อสรุปนำวิธีการแบบดั้งเดิมมาใช้เปรียบเทียบกับวิธีการพยากรณ์แบบใหม่โดยการใช้การเรียนรู้ของเครื่องอย่างหลากหลายทั้งนี้ก็เพื่อทราบถึงประสิทธิภาพของเทคนิคทั้งแบบดั้งเดิมและแบบใหม่ ในการพยากรณ์ราคาบิตคอยน์ในห้วงระยะเวลาที่มีความผันผวนสูงส่วนเกณฑ์ในการวัดประสิทธิภาพของแบบจำลองใช้ค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ และความคลาดเคลื่อนเฉลี่ยรากที่สอง

2. วิธีการวิจัย

วิธีดำเนินการวิจัยในงานวิจัยนี้ ประกอบด้วย 4 ขั้นตอน คือ การเตรียมข้อมูล การวิเคราะห์ข้อมูล การสร้างแบบจำลอง และ การวัดประสิทธิภาพแบบจำลอง

2.1. การเตรียมข้อมูล

ผู้วิจัยได้รวบรวมข้อมูล ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวัน และมูลค่าตลาดรวม หน่วยเป็น USD จากวันที่ 1 มกราคม ค.ศ. 2017 ถึง 31 ธันวาคม ค.ศ. 2019 จากเว็บไซต์ www.coinmarketcap.com [8] แล้วจึงทำการแปลงค่าต่าง ๆ หารด้วย 10^3 แต่ส่วนปริมาณการซื้อขาย และมูลค่าตลาดรวม นำมาหารด้วย 10^{10} เพื่อลดค่าความแปรปรวนของชุดข้อมูล [9], [10] จากการลดความแปรปรวนของชุดข้อมูลแล้ว ข้อมูลที่ได้เป็นไปตามที่แสดงใน Figure 1

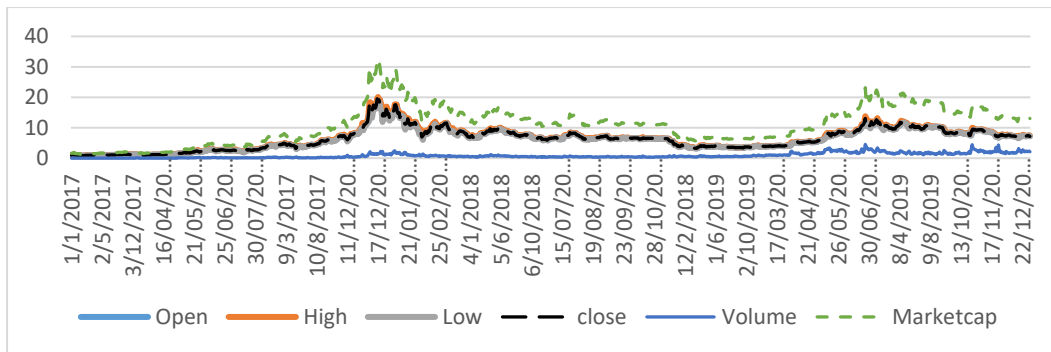


Figure 1 Bitcoin daily transactions

2.2. การวิเคราะห์ข้อมูลราคาปิดคอยน์

ในการวิเคราะห์ข้อมูลราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวัน และมูลค่าตลาดรวมของราคาปิดคอยน์นั้น ข้อมูลที่ได้เป็นข้อมูลอนุกรมเวลา [11] โดยข้อมูลจะเรียงกันด้วยเวลารายวัน เมื่อนำมาทดสอบด้วยวิธี Augmented Dickey–Fuller test (ADF) ซึ่งเป็นวิธีที่ใช้กันมากในการทดสอบความคงที่ของข้อมูลอนุกรมเวลาด้วยหลักการทางสถิติ [12] จากการทดสอบความคงที่ของข้อมูลเวลา พบว่าค่า p-values ของตัวแปรราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณซื้อขาย และ มูลค่าตลาดรวม ของแต่ละวัน เท่ากับ 0.14, 0.13, 0.18, 0.14, 0.51 และ 0.16 ตามลำดับ ซึ่งมากกว่า 0.05 หมายความว่าข้อมูลราคาปิดคอยน์ใน Figure 1 เป็นอนุกรมเวลาแบบไม่คงที่

2.3. การสร้างแบบจำลอง

การสร้างแบบจำลอง เป็นขั้นตอนการนำเอาชุดข้อมูลฝึกสอนที่เป็นอนุกรมเวลามาเป็นข้อมูลเข้าเพื่อสร้างแบบจำลองในการพยากรณ์ ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวัน และมูลค่าตลาดรวมของปิดคอยน์ ของราคาปิดคอยน์ในตลาดหุ้น ประกอบด้วยเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR

เทคนิค ARIMA เป็นวิธีการที่นำเอาสถิติถดถอยมาใช้ในการพยากรณ์อนุกรมเวลา โดยนำเอาข้อมูล

