

การพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม LSTM กับค่าเฉลี่ยเคลื่อนที่ที่ต่างกัน
Common Stock Price Prediction Using LSTM with Different Types of Moving Average

กัลยพร แก้วกาญจน์

จักรกฤษณ์ กลั่นเทศ

นายณัฐดนัย พรหมชัย

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมเครื่องกล หลักสูตรวิศวกรรมเครื่องกล

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2564

การพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม LSTM กับค่าเฉลี่ยเคลื่อนที่ที่ต่างกัน
Common Stock Price Prediction Using LSTM with Different Types of Moving Average

กัลยพร แก้วกาญจน์

จักรกฤษณ์ กลั่นเทศ

นายณัฐดนัย พรหมชัย

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมเครื่องกล หลักสูตรวิศวกรรมเครื่องกล

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2564

COMMON STOCK PRICE PREDICTION USING LSTM WITH DIFFERENT TYPES OF
MOVING AVERAGE

GALYAPORN GAEWGARN

CHAKKRIT KLANTET

NATDANAI PROMCHAI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF BACHELOR OF ENGINEERING IN MECHANICAL ENGINEERING
FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2021

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาานิพนธ์ปีการศึกษา 2564

สาขาวิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม LSTM กับค่าเฉลี่ยเคลื่อนที่ที่ต่างกัน

Common Stock Price Prediction Using LSTM with Different Types of Moving Average

ผู้จัดทำ

1. กัลยพร แก้วกาญจน์ รหัสประจำตัว 61010062
2. จักรกฤษณ์ กลั่นเทศ รหัสประจำตัว 61010127
3. นายณัฐดนัย พรหมชัย รหัสประจำตัว 61010330



อาจารย์ที่ปรึกษา

(ดร.บำรุง พวงเกิด)

การพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม LSTM กับค่าเฉลี่ยเคลื่อนที่ที่ต่างกัน

กัลยพร แก้วกาญจน์ 61010062

จักรกฤษณ์ กลิ่นเทศ 61010127

นายณัฐดนัย พรหมชัย 61010330

ดร.บำรุง พ่วงเกิด อาจารย์ที่ปรึกษา

ปีการศึกษา 2564

บทคัดย่อ

เนื่องจากปัจจุบันโลกได้พัฒนามาสู่ยุคที่ปัญญาประดิษฐ์ และการเรียนรู้ด้วยเครื่องจักรมีบทบาท ความสำคัญในหลายอุตสาหกรรม การนำแมชชีนเลิร์นนิ่งมาใช้ประโยชน์ในการทำงานเป็นจำนวนมาก และคาดว่า ในอนาคตปัญญาประดิษฐ์และการเรียนรู้ด้วยเครื่องจักรถูกนำมาทำหน้าที่แทนคนเป็นอันมาก และอาชีพวิศวกรก็ เป็นอีกหนึ่งอาชีพที่ต้องมีการคิดค้น สร้าง และพัฒนาตนเองอยู่ตลอดเวลาเพื่อออกแบบและแก้ปัญหามันที่ทำให้ ได้ การเริ่มศึกษาและทำความเข้าใจเรียนรู้ข้อมูลเกี่ยวกับปัญญาประดิษฐ์ จึงเป็นสิ่งที่น่าสนใจเป็นอย่างมากสำหรับ ชาววิศวกรในอนาคต ผู้ศึกษาได้เห็นถึงปัญหาและความสำคัญของการเรียนรู้ปัญญาประดิษฐ์ อีกทั้งการศึกษาการ ลงทุนก็เป็นอีกสิ่งหนึ่งที่มีความสำคัญต่อผู้คน เพราะในยุคปัจจุบันทุกคนล้วนได้รับผลกระทบจากเงินเฟ้อ ไม่ว่าจะ ราคาอาหาร สิ่งของเครื่องใช้ รวมไปถึงราคาน้ำมัน หากเรามีรายได้ที่จำกัดตามช่องทางเดียว อาจไม่เพียงพออีก ต่อไป ดังนั้น เราจึงมองเห็นความสำคัญในด้านการลงทุน และการนำ AIML มาปรับใช้เพื่อช่วยในการลงทุนและ เพื่อเป็นการศึกษาแมชชีนเลิร์นนิ่งเบื้องต้นไปในตัวด้วย

การศึกษามีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพการพยากรณ์ราคาหุ้นสามัญ โดยการใช้โครงข่ายประสาท เทียม Long Short Term Memory (LSTM) และเปรียบเทียบประสิทธิภาพการพยากรณ์ของโครงข่ายประสาท เทียม LSTM ที่ใช้ค่าเฉลี่ยเคลื่อนที่ในแต่ละแบบว่าแบบใดมีประสิทธิภาพการพยากรณ์ที่ดีที่สุด จากการใช้ข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย ซึ่งเป็นข้อมูลรายวันย้อนหลังเป็นเวลา 10 ปี ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020 สำหรับ PTT SPCG และ TTW

ผู้วิจัยได้ใช้โปรแกรม Google Colab ซึ่งเป็นบริการของ Google ที่ให้ผู้ใช้สามารถเขียนชุดคำสั่ง Python ได้ ในส่วนของขั้นตอนการเตรียมข้อมูลที่จำเป็นสำหรับการทดลอง ประกอบด้วยข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด ราคาที่เปลี่ยนแปลงไป และปริมาณการซื้อขาย รวมทั้งข้อมูลค่าเฉลี่ยทั้ง 3 แบบ คือ Simple moving average (SMA), Exponential moving average (EMA) และ Weighted moving average (WMA) และทำการแปลงข้อมูลให้มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อลดความซับซ้อนในการคำนวณทางคอมพิวเตอร์ จากนั้นแบ่งเป็นข้อมูลสำหรับการเรียนรู้ร้อยละ 80 และข้อมูลสำหรับการทดลองร้อยละ 20 หลังจากให้คอมพิวเตอร์เรียนรู้ข้อมูลส่วนแรก จากนั้นจึงนำโมเดลโครงข่ายประสาทเทียมมาทำนายข้อมูลในส่วนที่สอง หลังจากการทำนาย ผู้วิจัยได้แปลงข้อมูลที่อยู่ระหว่าง 0-1 นั้นกลับคืนสู่ข้อมูลเดิม และวัดผลความคลาดเคลื่อน

จากการศึกษาพบว่า สำหรับ PTT ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ SMA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.778616 สำหรับ SPCG ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ EMA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.116784 และสำหรับ TTW ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ SMA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.033334 แต่เมื่อเปรียบเทียบโดยรวมทั้งหมดการพยากรณ์โดยใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ WMA มีประสิทธิภาพสูงที่สุดหากมีการนำค่าเฉลี่ยเคลื่อนที่มาใช้ด้วย โดยมีค่า MSE เท่ากับ 0.342259 อีกทั้งยังพบอีกว่า เมื่อไม่มีการนำค่าเฉลี่ยเคลื่อนที่มาใช้เลย ประสิทธิภาพที่ได้นั้นดีขึ้น โดยมีค่า MSE เท่ากับ 0.341553

การศึกษาเพิ่มเติมมีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพในการพยากรณ์ราคาปิดในวันถัดไปของหุ้นสามัญ โดยการใช้ ราคาปิด ราคาเปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย จากหุ้นตัวอื่นร่วมกัน โดยใช้โครงข่ายประสาทเทียม Long Short Term Memory (LSTM) ร่วมกับข้อมูลย้อนหลัง 10 ปี ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020 สำหรับ PTT SPCG และ TTW

ผู้วิจัยได้ใช้โปรแกรม Google Colab ซึ่งเป็นบริการของ Google ที่ให้ผู้ใช้สามารถเขียนชุดคำสั่ง Python ได้ ในส่วนของขั้นตอนการเตรียมข้อมูลที่จำเป็นสำหรับการทดลอง ประกอบด้วยข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย ของหุ้นตัวอื่นจำนวน 2 ตัว โดยแปลงข้อมูลให้มีค่าอยู่ในช่วง 0 ถึง 1 เพื่อลดความซับซ้อนในการคำนวณทางคอมพิวเตอร์และแบ่งเป็นข้อมูลสำหรับการเรียนรู้ร้อยละ 80 และข้อมูลสำหรับการทดลองร้อยละ 20 หลังจากให้คอมพิวเตอร์เรียนรู้ข้อมูลส่วนแรก จากนั้นจึงนำโมเดลโครงข่ายประสาทเทียมมาทำนายข้อมูลในส่วนที่สอง หลังจากการทำนาย ผู้วิจัยได้แปลงข้อมูลที่อยู่ระหว่าง 0-1 นั้นกลับคืนสู่ข้อมูลเดิม และวัดผลความคลาดเคลื่อน

จากการศึกษาเพิ่มเติมพบว่าการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น SPCG และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 83.38899 การพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 58.03953 การพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ SPCG ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 1.885817



Common Stock Price Prediction Using LSTM with Different Types of Moving Average

Galyaporn Gaewgarn 61010062

Chakkrit Klantet 61010127

Natdanai Promchai 61010330

Dr.Bumroong Puangkird Advisor

Year 2021

Abstract

At present, the world has developed into an era where artificial intelligence. And machine learning plays an important role in many industries. Bringing machine learning to a large amount of work It is expected that in the future, artificial intelligence and machine learning will be used to replace many people. And the career of an engineer is another career that has to constantly innovate, create and develop oneself in order to design and solve work problems that can be done. Starting to study and understand, learn about artificial intelligence. It is very interesting for future engineers. The students saw the problems and importance of learning artificial intelligence. Moreover, investment education is another thing that is important to people. Because in today's era, everyone is affected by inflation. whether it is the price of food or appliances including the price of oil. If we have a limited income based on a single channel. It may not be enough anymore. Therefore, we see the importance of investing and adopting AIML to help with investment and as an introductory machine learning education.

This study aims to study the efficiency of forecasting common stock using Long Short Term Memory (LSTM) and compare the efficiency of the LSTM using moving averages for each model to whether there is any significant difference. And which one has the most efficient forecasting performance? The used data was daily which included open price, close price, high price, low price, and trading volume. It was 10 years from January 1, 2010, to December 31, 2020, for PTT, SPCG, and TTW.

We use Google Colab, a Google service that allows users to write Python commands. As part of the procedure for preparing the necessary data for the experiment, It contains opening price, closing price, high price, lowest price, changing price, and trading volume. It includes all three average data types, Simple moving average (SMA), Exponential moving average (EMA), and Weighted moving average (WMA). We normalize the data between 0 and 1 to simplify computing. It was then divided into 80 percent of the learning data and 20 percent of the experimental data. The neural network model was then used to predict the data in the second part. After the prediction, we converted the data range [0,1] returned to the original data and measure the discrepancy.

For PTT, the predictive efficiency of the LSTM combined with SMA was the best, with an MSE of 0.778616. For SPCG, the predictive efficiency of the LSTM combined with the EMA was the best, with an MSE of 0.116784. For TTW, the predictive efficiency of the LSTM combined with SMA was the best, with an MSE of 0.033334. But overall, forecasting using the LSTM network in combination with WMA was most effective if moving averages were used, with an MSE of 0.342259. When no moving averages are applied. The performance was improved with an MSE of 0.341553.

The objective of the further study was to determine the effectiveness of forecasting the next day close price of common stock using the closing price, the opening price, the highest price, the lowest price and the trading volume. from other stock together Using a Long Short Term Memory (LSTM) neural network with 10-year historical data from January 1, 2010 to December 31, 2020 for PTT, SPCG and TTW.

We used Google Colab, a Google service that allows users to write Python commands as part of the process of preparing the data needed for the experiment. It contains information, opening price, closing price, high price, lowest price and trading volume. of two other stocks by converting the data to values in the range 0 to 1 to simplify computational calculations and divide it into 80% learning data and 20% experimental data. The computer learns the first piece of information. The neural network model was then used to predict the data in the second part. after the prediction we converted the data between 0-1 back to the original data. and measure the discrepancy

From the further study, it was found that the 1-day forward closing price forecast of PTT using SPCG and TTW price and trading volume data has a mean square error of 83.38899. The 1-day forward closing price forecast of SPCG using PTT and TTW price and trading volume data has a mean square error of 58.03953. The 1-day forward closing price forecast of TTW using PTT and SPCG price and trading volume data has a mean square error of 1.885817.



กิตติกรรมประกาศ

การศึกษาค้นคว้าอิสระนี้สามารถสำเร็จลุล่วงได้ด้วยความกรุณาจาก อาจารย์ ดร.บำรุง พ่วงเกิด ที่ให้คำปรึกษาแนะนำในระหว่างการศึกษาค้นคว้า ตลอดจนตรวจทานเนื้อหาแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่อย่างดียิ่ง กระทั่งการศึกษาค้นคว้าอิสระเสร็จสมบูรณ์ไปด้วยดี ผู้ศึกษาจึงขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณ คุณพ่อทองสุข และ คุณแม่กัญญารัตน์ ที่คอยให้ความช่วยเหลือสนับสนุนทั้งด้านกำลังใจและกำลังทรัพย์ด้วยดีตลอดมา ทำให้การศึกษาต่อระดับปริญญาตรีนี้สำเร็จลุล่วงไปได้ด้วยดี ขอขอบคุณ วินวิน ต่งซือเฉิง ศิลปินผู้ที่เป็นพลังใจในการศึกษาปริญญาตรี และ ขอขอบคุณเพื่อนแท้ทุกท่านที่คอยให้ความสนับสนุนช่วยเหลือตลอดการศึกษาจนสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ คุณสถาพร งามเรืองพงศ์ ดร.นิเวศน์ เหมวชิรวรากร คุณปีเตอร์ ลินซ์ คุณวอร์เรน บัฟเฟตต์ และคุณมาร์ค มินเนอร์วินี ที่เป็นผู้สร้างแรงบันดาลใจในการศึกษาค้นคว้าอิสระนี้ และขอขอบพระคุณ คุณแม่ถนอมศรี ที่เป็นผู้สนับสนุนและให้ความช่วยเหลือในทุก ๆ เรื่อง

ขอบพระคุณ คุณชัยภัทร เนื่องคำมา คุณลุงลุงโฉลก สัมพันธ์รักษ์ ผู้เป็น ต้นแบบแนวทางในการศึกษา ทำให้ได้รู้จักโลกของการลงทุน และการเก็งกำไร ด้วยวิธีการที่ เป็นวิทยาศาสตร์ รู้จักการใช้แมชชีนเลิร์นนิ่ง ขอขอบคุณคุณ สเตฟาน เจนเซน ผู้เขียนหนังสือ Machine learning for algorithmic trading ให้รู้จักการใช้แมชชีนเลิร์นนิ่งอย่างละเอียด ขอขอบคุณคุณ แลร์รี จาก Youtube Part Time Larry ที่ช่วยสอนและเป็นแหล่งความรู้เกี่ยวกับการเขียนโปรแกรมเชิงการเงิน ขอขอบคุณคุณ Kong Ruksiam จาก Youtube KongRuksiam Official ที่อัปโหลดวีดีโอสอนการเขียนโค้ดในภาษาต่าง ๆ ให้ได้ศึกษาหาความรู้อย่างไม่มีการค่าใช้จ่าย

สารบัญ

หน้า

บทคัดย่อ.....	I
ABSTRACT.....	IV
สารบัญ	VIII
สารบัญตาราง	XI
สารบัญรูปภาพ.....	XII
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญ	1
1.2 วัตถุประสงค์และความมุ่งหมายการศึกษา	4
1.3 สมมติฐาน.....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ	4
1.5 ขอบเขตของงานวิจัย	5
1.6 ขั้นตอนการดำเนินงาน	5
บทที่ 2 แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง.....	6
2.1 หุ้น.....	6
2.1.1 เงินปันผล (DIVIDEND).....	7
2.1.2 ส่วนต่างราคา (CAPITAL GAIN/LOSS).....	7
2.1.3 การวิเคราะห์ปัจจัยพื้นฐาน (FUNDAMENTAL ANALYSIS)	7
2.1.4 การวิเคราะห์เชิงเทคนิค (TECHNICAL ANALYSIS)	10
2.1.5 การวิเคราะห์ข้อมูล (INFORMATION ANALYSIS)	14
2.2 การวิเคราะห์เชิงอนุกรมเวลา.....	15
2.2.1 การพยากรณ์แบบดั้งเดิมโดยใช้สถิติ	15
2.2.2 การพยากรณ์โดยใช้ MACHINE LEARNING.....	20

2.3	แมชชีนเลิร์นนิง	21
2.3.1	ประเภทต่าง ๆ ของ MACHINE LEARNING.....	23
2.4	โครงข่ายประสาทเทียม (ARTIFICIAL NEURON NETWORK : ANN).....	26
2.5	โครงข่ายประสาทเทียมแบบ LONG SHORT TERM MEMORY (LSTM).....	43
2.6	LOSS FUNCTION	47
2.7	งานวิจัยที่เกี่ยวข้อง.....	51
บทที่ 3 วิธีดำเนินการวิจัย		53
3.1	ระเบียบวิธีการศึกษา	53
3.2	การรวบรวมข้อมูล	53
3.3	วิธีการประมวลผลและวิเคราะห์ข้อมูล	54
3.4	การวัดประสิทธิภาพความแม่นยำ.....	56
3.5	วิธีการศึกษาเพิ่มเติม.....	56
บทที่ 4 ผลการวิเคราะห์ข้อมูล		58
4.1	ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วัน	58
4.2	ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วัน.....	67
4.3	ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วัน	75
4.4	การเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์.....	83
4.5	ผลการพยากรณ์ราคาปิดหุ้น PTT จากการศึกษาเพิ่มเติม.....	84
4.6	ผลการพยากรณ์ราคาปิดหุ้น SPCG จากการศึกษาเพิ่มเติม.....	88
4.7	ผลการพยากรณ์ราคาปิดหุ้น TTW จากการศึกษาเพิ่มเติม	92
บทที่ 5 สรุปผลการวิเคราะห์และข้อเสนอแนะ		96
5.1	ผลสรุปการศึกษา.....	96
5.2	ข้อสังเกตและข้อเสนอแนะ	97
บรรณานุกรม		98
ภาคผนวก		100

ภาคผนวก ก. กราฟดัชนีราคาหุ้น	101
ภาคผนวก ข. ตารางแสดงผล	110
ภาคผนวก ค. PYTHON CODE	117



สารบัญตาราง

ตารางที่	หน้า
2.1 เปรียบเทียบโครงข่ายทางชีววิทยาและโครงข่ายประสาทเทียม (NEGNEVITSKY, 2005).....	28
4.1 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคา.....	58
4.2 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ SMA.....	60
4.3 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ EMA.....	62
4.4 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ WMA.....	64
4.5 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคา	67
4.6 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ SMA.....	69
4.7 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ EMA.....	71
4.8 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ WMA.....	73
4.9 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคา	75
4.10 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ SMA	77
4.11 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ EMA	79
4.12 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ WMA.....	81
4.13 แสดงผลความคลาดเคลื่อนเฉลี่ยในการพยากรณ์	83
4.14 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาของหุ้น SPCG และ TTW	84
4.15 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาของหุ้น PTT และ TTW	88
4.16 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาของหุ้น PTT และ SPCG.....	92

สารบัญรูปภาพ

รูปที่	หน้า
1.1 มูลค่าหลักทรัพย์ตามราคาตลาดของตลาดหลักทรัพย์แห่งประเทศไทย.....	2
2.1 เส้นแนวโน้มขาขึ้น.....	13
2.2 เส้นแนวโน้มขาลง.....	14
2.3 เส้นแนวโน้มออกข้าง.....	14
2.4 TRADITIONAL PROGRAMMING & MACHINE LEARNING.....	21
2.5 LEARNING PHASE.....	22
2.6 INFERENCE FROM MODEL	22
2.7 ความสัมพันธ์ของระบบประสาททั้ง 4 ส่วน (NEGNEVITSKY, 2005).....	26
2.8 โครงสร้างของโครงข่ายประสาทเทียมโดยทั่วไป (NEGNEVITSKY, 2005).....	27
2.9 การทำงานของโครงข่ายประสาทในสมองเทียม (CHOOPAN RATTANAPOKA).....	28
2.10 แผนภาพของเซลล์ประสาท (NEGNEVITSKY, 2005).....	28
2.11 โครงข่ายประสาทเทียมอย่างง่าย (ANDERSON & MCNEILL, 1992).....	29
2.12 ACTIVATION FUNCTIONS OF A NEURON (NEGNEVITSKY, 2005).....	34
2.13 SINGLE LAYER PERCEPTRON (ESTEBON, 1997).....	35
2.14 เพอร์เซปตรอน (PERCEPTRON) (เกศรินดา สุศรีพ, 2554).....	36
2.15 MULTILAYER PERCEPTRON แบบมีชั้นซ่อน 2 ชั้น (NEGNEVITSKY, 2005).....	36
2.16 โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (NEGNEVITSKY, 2002).....	37
2.17 รูปแบบโครงข่ายประสาทเทียม RNN (OLAH, 2015).....	43
2.18 รูปแบบโครงข่ายประสาทเทียม LSTM (OLAH, 2015).....	44
2.19 สัญลักษณ์ที่ลูกใช้ในแผนภาพโครงข่ายประสาทเทียม (Olah, 2015).....	44
2.20 ขั้นตอนการตัดสินใจเลือกทั้งข้อมูล (OLAH, 2015).....	45
2.21 ขั้นตอนการตัดสินใจเก็บข้อมูล (OLAH, 2015).....	45
2.22 ขั้นตอนการอัปเดตสถานะเซลล์ (OLAH, 2015).....	46
2.23 ขั้นตอนการส่งออกเอาต์พุต (OLAH, 2015).....	46
2.24 MSE LOSS (Y-AXIS) VS. PREDICTIONS (X-AXIS) (PRINCE GROVER, 2018).....	48

2.25 MAE LOSS (Y-AXIS) VS. PREDICTIONS (X-AXIS) (PRINCE GROVER, 2018).....	49
2.26 การใช้ GRADIENT DESCENT ร่วมกับ LOSS FUNCTION (PCP, 2020).....	49
4.1 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 1	59
4.2 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 2	59
4.3 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 3	60
4.4 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 1	61
4.5 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 2	61
4.6 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 3	62
4.7 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 1	63
4.8 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 2	63
4.9 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 3	64
4.10 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 1	65
4.11 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 2	65
4.12 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 3	66
4.13 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 1	67
4.14 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 2	68
4.15 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 3	68
4.16 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 1	69
4.17 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 2	70
4.18 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 3	70
4.19 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 1	71
4.20 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 2	72
4.21 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 3	72
4.22 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 1	73
4.23 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 2	74
4.24 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 3	74
4.25 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 1	75
4.26 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 2	76

4.27 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 3.....	76
4.28 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 1	77
4.29 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 2	78
4.30 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 3	78
4.31 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 1	79
4.32 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 2	80
4.33 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 3	80
4.34 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 1	81
4.35 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 2	82
4.36 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 3	82
4.37 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (1)...	85
4.38 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (2)...	86
4.39 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (3)...	87
4.40 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (1)...	89
4.41 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (2)...	90
4.42 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (3)...	91
4.43 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (1)...	93
4.44 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (2)...	94
4.45 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (3)...	95