ย้อนหลังมาพยากรณ์ข้อมูลที่จะเกิดขึ้นในอนาคต จากการคำนวณค่าน้ำหนัก ซึ่งแตกต่างกันตามช่วงเวลา การสร้างแบบจำลองด้วยเทคนิคนี้มี 3 ขั้นตอน [13] เริ่มต้นด้วยการสร้างแบบจำลองด้วยข้อมูลถดถอยโดยใช้ข้อมูล ณ จุดหนึ่งของเวลา และข้อมูลก่อนหน้า ณ จุดเวลานั้น เรียกว่าขั้นตอน Autoregression (AR) ขั้นตอนถัดไปเป็นขั้นตอนการบูรณาการ (Integrated; I) โดยการหาค่าความแตกต่างของค่าจากสองจุดเวลาข้างต้น เพื่อให้ข้อมูลคงที่ ขั้นตอนสุดท้ายเป็นขั้นตอนการทำการสร้างแบบจำลองที่คำนวณจากค่าเฉลี่ยเคลื่อนที่ (Moving average; MA) โดยการนำเอาค่าที่ได้จากขั้นตอนที่ 2 และค่าผิดพลาดรวมกันเพื่อใช้ในการพยากรณ์ค่าในช่วงเวลาถัดไป

เทคนิค MLP เป็นเทคนิคที่นิยมนำมาใช้ในการจำแนกข้อมูล แต่เทคนิคนี้ก็สามารถนำมาใช้ในการวิเคราะห์เชิงถดถอย ในงานวิจัยได้นำเทคนิค MLP ไปใช้ในการวิเคราะห์เชิงถดถอยเพื่อพยากรณ์ข้อมูลอนุกรมเวลา โดยเทคนิค MLP เป็นเทคนิคหนึ่งในกลุ่มเทคนิคโครงข่ายประสาทเทียมที่ประกอบด้วย 3 ส่วน [14] คือ ส่วนของชั้นข้อมูลเข้า (Input layer) ส่วนของชั้นซ่อน (Hidden layer) และส่วนของชั้นข้อมูลออก (Output layer) ในแต่ละชั้นจะมีโหนดเซลล์ประสาท (Neuron) เชื่อมต่อกันแบบโครงข่าย หลักการทำงานของเทคนิค MLP เริ่มต้นด้วยชั้นข้อมูลเข้าจะทำการรับข้อมูลเข้าแล้วทำการส่งผ่านข้อมูลไปยังเส้นเชื่อม

ระหว่างโหนดไปยังส่วนของชั้นซ่อนเพื่อทำการคำนวณด้วยค่าน้ำหนักโดยในงานวิจัยนี้ได้ใช้อัตราการเรียนรู้ที่ 0.03 และค่าโมเมนตัม ของน้ำหนักที่ 0.2

เทคนิค RBF [15] เป็นเทคนิคที่นำไปวิเคราะห์เชิงถดถอย โครงสร้างของเทคนิคมีความคล้ายกับเทคนิค MLP แต่มีชั้นซ่อน ที่ใช้ฟังก์ชันเครือข่ายพื้นฐาน Gaussian radial basis function เพียงชั้นเดียวโดยเชื่อมต่อทุกโหนดในชั้นข้อมูลเข้า ทำให้ผลลัพธ์การพยากรณ์มีประสิทธิภาพ

เทคนิค SVR เป็นเทคนิคที่ใช้หาความสัมพันธ์เชิงเส้นได้ด้วยวิธีการถดถอยเชิงเส้น ซึ่งเหมาะกับข้อมูลเชิงเส้น และแบบไม่เป็นเชิงเส้น โดยการค้นหาฟังก์ชันที่ใกล้เคียงกับการกำหนดขอบเขตข้อมูลอินพุตเป็นตัวเลขจริงบนพื้นฐานของตัวอย่างชุดข้อมูลฝึกฝน และเมื่อมีการนำข้อมูลมาวิเคราะห์ด้วย SVR เส้น Hyperlane จะถูกใช้เป็นเส้นลักษณะแนวโน้มนการเคลื่อนที่ของความต่อเนื่องของข้อมูลที่เหมาะสมที่สุด โดยมีขอบเขตของเส้นตัดสินใจของชุดข้อมูลเป็นตัวกำหนดขอบเขต Hyperlane [16], [17] ในการศึกษาครั้งนี้ใช้ SVR แบบ nu-SVR

2.4. วัดประสิทธิภาพของแบบจำลอง

ในการวัดประสิทธิภาพของแบบจำลอง งานวิจัยนี้ได้นำหลักการเลื่อนหน้าต่าง [3] มาใช้ในการแบ่งข้อมูลเป็นชุดข้อมูลฝึกฝน และชุดข้อมูลทดสอบ จากข้อมูลราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวัน และมูลค่าตลาดรวม หน่วยเป็น USD จากวันที่ 1 มกราคม ค.ศ. 2017 ถึง 31 ธันวาคม ค.ศ. 2019 ผู้วิจัยได้ทำการทดลองทั้งสิ้น 12 รอบ โดยรอบแรก ข้อมูลจากเวลาเดือน มกราคม ค.ศ. 2017 ถึง ธันวาคม ค.ศ. 2018 เป็นชุดข้อมูลฝึกฝน และเดือน มกราคม ค.ศ. 2019 เป็นชุดข้อมูลทดสอบ จากนั้นทำการเลื่อนหน้าต่างไปหนึ่งเดือนจะได้ข้อมูลรอบที่ 2 โดยข้อมูลจากเดือน กุมภาพันธ์ ค.ศ. 2017 ถึง มกราคม ค.ศ. 2019 เป็นชุดข้อมูลฝึกสอน และชุดข้อมูลทดสอบ รอบที่เป็นข้อมูลของเดือน กุมภาพันธ์ ค.ศ. 2019 จากนั้นทำการเลื่อน

หน้าต่างไปรอบละ 1 เดือน จึงถึงรอบสุดท้ายซึ่งเป็นรอบที่ 12 เป็นข้อมูลจากเดือนธันวาคม ค.ศ. 2017 ถึง พฤศจิกายน ค.ศ.2019 เป็นชุดข้อมูลฝึกฝน ส่วนข้อมูลเดือนธันวาคม ค.ศ.2019 เป็นชุดข้อมูลทดสอบ ซึ่งสามารถแสดงหลักการ Sliding window ได้ดัง Figure 2

จากการทดลองทั้ง 12 รอบ จะได้ค่าคลาดเคลื่อนจากการพยากรณ์ของแต่ละแบบจำลองที่สร้างจากเทคนิค ARIMA, MLP, RBF และSVR มาทำการหาค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์ และความคลาดเคลื่อนเฉลี่ยรากที่สอง [18]

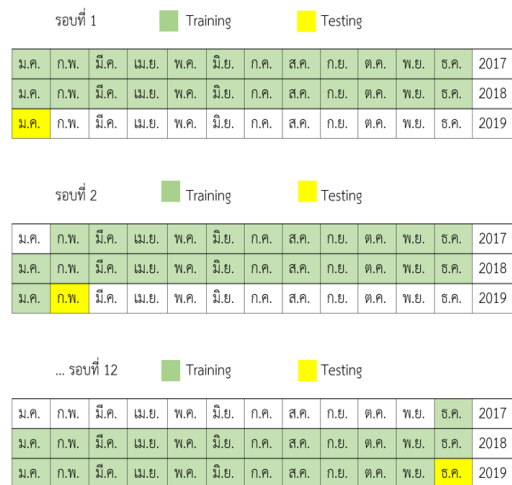


Figure 2 Sliding window

MAE คือ ค่าเฉลี่ยสมบูรณ์ของความแตกต่างระหว่างค่าพยากรณ์และค่าความจริง หากค่าที่ได้มีค่าน้อย แสดงว่าค่าพยากรณ์มีค่าใกล้เคียงค่าจริง (สมการที่ 1)

$$MAE = \frac{1}{N} \sum_{t=1}^N |e_t| \quad (1)$$

RMSE คือ ค่าเฉลี่ยรากที่สองของความแตกต่างระหว่างค่าจริงและค่าพยากรณ์ ซึ่งหากค่า RMSE มีค่าน้อย แสดงว่าการพยากรณ์นั้นใกล้เคียงกับค่าความเป็นจริง ค่า RMSE เป็นค่าที่อ่านค่าได้ง่ายกว่า MAE

เนื่องจากหน่วยของค่า RMSE มีหน่วยเดียวกันกับค่าที่แบบจำลองได้ทำการพยากรณ์ (สมการที่ 2)

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^n e_t^2} \quad (2)$$

โดยที่

$e_t = x_t - \hat{x}_t$ แทน ค่าคลาดเคลื่อนจากการ

พยากรณ์ ณ เวลา t

N แทนจำนวนเทอมทั้งหมดของ e_t

x_t แทนข้อมูล ณ เวลา t

$\hat{x}_t$ แทนค่าพยากรณ์ ณ เวลา t

การวัดประสิทธิภาพของแบบจำลองโดยนำวิธี MAE และ RMSE มาใช้ในการทดสอบแบบจำลองนั้น เป็นวิธีนำผลรวมค่าความคลาดเคลื่อนมาทำให้เป็นค่าบวก เนื่องจากค่าความคลาดเคลื่อน (Error) เป็นได้ทั้งค่าบวก และลบ เมื่อนำมารวมกันเพื่อหาค่าเฉลี่ยความคลาดเคลื่อน อาจทำให้ข้อมูลหักล้างกันเอง ดังนั้นการวัดประสิทธิภาพด้วยวิธี MAE และ RMSE เป็นวิธีที่สามารถหาค่าเฉลี่ยความคลาดเคลื่อนได้มีประสิทธิภาพเมื่อทำการทดสอบแบบจำลองในการพยากรณ์ข้อมูลที่เป็นอนุกรมเวลา

3. ผลการวิจัยและอภิปรายผลการวิจัย

ในการวิจัยครั้งนี้ โปรแกรม WEKA เวอร์ชัน 3.8.4 ใช้ในการสร้างแบบจำลองของเทคนิค MLP เทคนิค RBF และ เทคนิค SVR และโปรแกรม Python ใช้ Library statsmodels และใช้ Akaike information criterion (AIC) ในการสร้างแบบจำลองที่เหมาะสมของเทคนิค ARIMA เพื่อสร้างแบบจำลองการพยากรณ์ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวันและมูลค่าตลาดรวมของบิตคอยน์ ในตลาดหุ้น ด้วยเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR จากข้อมูลหลักการ Sliding window ในการแบ่งชุดข้อมูลออกเป็นชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบมาหาค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์ และความคลาดเคลื่อนเฉลี่ยรากที่สอง ตามลำดับ และในผลการทดลองนี้ยังแสดงค่าที่ได้จากการพยากรณ์ของแต่ละแบบจำลองอีกด้วย

ค่าความคลาดเคลื่อนเฉลี่ย เป็นค่าที่แสดงค่าความคลาดเคลื่อนสมบูรณ์ (MAE) ของแบบจำลองในการพยากรณ์ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวัน และมูลค่าตลาดรวมของราคาบิตคอยน์ในตลาดหุ้น ดังแสดงใน Figure 3