บทที่ 1 บทนำ

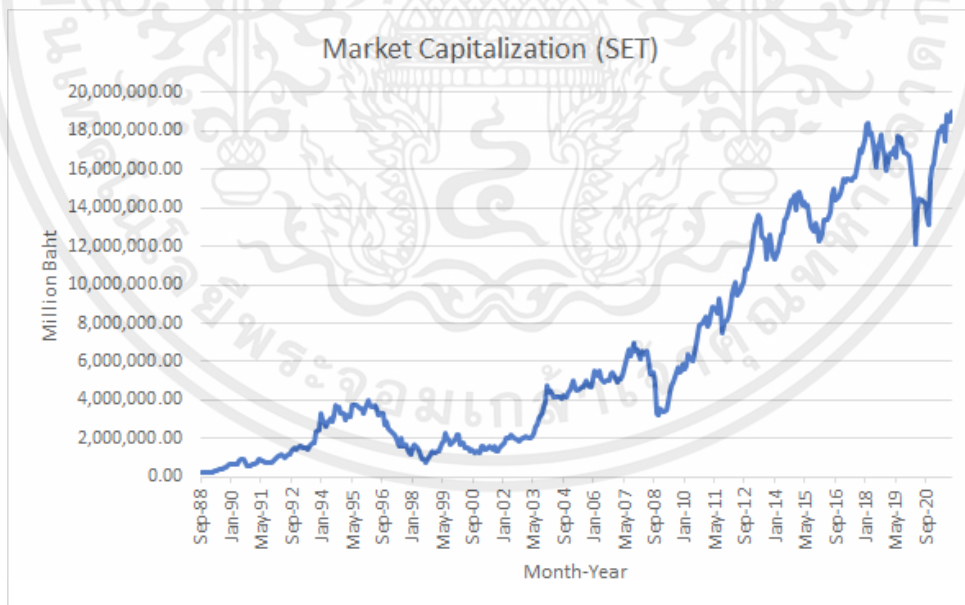
1.1 ที่มาและความสำคัญ

เนื่องจากปัจจุบันโลกได้พัฒนามาสู่ยุคที่ปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ด้วยเครื่องจักร (Machine Learning) มีบทบาทความสำคัญในหลายอุตสาหกรรม เช่น อุตสาหกรรมการผลิต การก่อสร้าง การแพทย์ การเกษตร และการเงิน เป็นต้น มีการนำแมชชีนเลิร์นนิงมาใช้ประโยชน์ในการทำงานเป็นจำนวนมาก และคาดว่าในอนาคตปัญญาประดิษฐ์และการเรียนรู้ด้วยเครื่องจักรถูกนำมาทำหน้าที่แทนคนเป็นอันมาก โลกในยุคต่อไป AIML สามารถเข้ามาแทนมนุษย์ในแทบทุกอุตสาหกรรม และอาชีพวิศวกรก็เป็นอีกหนึ่งอาชีพที่ต้องมีการคิดค้น สร้าง และพัฒนาตนเองอยู่ตลอดเวลาเพื่อออกแบบและแก้ปัญหางานที่ทำให้ได้ การเริ่มศึกษาและทำความเข้าใจเรียนรู้ข้อมูลเกี่ยวกับปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ด้วยเครื่องจักร (Machine Learning) จึงเป็นสิ่งที่น่าสนใจเป็นอย่างมากสำหรับชาววิศวกรในอนาคต ซึ่งการศึกษาเกี่ยวกับปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ด้วยเครื่องจักร (Machine Learning) ทำให้เราเข้าใจขั้นตอนวิธี (Algorithm) ของปัญญาประดิษฐ์และการเรียนรู้ด้วยเครื่องจักรอย่างลึกซึ้งพอที่สามารถนำไปแก้ไขปัญหาและแนะนำให้นักวิทยาศาสตร์ข้อมูล (Data scientist) นักวิเคราะห์ธุรกิจ (Business Analyst) นักวิเคราะห์ข้อมูล (Data Analyst) เลือกใช้ AIML ได้อย่างเหมาะสมกับประเภทและความซับซ้อนของปัญหาและข้อมูลที่ต้องการวิเคราะห์ ทำให้เข้าใจการปรับจูนค่าพารามิเตอร์ใน AIML และเข้าใจวิธีการหาค่าเหมาะสมที่สุด (Optimization) ใน AIML ให้ได้ผลลัพธ์ที่ดีที่สุดและแก้ไขปัญหาได้เมื่อเกิดปัญหา คาดว่าความต้องการวิศวกรปัญญาประดิษฐ์ในประเทศไทยเพิ่มมากขึ้นในอนาคต โดยเฉพาะอย่างยิ่งในภาคเอกชนที่มีการนำ AIML มาใช้งานจริงในธุรกิจเพิ่มขึ้นทุก ๆ วัน เพื่อเพิ่มความสามารถในการแข่งขันและความยั่งยืนทางธุรกิจ ซึ่งจำนวนของธุรกิจที่มีการนำ AIML มาใช้งานเพิ่มสูงขึ้นร้อยละ 270 ภายในระยะเวลา 4 ปี (Gartner, 2019)

ผู้ศึกษาได้เห็นถึงปัญหาและความสำคัญของการเรียนรู้ปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ด้วยเครื่องจักร (Machine Learning) เพราะในการเรียนการสอน ณ ปัจจุบันยังไม่แพร่หลายมากนัก จึงอยากศึกษาเกี่ยวกับวิธีการทำงานและกระบวนการความคิดที่จะทำการศึกษาเกี่ยวกับการเรียนรู้ปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ด้วยเครื่องจักร (Machine Learning) ขึ้นมา และเนื่องจากในปัจจุบัน การศึกษาการลงทุนก็เป็นอีกสิ่งหนึ่งที่มีความสำคัญต่อผู้คนเพราะในยุคปัจจุบัน ทุกคนล้วนได้รับผลกระทบจากเงินเพื่อ ไม่ว่าจะเป็นราคาอาหาร สิ่งของเครื่องใช้ รวมไปถึงราคาน้ำมัน หากเรามีรายได้ที่จำกัดตามช่องทางเดียว

อาจไม่เพียงพออีกต่อไป ดังนั้น เราจึงมองเห็นความสำคัญในด้านการลงทุน และการนำ AIML มาปรับใช้เพื่อช่วยในการลงทุน

ตลาดหลักทรัพย์เป็นสถานที่ซื้อขายหลักทรัพย์ เช่น ตราสารทุน ตราสารอนุพันธ์ เพื่อเป็นการระดมทุนให้กับบริษัทสำหรับใช้ในการขยายกิจการ โดยเสนอขายหุ้นบางส่วนของบริษัทให้นักลงทุน นับว่าเป็นตลาดทุนแห่งหนึ่งที่มีความนิยมและเป็นที่น่าสนใจแก่นักลงทุนอย่างมาก สังเกตได้จากมูลค่าหลักทรัพย์ตามราคาตลาดของตลาดหลักทรัพย์แห่งประเทศไทยที่เติบโตอย่างรวดเร็วในช่วงเวลา 10 ปีที่ผ่านมา ดังรูปที่ 1.1 นักลงทุนที่เข้ามาในตลาดต่างมีกลยุทธ์ในการลงทุนของตนเอง สามารถจำแนกออกเป็น 3 ประเภทคือ การลงทุนโดยวิเคราะห์ปัจจัยพื้นฐาน (Fundamental analysis) โดยวิเคราะห์จากงบการเงินของบริษัท เช่น หนี้สิน เงินปันผล กระแสเงินสด งบดุล และความสามารถในการแข่งขันในกลุ่มธุรกิจนั้น ๆ ของแต่ละบริษัท เพื่อหามูลค่าที่แท้จริง การลงทุนโดยวิเคราะห์ข้อมูล (Information analysis) ข้อมูลที่ใช้วิเคราะห์ถูกรวบรวมจากแหล่งสาธารณะ เช่น หนังสือพิมพ์ นิตยสาร กระดานข่าวอิเล็กทรอนิกส์ โดยพยากรณ์ตลาดจากแหล่งข้อมูลเหล่านี้ การลงทุนโดยวิเคราะห์เชิงเทคนิค (Technical analysis) เป็นการศึกษาพฤติกรรมตลาดผ่านการใช้แผนภูมิ อินดิเคเตอร์ (Indicator) เพื่อพยากรณ์แนวโน้มราคาในอนาคต



รูปที่ 1.1 มูลค่าหลักทรัพย์ตามราคาตลาดของตลาดหลักทรัพย์แห่งประเทศไทย

การพยากรณ์ (Forecast) เป็นการคาดการณ์แนวโน้มที่เกิดขึ้นในอนาคต โดยใช้ข้อมูลที่รวบรวมไว้ในอดีต สิ่งที่พยากรณ์มานั้นอาจตรงหรือแตกต่างกับสิ่งที่เกิดขึ้นจริงในอนาคต ในปัจจุบันยังไม่มีแนวคิด หรือวิธีการใดในการพยากรณ์มีความแม่นยำหนึ่งร้อยเปอร์เซ็นต์ แต่การพยากรณ์ยังเป็นสิ่งสำคัญต่อนักลงทุน เนื่องจากการพยากรณ์ช่วยในการวางแผนการลงทุน การเตรียมรับมือกับสิ่งที่อาจเกิดขึ้น และเป็นตัวช่วยในการตัดสินใจของผู้ลงทุน

โครงข่ายประสาทเทียม (Artificial neural networks) เป็นระบบประมวลผลข้อมูลที่มีโครงข่ายคล้ายกับโครงข่ายประสาททางชีววิทยา (Biology neural networks) ถูกพัฒนาในรูปแบบโมเดลทางคณิตศาสตร์ ในปัจจุบันมีการนำแนวคิดของโครงข่ายประสาทเทียมมาใช้ในการด้านการเงิน เช่น งานวิจัยของ Amod Merkute และ Tanuj Sarode ในหัวข้อ Forecasting Market Price of Stock using Artificial Neural Network มีการใช้ Neural network แบบ Multi Perceptron Model (MLP) เพื่อพยากรณ์ราคาตลาด

Long short term memory (LSTM) เป็นโครงข่ายชนิดหนึ่งของ Recurrent neural networks (RNN) ถูกพัฒนาเพื่อแก้ปัญหาของ RNN ที่ไม่สามารถเรียนรู้ความสัมพันธ์ในระยะยาวได้ เนื่องจากไม่มีหน่วยความจำ (Memory) ในส่วนของหน่วยความจำของ LSTM สามารถตัดสินใจได้ว่าเก็บข้อมูลใด ๆ หรือลบข้อมูลใด ๆ และข้อมูลใดจากหน่วยความจำที่ส่งผลต่อการพยากรณ์

ค่าเฉลี่ยเคลื่อนที่ (Moving average) เป็นการปรับค่าให้เรียบ เพื่อลดความแปรปรวนหรือลดทอนสัญญาณรบกวนจากข้อมูลที่นำมาเฉลี่ย ค่าเฉลี่ยเคลื่อนที่ที่สามารถแยกย่อยได้อีกตามลักษณะการถ่วงน้ำหนักของข้อมูลที่นำมาเฉลี่ย

ด้วยลักษณะที่ต่างกันในการถ่วงน้ำหนัก ผู้วิจัยคิดว่าหากมีการนำมาใช้ในฐานะอินพุตร่วมกับโครงข่ายประสาทเทียมแบบ LSTM ในการพยากรณ์ ประสิทธิภาพที่ได้อาจมีความแตกต่างกัน จึงส่งผลให้เกิดการศึกษานี้ขึ้น การศึกษานี้เป็นการพยากรณ์ราคาหุ้นสามัญในกลุ่มพลังงานคือ PTT SPCG TTW โดยใช้โครงข่ายประสาทเทียมแบบ LSTM ร่วมกับค่าเฉลี่ยเคลื่อนที่แบบ Simple moving average (SMA) Weight moving average (WMA) และ Exponential moving average (EMA) เพื่อเปรียบเทียบความแตกต่างและประสิทธิภาพในการพยากรณ์

1.2 วัตถุประสงค์และความมุ่งหมายการศึกษา

เพื่อศึกษาประสิทธิภาพการพยากรณ์ราคาหุ้นสามัญ โดยการใช้โครงข่ายประสาทเทียม Long short term memory (LSTM) จากการใช้ข้อมูลราคาหุ้นสามัญร่วมกับค่าเฉลี่ยเคลื่อนที่แบบ Simple moving average (SMA) Exponential moving average (EMA) และ Weighted moving average (WMA) และเปรียบเทียบประสิทธิภาพการพยากรณ์ของโครงข่ายประสาทเทียม Long short term memory (LSTM) จากการใช้ข้อมูลราคาหุ้นสามัญร่วมกับค่าเฉลี่ยเคลื่อนที่แบบใดมีประสิทธิภาพดีที่สุด และโดยการใช้โครงข่ายประสาทเทียม Long short term memory (LSTM) จากการใช้ข้อมูลราคาหุ้นสามัญร่วมกับค่าเฉลี่ยเคลื่อนที่ที่มีประสิทธิภาพดีกว่าการไม่ใช้ค่าเฉลี่ยเคลื่อนที่หรือไม่

1.3 สมมติฐาน

การพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม Long short term memory (LSTM) จากการใช้ข้อมูลราคาหุ้นสามัญร่วมกับค่าเฉลี่ยเคลื่อนที่ที่มีประสิทธิภาพดีกว่าการพยากรณ์ราคาหุ้นสามัญโดยใช้โครงข่ายประสาทเทียม Long short term memory (LSTM) จากการใช้เพียงข้อมูลราคาหุ้นสามัญ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นประโยชน์ต่อผู้ลงทุนในการช่วยวิเคราะห์ทิศทางในการลงทุน
2. ช่วยให้ผู้สามารถวางแผนการลงทุนตอบสนองกับการพยากรณ์ได้
3. ประหยัดเวลาในการวิเคราะห์การลงทุน
4. สามารถนำการพยากรณ์ไปประยุกต์ใช้กับตลาดทุนอื่นได้
5. ได้เรียนรู้เกี่ยวกับแมชชีนเลิร์นนิง

1.5 ขอบเขตของงานวิจัย

1. การศึกษาในครั้งนี้มุ่งศึกษาเฉพาะ ความแตกต่างของผลการพยากรณ์ที่ได้มาจาก ข้อมูลของค่าเฉลี่ยย้อนหลัง ที่แตกต่างกัน 3 แบบ คือ Simple moving average (SMA) Exponential moving average (EMA) และ Weighted moving average (WMA)
2. ผู้วิจัยได้ทำการศึกษาโดยอ้างอิงจากข้อมูลเฉพาะในช่วงเวลา ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020 รวมระยะเวลา 10 ปี
3. กลุ่มตัวอย่างที่ใช้ในการศึกษาคือ เฉพาะในหุ้นสามัญกลุ่มพลังงานของตลาดหลักทรัพย์แห่งประเทศไทย 3 บริษัทหลักทรัพย์คือ PTT SPCG TTW

1.6 ขั้นตอนการดำเนินงาน



บทที่ 2 แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2.1 หุ้น

ตราสารหุ้น หรือ หุ้น (Stock) คือสิ่งที่กิจการออกให้แก่ผู้ลงทุนเพื่อแสดงสิทธิ์การเป็นเจ้าของตามสัดส่วนที่ถือครอง ซึ่งผู้ถือหุ้นยังมีสัดส่วนในการถือหุ้นมากเท่าไร ก็ยังมีสิทธิ์ความเป็นเจ้าของกิจการมากขึ้นเท่านั้น โดยมีสิทธิ์ออกเสียงในการประชุมผู้ถือหุ้น และมีสิทธิ์ได้รับเงินปันผลเมื่อบริษัทมีกำไร ซึ่งหุ้นก็คือส่วนหนึ่งของกิจการนั่นเอง เพราะกิจการออกหุ้นมาให้แก่ผู้ลงทุนซื้อเพื่อระดมเงินทุนไปใช้ในการดำเนินกิจการในด้านต่าง ๆ โดยผู้ลงทุนที่ถือตราสารมีสถานะเป็น “เจ้าของกิจการ” หรือเรียกว่า ผู้ถือหุ้น ผู้ลงทุนสามารถเลือกลงทุนหุ้นแล้วเติบโตไปกับกิจการได้โดยที่ไม่จำเป็นต้องทำธุรกิจด้วยตนเอง แค่ซื้อหุ้น ติดตามข่าวสาร และใช้สิทธิ์ความเป็นเจ้าของ

ซึ่งโดยทั่วไปแล้วหุ้นแบ่งออกเป็น 2 ประเภท ได้แก่

1. หุ้นสามัญ (Common stock)

หุ้นสามัญ หมายถึง หุ้นที่กิจการที่เป็นบริษัทมหาชนจำกัด (บมจ.) ออกให้แก่ผู้ถือ เพื่อระดมเงินทุนหุ้นสามัญจากประชาชนทั่วไป ไปใช้ในวัตถุประสงค์ต่าง ๆ เช่น เพื่อการลงทุนขยายเพิ่มเติมเพื่อหมุนเวียนในกิจการ เพื่อใช้หนี้ ฯลฯ โดยผู้ถือหุ้นสามัญมีสิทธิ์ร่วมเป็นเจ้าของบริษัท มีสิทธิ์ในการออกเสียงลงมติในที่ประชุมผู้ถือหุ้นตามสัดส่วนของหุ้นที่ถือครองอยู่ ถ้ายิ่งถือครองมากก็มีสิทธิ์ในกิจการมาก นอกจากนี้ยังมีสิทธิ์ได้รับส่วนแบ่งจากกำไรของบริษัทตามนโยบายการจ่ายเงินปันผลอีกด้วย และมีโอกาสได้รับกำไรจากส่วนต่างของราคาเมื่อราคาหลักทรัพย์ปรับตัวสูงขึ้นตามศักยภาพของบริษัท

2. หุ้นบุริมสิทธิ (Preferred stock)

เป็นหุ้นที่ผู้ถือมีส่วนร่วมเป็นเจ้าของกิจการเช่นเดียวกับหุ้นสามัญ แต่ไม่มีสิทธิ์ในการออกเสียงลงมติในที่ประชุมผู้ถือหุ้น แต่เมื่อกิจการมีกำไรจากการดำเนินงาน ผู้ถือหุ้นบุริมสิทธิได้รับเงินปันผลในอัตราคงที่ ซึ่งมากหรือน้อยกว่าผู้ถือหุ้นสามัญก็ได้ ขณะเดียวกัน หากกิจการนั้นต้องเลิกกิจการ และมีการชำระบัญชีโดยการขายทรัพย์สิน ผู้ถือหุ้นบุริมสิทธิยังได้รับเงินคืนทุนก่อนผู้ถือหุ้นสามัญ หุ้นประเภทนี้มีไม่มากนักในตลาดหลักทรัพย์ มีการซื้อขายกันน้อย หรือภาษาเทคนิคเรียกว่ามีสภาพคล่องต่ำ

โดยผลตอบแทนจากการถือครองตราสารหุ้นโดยทั่วไปได้รับ 2 รูปแบบ แบ่งเป็น ส่วนของเงินปันผล และ ส่วนต่างราคา

2.1.1 เงินปันผล (Dividend)

หากการดำเนินงานในรอบปี บริษัทมีผลกำไรและสามารถนำผลกำไรที่ได้นี้มาแบ่งให้กับผู้ถือหุ้น เรียกว่า เงินปันผล ซึ่งผลตอบแทนส่วนนี้เป็นส่วนแบ่งจากกำไรของบริษัทที่ทำได้ตามนโยบายของบริษัท และเงินปันผลนี้ ขึ้นอยู่กับนโยบายของบริษัทอีกด้วย บริษัทสามารถจ่ายหรือไม่จ่ายเงินปันผลก็ได้ ถ้าหากบริษัทไม่จ่ายเงินปันผล ส่งผลให้บริษัทมีเงินสดกลับไปลงทุนขยายกิจการเพิ่มเติม ซึ่งผลกำไรที่นักลงทุนได้รับออกมามีในรูปแบบของ ราคาหุ้นหรือกำไรจากส่วนต่างราคาแทน แต่ส่วนมากบริษัทมีนโยบายแบ่งเงินปันผลเพื่อเป็นการให้มีผู้ลงทุนมาร่วม สนใจลงทุนกับบริษัท

2.1.2 ส่วนต่างราคา (Capital gain/Loss)

ผลตอบแทนส่วนนี้เกิดจากที่เราซื้อหุ้นในตอนที่เราซื้อและสามารถขายได้แพงกว่าราคาที่ซื้อเข้ามา ส่งผลให้เกิดกำไร (Capital gain) แต่ถ้าราคาที่เราซื้อหุ้น ถูกกว่าราคาที่เราขายได้ ส่งผลให้เกิดผลขาดทุน (Capital loss) ซึ่ง สามารถเป็นได้ทั้งการลงทุนและการเก็งกำไร มีโอกาสทั้งกำไรและขาดทุน ต้องใช้เทคนิคการวิเคราะห์ที่ค่อนข้างมาก

การลงทุนในหุ้นจำเป็นต้องมีความรู้หรือกลยุทธ์ เนื่องจากการลงทุนในหุ้นเป็นการลงทุนที่มีความเสี่ยงสูง มาก ดังที่กล่าวไปข้างต้นว่าสามารถจำแนกออกเป็น 3 ประเภท คือ การวิเคราะห์ปัจจัยพื้นฐาน การวิเคราะห์เชิง เทคนิค และการวิเคราะห์ข้อมูล

2.1.3 การวิเคราะห์ปัจจัยพื้นฐาน (Fundamental analysis)

การวิเคราะห์ปัจจัยพื้นฐาน เป็นแนวทางหรือกลยุทธ์ในการลงทุนที่มีความเสี่ยงต่ำกว่าการวิเคราะห์เชิง เทคนิค และการวิเคราะห์ข้อมูล เนื่องจาก เป็นการพยากรณ์แนวโน้มจากผลการดำเนินงานของบริษัทโดยตรง อีกทั้งยังวิเคราะห์ความสามารถในการแข่งขันของบริษัท ผู้บริหาร และโมเดลธุรกิจ

การวิเคราะห์ปัจจัยพื้นฐานเป็นการวิเคราะห์บริษัทด้วยข้อมูลทั้งในเชิงคุณภาพและปริมาณ โดยการนำข้อมูลตัวเลขจากงบการเงินมาหาอัตราส่วนต่าง ๆ อัตราส่วนเหล่านี้สามารถใช้เปรียบเทียบความสามารถของบริษัทที่ทำได้ในช่วงระยะเวลาที่ผ่านมา เพื่อเป็นการมองแนวโน้มของบริษัทในอนาคต อีกทั้งสามารถเปรียบเทียบกับบริษัทอื่น ๆ ที่อยู่ในกลุ่มอุตสาหกรรมเดียวกันว่าในช่วงระยะเวลาที่ผ่านมา บริษัทใดที่มีความสามารถในการแข่งขันมากกว่าหรือมีการดำเนินงานที่ดีกว่า โดยอัตราส่วนที่ใช้ในการเปรียบเทียบประกอบไปด้วย

2.1.3.1 กำไรต่อหุ้น (Earnings per share : EPS)

อัตราส่วนนี้เกิดจากการนำรายได้สุทธิของบริษัทหารด้วยจำนวนหุ้นที่อยู่ในมือผู้ถือหุ้น ทำให้เห็นถึงความสามารถของบริษัทในการสร้างกำไรให้กับแต่ละหุ้นที่ผู้ถือหุ้นถือครองอยู่ อีกทั้งสามารถนำมาเปรียบเทียบกับอัตราการเติบโตของกำไรในแต่ละปีได้ บริษัทที่มีอัตราการเติบโตของกำไรที่สม่ำเสมอหรือต่อเนื่องนับว่าเป็นบริษัทที่ดี เนื่องจากส่งผลให้ง่ายต่อการคาดการณ์กำไรของบริษัทในอนาคต แต่ทั้งนี้ทั้งนั้นก็ขึ้นอยู่กับแต่ละบริษัทว่าอยู่ในกลุ่มธุรกิจใด เช่น บริษัทกลุ่มอุตสาหกรรมยานยนต์มีช่วงที่มียอดขายและกำไรขึ้นลงเป็นประจำสม่ำเสมอ บริษัทอาจมีอัตราการเติบโตของกำไรที่ลดลงเนื่องจากเป็นช่วงขาลงของบริษัท ดังนั้น เราจึงไม่สามารถตัดสินได้ว่าบริษัทใดดำเนินการได้ดีหรือแย่จากการดูกำไรเพียงอย่างเดียว

2.1.3.2 อัตราส่วนทุนหมุนเวียน (Current ratio)

อัตราส่วนนี้เกิดจากการนำสินทรัพย์หมุนเวียนหารด้วยหนี้สินหมุนเวียน แสดงให้เห็นถึงความสามารถของบริษัทในการชำระหนี้ระยะสั้น ซึ่งส่งผลต่อสภาพคล่องในการดำเนินงานของบริษัท ยิ่งบริษัทใดมีอัตราส่วนทุนหมุนเวียนมากเท่าใด บริษัทนั้นก็มีความสามารถในการชำระหนี้มากเท่านั้น หากบริษัทมีอัตราส่วนทุนหมุนเวียนต่ำกว่าหนึ่ง แสดงว่าบริษัทกำลังประสบกับปัญหาในการจ่ายหนี้ในปัจจุบัน แต่ก็ไม่ได้หมายความว่าบริษัทนั้น ๆ ใกล้เคียงล้มละลาย เนื่องจากเจ้าหนี้ก็อาจยอมให้บริษัทเลื่อนการชำระหนี้หรือบางบริษัทอาจมีสินทรัพย์ซ่อนเร้นที่สามารถนำมาใช้ชำระหนี้แทนได้

2.1.3.3 อัตราผลตอบแทนผู้ถือหุ้น (Return on equity : ROE)

อัตราส่วนนี้เกิดจากการนำรายได้สุทธิของบริษัทหารด้วยส่วนของผู้ถือหุ้น แสดงให้เห็นถึงกำไรที่บริษัทสามารถทำได้จากเงินของผู้ถือหุ้น ยกตัวอย่างให้เห็นภาพ คือ เงินของผู้ลงทุนหนึ่งบาท บริษัทสามารถสร้างกำไรได้กี่บาท ถ้าหากบริษัทใดมีอัตราผลตอบแทนผู้ถือหุ้นสม่ำเสมอแสดงว่าบริษัทนั้นมีความสามารถในการแข่งขันกับบริษัทอื่น แต่ว่าบางบริษัทก็อาจมีอัตราผลตอบแทนผู้ถือหุ้นที่สูงได้เช่นกัน หากบริษัทมีการกู้ยืมเงินไปใช้ในการลงทุนมากขึ้น ดังนั้น หากบริษัทมีอัตราผลตอบแทนผู้ถือหุ้นสูงจากการเงินที่กู้ยืมมาในระดับที่สมเหตุสมผลถือว่าเป็นสิ่งที่ดี

2.1.3.4 อัตราผลตอบแทนจากสินทรัพย์ (Return on asset : ROA)

อัตราส่วนนี้เกิดจากการนำรายได้สุทธิของบริษัทหารด้วยสินทรัพย์ของบริษัท แสดงให้เห็นถึงกำไรที่บริษัทสามารถทำได้จากเงินที่ลงทุนไป ประกอบไปด้วยหนี้สินและส่วนของผู้ถือหุ้น อัตราส่วนนี้คล้ายกับอัตราส่วนก่อนหน้าแต่มีมุมมองที่มีต่อบริษัทกว้างกว่า ยกตัวอย่างให้เห็นภาพ คือ สินทรัพย์หนึ่งบาท บริษัทสามารถสร้างกำไรได้กี่บาท

2.1.3.5 อัตรากำไรสุทธิ (Net profit margin)

อัตราส่วนนี้เกิดจากการนำรายได้สุทธิของบริษัทหารด้วยรายได้ แสดงให้เห็นถึงความสามารถในการแข่งขันกับบริษัทอื่น เนื่องจากบางบริษัทพยายามรักษาหรือเพิ่มกำไรโดยการตัดราคาของสินค้าเพื่อให้สามารถแข่งขันกับบริษัทคู่แข่งได้ ซึ่งมันได้ผลแค่ในช่วงแรกเท่านั้น เพราะเมื่อมาถึงจุด ๆ หนึ่ง บริษัทไม่สามารถตัดราคาสินค้าลงได้อีกและในท้ายที่สุดกำไรของบริษัทก็หดตัวลง อีกทั้งอัตราส่วนนี้สามารถทำให้เห็นถึงความสามารถในการจัดการกับต้นทุนของบริษัท ถ้าหากบริษัทจัดการกับต้นทุนได้ดีอัตรากำไรสุทธิก็มีค่าเพิ่มสูงขึ้น แต่ถ้าบริษัทจัดการกับต้นทุนได้แย่ บริษัทก็มีอัตรากำไรสุทธิที่ต่ำ เนื่องจากต้นทุนที่สูง ส่งผลให้บริษัทสามารถสร้างกำไรที่สูงได้ยาก

จากอัตราส่วนที่กล่าวมาข้างต้นเป็นเพียงส่วนหนึ่งที่ใช้ในการเปรียบเทียบ ผู้ลงทุนสามารถสร้างอัตราส่วนเพิ่มเติมตามความสนใจ และมุมมองที่ผู้ลงทุนอยากเห็นจากบริษัทนั้น ๆ

2.1.4 การวิเคราะห์เชิงเทคนิค (Technical analysis)

การลงทุนทั้งในตราสารทุน ตราสารอนุพันธ์ หรือสกุลเงินดิจิทัล เราปฏิเสธไม่ได้ว่าการวิเคราะห์เชิงเทคนิคเข้ามามีส่วนเกี่ยวข้องในการพยากรณ์ตลาดในอนาคต โดยพื้นฐานของการวิเคราะห์เชิงเทคนิค คือการอนุมานถึงการเคลื่อนไหวของตลาดที่เกิดขึ้นในอนาคตอย่างสมเหตุสมผล แต่ไม่สามารถยืนยันได้ว่าสิ่งที่อนุมานมานั้นเกิดขึ้นจริงหรือตรงกันข้ามกับสิ่งที่เกิดในอนาคต

การวิเคราะห์เชิงเทคนิคเป็นการนำข้อมูลในอดีตมาศึกษาและวิเคราะห์ เช่น ราคาย้อนหลัง ปริมาณการซื้อขาย สัญญาซื้อขายล่วงหน้าคงค้าง เป็นต้น เพื่อพยากรณ์แนวโน้มตลาดในอนาคต อีกทั้งยังนำรูปแบบกราฟในอดีตและอินดิเคเตอร์ที่ถูกสร้างมาจากข้อมูลย้อนหลังมาพิจารณาร่วมด้วย และในปัจจุบันอินดิเคเตอร์มีให้เลือกใช้ได้อย่างหลากหลาย ดังนั้น การเลือกใช้อินดิเคเตอร์ขึ้นอยู่กับดุลยพินิจของผู้วิเคราะห์หรือผู้ลงทุน

2.1.4.1 ปริมาณการซื้อขาย

ปริมาณการซื้อขายถือเป็นปัจจัยหนึ่งที่ส่งผลต่อพฤติกรรมของราคา อันเนื่องมาจากกฎของอุปสงค์และอุปทาน เพราะจำนวนหุ้นของแต่ละบริษัทที่มีการซื้อขายอยู่ในตลาดหลักทรัพย์ไม่ได้เป็นจำนวนหุ้นทั้งหมดของบริษัทที่มีอยู่ แต่เป็นเพียงหุ้นบางส่วนที่เจ้าของของบริษัทดั้งเดิมเสนอขายออกมาเพื่อเป็นการเพิ่มทุนให้กับบริษัทด้วยเหตุนี้ การที่หุ้นที่มีการซื้อขายอยู่มีจำนวนจำกัด หากอุปสงค์หรืออุปทานเปลี่ยนไปก็ส่งผลให้ราคาเคลื่อนไหวได้เช่นกัน

ความสัมพันธ์ระหว่างราคาและปริมาณการซื้อขายมีพฤติกรรมดังนี้

1. ในช่วงที่ตลาดมีแนวโน้มสูงขึ้น สามารถสังเกตเห็นได้ว่าปริมาณการซื้อขายเพิ่มสูงขึ้น เนื่องจากอุปสงค์เพิ่มสูงขึ้น หรือในช่วงที่กราฟราคามีการกลับตัวขึ้นก็มีปริมาณการซื้อขายที่เพิ่มสูงขึ้นเช่นกัน เนื่องจากอุปสงค์สูงขึ้นหรืออุปทานลดลง
2. ในช่วงที่ตลาดมีแนวโน้มลดลง สามารถสังเกตเห็นได้ว่าปริมาณการซื้อขายลดลง เนื่องจากอุปสงค์ลดลง หรือในช่วงที่กราฟราคามีการกลับตัวลงก็มีปริมาณการซื้อขายที่ลดลงเช่นกัน เนื่องจากอุปสงค์ลดลงหรืออุปทานสูงขึ้น

3. ในช่วงตลาดที่มีการเคลื่อนไหวขึ้นลงในช่วงแคบ ๆ ปริมาณการซื้อขายลดลง เนื่องจากทิศทางของตลาดไม่ชัดเจน ส่งผลให้นักลงทุนส่วนมากยังคงเฝ้ามองทิศทางของตลาดและรอจนกว่าตลาดมีทิศทางไปในทิศทางใดทิศทางหนึ่ง

ด้วยความสัมพันธ์ดังกล่าว ผู้ลงทุนสามารถนำไปประยุกต์ใช้ในการพยากรณ์แนวโน้มของตลาดในอนาคต หรือหาสาเหตุกับสิ่งที่เกิดขึ้นในอดีตได้

ในงานวิจัยนี้ ผู้วิจัยสนใจและมุ่งเน้นไปที่อินดิเคเตอร์ตัวที่มีลักษณะสะท้อนมาจากราคาของตลาดโดยตรง อีกทั้งยังมีโครงสร้างที่ไม่ซับซ้อน และง่ายต่อการทำความเข้าใจ ซึ่งก็คือค่าเฉลี่ยเคลื่อนที่ (Moving Average)

2.1.4.2 ค่าเฉลี่ยเคลื่อนที่ (Moving average)

ค่าเฉลี่ยเคลื่อนที่ที่ถูกสร้างมาเพื่อลดความแปรปรวนหรือลดทอนสัญญาณรบกวนในตลาดที่มีประสิทธิภาพ (John J. Murphy, 1999) โดยงานวิจัยของคุณสุพร พรชัย และ ผศ. ดร. วรณรพี บานชื่นวิจิตร (2013) ในหัวข้อ “ความมีประสิทธิภาพของตลาดหลักทรัพย์แห่งประเทศไทย: กรณีศึกษาบริษัท ไทยรุ่งยูเนี่ยนคาร์ จำกัด (มหาชน)” ได้กล่าวไว้ว่า ตลาดที่มีประสิทธิภาพคือตลาดที่มีผู้ลงทุนจำนวนมาก และผู้ลงทุนแต่ละคนได้รับข่าวสารอย่างครบถ้วนสมบูรณ์ ไร้ซึ่งข้อจำกัด

เนื่องจากสภาวะตลาดในบางครั้งมีปัจจัยภายนอกเช่น อารมณ์ของผู้ลงทุน ซึ่งส่งผลให้ราคาเกิดการขยับหรือกระโดดจากราคาก่อนหน้ามาก และกลับมาในระดับใกล้เคียงหรือเท่าเดิมในภายหลัง ดังนั้นการนำค่าเฉลี่ยเคลื่อนที่เข้ามามีส่วนในการวิเคราะห์เชิงเทคนิค ส่งผลให้เราสามารถมองเห็นภาพรวมและแนวโน้มของตลาดได้ชัดเจนขึ้น

หลักการคิดพื้นฐานของการหาค่าเฉลี่ยคือ นำข้อมูลในอดีตที่ใกล้ปัจจุบันมากที่สุด ในที่นี้เราใช้ราคาปิดย้อนหลังในการหาค่าเฉลี่ย นำราคาปิดย้อนหลังมาหารด้วยจำนวนของชุดข้อมูลหรือระยะเวลาย้อนหลังของราคาปิด ทำให้เราได้ค่าเฉลี่ยค่าหนึ่ง เมื่อเวลาผ่านไปทำให้ราคาปิดที่มาใช้หาคำนวณเปลี่ยนไปและส่งผลให้เกิดค่าเฉลี่ยค่าใหม่ด้วย

ในงานวิจัยนี้ ผู้วิจัยเลือกค่าเฉลี่ยเคลื่อนที่ทั้งหมด 3 แบบใช้ร่วมกับโครงข่ายประสาทเทียมแบบ LSTM การเลือกตั้งกล่าวเกิดจากความสนใจของผู้วิจัยเอง ประกอบไปด้วย

1. Sample Moving Average (SMA)

หลักการคิดของ Sample moving average (SMA) เป็นไปตามดังที่กล่าวไว้ในข้างต้น โดยทุกราคาที่ใช้หาค่าเฉลี่ยถูกให้น้ำหนักเท่ากันหมด ดังสมการที่ (2.1)

$$SMA_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.1)$$

โดย SMA_n คือ Sample moving average ที่ใช้ราคาย้อนหลัง n วัน และ X_i เป็นราคาปิดย้อนหลัง ณ เวลาที่ i

2. Weight Moving Average (WMA)

หลักการคิดของ Weight moving average (WMA) ถูกพัฒนามาจาก Sample moving average (SMA) ซึ่งเกิดจากมุมมองที่ว่า SMA ให้สัญญาณช้าเกินไป และไม่สมเหตุผลในการให้น้ำหนักของราคาปิดย้อนหลังทุกราคาเท่ากัน เนื่องจากราคาปิดในแต่ละวันมีนัยสำคัญไม่เท่ากันตามความเข้าใกล้ปัจจุบัน ราคาปิดใกล้ปัจจุบันมากที่สุดควรให้น้ำหนักมากที่สุด และลดหลั่นกันไปตามความห่างจากปัจจุบันของราคาปิด ดังสมการที่ (2.2)

$$WMA_n(X)_i = \sum_{l=1}^n \frac{X_{i-l+1}}{\frac{n^2-n}{2+n}} (n-l+1) \quad (2.2)$$

โดย WMA_n คือ Weight moving average ที่ใช้ราคาย้อนหลัง n วัน และ X_i เป็นราคาปิดย้อนหลัง ณ เวลาที่ i

3. Exponential Moving Average (EMA)

หลักการคิดของ Exponential moving average (EMA) เป็นการให้น้ำหนักราคาปิดย้อนหลังแตกต่างกัน ขึ้นอยู่กับความเป็นปัจจุบันของราคาปิด คล้ายกับ WMA แต่แตกต่างกันตรงที่ราคาปิดที่เข้าใกล้อดีตถูกลดน้ำหนักลงแบบทวีคูณ ดังสมการที่ (2.3)

$$EMA_n(X)_i = \alpha X_i + (1 - \alpha)EMA_n(X)_{i-1} \quad (2.3)$$

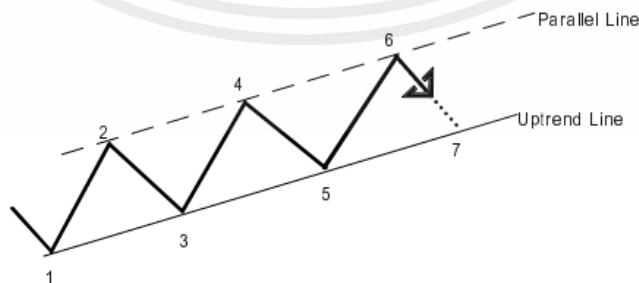
โดย EMA_n คือ Exponential moving average ที่ใช้ราคาปิดย้อนหลัง n วัน และ X_i เป็นราคาปิดย้อนหลัง ณ เวลาที่ i และค่าที่เหมาะสมที่สุดสำหรับ α เท่ากับ $\frac{2}{n+1}$ (Jack K. Hutson, 1984)

ด้วยลักษณะแนวคิดที่แตกต่างกันเช่นนี้ ส่งผลให้เส้นค่าเฉลี่ยที่เกิดขึ้นมีความไวในการแสดงผลที่แตกต่างกัน

2.1.4.3 เส้นแนวโน้ม (Trend line)

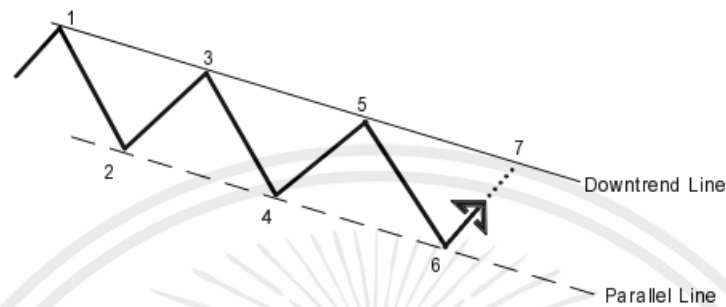
เส้นแนวโน้มคือการลากเส้นตรงเพื่อใช้เป็นจุดอ้างอิงสำหรับการวิเคราะห์การเคลื่อนไหวของราคา โดยเส้นแนวโน้มแบ่งออกเป็น 3 ประเภท เส้นแนวโน้มขาขึ้น (Uptrend line) เส้นแนวโน้มขาลง (Downtrend line) และเส้นแนวโน้มออกข้าง (Sideways)

เส้นแนวโน้มขาขึ้น (Uptrend line) คือการลากเส้นตรงโดยอ้างอิงจากจุดราคาต่ำสุดที่เคลื่อนที่สูงขึ้น มีลักษณะความชันเป็นบวกเมื่อเทียบกับเวลาแสดงถึงการรักษาแนวโน้มของราคาที่กำลังเคลื่อนที่สูงขึ้น ดังรูปที่ 2.1



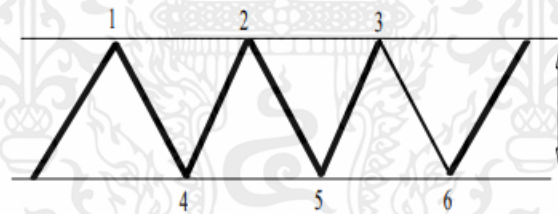
รูปที่ 2.1 เส้นแนวโน้มขาขึ้น

เส้นแนวโน้มขาลง (Downtrend line) คือการลากเส้นตรงโดยอ้างอิงจากจุดราคาสูงสุดที่เคลื่อนตัวต่ำลง มีลักษณะความชันเป็นลบเมื่อเทียบกับเวลาแสดงถึงการรักษาแนวโน้มของราคาที่กำลังเคลื่อนที่ลง ดังรูปที่ 2.2



รูปที่ 2.2 เส้นแนวโน้มขาลง

เส้นแนวโน้มออกข้าง (Sideways) คือการเคลื่อนไหวอย่างไม่มีทิศทาง คือไม่ขึ้นอย่างเด่นชัด หรือไม่ลงอย่างเด่นชัด กรอบของราคาเรียบ จุดสูงสุดอยู่ในระดับใกล้เคียงกัน และจุดต่ำสุดอยู่ในระดับใกล้เคียงกัน ดังรูปที่ 2.3



รูปที่ 2.3 เส้นแนวโน้มออกข้าง

2.1.5 การวิเคราะห์ข้อมูล (Information analysis)

การวิเคราะห์แบบนี้เป็นการวิเคราะห์จากข้อมูลที่รวบรวมมาจากแหล่งข้อมูลสาธารณะ เช่น นิตยสาร หนังสือพิมพ์ สำนักข่าว และกระดานข่าวออนไลน์ แล้วทำการพยากรณ์แนวโน้มจากข้อมูลที่ถูกรวบรวมมา

การวิเคราะห์ดังกล่าวไม่มีประสิทธิภาพมากนัก เนื่องจากข้อมูลที่นำมาใช้ในการวิเคราะห์มีผลต่อการขับเคลื่อนของราคาช้ากว่าเมื่อเทียบกับข้อมูลที่ไม่ถูกเปิดเผยต่อสาธารณะที่มีความอ่อนไหวสูง และส่งผลให้ราคาที่มีการเคลื่อนไหวอย่างมีนัยสำคัญ

จากการไม่เปิดเผยข้อมูลต่อสาธารณะ ส่งผลให้คนในได้รับประโยชน์ทางการเงินอย่างมากเมื่อเทียบกับคนนอกที่รับรู้ช้ากว่า ดังนั้น จึงไม่ใช่เรื่องดีที่พยากรณ์แนวโน้มด้วยข้อมูลจากแหล่งสาธารณะ

2.2 การวิเคราะห์เชิงอนุกรมเวลา

“การพยากรณ์ (Forecasting) หมายถึงการคาดคะเนการทำนายการเกิดเหตุการณ์ หรือสภาพการณ์ต่าง ๆ ในอนาคต โดยการพยากรณ์ทำการศึกษาแนวโน้มและรูปแบบการเกิดเหตุการณ์ จากข้อมูลในอดีต และ/หรือ ใช้ความสามารถ ความรู้ ประสบการณ์ และดุลยพินิจของผู้พยากรณ์”

(นิภา นิรุติกุล, 2551)

การวิเคราะห์อนุกรมเวลา เป็นวิธีการทางสถิติที่ใช้ข้อมูลในอดีต เพื่อวิเคราะห์หาความสัมพันธ์ ระหว่างค่าที่สังเกตที่เป็นลำดับตามเวลาเพื่อสร้าง และใช้แบบจำลองนั้นในการพยากรณ์ค่า สังเกตในอนาคตจากความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ที่ถูกจำลองขึ้น โดยการวิเคราะห์อนุกรมเวลาโดยวิธีทางสถิติ สามารถแบ่งได้เป็น 2 กลุ่มใหญ่คือ วิธีการเชิงเส้น (Linear Approach) เช่น AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) และ กลุ่มที่ 2 คือ วิธีการแบบไม่เป็นเชิงเส้น (Non-Linear Approach) เช่น GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY (GARCH)

2.2.1 การพยากรณ์แบบดั้งเดิมโดยใช้สถิติ

1. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODELS

ข้อสันนิษฐานสำหรับแบบจำลอง ARIMA คือ ข้อมูลอนุกรมเวลาใด ๆ สามารถแสดงเป็นผลรวมขององค์ประกอบที่แตกต่างกันสองส่วน คือ ส่วนที่ชี้เฉพาะ และส่วนที่เป็นแบบสุ่ม รวมกันเป็นฟังก์ชันของเวลา โดยสัญญาณอย่างสุ่มถูกเพิ่มเข้าไปในสัญญาณชี้เฉพาะ ทำให้เกิดพฤติกรรมการสุ่มของการเคลื่อนไหวของข้อมูล

อนุกรมเวลา โดย แบบจำลอง ARIMA พยายามกำจัดสัญญาณรบกวนออกจากข้อมูลอนุกรมเวลา เพื่อลดข้อมูล ความผิดพลาด

แบบจำลองการถดถอยอัตโนมัติ (Autoregressive Model : AR) เป็นตัวแปรตามที่ขึ้นกับความล่าช้า AR คือส่วนของอนุกรมเวลา Y_t ซึ่งมีค่าที่ขึ้นกับการจัดกลุ่มเชิงเส้นของค่าก่อนหน้า ซึ่งกำหนดค่าสูงสุดความล่าช้า (P) และยังมีค่าความผิดพลาดโดยพลการ ε_t ดังสมการต่อไปนี้

$$\text{First-order AR: } \hat{Y}_t = \alpha + b_1 Y_{t-1} \quad (2.4)$$

$$\text{Second-order AR: } \hat{Y}_t = \alpha + b_1 Y_{t-1} + b_2 Y_{t-2} \quad (2.5)$$

$$\text{Third-order AR: } \hat{Y}_t = \alpha + b_1 Y_{t-1} + b_2 Y_{t-2} + b_3 Y_{t-3} \quad (2.6)$$

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t \quad (2.7)$$

โดย φ_t คือค่าคงที่

ค่าเฉลี่ยเคลื่อนที่ (Moving Average : AR) คือส่วนของอนุกรมเวลา Y_t ซึ่งเป็นค่าที่สังเกตได้ในรูปของ ข้อผิดพลาดแบบสุ่มและการจัดกลุ่มเชิงเส้นของเงื่อนไขข้อผิดพลาดโดยพลการก่อนหน้านั้น ถึงความล่าช้าสูงสุดที่ อธิบายไว้ (q)

$$\text{First order MA(1): } \hat{Y}_t = \gamma + d_0 u_t + d_1 u_{t-1} \quad (2.8)$$

$$\text{Second order MA(2): } \hat{Y}_t = \gamma + d_0 u_t + d_1 u_{t-1} + d_2 u_{t-2} \quad (2.9)$$

$$\text{Third order MA(2): } \hat{Y}_t = \gamma + d_0 u_t + d_1 u_{t-1} + d_2 u_{t-2} + d_3 u_{t-3} \quad (2.10)$$

$$y_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q} \quad (2.11)$$

โดย θ_t เป็นค่าคงที่

การรวม (Integration : I) ข้อมูลอนุกรมเวลานั้นมักมีค่าไม่คงที่ และเพื่อที่จะทำให้ข้อมูล อนุกรมเวลาให้ คงที่นั้น อนุกรมนั้นต้องแตกต่าง กระบวนการนี้รู้จักกันในชื่อ Integration part (I) และอันดับของความแตกต่าง แทนด้วยเครื่องหมาย d ความแตกต่างจะกำจัดสัญญาณตามเวลาซึ่งมีแนวโน้มและ ตามช่วงเวลาดังนั้นอนุกรมนี้จะ มีการรบกวนและส่วนประกอบที่ผิดปกติ

Integral d Value

Formula(Y_t)

$$d = 0$$

$$Y_t$$

$$(2.12)$$

$$d = 1 \qquad Y_t - Y_1 \qquad (2.13)$$

$$d = 2 \qquad Y_t - 2Y_1 + Y_{t-2} \qquad (2.14)$$

2. AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY (ARCH) MODELS

เป็นโมเดลแบบจำลองทางสถิติสำหรับข้อมูลอนุกรมเวลา que อธิบายความแปรปรวนของข้อมูลความผิดพลาดในปัจจุบัน หรือนวัตกรรมที่เป็นฟังก์ชันของขนาดจริงของเงื่อนไขข้อผิดพลาดของช่วงเวลาก่อนหน้า