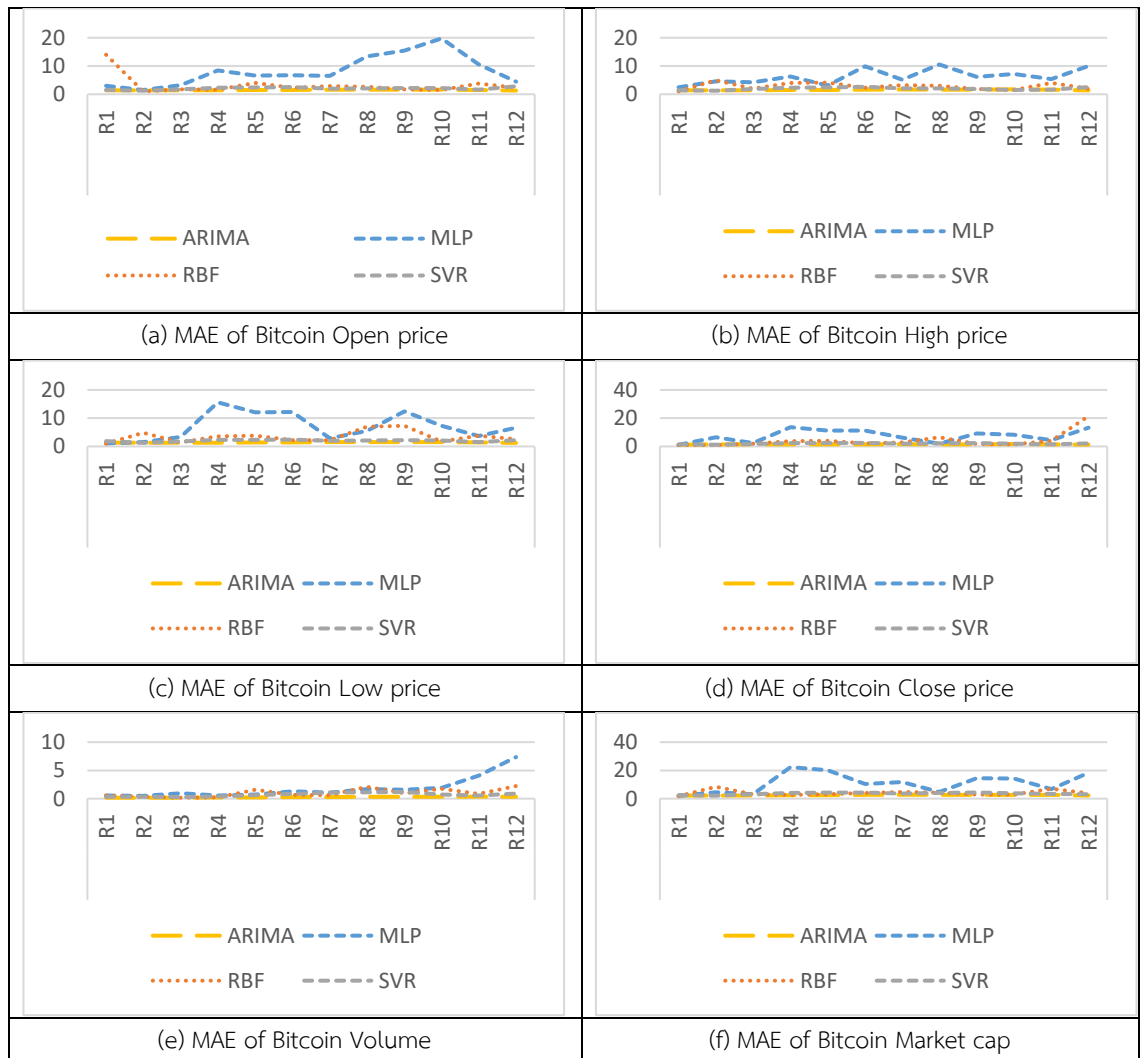


Figure 3 MAE values of the Bitcoin forecasting models

Figure 3 แสดงค่าความคลาดเคลื่อนเฉลี่ยของแบบจำลองการพยากรณ์ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวันและมูลค่าตลาดรวมของราคาปิดคอยน์ในตลาดหุ้นที่สร้างจากเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR ผลปรากฏจาก (a) ราคาเปิดตลาด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 8.31 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย 1.49 ซึ่งน้อยที่สุด (b) ราคาสูงสุด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 6.25 ซึ่งมากที่สุด

และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย 1.55 ซึ่งน้อยที่สุด (c) ราคาต่ำสุด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 7.02 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย 1.42 ซึ่งน้อยที่สุด (d) ราคาปิดตลาด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 7.48 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย 1.49 ซึ่งน้อยที่สุด (e) ปริมาณซื้อขายในแต่ละวัน เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ย เท่ากับ 1.85 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย

0.27 ซึ่งมีย่าน้อยที่สุด และ (f) มูลค่าตลาดรวม เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยเท่ากับ 11.06 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ย 2.54 ซึ่งมีย่าน้อยที่สุด

ค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง (RMSE) เป็นค่าเฉลี่ยรากที่สองของความแตกต่างระหว่างค่าพยากรณ์และค่าความจริง ของแบบจำลองในการพยากรณ์ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวันและมูลค่าตลาดรวมของราคาปิดคอยน์ในตลาดหุ้น ได้ดัง Figure 4