เนื่องจากสำหรับบางข้อมูล ค่าความแปรปรวนของค่าความผิดพลาดไม่ใช่ฟังก์ชันของตัวแปรอิสระ แต่เป็นค่าที่เปลี่ยนแปลงไปตามช่วงเวลา เช่น ข้อมูลราคาหุ้น ที่บางช่วงเวลามีความผันผวนสูง และ บางช่วงเวลามีความผันผวนต่ำ

การหาค่าเฉลี่ยและการหาความแปรปรวนของข้อมูลอนุกรมเวลาไปพร้อมกันนั้น การพยากรณ์อย่างมีเงื่อนไขมีความแม่นยำเหนือกว่า การพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) เขียนได้เป็นสมการที่ 2.15

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_t \qquad (2.15)$$

และต้องการพยากรณ์อย่างมีเงื่อนไขของ X_{t+1} ดังสมการที่ 2.5

$$E_t X_{t+1} = a_0 + a_1 X_t \qquad (2.16)$$

ถ้าเราใช้ค่าเฉลี่ยอย่างมีเงื่อนไขในการพยากรณ์ X_{t+1} ค่าความคลาดเคลื่อนของความแปรปรวน อย่างมีเงื่อนไขที่พยากรณ์ ดังสมการที่ 2.17

$$E_t [(X_{t+1} - a_0 - a_1 X_t)^2] = E_t \varepsilon_{t+1}^2 = \sigma^2 \qquad (2.17)$$

ถ้าเปลี่ยนไปใช้การพยากรณ์อย่างไม่มีเงื่อนไข ผลที่ได้เป็นค่าเฉลี่ย ในช่วง Long-Run ของลำดับ $\{X_t\}$ ซึ่งเท่ากับ $\frac{a_0}{1-a_1}$ ได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขตามสมการที่ 2.7

$$E \left[\left(X_{t+1} - \frac{a_0}{1-a_1} \right)^2 \right] = E[(\varepsilon_{t+1} + a_1\varepsilon_t + a_1^2\varepsilon_{t-1} + \dots)^2] = \frac{\sigma^2}{(1+a_1)^2} \quad (2.18)$$

เนื่องจาก $\frac{1}{(1-a_1)^2} > 1$ ดังนั้นค่าความแปรปรวนจากการพยากรณ์แบบไม่มีเงื่อนไข จึงมีค่าสูงกว่าความแปรปรวนของการพยากรณ์แบบมีเงื่อนไขในลักษณะเดียวกัน ถ้าความแปรปรวนของ $\{\varepsilon_t\}$ ไม่เป็นค่าคงที่ สามารถประมาณค่าแนวโน้มของการแปรปรวนโดยใช้ ARIMA Model โดยให้ $\{\varepsilon_t\}$ แทนส่วนที่เหลือที่ได้จากการประมาณการค่าความแปรปรวนแบบเงื่อนไขของ X_{t+1} ได้สมการที่ 2.19

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + \dots + a_q\varepsilon_{t-q}^2 + V_t = \text{White noise process} \quad (2.19)$$

ถ้าค่าของ a_1, a_2, \dots, a_q มีค่าเท่ากับ 0 ค่าความแปรปรวนที่ประมาณค่ามาได้มีค่าคงที่หรือหรือคงตัว a_0 หรืออีกนัยหนึ่งคือค่าความแปรปรวนอย่างมีเงื่อนไข ดังสมการที่ (2.20)

$$E\varepsilon_{t+1}^2 = a_0 + a_1\varepsilon_t^2 + a_2\varepsilon_{t-1}^2 + \dots + a_q\varepsilon_{t+1-q}^2 \quad (2.20)$$

สมการนี้เรียกว่า ARCH Model และสมการถัดมาเป็น ARCH(q) โดยค่า $E\varepsilon_{t+1}^2$ หรือ

σ_{t+1}^2 ประกอบด้วยองค์ประกอบ 2 ส่วน คือ ความผันผวนและค่าคงที่ในคาบเวลาที่ผ่านมาเขียนได้เป็นส่วนที่เหลือกำลังสองของคาบในอดีต ส่วนค่าสัมประสิทธิ์ (a_1, a_2, \dots, a_q) สามารถหาได้โดยใช้วิธี Maximum likelihood

3. GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY (GARCH) MODELS

แบบจำลองนี้ถูกพัฒนามาจาก ARCH เป็นแบบจำลองที่เกิดตามหลังจาก ARMA โดยนำ ARMA(p,q) มาใช้วิเคราะห์ความแปรปรวนของข้อมูลอนุกรมเวลา โดยที่ความแปรปรวนของ v_t เท่ากับ σ_1^2 ซึ่งสามารถคำนวณได้จากสมการที่ 2.21

$$\sigma_1^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^p \beta_i \sigma_{t-1}^2 = 1, \varepsilon_1 = v_1 \int \sigma_1^2 \quad (2.21)$$

เมื่อ v_t คือ White Noise process ที่เป็นค่าอิสระจากเหตุการณ์ในอดีต \mathcal{E}_{t-1}
 การหาความแปรปรวนแบบมีเงื่อนไข สามารถหาได้จากสมการที่ 2.22 โดยที่ค่าเฉลี่ยแบบมีเงื่อนไขของ \mathcal{E}_t มีค่าเท่ากับศูนย์

$$E_{t-1}\varepsilon_t^2 = \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (2.22)$$

จากที่กล่าวมาข้างต้น หากเรากำหนดให้ β_1 มีค่าเท่ากับศูนย์ เห็นได้ว่า GARCH(p,q) มีค่าเท่ากับ ARCH(q) ซึ่งเป็นคุณสมบัติที่สำคัญของแบบจำลอง GARCH ที่ค่าความแปรปรวนแบบมีเงื่อนไขของค่า X_t ถูกสร้างขึ้นจากกระบวนการของ ARMA

4. Simple Exponential Smoothing (SES) MODELS

เป็นการปรับข้อมูลให้เรียบแบบไม่มีตัวแปรของข้อมูลที่ไม่มีแนวโน้มชัดเจน โดยแนวคิดหลักคือการใช้ค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนักรวมทั้งน้ำหนักที่ลดลงแบบทวีคูณ ซึ่งถูกกำหนดให้สูงขึ้นต่อการสังเกตล่าสุด SES เป็นการทำนายระดับในอนาคตโดยให้น้ำหนักกับข้อมูลที่ใกล้เวลาปัจจุบันมากที่สุด ดังสมการที่ 2.23

$$F_{t+k} = L_t \quad (2.23)$$

โดยสมการที่ 2.13 เป็นสมการสำหรับอัปเดตค่าระดับ L_t

$$L_t = \alpha Y_t + (1 - \alpha)L_{t-1} \quad (2.24)$$

ทำให้ได้ L_t และอัปเดตระดับก่อนหน้า L_{t-1} โดยการรวมข้อมูลจากจุดข้อมูลล่าสุด Y_t นี้คือค่าเฉลี่ยถ่วงน้ำหนัก โดยที่ α และ $1-\alpha$ เป็นน้ำหนัก ค่าคงที่สำหรับการปรับเปรียบคือ α และ ค่า α มีค่าระหว่าง $0 \leq \alpha \leq 1$ โดยสำหรับข้อมูลแรก L_1 เป็นไปดังสมการที่ 2.25

$$F_1 = L_1 = Y_1 \quad (2.25)$$

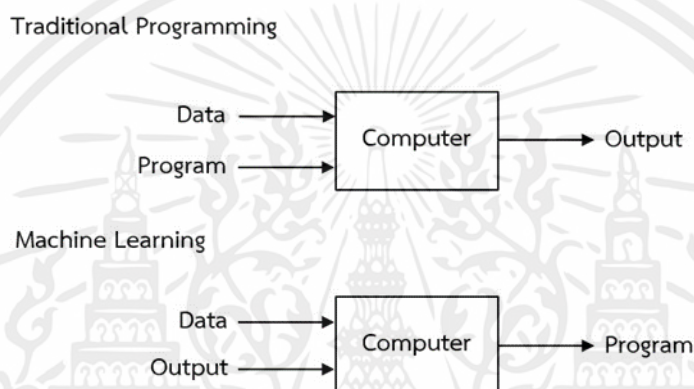
2.2.2 การพยากรณ์โดยใช้ MACHINE LEARNING

ปัจจุบัน เมื่อเทคโนโลยีเปลี่ยนไป และมี Data Science เกิดขึ้น ทำให้การทำการพยากรณ์ มีแนวทางที่ละเอียดขึ้น โดยมีการใช้ Machine Learning เข้ามาทำนาย ซึ่งเราสามารถใช้ประโยชน์จากการเรียนรู้ของเครื่อง เพื่อช่วยในการพยากรณ์ข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบต่าง ๆ โดยผู้วิจัยได้กล่าวรายละเอียดไว้ในหัวข้อถัดไป



2.3 แมชชีนเลิร์นนิง

Machine learning คือ การทำให้ระบบคอมพิวเตอร์เรียนรู้ได้ด้วยตนเอง โดยใช้ข้อมูล ซึ่งแตกต่างกับการเขียนโปรแกรมทั่วไป เพราะในส่วนของ Traditional programming ต้องใส่ข้อมูลและโปรแกรมเข้าไปเพื่อให้ได้ Output ส่วน Machine learning นั้นไม่ได้ใส่โปรแกรมคำตอบ แต่ใส่ชุดข้อมูลและผลลัพธ์เข้าไป เพื่อให้หาโปรแกรมที่นำไปทำนายคำตอบในอนาคตได้ว่าถ้าใส่ Input เข้าไปแบบนี้ Output ที่ออกมาเป็นอะไร ดังรูปที่ 2.4



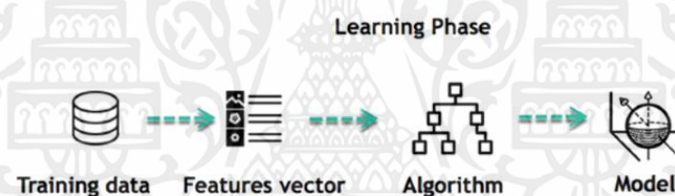
รูปที่ 2.4 TRADITIONAL PROGRAMMING & MACHINE LEARNING

แมชชีนเลิร์นนิงเรียนรู้จากสิ่งที่ส่งเข้าไปกระตุ้น แล้วจดจำเอาไว้เป็นเหมือนสมอง ส่งผลลัพธ์ออกมาเป็นตัวเลข หรือ Code สามารถเอาไปใช้งานได้หลายรูปแบบ ต้องอาศัยกลไกที่เป็นโปรแกรม หรือเรียกว่าอัลกอริทึม (Algorithm) ที่มีหลากหลายแบบ โดยมี Data scientist เป็นผู้ออกแบบ หนึ่งในอัลกอริทึมที่ได้รับความนิยมสูง คือ Deep learning ซึ่งถูกออกแบบมาให้ใช้งานได้ง่าย และประยุกต์ใช้ได้หลายลักษณะงาน อย่างไรก็ตาม ในการทำงานจริง Data scientist จำเป็นต้องออกแบบตัวแปรต่าง ๆ ทั้งในตัวของ Deep learning และต้องหาอัลกอริทึมอื่น ๆ มาเป็นคู่เปรียบเทียบกับ เพื่อมองหา Algorithm ที่เหมาะสมที่สุดในการใช้งานจริง ซึ่งตัวแมชชีนเลิร์นนิงนั้นมีการเรียนรู้เหมือนกับมนุษย์ มนุษย์เรียนรู้จากประสบการณ์ ยิ่งมีความรู้มาก ก็ยิ่งง่ายต่อการพยากรณ์ถึงสิ่งที่เกิดขึ้นในอนาคต โดยลองเปรียบเทียบเมื่อพวกเขาประสบกับเหตุการณ์ที่ไม่เคยเจอมาก่อน มีความเป็นไปได้ที่ความสำเร็จนั้นลดลงกว่าเหตุการณ์ที่เคยเจอมาแล้ว ตัวแมชชีนสามารถถูกฝึก (Train) ได้ในรูปแบบเดียวกัน เพื่อเพิ่มความแม่นยำในการพยากรณ์เหตุการณ์ต่าง ๆ ที่เกิดขึ้น แมชชีนสามารถมองเห็นตัวอย่างที่เราต้องการให้เห็นได้ เมื่อพวกเราให้ตัวอย่างที่เหมือนกันให้ตัวแมชชีนศึกษาดู มันสามารถค้นหาผลลัพธ์นั้นจนพบได้

อย่างไรก็ตามตัวแมชชีนเลิร์นนึงก็เหมือนกับมนุษย์ ถ้าตัวแมชชีนถูกสั่งให้ค้นหาในสิ่งที่ไม่เคยถูกฝึกมาก่อน มันก็ยากที่แมชชีนนั้นสามารถค้นพบสิ่งนั้นได้ จึงจำเป็นต้องมีส่วนที่ใช้เพื่อฝึกฝนและส่วนที่ทดสอบตัวแมชชีนเลิร์นนึง

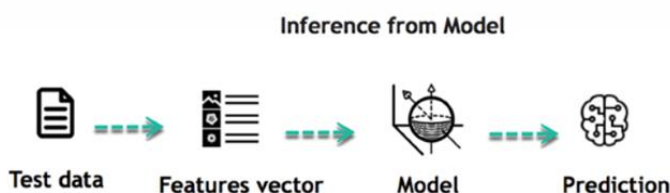
วัตถุประสงค์หลักของ Machine learning คือการเรียนรู้ (Learning) และการอนุมาน (Inference) สิ่งแรกคือ Machine เรียนรู้ผ่านการค้นพบรูปแบบหรือแบบแผนซ้ำ ๆ เดิม ๆ การค้นพบรูปแบบนี้ต้องอาศัยการมีอยู่ของข้อมูล (Data) และส่วนหนึ่งที่สำคัญคือนักวิทยาศาสตร์ข้อมูล (Data scientist) โดย Data scientist เลือกใช้ข้อมูลอย่างระมัดระวังว่าข้อมูลไหนสนับสนุนกับเครื่อง คุณสมบัติต่าง ๆ ที่ถูกใช้เพื่อแก้ปัญหาถูกเรียกว่า Feature vector

บางครั้งแมชชีนก็ใช้อัลกอริทึมแปลก ๆ เพื่อให้ง่ายต่อชีวิตจริงและตัดแปลงสิ่งที่ค้นพบให้กลายเป็นโมเดล (Model) ดังนั้น ขั้นตอนการเรียนรู้ (Learning stage) นั้นถูกใช้เพื่ออธิบายเกี่ยวกับข้อมูลและรวบรวมให้มันกลายเป็น Model ขึ้น ดังรูปที่ 2.5



รูปที่ 2.5 LEARNING PHASE

เมื่อ Model ถูกสร้างขึ้นมา มันถูกทดสอบคุณภาพการทำงานของมันกับข้อมูลที่ไม่เคยพบเจอมาก่อน โดยข้อมูลที่แปลกตาเหล่านั้นถูกแปลงให้กลายเป็น Feature Vector และถูกรวบรวมให้กลายเป็น Model และทำการพยากรณ์ออกมาได้ ซึ่งแมชชีนเลิร์นนึงไม่จำเป็นต้องเพิ่มเติมกฎใด ๆ ลงไปหรือ Train สิ่งใด ๆ เพิ่มเติมเข้าไปใน Model ผู้วิจัยสามารถใช้ Model ก่อนหน้านี้นี้มา Train เพื่อทำการอนุมานข้อมูลใหม่ ๆ ต่อไปได้ ดังรูปที่ 2.6



รูปที่ 2.6 INFERENCE FROM MODEL

ซึ่งสามารถสรุปหน้าที่ของโปรแกรม Machine learning ได้ตามขั้นตอนดังนี้

1. นิยามคำถาม (Define a question)
2. เก็บข้อมูล (Collect data)
3. แสดงผลของข้อมูล (Visualize data)
4. ฝึก/ชดเชยอัลกอริทึม (Train algorithm)
5. ทดสอบอัลกอริทึม (Test algorithm)
6. เก็บข้อมูลข้อเสนอแนะ (Collect feedback)
7. ปรับปรุงอัลกอริทึม (Refine the algorithm)
8. ทำข้อ 4.-7. วนไปจนกระทั่งได้ผลลัพธ์ที่พึงพอใจ (Loop 4-7 until the results are satisfying)
9. ใช้ model ในการพยากรณ์ (Use the model to make a prediction)

ถ้าแมชชีนเลิร์นนิงมาถึงจุดที่อัลกอริทึมนั้นดีพอจนเกิดข้อสรุปที่ถูกต้องแล้วให้ลองเอาไปประยุกต์ใช้กับข้อมูลชุดใหม่ ๆ เพื่อเป็นการทดสอบว่า Model มีความแม่นยำในการพยากรณ์ข้อมูลมากน้อยแค่ไหน

2.3.1 ประเภทต่าง ๆ ของ Machine learning

Machine learning ถูกแบ่งออกเป็นการเรียนรู้ได้ 3 ประเภทใหญ่ ๆ ได้แก่ การเรียนรู้แบบมีผู้สอน (Supervised learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) และ การเรียนรู้และเปลี่ยนไปตามสิ่งแวดล้อม (Reinforcement learning)

Supervised Learning คือ การเรียนรู้แบบมีผู้สอน เป็นกลุ่มของอัลกอริทึมที่เน้นสอนคอมพิวเตอร์โดยการศึกษาจากข้อมูลตัวอย่าง เพื่อให้คอมพิวเตอร์สามารถหาคำตอบของปัญหาได้ด้วยตัวเอง หลังจากเรียนรู้จากชุดข้อมูลตัวอย่างที่ได้ป้อนให้ไปแล้วระยะหนึ่ง อัลกอริทึมจำเป็นต้องใช้ข้อมูลในส่วนสำหรับการ Train (Training data) และส่วนที่รับกลับมาเพื่อปรับปรุง (Feedback) จากมนุษย์เพื่อเรียนรู้ความสัมพันธ์ระหว่างข้อมูลที่ถูกป้อนเข้ามาสู่ข้อมูลที่ออกไป โดยสามารถใช้ Supervised learning เมื่อผลลัพธ์ของข้อมูลเป็นสิ่งที่รู้อยู่แล้ว อัลกอริทึมนี้ก็สามารถทำนายข้อมูลใหม่ได้

โดยหลักการของ Supervised learning มีทั้งหมด 2 รูปแบบคือ

1. การจำแนกหมวดหมู่ (Classification) - เป็นการพยายามพยากรณ์คำตอบที่เป็น Discrete output หรือคำตอบที่ไม่ต่อเนื่องกัน ผลลัพธ์คือการจำแนกหมวดหมู่ มีเป้าหมายเป็นชนิดของข้อมูล เป็นการจำแนกข้อมูลออกเป็นประเภทต่าง ๆ ตามที่ Label ได้กำหนดไว้ โดย Machine learning ประเภท Classification ให้คำตอบเป็น Label / Class เท่านั้น ไม่สามารถให้คำตอบที่นอกเหนือจาก Label ในชุดฝึกฝน หรือออกมาเป็นตัวเลขที่ผ่านการคำนวณได้นั้นเอง โดย ML model สำหรับงาน Classification ที่เด่น ๆ ประกอบด้วย KNN SVM Logistic regression Decision tree เป็นต้น

2. การถดถอย (Regression) - เป็นการพยายามพยากรณ์คำตอบที่เป็น Continuous output หรือคำตอบที่ต่อเนื่องกัน เมื่อผลลัพธ์มีค่าต่อเนื่องกัน เป็นงานของ Regression ที่เข้ามาช่วยแก้ปัญหา การวิเคราะห์การถดถอย เป็นวิธีการทางสถิติใช้เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป ประกอบด้วย ตัวแปรที่เราทราบค่า เรียกว่าตัวประมาณการหรือตัวแปรต้น (Predictor, Independent variable, X) และตัวแปรที่เราต้องการทราบค่า เรียกว่าตัวตอบสนองหรือตัวแปรตาม (Response, Dependent variable, Y) ว่าเป็นตัวแปรที่เป็นปัจจัยหรือเป็นเหตุผลของกันและกันหรือไม่ยกตัวอย่างเช่น นักวิเคราะห์ด้านการเงินคนหนึ่งอาจต้องการทำนายมูลค่าของหุ้นโดยดูจาก Feature ต่าง ๆ เช่น ส่วนได้ส่วนเสีย (Equity) สถานะของหุ้นก่อนหน้านี้ และดัชนีเศรษฐกิจมหภาค ระบบถูก Train เพื่อประเมินราคาของหุ้นด้วยความผิดพลาดที่น้อยที่สุด

Unsupervised learning คือ การเรียนรู้แบบไม่มีผู้สอนโดยที่เครื่องจักรนั้น ๆ ทำการเรียนรู้สิ่งต่าง ๆ ด้วยตัวเอง เป็นการสร้างโมเดลโดยอัลกอริทึมตรวจสอบเฉพาะข้อมูลที่ป้อนเข้ามาเท่านั้น (Input) โดยปราศจากการให้ผลลัพธ์ที่เกิดขึ้นจริง (ไม่มี Target) ระบบสามารถนำอินพุตไปวิเคราะห์และสร้างแบบแผนจากข้อมูลที่ได้รับเข้าไป เช่น การสำรวจข้อมูลประชากรเพื่อหาแบบแผน (Pattern) ของข้อมูลนั้น Unsupervised learning เป็นกลุ่ม Algorithm ที่ไม่มี Label หรือการสอนอย่างชัดเจนว่าถ้าทำงานแล้วได้ผลลัพธ์แบบนี้หมายถึงถูกหรือผิดมักนำไปใช้ในการแนะนำผลิตภัณฑ์ต่าง ๆ หรือคัดเลือกข้อมูลหรือผลิตภัณฑ์ การใช้งานหลักมี 2 รูปแบบคือ

- 1. Dimensionality Reduction** คือ การลดมิติของข้อมูล เพื่อลดความซับซ้อนก่อนนำข้อมูลไปใช้ต่อหรือเพื่อแสดงผลในรูปแบบกราฟ อัลกอริทึมสำหรับงานประเภท Dimensionality reduction เช่น PCA/T-SNE
- 2. Clustering** คือ การจัดกลุ่มข้อมูลตามคุณลักษณะ เช่น การจัดกลุ่มหุ้น 6 ประเภทของปีเตอร์ ลินซ์ แบ่งกลุ่มคุณลักษณะเฉพาะของหุ้นออกเป็นทั้งหมด 6 กลุ่ม ประกอบด้วย หุ้นโตช้า หุ้นแข็งแกร่ง หุ้นโตเร็ว หุ้นวัฏจักร หุ้นฟื้นตัว และหุ้นทรัพย์สินมาก

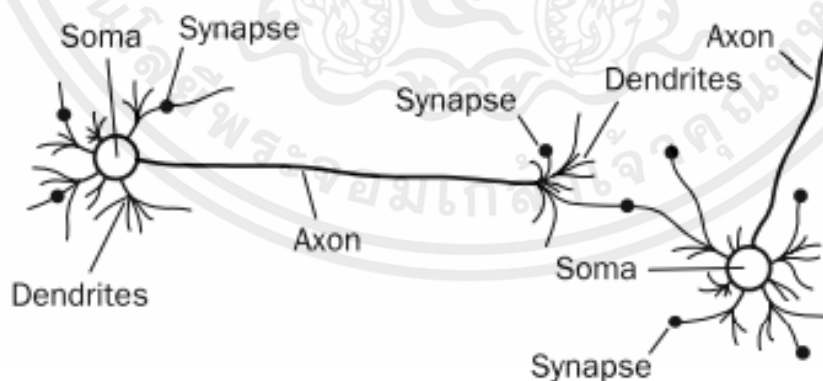
อัลกอริทึมสำหรับงานประเภท Clustering ประกอบด้วย การแบ่งกลุ่มข้อมูลแบบเคมีน (K-means clustering) แบบจำลองผสมของ Gauss (Gaussian mixture model) การแบ่งกลุ่มข้อมูลตามลำดับชั้น (Hierarchical clustering) ระบบให้คำแนะนำ (Recommender system)

Reinforcement learning คือ การเรียนรู้จากการลองผิดลองถูกจากสถานการณ์ในอดีตหรือระบบจำลอง เพื่อพยายามค้นหาแนวทางการรับมือหรือแก้ไขกับปัญหาที่พบเจอให้ดีที่สุด เพื่อพัฒนาระบบการตัดสินใจให้ดียิ่งขึ้น ซึ่งนำไปสู่ผลลัพธ์ที่มีประสิทธิภาพ Reinforcement learning เป็นวิธีการเรียนรู้แบบหนึ่ง ที่การเรียนรู้เกิดมาจากการปฏิสัมพันธ์ (Interaction) ระหว่างผู้เรียนรู้ (Agent) กับสิ่งแวดล้อม (Environment) จากนั้น Agent สามารถปฏิสัมพันธ์กับสิ่งแวดล้อมโดยใช้การกระทำบางอย่าง (Action) หลังจากที่ใช้ Action ไปแล้ว Agent สามารถได้ Reward กลับมา เป็นแนวทางการเรียนรู้ของ AI ซึ่งมีลักษณะที่เหมือนกับการเรียนรู้ของมนุษย์เพราะมนุษย์มีการเรียนรู้และเปลี่ยนไปตามสิ่งแวดล้อม Reinforcement learning ก็เป็นแบบนี้เช่นกัน มีตัวอย่างเช่น AlphaGo การพัฒนาระบบผู้เล่นอัตโนมัติให้ชนะผู้เล่นระดับโลก โดย AlphaGo ถูกสอนให้เล่นด้วยวิธี Reinforcement learning

2.4 โครงข่ายประสาทเทียม (Artificial neuron network : ANN)

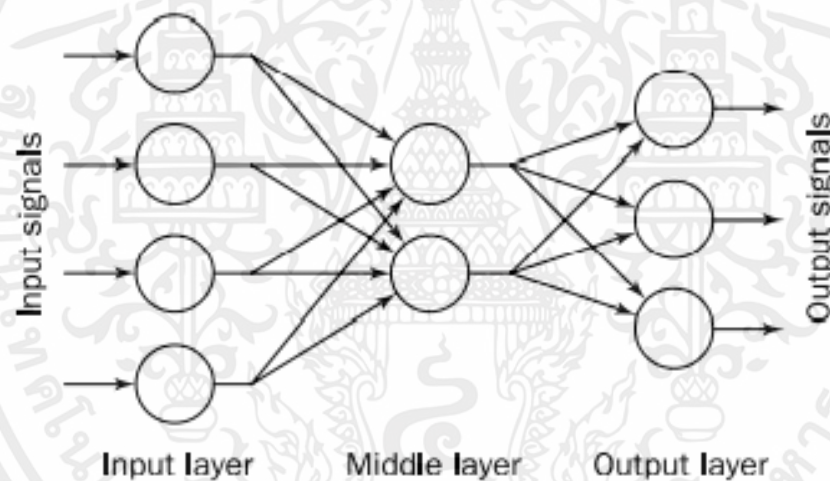
โครงข่ายประสาทเทียมเป็นศาสตร์แขนงหนึ่งของปัญญาประดิษฐ์ (Artificial intelligence : AI) มีรูปแบบโครงสร้างและการทำงานของการทำงานประมวลผลเหมือนกับสมองของสิ่งมีชีวิตที่มีการเปลี่ยนแปลงตัวเองต่อการตอบสนองของอินพุตในรูปแบบต่าง ๆ ซึ่งโครงข่ายประสาทเทียม คือ ระบบคอมพิวเตอร์จากโมเดลทางคณิตศาสตร์ เพื่อจำลองหรือลอกเลียนแบบการทำงานของโครงข่ายประสาทที่อยู่ในสมองของสิ่งมีชีวิต สามารถเรียนรู้งานที่มอบหมายได้ด้วยการเรียนรู้จากตัวอย่างโดยไม่ถูกโปรแกรมด้วยกฎเกณฑ์ตายตัว แต่รูปแบบการประมวลผลสารสนเทศมีความซับซ้อนมากกว่าโดยได้นำตัวเลขจำนวนมาก ๆ จำนวนหนึ่งมาคำนวณซึ่งกันและกันเป็นโครงข่าย ซึ่งทำให้แก้ปัญหาเฉพาะบางปัญหาได้ เหมือนสมองของมนุษย์ที่สามารถเรียนรู้ได้จากตัวอย่างและประสบการณ์ ด้วยเหตุนี้โครงข่ายประสาทเทียมจึงสามารถนำไปประยุกต์ใช้กับงานหลายด้านได้อย่างมีประสิทธิภาพ โปรแกรมประยุกต์ที่ใช้เทคโนโลยีโครงข่ายประสาทเทียม ได้แก่ งานที่เกี่ยวข้องกับการจดจำรูปแบบ (Pattern recognition) การจัดหมวดหมู่ข้อมูล (Data classification) ตลอดจนถึงกระบวนการเรียนรู้ (Learning process)

แนวคิดเริ่มต้นของเทคนิคนี้ได้มาจากการศึกษาโครงข่ายไฟฟ้าชีวภาพในสมอง (Bioelectric network) ซึ่งประกอบด้วยหน่วยประมวลผลพื้นฐานหรือเซลล์ประสาท (Nerve cells or neurons) จำนวนมากที่ทำงานร่วมกัน เซลล์ประสาทตามธรรมชาติมีส่วนประกอบพื้นฐาน 4 ส่วน เรียกตามชื่อทางชีววิทยาดังนี้ เดนไดรต์ (Dendrites) โซมา (Soma) แอ็คซอน (Axon) และไซแนปส์ (Synapses) ดังรูปที่ 2.7



รูปที่ 2.7 ความสัมพันธ์ของระบบประสาททั้ง 4 ส่วน (NEGNEVITSKY, 2005)

จากรูปที่ 2.7 เดนไดรท์เป็นส่วนที่ขยายออกมาจากโซมา มีลักษณะคล้ายเส้นขนทำหน้าที่เปรียบเสมือนเป็นช่องทางนำเข้าของค่าอินพุต เดนไดรท์ทำการรับอินพุตผ่านไซแนปส์ของเซลล์ประสาทอื่น โดยที่โซมาคอยประมวลผลสัญญาณไฟฟ้าที่รับเข้ามาอยู่ตลอดเวลา แล้วส่งผลการทำงานเป็นเอาต์พุตออกไปให้เซลล์ประสาทอื่นผ่านทางแอกซอนและไซแนปส์ หรือปลายในการส่งกระแสประสาท และจุดประสานประสาท ซึ่งการทำงานร่วมกันของส่วนประกอบเหล่านี้ ทำให้มนุษย์สามารถเรียนรู้จดจำ และระลึกถึงสิ่งที่จำไว้ได้ โดยที่รูปแบบของความทรงจำสำหรับเรื่องใดเรื่องหนึ่งที่เกิดขึ้นในสมองเกิดจากการที่โซมาตัวหนึ่งทำงานร่วมกับโซมาอีกหลายตัวซึ่งหมายความว่าความจำในแต่ละเรื่องมักเกิดจากการเปลี่ยนแปลงของโซมาหลาย ๆ ตัว และเกิดขึ้นในส่วนใดส่วนหนึ่งของสมอง โดยที่แต่ละส่วนของสมองแบ่งหน้าที่และความถนัดในเรื่องที่ต่างกันออกไป โครงสร้างของโครงข่ายประสาทเทียมโดยทั่วไปแสดงดังรูปที่ 2.8 และการทำงานของโครงข่ายประสาทในสมองเทียมแสดงดังรูปที่ 2.9

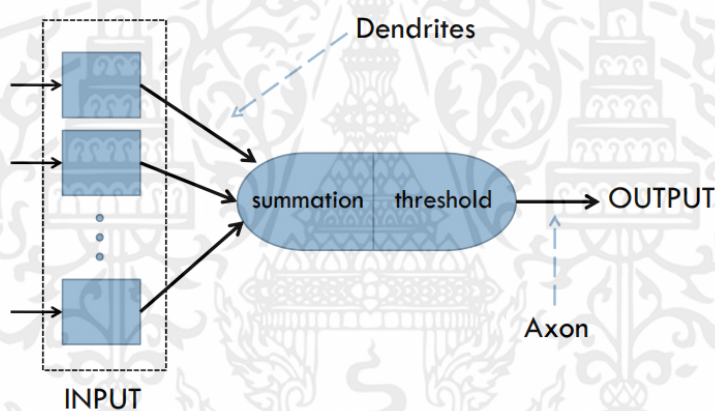


รูปที่ 2.8 โครงสร้างของโครงข่ายประสาทเทียมโดยทั่วไป (NEGNEVITSKY, 2005)

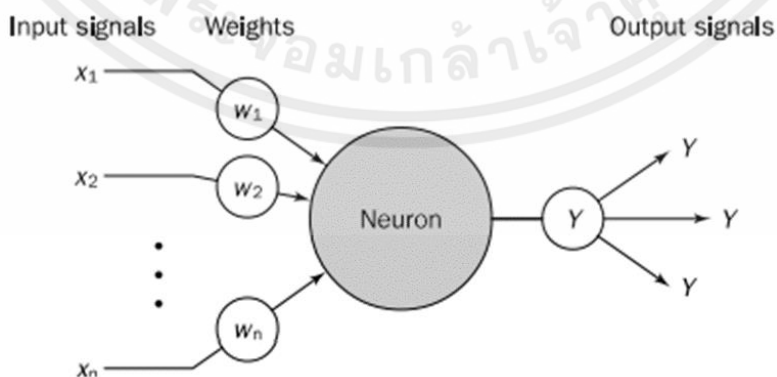
และเมื่อทำการเปรียบเทียบระหว่างโครงข่ายทางชีววิทยาและโครงข่ายประสาทเทียมได้ผลลัพธ์ตามตารางที่ 2.1

ตารางที่ 2.1 เปรียบเทียบโครงข่ายทางชีววิทยาและโครงข่ายประสาทเทียม (NEGNEVITSKY, 2005)

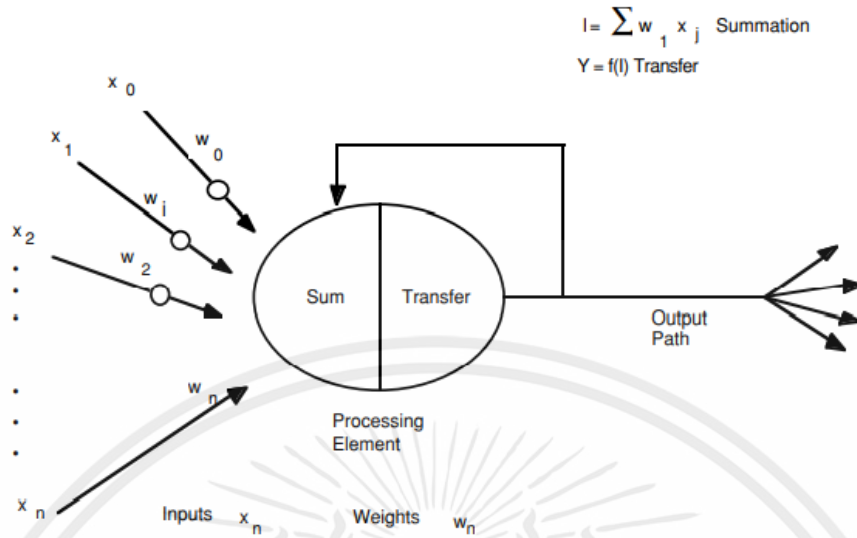
Biological neural network	Artificial neural network
Soma	Neuron
Dendrite	Input
Axon	Output
Synapse	Weight



รูปที่ 2.9 การทำงานของโครงข่ายประสาทในสมองเทียม (CHOOPAN RATTANAPOKA)



รูปที่ 2.10 แผนภาพของเซลล์ประสาท (NEGNEVITSKY, 2005)



รูปที่ 2.11 โครงข่ายประสาทเทียมอย่างง่าย (ANDERSON & MCNEILL, 1992)

จากคำอธิบายหลักการทำงานของเซลล์ประสาทจากตารางที่ 2.1 สามารถแบ่งคุณสมบัติของโครงข่ายประสาทเทียมออกเป็น 2 ประเภท คือ การเรียนรู้ (Learning) และการจดจำหรือระลึกถึง (Recall) โดยการเรียนรู้ นั้น โครงข่ายประสาทเทียมสามารถเรียนรู้ได้จากการฝึกสอน (Training) ที่ถูกป้อนเข้าไปเพื่อให้โครงข่ายประสาทเทียมเรียนรู้ ส่วนการจดจำหรือระลึกถึงนั้น โครงข่ายประสาทเทียมสามารถระลึกได้ทั้งการฝึกสอนและการทดสอบ (Testing) โดยทั่วไปแล้วข้อมูลสำหรับการฝึกสอนและการทดสอบมีความแตกต่างกัน

จากรูปที่ 2.10 แสดงให้เห็นว่าโครงข่ายประสาทเทียมโดยทั่วไปประกอบด้วยลำดับชั้นของเลเยอร์ และเซลล์ประสาทในเครือข่ายถูกจัดเรียงตามชั้นเหล่านี้ เซลล์ประสาทที่เชื่อมต่อกับสภาพแวดล้อมภายนอกสร้างชั้นอินพุต (Input) และเอาต์พุต (Output) น้ำหนัก (Weights) ได้รับการปรับเปลี่ยนแก้ไขเพื่อให้พฤติกรรมอินพุตและเอาต์พุตของเครือข่ายสอดคล้องกับสภาพแวดล้อม เซลล์ประสาทแต่ละเซลล์เป็นหน่วยประมวลผลเบื้องต้น มันมีวิธีการคำนวณระดับการเปิดใช้งานโดยพิจารณาจากอินพุตต่าง ๆ และน้ำหนักเชิงตัวเลข ในการสร้างโครงข่ายประสาทเทียมเราต้องตัดสินใจก่อนว่าใช้เซลล์ประสาทจำนวนเท่าใด และเชื่อมต่อเซลล์ประสาทเพื่อสร้างเครือข่ายอย่างไร จากนั้นเราจึงค่อยตัดสินใจว่าใช้อัลกอริทึมการเรียนรู้ใด และขั้นสุดท้ายจึงฝึกโครงข่ายประสาท

จากรูปที่ 2.11 เป็นการนำเอาหลักการของเซลล์ประสาทมาเขียนแบบจำลองใหม่ ได้โครงข่ายประสาทเทียมอย่างง่ายดังแสดงในรูป โดยแทนค่าอินพุตตั้งแต่ x_0 จนถึง x_n และแต่ละอินพุตถูกคูณด้วยน้ำหนักการเชื่อมต่อ หรือค่าน้ำหนัก (Weight) ซึ่งแทนด้วย w_0 จนถึง w_n ซึ่งในกรณีที่ง่ายที่สุดคือ $x_0w_0, x_1w_1, \dots, x_nw_n$ ถูกนำมารวมกันและส่งผ่านไปยังฟังก์ชันการถ่ายโอน (Transfer function) เพื่อสร้างผลลัพธ์ก่อนส่งออกไป

โดยโครงข่ายประสาทเทียมอย่างง่ายนี้ สามารถนำไปประยุกต์ใช้กับโครงข่ายชนิดอื่นได้ เช่น การจดจำข้อความ การค้นหาคำพูด และการแปลความหมายรูปภาพ เป็นต้น

องค์ประกอบหลักของเซลล์ประสาทเทียม

ส่วนนี้อธิบายองค์ประกอบหลักเจ็ดประการที่ประกอบกันเป็นเซลล์ประสาทเทียม ซึ่งส่วนประกอบเหล่านี้สามารถใช้ได้ไม่ว่าเซลล์ประสาทรุ่นนั้นใช้สำหรับอินพุต เอาต์พุต หรืออยู่ในเลเยอร์ที่ซ่อนอยู่อย่างใดอย่างหนึ่ง

1. ตัวประกอบการถ่วงน้ำหนัก (Weighting factors)

โดยปกติเซลล์ประสาทได้รับอินพุตจำนวนมากพร้อมกัน แต่ละอินพุตมีน้ำหนักสัมพัทธ์แตกต่างกัน กล่าวคือ มีนัยสำคัญที่ส่งผลต่อเอาต์พุตต่างกัน ลักษณะดังกล่าวเหมือนกับความแรงของไซแนปส์ที่แตกต่างกันในทางชีววิทยา บางอินพุตมีความสำคัญมากกว่าอินพุตอื่น ๆ ดังนั้นจึงมีผลต่อการประมวลผลมากกว่า

การถ่วงน้ำหนักเป็นค่าสัมประสิทธิ์แบบปรับค่าได้ภายในโครงข่ายที่ถูกกำหนดด้วยความเข้มของสัญญาณอินพุต ซึ่งถูกบันทึกโดยเซลล์ประสาทเทียม เป็นการวัดความแรงในการเชื่อมของอินพุต ความแรงสามารถปรับแก้ได้ตามความหลากหลายของชุดข้อมูลทดสอบ และโครงสร้างเฉพาะของโครงข่ายผ่านกฎการเรียนรู้

2. ฟังก์ชันผลรวม (Summation function)

ขั้นตอนแรกในการดำเนินการประมวลผล คือคำนวณผลรวมน้ำหนักของอินพุตทั้งหมด ในทางคณิตศาสตร์อินพุตและน้ำหนักที่สอดคล้องกับอินพุตเป็นเวกเตอร์ที่สามารถเขียนในรูป (x_0, x_1, \dots, x_n) และ (w_0, w_1, \dots, w_n) สัญญาณอินพุตทั้งหมดคือผลคูณภายในของทั้ง 2 เวกเตอร์ ฟังก์ชันผลรวมอย่างง่ายนี้ ถูกค้นพบจากการคูณแต่ละองค์ประกอบของเวกเตอร์ x กับองค์ประกอบของเวกเตอร์ w ได้ $Input_0 = x_0w_0, Input_1 = x_1w_1, \dots, Input_n = x_nw_n$ แล้วบวกผลคูณทั้งหมดเข้าด้วยกันได้ $Input_0 + Input_1 + \dots + Input_n$ ผลลัพธ์ที่ได้เป็นเลขตัวเดียว

ในเชิงเรขาคณิต ผลคูณภายในของ 2 เวกเตอร์สามารถพิจารณาเป็นตัววัดความคล้ายคลึงได้ ถ้าเวกเตอร์ชี้ไปในทิศทางเดียวกันส่งผลให้ผลคูณภายในมีค่าสูงสุด แต่ถ้าเวกเตอร์ชี้ไปในทิศทางตรงข้ามกันส่งผลให้ผลคูณภายในมีค่าต่ำสุด ฟังก์ชันผลรวมสามารถทำให้ซับซ้อนขึ้นได้ อินพุตและค่าสัมประสิทธิ์การถ่วงน้ำหนักสามารถรวมกันได้หลายวิธีก่อนส่งต่อไปยังฟังก์ชันถ่ายโอน 23 ฟังก์ชัน นอกเหนือจากการรวมผลลัพธ์อย่างง่าย ฟังก์ชันผลรวมสามารถเลือกค่าต่ำสุด ค่าสูงสุด ฐานนิยม ผลลัพธ์ และ Normalizing algorithms

ฟังก์ชันผลรวมสามารถมีกระบวนการเพิ่มเติมที่ประยุกต์ใช้กับผลลัพธ์ก่อนส่งผ่านไปยังฟังก์ชันการถ่ายโอน กระบวนการนี้ถูกเรียกว่า ฟังก์ชันการกระตุ้นการทำงาน จุดประสงค์ของการกระตุ้นการทำงานคือต้องการให้ผลรวมเปลี่ยนแปลงตามเวลา นอกจากนี้ ฟังก์ชันดังกล่าวมีแนวโน้มเป็นองค์ประกอบของโครงข่ายแทนการเป็นองค์ประกอบของการประมวล

3. ฟังก์ชันการถ่ายโอน (Transfer Function)

ผลลัพธ์ของฟังก์ชันผลรวมถูกแปลงเป็นเอาต์พุตผ่านกระบวนการอัลกอริทึมที่เรียกว่าฟังก์ชันการถ่ายโอน ในฟังก์ชันการถ่ายโอน ผลรวมทั้งหมดถูกเปรียบเทียบกับเกณฑ์เพื่อกำหนดเอาต์พุตของเซลล์ประสาท ถ้าผลรวมมีค่ามากกว่าเกณฑ์ส่งผลให้เกิดสัญญาณ แต่ถ้าผลรวมน้อยกว่าเกณฑ์ส่งผลให้ไม่เกิดสัญญาณ

เกณฑ์หรือฟังก์ชันการถ่ายโอนไม่อยู่ในรูปเชิงเส้น เนื่องจากเอาต์พุตเป็นสัดส่วนอย่างง่ายกับอินพุต ส่งผลให้การใช้ฟังก์ชันเชิงเส้นไม่เกิดประโยชน์ ซึ่งเป็นปัญหาเริ่มแรกของโมเดลโครงข่าย

ฟังก์ชันการถ่ายโอนเป็นสิ่งที่ง่ายที่ขึ้นอยู่กับค่าของผลรวมว่าเป็นบวกหรือลบ โครงข่ายสามารถเอาต์พุตออกมาเป็น 0 และ 1 หรือ 1 และ -1 หรือชุดค่าผสมตัวเลข กล่าวได้ว่าฟังก์ชันการถ่ายโอนเป็น Hard Limiter หรือฟังก์ชันขั้นบันได

Ramping Function สามารถสะท้อนอินพุตภายในช่วงที่กำหนดและทำตัวเป็น Hard limiter นอกช่วงที่กำหนด มันเป็นฟังก์ชันเชิงเส้นที่ถูกตัดให้เหลือค่าต่ำสุดและค่าสูงสุดที่เส้นกำกับ

Sigmoid หรือเส้นโค้งรูปตัวเอส เป็นเส้นโค้งที่เข้าใกล้ค่าต่ำสุดและค่าสูงสุดที่เส้นกำกับ มันถูกเรียกว่า Sigmoid เมื่ออยู่ในช่วง 0 ถึง 1 และถูกเรียกว่า Hyperbolic Tangent เมื่ออยู่ในช่วง -1 ถึง 1 ในทางคณิตศาสตร์ เส้นโค้งเหล่านี้นับว่าเป็นทั้งฟังก์ชันและอนุพันธ์ที่ต่อเนื่องกัน

ก่อนการใช้ฟังก์ชันการถ่ายโอนมีการเพิ่มสัญญาณรบกวนแบบสุ่มที่กระจายอย่างสม่ำเสมอ แพล่งที่มาและปริมาณของสัญญาณรบกวนถูกกำหนดด้วยโหมดการเรียนรู้ของแบบโครงข่ายที่กำหนดไว้ในทางชีววิทยา สัญญาณรบกวนเหล่านี้คือ อุณหภูมิ ด้วยอุณหภูมิที่สูงหรือต่ำเกินไป ส่งผลให้ความสามารถในการคิดของมนุษย์ได้รับผลกระทบ กระบวนการเหล่านี้ถูกจำลองโดยการเพิ่มสัญญาณรบกวนในระดับต่าง ๆ ในผลลัพธ์ของผลรวม ส่งผลให้ฟังก์ชันการถ่ายโอนมีความคล้ายคลึงกับสมองมากขึ้น

4. การปรับขนาดและการจำกัด (Scaling and limiting)

หลังจากกระบวนการฟังก์ชันการถ่ายโอน ผลลัพธ์สามารถถูกส่งผ่านไปยังกระบวนการเพิ่มเติมที่มีการปรับขนาดและมีการจำกัด ขนาดเกิดจากการคูณกันระหว่างตัวประกอบมาตราส่วนกับค่าการถ่ายโอน และบวกค่าออฟเซตเพิ่ม การจำกัดเป็นกลไกที่ใช้ประกันว่าผลลัพธ์ของขนาดไม่เกินขอบเขตบนและล่างการปรับขนาดและการจำกัดถูกใช้ในโทโพโลยีเพื่อทดสอบแบบจำลองประสาททางชีววิทยา

5. ฟังก์ชันเอาต์พุต (Output function)

ในแต่ละส่วนของการประมวลผลได้รับอนุญาตให้ส่งสัญญาณเพียงสัญญาณเดียว แต่สามารถส่งไปยังเซลล์ประสาทอื่น ๆ ได้หลายแสนเซลล์ ลักษณะเช่นนี้เหมือนกับเซลล์ประสาทในทางชีววิทยาที่มีอินพุตจำนวนมากแต่มีเอาต์พุตเพียงหนึ่งเดียว เอาต์พุตเปรียบได้กับผลลัพธ์ของฟังก์ชันการถ่ายโอน อย่างไรก็ตาม มีโครงข่ายโทโพโลยีที่ปรับเปลี่ยนผลลัพธ์ของการถ่ายโอน เพื่อรวบรวมการแข่งขันขององค์ประกอบการประมวลผลที่อยู่ใกล้เคียงกัน เซลล์ประสาทได้รับอนุญาตให้แข่งขันซึ่งกันและกัน การแข่งขันสามารถเกิดขึ้นในชั้นหนึ่ง หรือทั้งสองชั้น ชั้นแรก การแข่งขันเป็นตัวกำหนดว่าประสาทเทียมตัวใดได้ทำงานหรือให้ผลลัพธ์ออกมา ชั้นต่อมา การแข่งขันเป็นตัวกำหนดว่าอินพุตใดมีส่วนในกระบวนการเรียนรู้หรือการปรับตัว