Figure 4 แสดงค่าความคลาดเคลื่อนเฉลี่ยรากที่สองของแบบจำลองการพยากรณ์ราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณการซื้อขายในแต่ละวันและมูลค่าตลาดรวมของราคาปิดคอยน์ในตลาดหุ้น ที่สร้างจากเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR ผลปรากฏจาก (a) ราคาเปิดตลาด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 9.75 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สองเท่ากับ 2.38 ซึ่งมีย่าน้อยที่สุด (b) ราคาสูงสุด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 7.63 ซึ่งมากที่สุด และเทคนิค SVR มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง 2.27 ซึ่งมีย่าน้อยที่สุด (c) ราคาต่ำสุด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 8.22 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง 2.25 ซึ่งมีย่าน้อยที่สุด (d) ราคาปิดตลาด เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 10.14 ซึ่งมากที่สุด และเทคนิค SVR มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง 2.23 ซึ่งมีย่าน้อยที่สุด (e) ปริมาณซื้อขายในแต่ละวัน

เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 2.23 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง 0.43 ซึ่งมีย่าน้อยที่สุด (f) มูลค่าตลาดรวม เทคนิค MLP มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง เท่ากับ 15.89 ซึ่งมากที่สุด และเทคนิค ARIMA มีค่าความคลาดเคลื่อนเฉลี่ยรากที่สอง 4.03 ซึ่งมีย่าน้อยที่สุด

จากผลการทดลอง ผู้วิจัยได้นำเสนอผลของราคาจริงของราคาเปิดตลาด ราคาสูงสุด ราคาต่ำสุด ราคาปิดตลาด ปริมาณซื้อขายในแต่ละวัน และมูลค่าตลาดรวม และราคาที่ได้จากการพยากรณ์จากแบบจำลองที่สร้างจากเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR ดังแสดงใน Figure 5

Figure 5 แสดงผลค่าการพยากรณ์ ของแบบจำลองในรอบที่ 12 มาเปรียบเทียบกับค่าจริง (Actual) โดย ค่าจริง คือเส้นสีฟ้าที่เป็นเส้นทึบ พบว่าค่าพยากรณ์ที่มีลักษณะเข้าใกล้หรือจะเป็นเส้นเดียวกันกับค่าจริงมากที่สุด คือ ผลการพยากรณ์ของแบบจำลองที่สร้างจากเทคนิค ARIMA ซึ่งแทนด้วยเส้นปะสีเขียว รองลงมาคือ เส้นค่าพยากรณ์ของแบบจำลองที่สร้างจากเทคนิค SVR ต่อมาเป็นแบบจำลองที่สร้างจากเทคนิค RBF และสุดท้ายที่ห่างจากค่าจริงมากที่สุดจะเป็นเส้นค่าพยากรณ์ของแบบจำลองจากเทคนิค MLP