6. Error function and back-propagated value

ในโครงข่ายการเรียนรู้มีการคำนวณความแตกต่างระหว่างเอาต์พุตล่าสุดกับเอาต์พุตที่ต้องการ ข้อผิดพลาดดิบนี้ถูกแปลงโดยฟังก์ชันข้อผิดพลาดที่ตรงกับสถาปัตยกรรมโครงข่ายโดยเฉพาะ สถาปัตยกรรมพื้นฐานใช้ข้อผิดพลาดนี้โดยตรง ความผิดพลาดของประสาทเทียมถูกแพร่กระจายไปยังฟังก์ชันการเรียนรู้ขององค์ประกอบการประมวลผลอื่น ๆ ในส่วนของความผิดพลาดนี้ถูกเรียกว่าความผิดพลาดล่าสุด

7. ฟังก์ชันการเรียนรู้ (Learning function)

วัตถุประสงค์ของฟังก์ชันการเรียนรู้คือการปรับเปลี่ยนน้ำหนักการเชื่อมต่อตัวแปรบนอินพุตขององค์ประกอบการประมวลผลแต่ละรายการตามอัลกอริทึมที่อิงกับระบบประสาท โดยกระบวนการนี้เปลี่ยนน้ำหนักของการเชื่อมต่ออินพุตเพื่อให้ได้ผลลัพธ์ที่ต้องการ เรียกอีกอย่างว่าฟังก์ชันการปรับตัว

การเรียนรู้ของโครงข่ายประสาทเทียม

การเรียนรู้ของโครงข่ายประสาทเทียมเกิดจากเซลล์ประสาทเชื่อมต่อกันด้วยลิงก์ และแต่ละลิงก์มีน้ำหนักเป็นตัวเลขที่เกี่ยวข้องกัน ซึ่งน้ำหนักเป็นวิธีการพื้นฐานของหน่วยความจำระยะยาวใน ANNs แสดงความแข็งแกร่งหรือกล่าวอีกนัยหนึ่งได้ถึงความสำคัญของการป้อนข้อมูลของเซลล์ประสาทแต่ละเซลล์ โครงข่ายประสาทเรียนรู้ผ่านการปรับน้ำหนักเหล่านี้ซ้ำ ๆ ซึ่งการเรียนรู้ทั้งหมดนี้เกิดจากการคำนวณระหว่างค่าอินพุต และค่าน้ำหนักที่ถูกปรับแก้ด้วยกระบวนการทำซ้ำนั่นเอง

สามารถแบ่งรูปแบบการเรียนรู้ของโครงข่ายประสาทเทียมออกเป็น 3 ประเภท ได้ดังนี้

1. Supervised learning

โครงข่ายประสาทเทียมประเภทนี้ได้รับการฝึกภายใต้การดูแลแบบมีผู้สอน โดยฝึกจากข้อมูลอินพุตและเอาต์พุตเป้าหมาย โดยทำการเปรียบเทียบเอาต์พุตจริงของโครงข่ายประสาทเทียมกับเอาต์พุตที่ต้องการ ซึ่งขณะที่ฝึกนั้นได้เอาต์พุตจริงที่มีความแตกต่างจากเอาต์พุตที่ต้องการ หรือก็คือค่าความคลาดเคลื่อน ส่งผลให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ของข้อมูลโดยทำการปรับค่าน้ำหนักและค่าไบแอส (Bias) เพื่อให้ได้ผลลัพธ์ที่แม่นยำหรือคลาดเคลื่อนน้อยที่สุด

2. Unsupervised learning

โครงข่ายประสาทเทียมประเภทนี้เป็นการเรียนรู้แบบไม่มีผู้สอนหรือเป็นการเรียนรู้ด้วยตนเอง เกิดการเรียนรู้ของข้อมูลโดยการฝึกจากข้อมูลอินพุตเท่านั้น ซึ่งทำการจัดแบ่งกลุ่มข้อมูลให้เป็นระดับต่างๆ โดยใช้หลักทางสถิติสำหรับการเรียนรู้ในการปรับค่าน้ำหนักและค่าไบแอส (Bias) เพื่อหาผลลัพธ์หรือเอาต์พุตออกมา

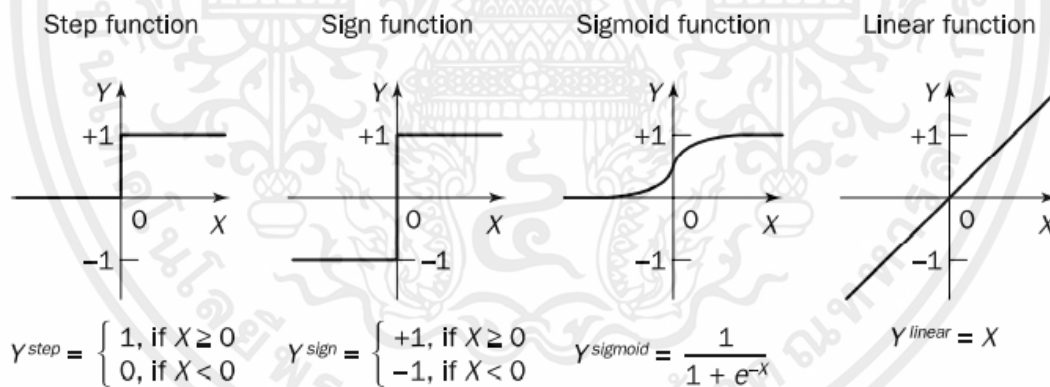
3. Reinforcement learning

เป็นการเรียนรู้แบบผสมผสานทั้ง Supervised learning และ Unsupervised learning ซึ่งมีความซับซ้อนมากกว่าและใช้เวลาในการฝึกมากกว่าการเรียนรู้สองประเภทข้างต้น จึงทำให้ยังไม่เป็นที่นิยม ถึงแม้การเรียนรู้ประเภทนี้ได้ผลลัพธ์หรือเอาต์พุตที่ดีที่สุดจากการเรียนรู้ด้วยตัวเอง แต่ก็ไม่ได้ดีที่สุดสำหรับผู้ใช้งาน เพราะตัวโครงข่ายประสาทเทียมไม่ทราบว่าจะดีที่สุดคือเท่าใด

The perceptron

จุดเริ่มต้นของ Perceptron เกิดขึ้นในปี 1958 Frank Rosenblatt ได้แนะนำอัลกอริทึมการฝึกอบรมที่ให้ขั้นตอนแรกสำหรับการฝึก ANN อย่างง่าย: a perceptron (Rosenblatt, 1958) เพอร์เซ็ปตรอนเป็นรูปแบบที่ง่ายที่สุดของโครงข่ายประสาทเทียม ประกอบด้วยเซลล์ประสาทเซลล์เดียวที่ปรับน้ำหนักไซแนปส์ได้ ตัวรวมเชิงเส้นและฮาร์ดลิมิตเตอร์ (Hard Limiter) ที่เป็นฟังก์ชันชนิดหนึ่งที่กระตุ้นการทำงานของเซลล์ประสาท มีลักษณะคล้ายกับ Step Function และ Sign Function โดยผลรวมถ่วงน้ำหนักของอินพุตถูกนำไปใช้กับฮาร์ดลิมิตเตอร์ ซึ่งถ้าอินพุตเป็นบวกสามารถสร้างเอาต์พุตเท่ากับ +1 แต่ถ้าหากเป็นลบสามารถสร้างเอาต์พุตเท่ากับ -1 มักใช้ในเซลล์ประสาทการตัดสินใจสำหรับการจำแนกประเภท และการจัดจำรูปแบบงาน

โดยฟังก์ชันที่กระตุ้นการทำงานของเซลล์ประสาท (Activation function) หรือที่เรียกอีกชื่อว่า ฟังก์ชันถ่ายโอน (Transfer function) มีหลายแบบ แต่ฟังก์ชันที่ได้รับการทดสอบแล้วว่าใช้งานได้จริงและพบบ่อยมีอยู่ 4 ตัวเลือกคือ ฟังก์ชันขั้นบันได (Step function) ฟังก์ชันไซน์ (Sign function) ฟังก์ชันซิกมอยด์ (Sigmoid function) และฟังก์ชันเชิงเส้น (Linear function) แสดงดังรูปที่ 2.12



รูปที่ 2.12 ACTIVATION FUNCTIONS OF A NEURON (NEGNEVITSKY, 2005)

ฟังก์ชันซิกมอยด์ เปลี่ยนรูปของหน่วยรับเข้าที่มีค่าตั้งแต่บวกอนันต์ถึงลบอนันต์ ให้เป็นค่าที่เป็นไปได้ในช่วง 0 และ 1 ฟังก์ชันนี้ใช้ในโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation: BP)

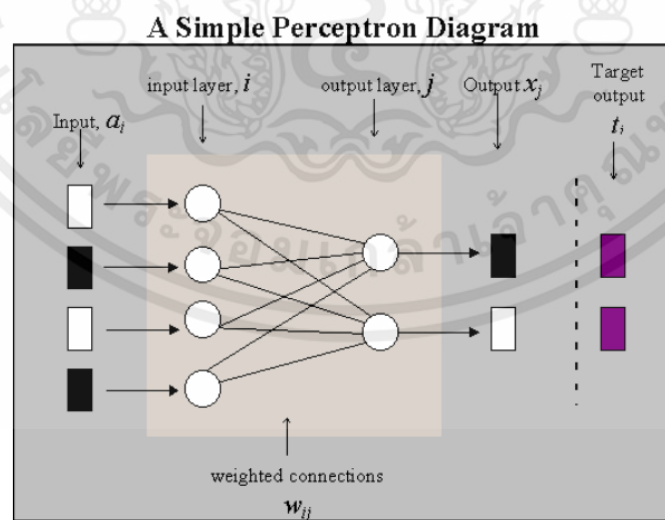
ฟังก์ชันเชิงเส้น สามารถให้ค่าหน่วยรับเข้าเท่ากับน้ำหนักของหน่วยรับเข้า โดยเซลล์ประสาทที่ใช้ฟังก์ชันแบบเชิงเส้นถูกใช้ในการประมาณเป็นจำนวนมาก

ประเภทของ Perceptron

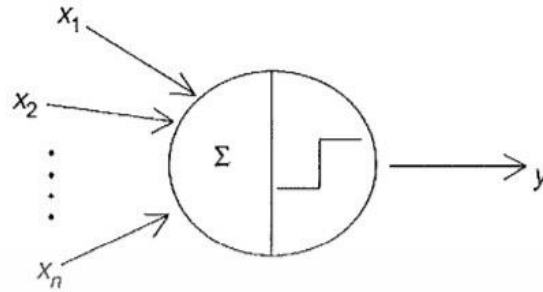
เพอร์เซปตรอน (Perceptron) เป็นข่ายงานประสาทเทียมแบบง่าย มีหน่วยเดียวที่จำลองลักษณะของเซลล์ประสาท ดังรูปที่ 2.14 เมื่อนำ Perceptron และแบบจำลองเซลล์ประสาท มาเขียนให้อยู่ในรูปของโครงข่ายประสาทเทียมสามารถแบ่งออกได้เป็น 2 ประเภทดังนี้

1. Single layer perceptron

เป็นโครงข่ายประสาทอย่างง่ายที่มีโครงสร้างไม่ซับซ้อน และเป็นพื้นฐานของการคำนวณโครงข่ายประสาทแบบอื่น ๆ ประกอบด้วยชั้นอินพุตและเอาต์พุตเท่านั้น และผลลัพธ์ได้จากผลรวมของข้อมูลอินพุตและค่าน้ำหนักของแต่ละจุดที่เชื่อมโยงกัน หากค่าที่คำนวณได้มากกว่าค่าของเทรชโอด (Threshold) θ ค่าที่ได้เรียกว่า แอกติเวท (Activate) มีค่าเท่ากับ 1 แต่หากค่าที่คำนวณได้น้อยกว่าค่าของเทรชโอดค่าที่ได้เรียกว่า ดีแอกติเวท (Deactivate) ซึ่งมีค่าเท่ากับ -1 การคำนวณของเพอร์เซปตรอน (Perceptron) สามารถใช้ค่าอื่น ๆ ได้ซึ่งค่าเทรชโอดขึ้นอยู่กับค่าแอกติเวท และดีแอกติเวทนั้น ๆ ค่าหน่วยส่งออก (Output) เป็น 1 และ -1 และค่าเทรชโอดเป็น 0 ซึ่งทำให้การคำนวณเร็วกว่าการใช้ตัวเลขอื่น ๆ ขั้นตอนวิธีที่ใช้ในการเรียนรู้เรียกว่า กฎเดลตา (Delta rule) ซึ่งสามารถคำนวณค่าคลาดเคลื่อนระหว่างการคำนวณค่าหน่วยส่งออกทั้งหมดกับข้อมูลหน่วยส่งออกที่เป็นตัวอย่างได้ และสามารถนำค่าคลาดเคลื่อนเหล่านี้มาปรับค่าน้ำหนักได้ เพอร์เซปตรอนที่มีชั้นเดียวสามารถแยกแยะรูปแบบต่าง ๆ ได้ แต่ยังมีข้อจำกัดอยู่มาก ซึ่งงานวิจัยต่อ ๆ มาได้พัฒนานำค่าจำนวนจริงมาใช้ซึ่งให้ผลลัพธ์ที่ดีกว่า โครงสร้างการทำงานของ Single Layer Perceptron แสดงดังรูปที่ 2.13



รูปที่ 2.13 SINGLE LAYER PERCEPTRON (ESTEBON, 1997)

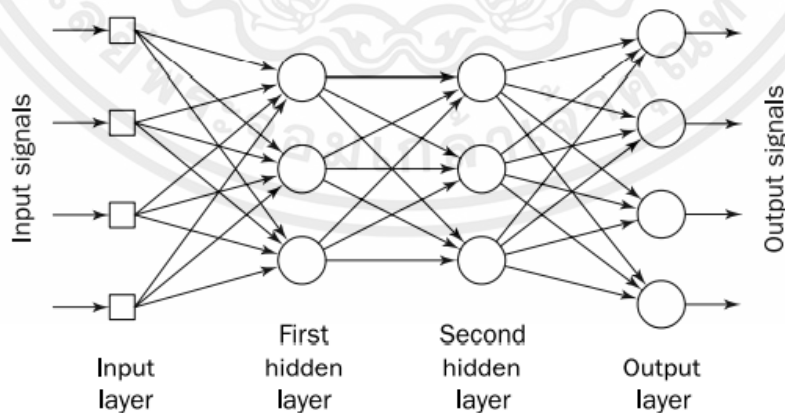


รูปที่ 2.14 เพอร์เซปตรอน (PERCEPTRON) (เกศรินดา สุครีพ, 2554)

2. Multilayer perceptron

เป็นโครงข่ายหลายชั้น (เพอร์เซปตรอน มีมากกว่า 1 ชั้น) ประกอบด้วยชั้นอินพุต ชั้นกลางหรือชั้นซ่อน (Middle or hidden layer) และชั้นเอาต์พุต โดยที่ชั้น Hidden layer สามารถมีได้มากกว่าหนึ่งชั้น ลักษณะโครงข่ายแบบนี้ประกอบด้วยชั้นคำนวณหลายชั้น ซึ่งแต่ละชั้นมีการคำนวณแบบไปข้างหน้า (Feedforward) ซึ่งมีหลักการทำงานคือ เมื่อแต่ละชั้นรับค่าและคำนวณค่าผลรวมของข้อมูลนำเข้าและค่าน้ำหนักของแต่ละจุดที่เชื่อมโยงกันแล้วส่งต่อค่าเหล่านั้นไปยังจุดเชื่อมต่อในชั้นถัดไป และได้ผลลัพธ์ออกมาในชั้นสุดท้าย โครงข่ายลักษณะนี้ใช้ฟังก์ชันซิกมอยด์ (Sigmoid function) เป็นฟังก์ชันกระตุ้น (Activation function) โครงสร้างของ Multilayer Perceptron แสดงดังรูปที่ 2.15

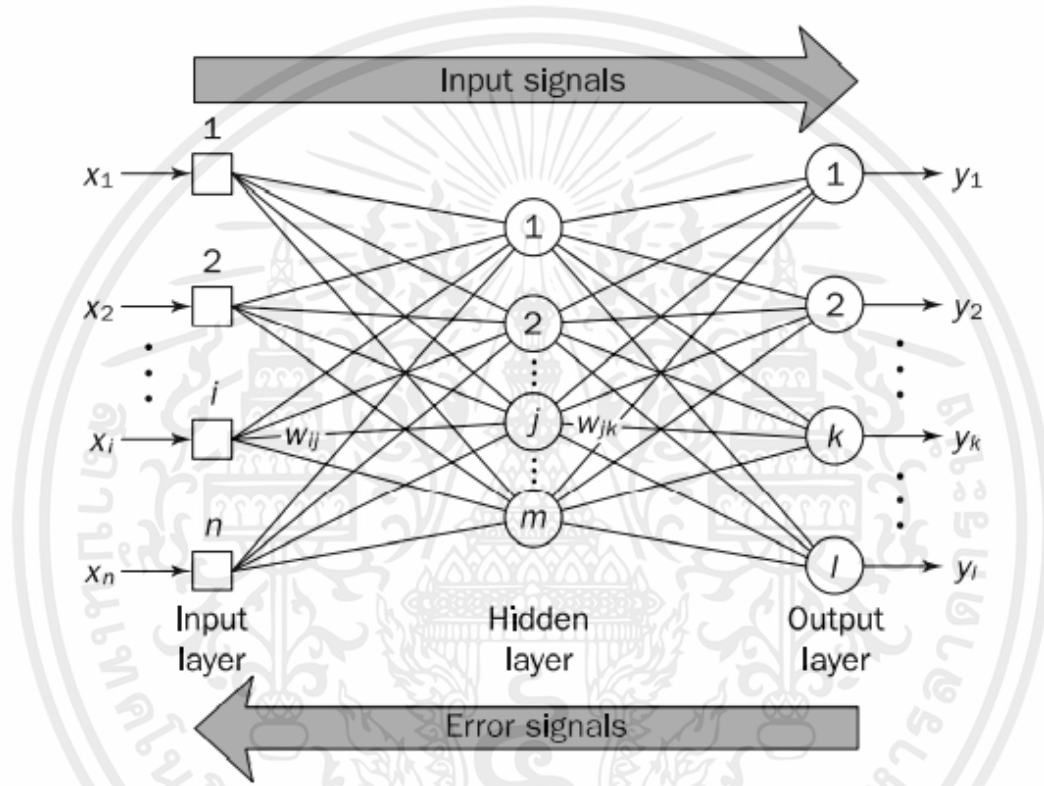
โครงข่ายแบบหลายชั้นสามารถนำมาใช้ในเทคนิคการเรียนรู้หลายแบบ แบบหนึ่งที่ใช้กันมากคือ โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation: BP) ซึ่งการเรียนรู้นี้ถูกกล่าวถึงในหัวข้อถัดไป



รูปที่ 2.15 MULTILAYER PERCEPTRON แบบมีชั้นซ่อน 2 ชั้น (NEGNEVITSKY, 2005)

โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation)

เป็นโครงข่ายที่มีชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุต โดยที่โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับมีการคำนวณค่าความคลาดเคลื่อนหรือค่าความผิดพลาด (Error) ที่เกิดจากการเปรียบเทียบระหว่างเอาต์พุตที่คำนวณได้กับเอาต์พุตที่ต้องการให้เป็น เรียกอีกอย่างได้ว่า เป็นการเรียนรู้จากตัวอย่างที่เข้ามา



รูปที่ 2.16 โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (NEGNEVITSKY, 2002)

จากรูปที่ 2.16 แสดงให้เห็นรูปแบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ สามารถอธิบายได้ดังนี้ สัญญาณอินพุต x_1, x_2, \dots, x_n ส่งผ่านโครงข่ายจากทางซ้ายไปทางขวา โดยสัญลักษณ์ w_{ij} คือ น้ำหนักสำหรับการเชื่อมต่อระหว่างเซลล์ประสาท i ในชั้นอินพุต (Input layer) กับเซลล์ประสาท j ในชั้นซ่อน (Hidden layer) และสัญลักษณ์ w_{jk} คือ น้ำหนักระหว่างเซลล์ประสาท j ในชั้นซ่อน กับเซลล์ประสาท k ในชั้นเอาต์พุต (Output layer)

โดยที่ชั้นซ่อน (Hidden layer) ทำหน้าที่ในการซ่อนเอาต์พุตที่ต้องการไว้ โดยที่หลักการทำงานที่แท้จริงหรือพฤติกรรมภายใน ไม่สามารถทราบได้จากอินพุตหรือเอาต์พุต อีกทั้งไม่มีวิธีที่ชัดเจนสำหรับการหาเอาต์พุตที่ชั้นซ่อนทำได้ กล่าวคือ เอาต์พุตที่ได้จากชั้นซ่อนถูกกำหนดด้วยตัวมันเอง (Negnevitsky, 2002)

สำหรับการเผยแพร่สัญญาณความผิดพลาด (Error signals) เริ่มต้นที่ชั้นเอาต์พุตและย้อนกลับไปชั้นซ่อน ซึ่งส่งผ่านโครงข่ายจากทางขวาไปทางซ้าย โดยสัญญาณความผิดพลาดที่เอาต์พุตของเซลล์ประสาท k ที่การทำซ้ำ p ถูกกำหนดโดยสมการที่ (2.26)

$$e_k(p) = y_{d,k}(p) - y_k(p) \quad (2.26)$$

โดยที่ $y_{d,k}(p)$ คือ เอาต์พุตที่ต้องการของเซลล์ประสาท k ที่การทำซ้ำ p

เซลล์ประสาท k ซึ่งอยู่ในเลเยอร์เอาต์พุตนั้นมาพร้อมกับเอาต์พุตที่ต้องการของมันเอง ดังนั้นจึงใช้ขั้นตอนที่ตรงไปตรงมาในการอัปเดตค่าน้ำหนัก w_{jk} ซึ่งกฎสำหรับการอัปเดตน้ำหนักที่เลเยอร์เอาต์พุตนั้นคล้ายกับกฎการเรียนรู้ของ Perceptron ของสมการที่ (2.27)

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \quad (2.27)$$

โดยที่ $\Delta w_{jk}(p)$ คือ ค่าการปรับน้ำหนัก

เมื่อเราพิจารณาการแก้ไขน้ำหนักของเพอร์เซปตรอน เราใช้สัญญาณอินพุต x_i แต่ในโครงข่ายหลายชั้น อินพุตของเซลล์ประสาทในเลเยอร์เอาต์พุตแตกต่างจากอินพุตของเซลล์ประสาทในเลเยอร์อินพุต และเนื่องจากไม่สามารถใส่สัญญาณอินพุต x_i ได้ ในการกำหนดและปรับน้ำหนักในโครงข่ายประสาทเทียมหลายชั้นจึงใช้เอาต์พุตของเซลล์ประสาท j ในเลเยอร์ y_j ที่ซ่อนอยู่ แทนที่อินพุต x_i สามารถคำนวณหาค่าน้ำหนักที่ปรับปรุงได้จากสมการที่ (2.28)

$$\Delta w_{jk}(p) = \alpha \times y_j(p) \times \delta_k(p) \quad (2.28)$$

โดยที่ $\delta_k(p)$ คือ error gradient ที่เซลล์ประสาท k ในเลเยอร์เอาต์พุตที่การวนซ้ำ p

Error gradient ถูกกำหนดเป็นอนุพันธ์ของฟังก์ชันการกระตุ้น คูณด้วยข้อผิดพลาดที่เอาต์พุตของเซลล์ประสาท สามารถหา $\delta_k(p)$ ได้จากสมการที่ (2.29)

$$\delta_k(p) = \frac{\partial y_k(p)}{\partial X_k(p)} \times e_k(p) \quad (2.29)$$

โดยที่ $y_k(p)$ คือ เอาต์พุตของเซลล์ประสาท k ที่การวนซ้ำ p และ $X_k(p)$ คืออินพุตที่ถ่วงน้ำหนักสุทธิไปยังเซลล์ประสาท k ในการวนซ้ำเดียวกัน สำหรับการเปิดใช้งานฟังก์ชันซิกมอยด์แสดงดังสมการที่ (2.30)

$$\delta_k(p) = \frac{\partial \left\{ \frac{1}{1+e^{-X_k(p)}} \right\}}{\partial X_k(p)} \times e_k(p) = \frac{e^{-X_k(p)}}{\{1+e^{-X_k(p)}\}^2} \times e_k(p) \quad (2.30)$$

ทำให้ได้

$$\delta_k(p) = y_k(p) \times [1 - y_k(p)] \times e_k(p) \quad (2.31)$$

โดยที่

$$y_k(p) = \frac{1}{1+e^{-X_k(p)}} \quad (2.32)$$

ดังนั้นได้

$$\delta_k(p) = y_k(p) \times \left[1 - \frac{1}{1+e^{-X_k(p)}} \right] \times e_k(p) \quad (2.33)$$

การกำหนดค่าน้ำหนักของเซลล์ประสาทที่ปรับปรุงในเลเยอร์ที่ซ่อนอยู่ (Hidden layer)

ในการคำนวณการแก้ไขน้ำหนักสำหรับเลเยอร์ที่ซ่อนอยู่ เราสามารถใช้สมการเดียวกันกับเอาต์พุตเลเยอร์ ดังสมการที่ (2.5)

$$\Delta w_{ij}(p) = \alpha \times x_i(p) \times \delta_j(p) \quad (2.34)$$

โดยที่ $\delta_j(p)$ แสดงถึง Error gradient ที่เซลล์ประสาท j ในเลเยอร์ที่ซ่อนอยู่

$$\delta_j(p) = y_j(p) \times [1 - y_j(p)] \times \sum_{k=1}^l \delta_k(p) w_{jk}(p) \quad (2.35)$$

$$y_j(p) = \frac{1}{1 + e^{-X_j(p)}} \quad (2.36)$$

$$X_j(p) = \sum_{i=1}^n x_i(p) \times w_{ij}(p) - \theta_j \quad (2.37)$$

โดยที่ l คือจำนวนเซลล์ประสาทในชั้นเอาต์พุต และ n คือจำนวนเซลล์ประสาทในชั้นอินพุต
ขั้นตอนการเรียนรู้ของโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ

สามารถแบ่งออกได้เป็น 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 กำหนดค่าเริ่มต้น (Initialisation)

ตั้งค่าน้ำหนักและระดับเกณฑ์ทั้งหมดของเครือข่ายเป็นตัวเลขสุ่มที่กระจายอย่างสม่ำเสมอภายในช่วงขนาดเล็ก (Haykin, 1999)

$$\left(-\frac{2.4}{F_i}, +\frac{2.4}{F_i}\right) \quad (2.38)$$

โดยที่ F_i คือจำนวนอินพุตของเซลล์ประสาท i ทั้งหมดในเครือข่าย การกำหนดค่าเริ่มต้นน้ำหนักทำบนพื้นฐานของเซลล์ประสาทโดยเซลล์ประสาท

ขั้นตอนที่ 2 กระตุ้นการใช้งาน (Activation)

กระตุ้นการใช้งานโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ โดยใช้อินพุต $x_1(p), x_2(p), \dots, x_n$ และ เอาต์พุตที่ต้องการ $y_{d,1}(p), y_{d,2}(p), \dots, y_{d,n}(p)$

2.1 คำนวณผลลัพธ์ที่แท้จริงของเซลล์ประสาทในชั้นที่ซ่อนอยู่ จากสมการต่อไปนี้

$$y_j(p) = \text{sigmoid}[\sum_{i=1}^n x_i(p) \times w_{ij}(p) - \theta_j] \quad (2.39)$$

โดยที่ n คือจำนวนอินพุตของเซลล์ประสาท j ในเลเยอร์ที่ซ่อนอยู่ และ *Sigmoid* คือฟังก์ชันที่กระตุ้นการทำงานของเซลล์ประสาท (Sigmoid activation function)

2.2 คำนวณผลลัพธ์ที่แท้จริงของเซลล์ประสาทในชั้นเอาต์พุต จากสมการต่อไปนี้

$$y_k(p) = \text{sigmoid}[\sum_{j=1}^m x_{jk}(p) \times w_{jk}(p) - \theta_k] \quad (2.40)$$

โดยที่ m คือจำนวนอินพุตของเซลล์ประสาท k ในเลเยอร์เอาต์พุต

ขั้นตอนที่ 3 การปรับปรุณค่าน้ำหนัก (Weight training)

การปรับปรุณค่าน้ำหนักในโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ สามารถคำนวณได้จาก

3.1 คำนวณ error gradient สำหรับเซลล์ประสาทในชั้นเอาต์พุต

$$\delta_k(p) = y_k(p) \times [1 - y_k(p)] \times e_k(p) \quad (2.41)$$

โดยที่

$$e_k(p) = y_{d,k}(p) - y_k(p) \quad (2.42)$$

ซึ่งสามารถคำนวณค่าน้ำหนักที่ปรับปรุงได้จาก

$$\Delta w_{jk}(p) = \alpha \times y_j(p) \times \delta_k(p) \quad (2.43)$$

และสามารถหาค่าน้ำหนักที่ปรับปรุงที่เซลล์ประสาทเอาต์พุตได้จาก

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \quad (2.44)$$

3.2 คำนวณ error gradient สำหรับเซลล์ประสาทในชั้นซ่อน

$$\delta_j(p) = y_j(p) \times [1 - y_j(p)] \sum_{k=1}^l \delta_k(p) \times w_{jk}(p) \quad (2.45)$$

ซึ่งสามารถคำนวณค่าน้ำหนักที่ปรับปรุงได้จาก

$$\Delta w_{ij}(p) = \alpha \times x_i(p) \times \delta_j(p) \quad (2.46)$$

และสามารถหาค่าน้ำหนักที่ปรับปรุงที่เซลล์ประสาทที่ซ่อนอยู่ได้จาก

$$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p) \quad (2.47)$$

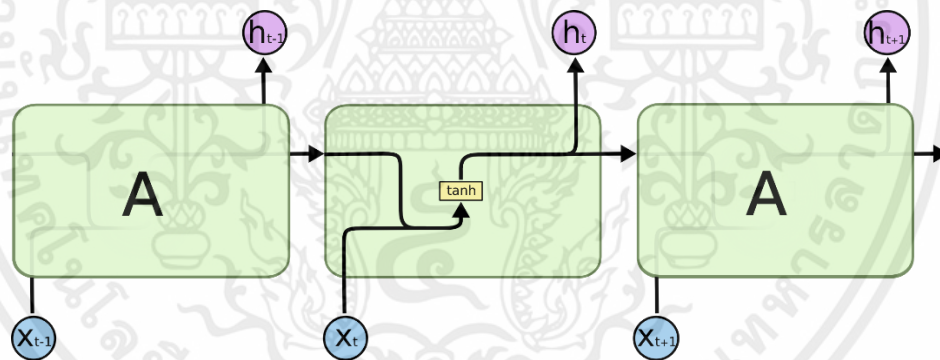
ขั้นตอนที่ 4 การทำซ้ำ (Iteration)

สามารถทำได้โดยการเพิ่มจำนวนการทำซ้ำ p ไปขึ้นละหนึ่งครั้ง จากนั้นกลับไปเริ่มทำที่ขั้นตอนที่ 2 และทำขั้นตอนซ้ำ ๆ จนกระทั่งเกณฑ์ความผิดพลาดที่เลือกเป็นที่น่าพอใจ

2.5 โครงข่ายประสาทเทียมแบบ Long short term memory (LSTM)

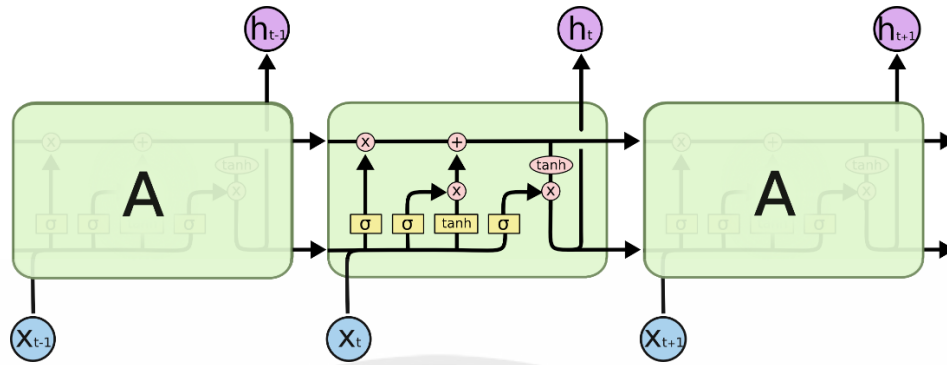
โครงข่ายประสาทเทียม LSTM เป็น Recurrent neural networks หรือ RNN ชนิดพิเศษที่ถูกพัฒนาโดย Hochreiter & Schmidhuber (1997) เพื่อแก้ปัญหา Gradient เนื่องจากมีการผ่านเลเยอร์หลายชั้น ส่งผลให้ขนาดหรือแอมพลิจูดของมันลดลง

โครงข่ายประสาทเทียม RNN ทั้งหมดมีรูปแบบเป็นลูกโซ่ของโมดูลการทำซ้ำของโครงข่ายประสาท โดยพื้นฐานของ RNN โมดูลการทำซ้ำนี้มีโครงสร้างที่เรียบง่าย เช่น มีเลเยอร์ tanh เพียงเลเยอร์เดียว ดังรูปที่ 2.17



รูปที่ 2.17 รูปแบบโครงข่ายประสาทเทียม RNN (OLAH, 2015)

โครงข่ายประสาทเทียม LSTM มีโครงสร้างเหมือนกับลูกโซ่นี้ แต่โมดูลการทำซ้ำนั้นแตกต่างออกไป เนื่องจากไม่ได้มีเพียงเลเยอร์เดียว แต่มีถึง 4 เลเยอร์ ดังรูปที่ 2.18



รูปที่ 2.18 รูปแบบโครงข่ายประสาทเทียม LSTM (OLAH, 2015)

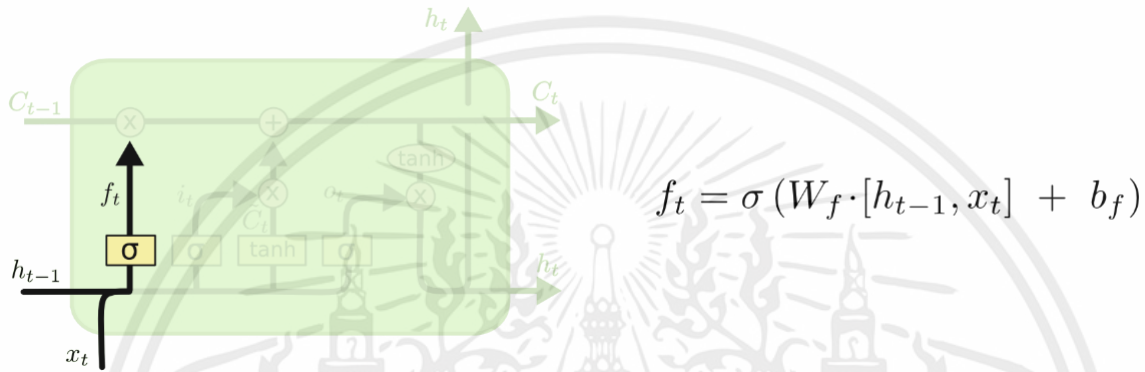


รูปที่ 2.19 สัญลักษณ์ที่用在แผนภาพโครงข่ายประสาทเทียม (OLAH, 2015)

จากรูปที่ 2.19 เป็นสัญลักษณ์ที่用在แผนภาพโครงข่ายประสาทเทียม วงกลมสีชมพูเป็นการแสดงการดำเนินการตามจุด เช่น การบวกเวกเตอร์ ส่วนสี่เหลี่ยมสีเหลืองเป็นเลเยอร์ของโครงข่ายประสาท Vector transfer เป็นเวกเตอร์ที่ออกจากเอาต์พุตของโหนดหนึ่งไปยังอินพุตของอีกโหนดหนึ่ง Concatenate เป็นการต่อกันของเวกเตอร์ Copy แสดงถึงการถูกคัดลอกและทำสำเนาไปยังตำแหน่งต่อไป

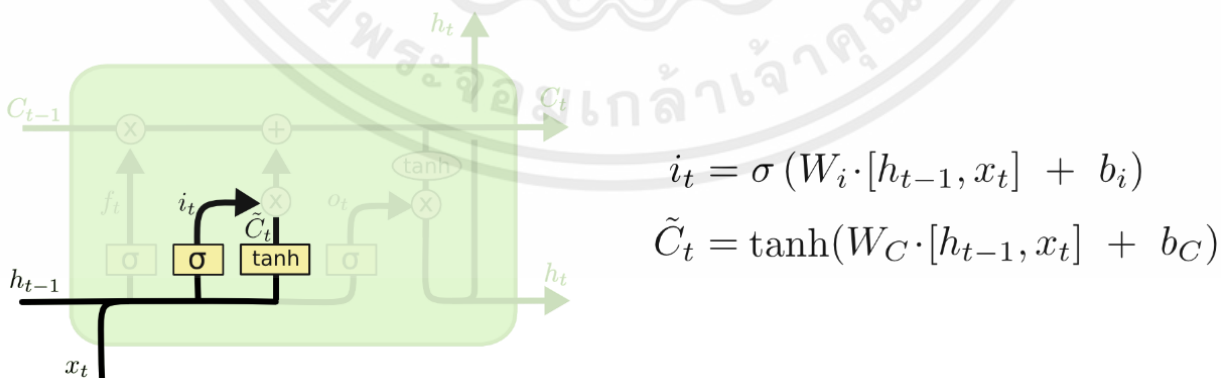
ขั้นตอนการทำงานของ LSTM แบ่งออกเป็น 4 ขั้นตอน

1. ในขั้นตอนแรกเป็นการตัดสินใจว่าเลือกทิ้งข้อมูลใดออกจากสถานะเซลล์ การตัดสินใจนี้ถูกทำโดย เลเยอร์ Sigmoid ที่ถูกเรียกว่า Forget gate layer โดยดูที่ h_{t-1} และ x_t เอาต์พุตที่ได้เป็นตัวเลข ระหว่าง 0 ที่หมายถึงกำจัดค่าทิ้ง และ 1 ที่หมายถึงเก็บค่าไว้ในสถานะเซลล์ C_{t-1} ดังรูปที่ 2.20



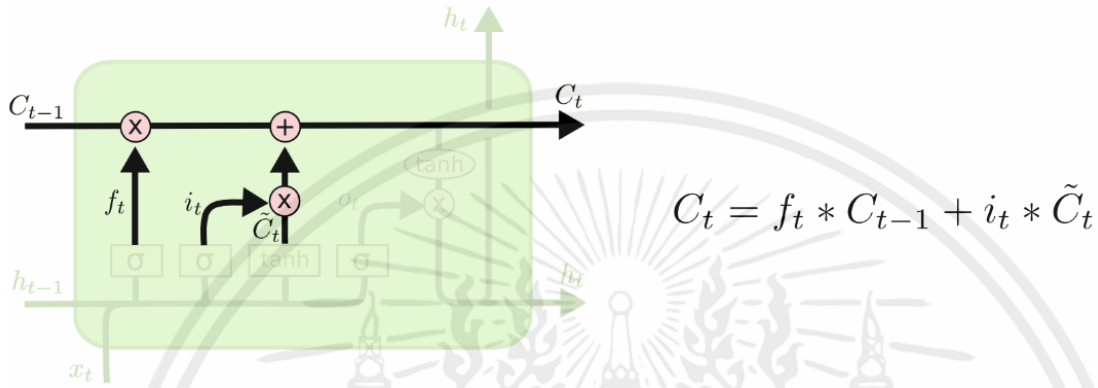
รูปที่ 2.20 ขั้นตอนการตัดสินใจเลือกทิ้งข้อมูล (OLAH, 2015)

2. ในขั้นตอนนี้เป็นการตัดสินใจเก็บข้อมูลใหม่เข้าไปในสถานะเซลล์ โดยมี 2 ส่วนคือ เลเยอร์ Sigmoid ที่เรียกว่า Input gate layer ทำหน้าที่ตัดสินใจว่าค่าใดที่ทำการอัปเดตต่อไป และเลเยอร์ tanh สามารถสร้างเวกเตอร์ของค่าตัวเลือกใหม่ \tilde{C}_t และเพิ่มไปในสถานะ จากนั้นรวมสองส่วนนี้เข้าด้วยกัน ดังรูปที่ 2.21



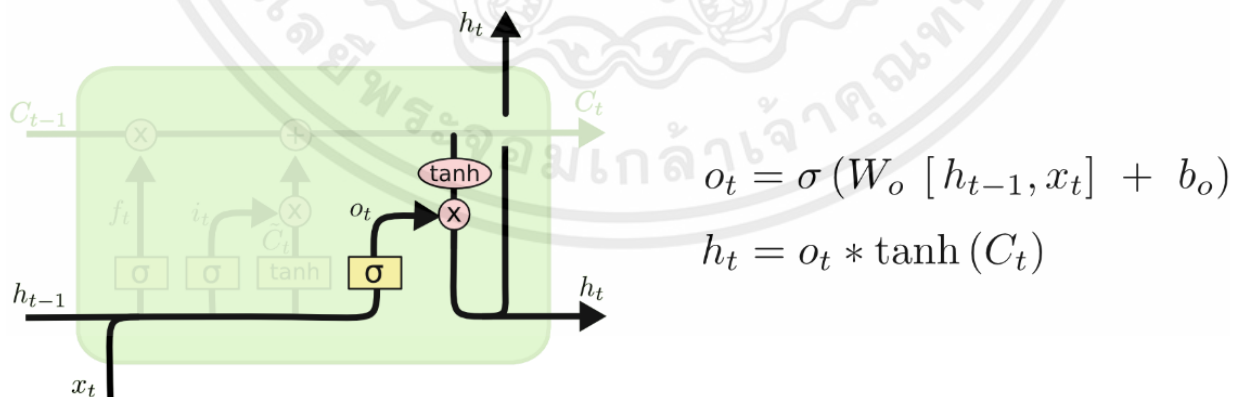
รูปที่ 2.21 ขั้นตอนการตัดสินใจเก็บข้อมูล (OLAH, 2015)

3. ในขั้นตอนนี้เรารวมเอาต์พุตที่ได้จาก 2 ขั้นตอนก่อนหน้าเพื่อสร้างการอัปเดตสถานะเซลล์ C_{t-1} ไปเป็นสถานะเซลล์ใหม่ C_t เราคูณสถานะเก่าด้วย f_t และบวก $i_t * \tilde{C}_t$ เพิ่มเข้าไป เอาต์พุตที่ได้เป็นค่าตัวเลือกใหม่ โดยขนาดขึ้นอยู่กับจำนวนที่เราตัดสินใจอัปเดตค่าสถานะแต่ละค่า ดังรูปที่ 2.22



รูปที่ 2.22 ขั้นตอนการอัปเดตสถานะเซลล์ (OLAH, 2015)

4. ขั้นตอนสุดท้ายเป็นการตัดสินใจว่าส่งออกเอาต์พุตเป็นอะไร โดยเอาต์พุตที่ได้ขึ้นอยู่กับสถานะเซลล์ที่ทำการกรอง ชั้นแรก เราเรียกใช้ไลเยอร์ Sigmoid ซึ่งตัดสินใจว่าส่งออกส่วนใดของสถานะเซลล์ เราส่งค่า C_t ผ่าน \tanh เพื่อให้ค่าอยู่ระหว่าง -1 ถึง 1 แล้วคูณด้วยเอาต์พุตของ Sigmoid gate ดังนั้นเราส่งเอาต์พุตเฉพาะส่วนที่มีการตัดสินใจ ดังรูปที่ 2.23



รูปที่ 2.23 ขั้นตอนการส่งออกเอาต์พุต (OLAH, 2015)

2.6 Loss function

อัลกอริทึมทั้งหมดในแมชชีนเลิร์นนิงจำเป็นต้องอาศัยการย่อหรือขยายฟังก์ชันที่เรียกว่า Objective function กลุ่มของฟังก์ชันที่ถูกล่อเรียกว่า Loss function ซึ่งเป็นตัววัดความสามารถในการพยากรณ์ของโมเดลว่าดีเพียงใด เมื่อเทียบกับผลลัพธ์ที่คาดหวัง หนึ่งในวิธีที่ใช้หาจุดต่ำสุดของฟังก์ชันคือ Gradient descent

การเลือกใช้ Loss function กับข้อมูลประเภทใดประเภทหนึ่งขึ้นอยู่กับปัจจัยหลายอย่าง เช่น การปรากฏตัวของค่าผิดปกติ อัลกอริทึมการเลือกของแมชชีนเลิร์นนิง ประสิทธิภาพทางด้านเวลาของ Gradient descent ความง่ายในการหาอนุพันธ์ และความมั่นใจในการพยากรณ์

Loss function สามารถแบ่งออกเป็น 2 ประเภทใหญ่ ๆ คือ Classification loss และ Regression loss โดยในงานวิจัยนี้ ผู้วิจัยให้ความสนใจไปที่ Regression loss ซึ่งแบ่งออกเป็น 2 แบบตามความสนใจของผู้วิจัย

1. Mean squared error (MSE) หรือ L2 Loss

Mean square error หรือค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เป็นการนำกำลังสองมาใช้ เพื่อให้การวัดค่าความคลาดเคลื่อนได้ผลลัพธ์เป็นบวกเสมอโดยที่ข้อผิดพลาดลดลงเมื่อข้อผิดพลาดเข้าใกล้ศูนย์ MSE สามารถประเมินคุณภาพของตัวทำนาย หรือตัวประมาณ ซึ่งค่า MSE สามารถคำนวณได้ ดังสมการที่ (2.48) และ (2.49)

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.48)$$

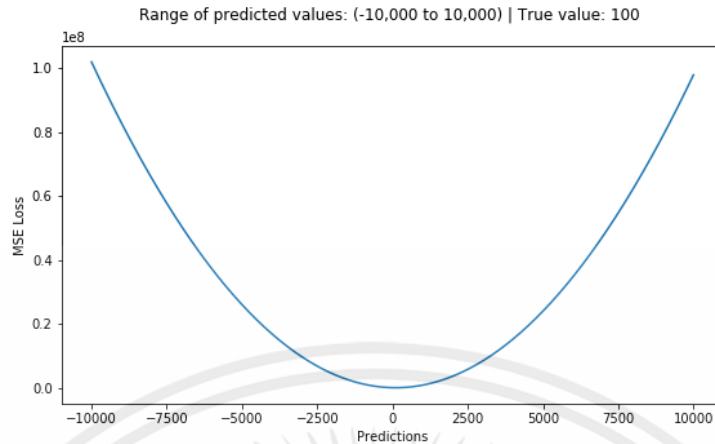
หรือ

$$MSE = \frac{1}{n} \sum_{t=1}^n (Error)^2 \quad (2.49)$$

โดยที่ Y_i คือ ค่าที่ทำนาย (prediction)

\hat{Y}_i คือ ค่าจริง (actual)

n คือ จำนวนข้อมูล



รูปที่ 2.24 MSE LOSS (Y-AXIS) VS. PREDICTIONS (X-AXIS) (PRINCE GROVER, 2018)

จากรูปที่ 2.24 เป็นการพล็อต MSE function ที่มีค่าเป้าหมายจริงเท่ากับ 100 และค่าที่พยากรณ์อยู่ระหว่าง -10,000 ถึง 10,000

2. Mean absolute error (MAE) หรือ L1 Loss

Mean absolute error (MAE) หรือค่าความคลาดเคลื่อนสัมบูรณ์ เป็น Loss function ที่ใช้ Regression model เป็นผลรวมสัมบูรณ์ของผลต่างระหว่างค่าเป้าหมายและค่าจากการพยากรณ์ ดังนั้นมันสามารถวัดขนาดเฉลี่ยสำหรับความผิดพลาดของกลุ่มข้อมูลพยากรณ์ ค่า MAE สามารถคำนวณได้ดังสมการที่ (2.50) และ (2.51)

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_i - \hat{Y}_i| \quad (2.50)$$