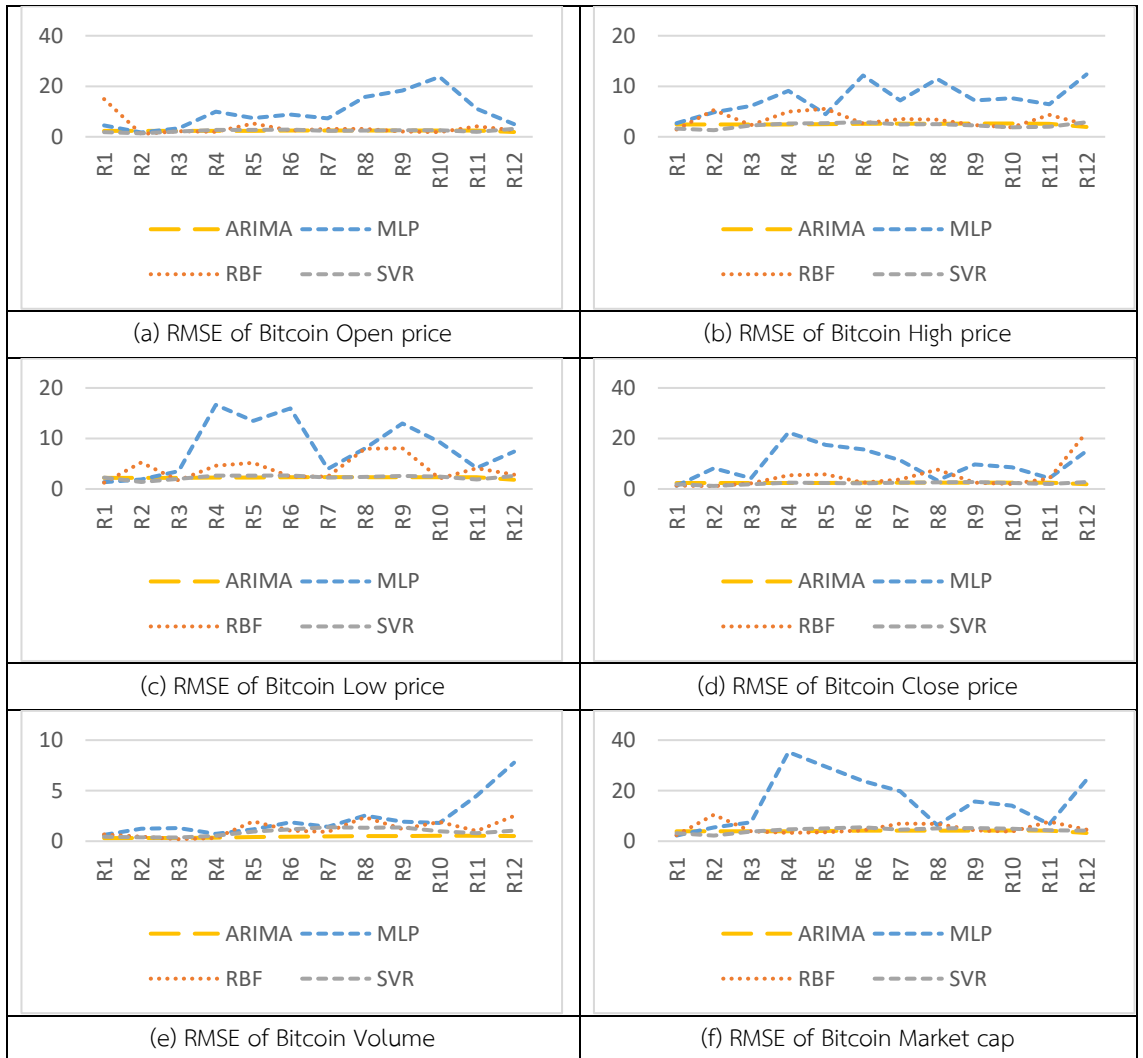


Figure 4 RMSE values of the Bitcoin forecasting models

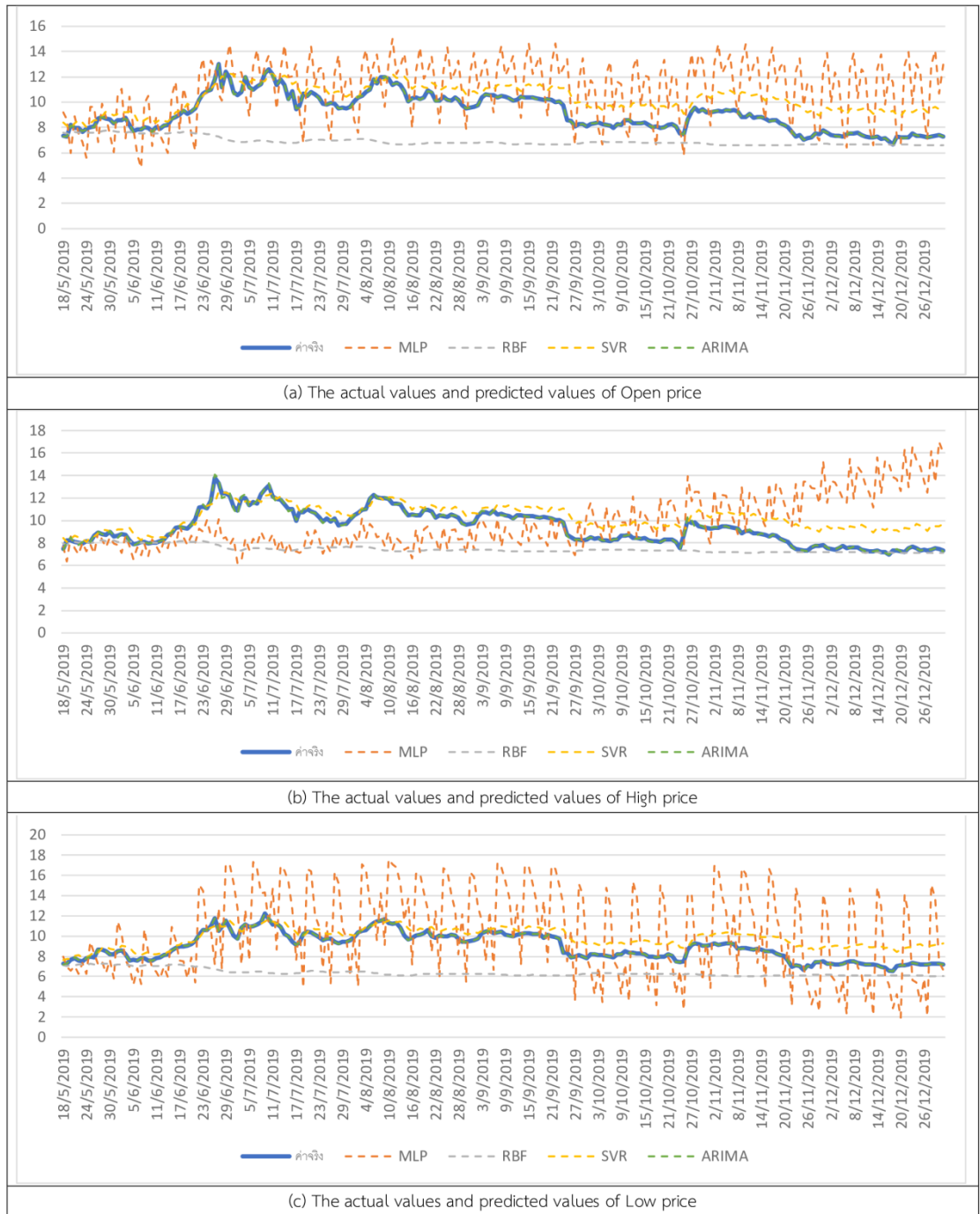


Figure 5 Actual prediction of the Bitcoin forecasting models

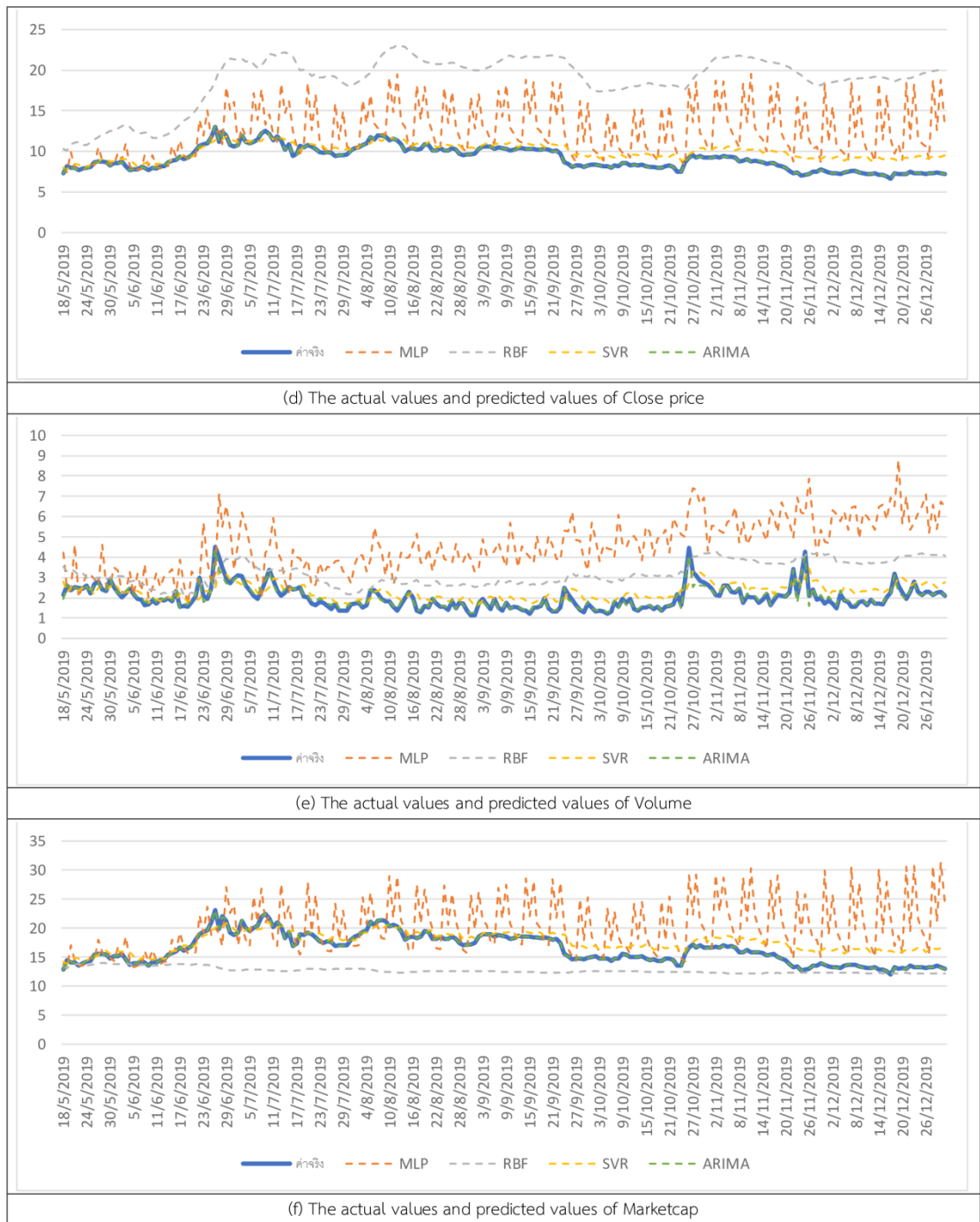


Figure 5 Actual prediction of the Bitcoin forecasting models (continued)

4. บทสรุป

ข้อมูลราคาของบิตคอยน์ทั้ง 6 ตัวแปร ซึ่งเป็นรายวัน ตั้งแต่วันที่ 1 มกราคม ค.ศ. 2017 ถึงวันที่ 31 ธันวาคม ค.ศ. 2019 นั้น เมื่อนำมาศึกษาและเปรียบเทียบประสิทธิภาพในการพยากรณ์ราคาบิตคอยน์ กลับพบว่า ทั้งเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR ล้วนมีประสิทธิภาพที่ดีในการพยากรณ์ราคาบิตคอยน์ที่มีความผันผวนสูงได้ดี แต่หากเปรียบเทียบเทคนิค ทั้ง 4 นี้ กลับพบว่าแบบจำลอง ในการทดสอบระบบแบบวิธีความคลาดเคลื่อนเฉลี่ยสัมบูรณ์และความคลาดเคลื่อนเฉลี่ยรากที่สอง โดยมีค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์เท่ากับ 1.49, 1.55, 1.42, 1.49, 0.27 และ 2.54 ในการวัดประสิทธิภาพวิธีคลาดเคลื่อนเฉลี่ยสัมบูรณ์รากที่สอง เท่ากับ 2.38, 2.48, 2.25, 2.38, 0.43 และ 4.03 เปอร์เซ็นต์ ตามลำดับ และเมื่อการเปรียบเทียบในการพยากรณ์กับค่าจริง กลับพบว่าแบบจำลองจากเทคนิค ARIMA มีผลการพยากรณ์ที่ใกล้เคียงค่าจริงมากที่สุด โดยค่าพยากรณ์ที่ได้เมื่อนำมาสร้างกราฟเปรียบเทียบ ทำให้เห็นว่าทุกช่วงเวลาของกราฟที่แบบจำลองจากเทคนิค ARIMA พยากรณ์ราคาออกมา ทิศทางของกราฟคล้ายกันกับราคาจริงเป็นอย่างมาก และมีประสิทธิภาพในการพยากรณ์ ซึ่งโดยธรรมชาติของเทคนิค ARIMA นั้น มีหลักสำคัญ คือการลดความแปรปรวนให้ได้มากที่สุด ทำให้ชุดข้อมูลอนุกรมเวลาที่มีความเป็น stationary โดยทำการ Integrated ข้อมูล เพื่อให้ได้ค่าความเชื่อมั่น (confidence) มากถึง 95 เปอร์เซ็นต์ ก่อนที่จะมีการพยากรณ์ข้อมูลอนุกรมเวลา ดังนั้นเมื่อทำการทดลองโดยใช้เทคนิค ARIMA มาสร้างแบบจำลอง ซึ่งสอดคล้องกับงานวิจัยของ Munim และคณะ [19] ที่ทำการพยากรณ์ราคาบิตคอยน์ และผลออกมาว่า เทคนิค ARIMA ให้ประสิทธิภาพในการพยากรณ์ราคาบิตคอยน์ดีที่สุด

ดังนั้น ถึงแม้ว่าทั้งเทคนิค ARIMA เทคนิค MLP เทคนิค RBF และเทคนิค SVR มีประสิทธิภาพในการพยากรณ์ราคาของบิตคอยน์ได้ดี แต่เนื่องจาก

ประสิทธิภาพของเทคนิค ARIMA ดีกว่าทุกเทคนิค เนื่องจาก เทคนิค ARIMA มีการลดความแปรปรวนของข้อมูลในการทำ Integrated ข้อมูล ซึ่งเหมาะแก่การพยากรณ์ข้อมูลอนุกรมเวลาที่มีความผันผวนสูงมาก อย่างเช่น บิตคอยน์ เป็นต้น

5. กิตติกรรมประกาศ

งานวิจัยชิ้นนี้ได้รับการสนับสนุนจากคณาจารย์คณะวิทยาการสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ ในการช่วยผลักดันและชี้แนะงานวิจัยนี้จนประสบความสำเร็จ

6. References

- [1] Crosby, M. and et al. 2016. Blockchain technology: Beyond bitcoin. **Applied Innovation Review.** 2: 6-19.
- [2] Bohme, R. and et al. 2015. Bitcoin: Economics, technology, and governance. **Journal of Economic Perspectives.** 29(2): 213-238.
- [3] Lekkla, S. and Thongkam, J. 2018. Forecasting the trend of foreign exchange rates using time series analysis techniques. **Journal of Technology Management Rajabhat Maha Sarakham University.** 5(2): 94-103. (in Thai)
- [4] Cibikdiken, A.O. and Karakoyun, E.S. 2018. Comparison of arima time series model and lstm deep learning algorithm for bitcoin price forecasting. In: **Proceedings of the 13th Multidisciplinary Academic Conference**, 12-13 October 2018. Prague, Czech Republic.

- [5] Chentawat, R., Kerdprasop, K. and Kerdprasop, N. 2017. The demand forecasting of electrical power distribution units by support vector regression with 3 folds cross-validation. **Journal of Science and Technology, Ubon Ratchathani University**. 19(1): 215-232. *(in Thai)*
- [6] Jordan, M.I. and Mitchell, T.M. 2015. Machine learning: Trends, perspectives, and prospects. **Science**. 349(6245): 255-260.
- [7] Khedmati, M., Seifi, F. and Azizi, M.J. 2020. Time series forecasting of Bitcoin price based on autoregressive integrated moving average and machine learning approaches. **International Journal of Engineering**. 33(7): 1293-1303.
- [8] CoinMarketCap. 2020. **CoinMarketCap**. <https://coinmarketcap.com>. Accessed 31 December 2020.
- [9] Bataineh, M. and Marler, T. 2017. Neural network for regression problems with reduced training sets. **Neural Networks**. 95: 1-9.
- [10] Struga, K. and Qirici, O. 2018. Bitcoin price prediction with neural networks. In: **Proceedings of the Recent Trends and Applications in Computer Science and Information Technology**, 23-24 November 2018. Tirana, Albania.
- [11] Carlberg, C. 2016. **Regression Analysis Microsoft Excel**. Indianapolis, IN: Que Publishing.
- [12] Detthamrong, U. and Chansanam, W. 2018. Time series forecasting stock closing prices of listed company using ARIMA model. **Journal of Business, Economics and Communications**. 13(2): 57-72. *(in Thai)*
- [13] Chaiyadecha, S. 2020. **ARIMA model in Python**. <https://lengyi.medium.com>. Accessed 12 April 2020. *(in Thai)*
- [14] Soonthornwiphat, S. 2015. **Application of Artificial Neural Network for Biomass Requirement Management**. M.Sc. Thesis, Silpakorn University. *(in Thai)*
- [15] Yuk W. and Thongkam, J. 2018. Comparison of time series techniques for predicting gold and oil prices. **RMUTI Journal Science and Technology**. 11(2): 154-167. *(in Thai)*
- [16] Sanguansat, P. 2019. **Artificial Intelligence Machine Learning Python Edition**. Nonthaburi: Meb Corporation. *(in Thai)*
- [17] Pimpaporn, W. 2018. The Development of a Predication Model for Academic Achievement of Students During a Learning Process by Using Data mining Techniques. **Sripatum Review of Science Technology**. 10(1): 7-17. *(in Thai)*
- [18] Thanapala, D., Charoensiri, T. and Sophonpimol, C. 2016. Forecasting of Factory Pineapple Prices with Box-Jenkins Method. **Burapha Science Journal**. 21(1): 110-118. *(in Thai)*

- [19] Munim, Z.H., Shakil, M.H. and Alon, I.
2019. Next-day bitcoin price forecast.
**Journal of Risk and Financial
Management.** 12(2): 103.