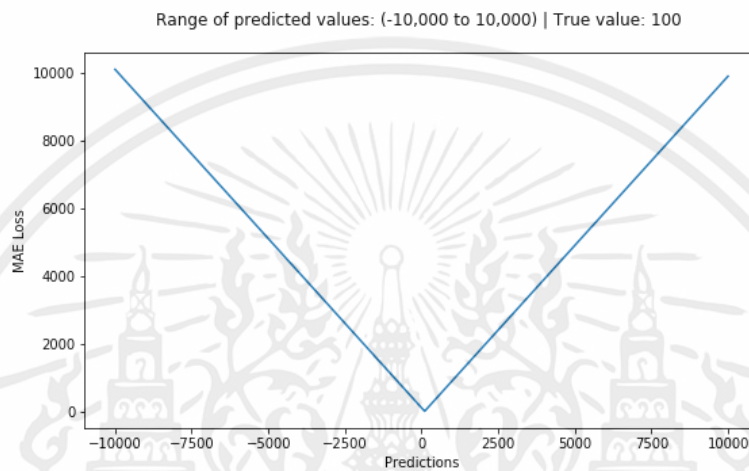
หรือ

$$MAE = \frac{1}{n} \sum_{t=1}^n |Error| \quad (2.51)$$

โดยที่ \hat{Y}_i คือ ค่าที่ทำนาย (prediction)

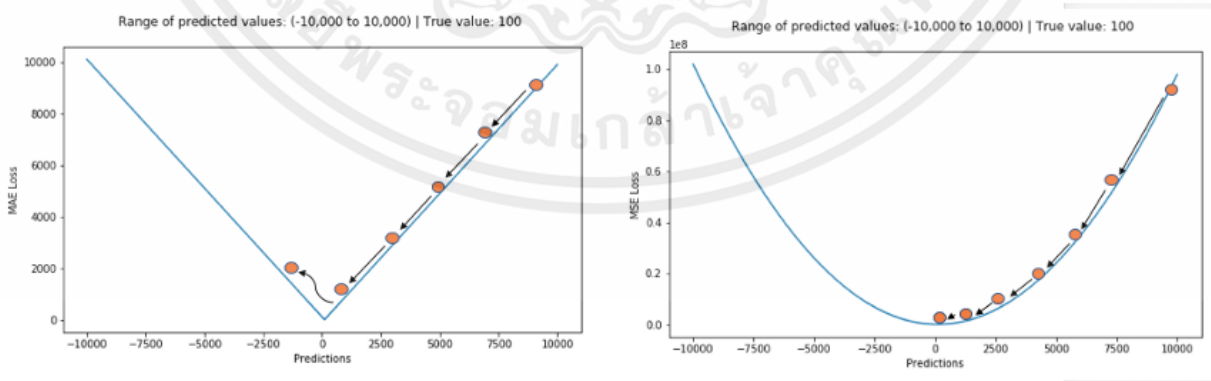
Y_i คือ ค่าจริง (actual)

n คือ จำนวนข้อมูล



รูปที่ 2.25 MAE LOSS (Y-AXIS) VS. PREDICTIONS (X-AXIS) (PRINCE GROVER, 2018)

จากรูปที่ 2.25 เป็นการพล็อต MAE function ที่มีค่าเป้าหมายจริงเท่ากับ 100 และค่าที่พยากรณ์อยู่ระหว่าง -10,000 ถึง 10,000



รูปที่ 2.26 การใช้ GRADIENT DESCENT ร่วมกับ LOSS FUNCTION (PCP, 2020)

จากรูปที่ 2.26 เห็นได้ว่าการใช้ Gradient descent ร่วมกับ MAE นั้นมีผลต่างของค่าความผิดพลาดที่เท่า ๆ กัน ส่งผลให้ค่าที่อัปเดตเปลี่ยนแปลงไปแบบคงที่ และเป็นผลสืบเนื่องให้ไม่สามารถหาค่าที่มีความผิดพลาดต่ำที่สุดได้ แต่การใช้ Gradient descent ร่วมกับ MSE สามารถหาค่าที่มีความผิดพลาดต่ำที่สุดได้ เนื่องจากผลต่างของค่าความผิดพลาดเปลี่ยนแปลงไปเรื่อย ๆ และค่อย ๆ ลดลงเมื่อเข้าใกล้จุดต่ำสุด ส่งผลให้ MSE เหมาะสมกับการใช้ร่วมกับ Neural network

ดังนั้น ผู้วิจัยจึงตัดสินใจเลือก Mean square error ในการวัดความคลาดเคลื่อนจากการพยากรณ์



2.7 งานวิจัยที่เกี่ยวข้อง

ปรเมษฐ์ ชัยกร และธนภัทร (2017) ได้ศึกษาการประยุกต์ใช้โมเดลการเรียนรู้แบบรวมกลุ่ม เพื่อพยากรณ์แนวโน้มของราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย จากการใช้ข้อมูลราคาซื้อขายหลักทรัพย์รายวัน ประกอบด้วย ราคาเปิด ราคาสูงสุด ราคาต่ำสุด ราคาปิด ปริมาณการซื้อขายแบบรายวัน และข้อมูลดัชนีชี้วัดทางเทคนิค ประกอบด้วย Stochastic oscillator (STO) Relative strength index (RSI) และ On balance volume (OBV) ใช้ข้อมูลจากหลักทรัพย์ 42 หลักทรัพย์ที่อยู่ในดัชนี SET50 ตั้งแต่วันที่ 1 มกราคม 2011 ถึงวันที่ 31 ธันวาคม 2016 โดยแบ่งเป็นกลุ่มข้อมูลสำหรับเรียนรู้ 24 เดือน กลุ่มข้อมูลสำหรับทดสอบ 6 เดือน และกลุ่มข้อมูลประเมินผล 6 เดือน ข้อมูลเหล่านี้ถูกใช้ร่วมกับ 4 เทคนิคในการสร้างโมเดล ประกอบด้วย Random forest (RF) Support vector machine (SVM) K nearest neighbors (KNN) และ Naive bayes (NB) จากการศึกษาพบว่า Ensemble model แบบถ่วงน้ำหนักมีประสิทธิภาพในการพยากรณ์สูงสุด และมีแนวโน้มเพิ่มขึ้นตามจำนวนวันที่ถือ

Boonpeng and Jeatrakul (2016) ได้ศึกษาระบบช่วยเหลือในการตัดสินใจสำหรับการลงทุนในตลาดหลักทรัพย์โดยใช้โครงข่ายประสาทเทียม OAA โดยพยากรณ์จากการใช้โมเดล 3 แบบเปรียบเทียบกัน ประกอบด้วย Traditional neural network model One-against-one technique (OAO) และ One-against-all (OAA) จากการใช้ข้อมูลราคาซื้อขายหลักทรัพย์รายวัน ประกอบด้วย ราคาเปิด ราคาสูงสุด ราคาต่ำสุด ราคาปิด ปริมาณการซื้อขายแบบรายวัน และข้อมูลดัชนีชี้วัดทางเทคนิค ประกอบด้วย Stochastic oscillator (STO) Relative strength index (RSI) และ Moving average convergence/divergence (MACD) ใช้ข้อมูลจากตลาดหลักทรัพย์แห่งประเทศไทยตั้งแต่วันที่ 3 มกราคม 2007 ถึง 29 สิงหาคม 2014 โดยแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 80 และข้อมูลสำหรับทดสอบร้อยละ 20 จากการศึกษาพบว่าประสิทธิภาพในการพยากรณ์เมื่อเปรียบเทียบระหว่างโมเดลทั้ง 3 แบบ OAA-NN มีประสิทธิภาพสูงสุด และจากการเปรียบเทียบ OAA-NN กับดัชนีชี้วัดทางเทคนิคในการสร้างผลตอบแทนจากสัญญาณซื้อขายหลักทรัพย์ พบว่า OAA-NN ให้อัตราผลตอบแทนสูงสุด

Chaigusin Chirathamjaree และ Clayden (2008) ได้ศึกษาการใช้โครงข่ายประสาทเทียมในการพยากรณ์ดัชนีตลาดหลักทรัพย์แห่งประเทศไทย โดยเปรียบเทียบประสิทธิภาพในการพยากรณ์ของโครงข่ายประสาทเทียมที่มีจำนวนชั้นของโครงข่ายแตกต่างกัน ประกอบด้วยโครงข่ายที่มีจำนวน 3 4 และ 5 ชั้น จากการศึกษาพบว่า โมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 3 ชั้น คือ 7-3-1 โมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 4 ชั้น คือ 7-7-3-1 และโมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 5 ชั้น คือ 7-13-7-3-1 วัดประสิทธิภาพโดยใช้ MAPE ได้ 1.26594 1.14719 และ 1.14578 9 ตามลำดับ

Vijh Chandola Tikkiwal และ Kumar (2020) ได้ศึกษาการพยากรณ์ราคาปิดหุ้นโดยใช้แมชชีนเลิร์นนิง จากการใช้ข้อมูล High minus Low price (H-L) Close minus Open price (O-C) Price's seven days' moving average, Price's fourteen days' moving average, Price's twenty one days' moving average และ Price's standard deviation for the past seven days จากบริษัท Nike Goldman Sachs Johnson and Johnson Pfizer และ JP Morgan Chase and Co. ตั้งแต่วันที่ 4 พฤษภาคม 2009 ถึง 4 พฤษภาคม 2019 ข้อมูลดังกล่าวถูกนำมาใช้ในการพยากรณ์ด้วย ANN และ RF จากการศึกษาพบว่าโมเดล ANN มีประสิทธิภาพในการพยากรณ์มากกว่า RF

Sanboon, Keatruangkamala and Jaiyen (2019) ได้ศึกษาการพยากรณ์ค่านำการซื้อขายในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้ LSTM จากการใช้ข้อมูลราคาซื้อขายหลักทรัพย์รายวัน ประกอบด้วย ราคาสูงสุด ราคาต่ำสุด ราคาเปิด และราคาปิดแบบรายวัน จากบริษัท AOT, BDMS, BEAUTY, CPN และ SCC ตั้งแต่วันที่ 5 มกราคม 2015 ถึง 29 ธันวาคม 2017 โดยแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 80 และข้อมูลสำหรับทดสอบร้อยละ 20 จากการศึกษาพบว่าประสิทธิภาพในการพยากรณ์ของแต่ละบริษัท โมเดล LSTM มีประสิทธิภาพในการพยากรณ์สูงสุดเมื่อเทียบกับ SVM, Logistic Regression, RF, Decision Tree (Entropy), Decision Tree (Gini), KNN และ MLP

Moghae and Hamiche (2020) ได้ศึกษาการพยากรณ์ราคาตลาดหลักทรัพย์โดยใช้ LSTM RNN จากการใช้ข้อมูลราคาเปิดจาก 2 หลักทรัพย์ที่อยู่ในตลาดหลักทรัพย์นิวยอร์กคือ GOOGL และ NKE โดยที่ GOOGL ใช้ข้อมูลตั้งแต่วันที่ 19 สิงหาคม 2004 ถึง 19 ธันวาคม 2019 และ NKE ใช้ข้อมูลตั้งแต่วันที่ 4 มกราคม 2010 ถึง 19 ธันวาคม 2019 โดยแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 80 และข้อมูลทดสอบร้อยละ 20 และ Epochs สำหรับข้อมูลสำหรับทดลองเท่ากับ 12, 25, 50 และ 100 จากการศึกษาพบว่าความแม่นยำในการพยากรณ์เพิ่มขึ้นหาก Epochs เพิ่มขึ้น

ปิยศรีณีย์ สมพร และบำรุง (2021) ได้ศึกษาการพยากรณ์ดัชนีราคาตลาดหลักทรัพย์โดยใช้ Wavelet Transform, LSTM และ Bi-LSTM จากการใช้ข้อมูลราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขายแบบรายวัน ตั้งแต่วันที่ 1 มกราคม 1995 ถึง 31 สิงหาคม 2018 สำหรับ S&P 500, Nikkei 225 และ NASDAQ Composite และการใช้ข้อมูลตั้งแต่วันที่ 1 มกราคม 1987 ถึง 31 สิงหาคม 2018 สำหรับ Hang Seng index โดยแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 70 และข้อมูลทดสอบร้อยละ 30 จากการศึกษาพบว่าโครงข่ายประสาทเทียม Bi-LSTM ให้ประสิทธิภาพที่ดีกว่าเมื่อเทียบกับโครงข่ายประสาทเทียม LSTM และการใช้ Wavelet transform ส่งผลให้โครงข่ายประสาทเทียมสามารถเรียนรู้ได้รวดเร็วยิ่งขึ้น

บทที่ 3 วิธีดำเนินการวิจัย

3.1 ระเบียบวิธีการศึกษา

ตัวแปรต้น

Simple moving average (SMA), Exponential moving average (EMA) และ Weighted moving average (WMA)

ตัวแปรตาม

ราคาปิดที่ถูกพยากรณ์ในวันถัดไป (Predicted Price)

ตัวแปรควบคุม

ข้อมูลราคาเปิด(O) ราคาสูงสุด(H) ราคาต่ำสุด(L) ราคาปิด(C) ปริมาณการซื้อขาย(Volume) และ ราคาปิดที่เปลี่ยนแปลงรายวัน (Daily price change)

ในการศึกษาความสามารถพยากรณ์ราคาหลักทรัพย์ในกลุ่มพลังงานคือ PTT SPCG TTW โดยใช้ LSTM และใช้ข้อมูลราคา และปริมาณการซื้อขายร่วมกับการใช้เส้นค่าเฉลี่ยทั้ง 3 แบบคือ Simple moving average (SMA), Exponential moving average (EMA) และ Weighted moving average (WMA) และวัดผลความคลาดเคลื่อนว่ามีความแตกต่างกันหรือไม่ อย่างไร

3.2 การรวบรวมข้อมูล

การศึกษานี้ได้ใช้ข้อมูลราคาหลักทรัพย์จาก YAHOO FINANCE โดยได้รวบรวมข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย โดยใช้เป็นข้อมูลรายวัน ย้อนหลังเป็นเวลา 10 ปี ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020

3.3 วิธีการประมวลผลและวิเคราะห์ข้อมูล

ในขั้นตอนนี้ ผู้วิจัยได้ใช้ โปรแกรม Google Colab ซึ่งเป็นบริการของ Google ที่ให้บริการโฮสต์โปรแกรม Jupyter Notebook สำหรับภาษา Python

การเตรียมข้อมูล

ภายหลังจากนำเข้าข้อมูลราคาและปริมาณการซื้อขายของหุ้นสามัญ ผู้วิจัยได้ทำการเตรียมข้อมูลที่จำเป็นในการทดสอบโดยเตรียมข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด ผนวกรวมกับข้อมูลค่าเฉลี่ยย้อนหลังที่คำนวณมาจากราคาปิด นั่นคือ Simple moving average (SMA) Exponential moving average (EMA) และ Weighted moving average (WMA) ที่ได้มาจากการคำนวณที่ได้มาจากไลบรารี TA-LIB ซึ่งเป็นไลบรารีที่ช่วยอำนวยความสะดวก ในการคำนวณข้อมูลในการวิเคราะห์ทางเทคนิค โดยเลือกใช้ค่าเฉลี่ยย้อนหลัง 5 วัน หรือจำนวนวันที่ตลาดเปิด ใน 1 สัปดาห์

การแปลงข้อมูล

เนื่องจากข้อมูลที่ใช้ในการวิเคราะห์มีค่าตั้งแต่หลักหน่วยจนถึงหลักร้อย ส่งผลให้เกิดความซับซ้อนในการคำนวณของระบบคอมพิวเตอร์ ผู้ศึกษาจึงแปลงข้อมูลให้มีค่าอยู่ระหว่าง 0 ถึง 1

การเรียนรู้ การทดสอบ

ในขั้นตอนนี้ ผู้ศึกษาได้แบ่งข้อมูลออกเป็น 2 ช่วง ช่วงที่ 1 คือข้อมูลที่ถูกใช้เพื่อการเรียนรู้ของโครงข่ายประสาทเทียม LSTM เพื่อนำไปทดสอบความสามารถในการทำนายกับข้อมูลชุดที่ 2 ซึ่งเป็นข้อมูลที่โครงข่ายประสาทเทียม LSTM ยังไม่เคยผ่านประสบการณ์มา โดยแบ่งเป็นสัดส่วน 80:20 ตามลำดับ

ตัวแปรที่ใช้

Lookback = 5

จากงานวิจัยของ Saud & shakya (2020) กล่าวว่า ค่าที่เหมาะสมกับโครงข่ายประสาทเทียมชนิด LSTM คือน้อยกว่า 5 และจากการศึกษาวิจัยของ ปิยศรีนัย ลิขิตาภรณ์ (2021) ในการทดสอบค่า lookback โดยใช้ค่าที่ทดสอบคือจำนวนเต็ม 1 ถึง 5 กับดัชนี S&P500 โดยค่า Lookback = 5 ให้ผลค่า Average RMSE น้อยที่สุด

Hidden unit = 7-13-7-3-1

จากงานวิจัยของ Chaigusin Chirathamjaree และ Clayden (2008) ได้ศึกษาการใช้โครงข่ายประสาทเทียมในการพยากรณ์ดัชนีตลาดหลักทรัพย์แห่งประเทศไทย โดยเปรียบเทียบประสิทธิภาพในการพยากรณ์ของโครงข่ายประสาทเทียมที่มีจำนวนชั้นของโครงข่ายแตกต่างกัน ประกอบด้วยโครงข่ายที่มีจำนวน 3, 4 และ 5 ชั้น จากการศึกษาพบว่า โมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 3 ชั้น คือ 7-3-1 โมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 4 ชั้น คือ 7-7-3-1 และโมเดลที่ดีที่สุดสำหรับโครงข่ายประสาทเทียม 5 ชั้น คือ 7-13-7-3-1 วัดประสิทธิภาพโดยใช้ MAPE ได้ 1.26594, 1.14719 และ 1.145789 ตามลำดับ

Epochs = 100 (จำนวนรอบการทำซ้ำที่ 100 ครั้ง)

จากการวิจัยของ Moghae และ Hamiche (2020) ได้ศึกษาการพยากรณ์ราคาตลาดหลักทรัพย์โดยใช้ LSTM RNN จากการใช้ข้อมูลราคาเปิดจาก 2 หลักทรัพย์ที่อยู่ในตลาดหลักทรัพย์นิวยอร์กคือ GOOGL และ NKE โดยที่ GOOGL ใช้ข้อมูลตั้งแต่วันที่ 19 สิงหาคม 2004 ถึงวันที่ 19 ธันวาคม 2019 และ NKE ใช้ข้อมูลตั้งแต่วันที่ 4 มกราคม 2010 ถึงวันที่ 19 ธันวาคม 2019 โดยแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 80 และข้อมูลทดสอบร้อยละ 20 และ Epochs สำหรับข้อมูลสำหรับทดลองเท่ากับ 12 25 50 และ 100 จากการศึกษาพบว่าความแม่นยำในการพยากรณ์เพิ่มขึ้นหาก Epochs เพิ่มขึ้น ดังนั้น ผู้วิจัยจึงเลือกใช้ ค่า Epochs = 100 เนื่องจากให้ผลที่น่าพอใจและไม่ใช้เวลานานจนเกินไป

Loss = Mean squared error

การใช้ Gradient descent ร่วมกับ MAE นั้นมีผลต่างของค่าความผิดพลาดที่เท่า ๆ กันส่งผลให้ค่าที่อัปเดตเปลี่ยนแปลงไปแบบคงที่ และเป็นผลสืบเนื่องให้ไม่สามารถหาค่าที่มีความผิดพลาดต่ำที่สุดได้ แต่การใช้ Gradient descent ร่วมกับ MSE สามารถหาค่าที่มีความผิดพลาดต่ำที่สุดได้ เนื่องจากผลต่างของค่าความผิดพลาดเปลี่ยนแปลงไปเรื่อย ๆ และค่อย ๆ ลดลงเมื่อเข้าใกล้จุดต่ำสุด ส่งผลให้ MSE เหมาะสมกับการใช้ร่วมกับ Neural Network

3.4 การวัดประสิทธิภาพความแม่นยำ

หลังจากที่ผ่านกระบวนการเรียนรู้เกี่ยวกับการใช้โมเดล LSTM เพื่อใช้ในการพยากรณ์ราคาหุ้นล่วงหน้า 1 วัน และทำการทดสอบโมเดลนี้เสร็จสิ้น การวัดประสิทธิภาพของโมเดล LSTM สำหรับการศึกษาได้ใช้วิธี Mean square error ในการวัดความคลาดเคลื่อนจากการพยากรณ์ เนื่องจากข้อมูลราคาปิดที่พยากรณ์เป็นข้อมูลประเภทอนุกรมเวลา ทำให้ค่าความคลาดเคลื่อนที่เกิดขึ้นเป็นข้อมูลที่ไม่ได้มีการแจกแจงแบบปกติ ซึ่งเป็นข้อกำหนดสำหรับการทดสอบทางสถิติอาจส่งผลให้การวิเคราะห์ไม่น่าเชื่อถือ จึงเลือกใช้ Mean square error เนื่องจากเป็นที่นิยมสำหรับการวิเคราะห์แบบข้อมูลอนุกรมเวลาและสามารถสะท้อนความแปรปรวนของความคลาดเคลื่อนได้ดี

3.5 วิธีการศึกษาเพิ่มเติม

การศึกษาเพิ่มเติมนี้เป็นการทดลองการพยากรณ์ราคาปิดของหุ้นตัวหนึ่งจากการนำข้อมูลของหุ้นตัวอื่นมาใช้เป็นอินพุตของโมเดล LSTM โดยในการทดลองนี้เราทำการพยากรณ์ราคาปิดหุ้นของ PTT, SPCG และ TTW โดยที่การพยากรณ์ราคาปิดหุ้น PTT เรากำหนดอินพุตโดยใช้ข้อมูลราคาจาก SPCG และ TTW การพยากรณ์ราคาปิดหุ้น SPCG เรากำหนดอินพุตโดยใช้ข้อมูลราคาจาก PTT และ TTW และการพยากรณ์ราคาปิดหุ้น TTW เรากำหนดอินพุตโดยใช้ข้อมูลราคาจาก PTT และ SPCG

ตัวแปรต้น

ข้อมูลราคาเปิด (O) ราคาสูงสุด (H) ราคาต่ำสุด (L) ราคาปิด (C) และปริมาณการซื้อขาย (Volume)

ตัวแปรตาม

ราคาปิดที่ถูกพยากรณ์ในวันถัดไป (Predicted Price)

ตัวแปรควบคุม

ระยะเวลา (วัน)

การศึกษานี้ได้ใช้ข้อมูลราคาหลักทรัพย์จาก YAHOO FINANCE โดยได้รวบรวมข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย โดยใช้เป็นข้อมูลรายวัน ย้อนหลังเป็นเวลา 10 ปี ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020

ผู้วิจัยได้ใช้ โปรแกรม Google Colab ซึ่งเป็นบริการของ Google ที่ให้บริการโฮสต์ โปรแกรม Jupyter Notebook สำหรับภาษา Python

การเตรียมข้อมูล

ภายหลังจากนำเข้าข้อมูลราคาและปริมาณการซื้อขายของหุ้นสามัญ ผู้วิจัยได้ทำการเตรียมข้อมูลที่จำเป็นในการทดสอบโดยเตรียมข้อมูลราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย

การแปลงข้อมูล

เนื่องจากข้อมูลที่ใช้ในการวิเคราะห์มีค่าตั้งแต่หลักหน่วยจนถึงหลักร้อย ส่งผลให้เกิดความซับซ้อนในการคำนวณของระบบคอมพิวเตอร์ ผู้ศึกษาจึงแปลงข้อมูลให้มีค่าอยู่ระหว่าง 0 ถึง 1

การเรียนรู้ การทดสอบ

ในขั้นตอนนี้ ผู้ศึกษาได้แบ่งข้อมูลออกเป็น 2 ช่วง ช่วงที่ 1 คือข้อมูลที่ถูกใช้เพื่อการเรียนรู้ของโครงข่ายประสาทเทียม LSTM เพื่อนำไปทดสอบความสามารถในการทำนายกับข้อมูลชุดที่ 2 ซึ่งเป็นข้อมูลที่โครงข่ายประสาทเทียม LSTM ยังไม่เคยผ่านประสบการณ์มา โดยแบ่งเป็นสัดส่วน 80:20 ตามลำดับ

ตัวแปรที่ใช้

Lookback = 5

Hidden unit = 10-13-7-3-1

Epochs = 100 (จำนวนรอบการทำซ้ำที่ 100 ครั้ง)

Loss = Mean squared error

บทที่ 4 ผลการวิเคราะห์ข้อมูล

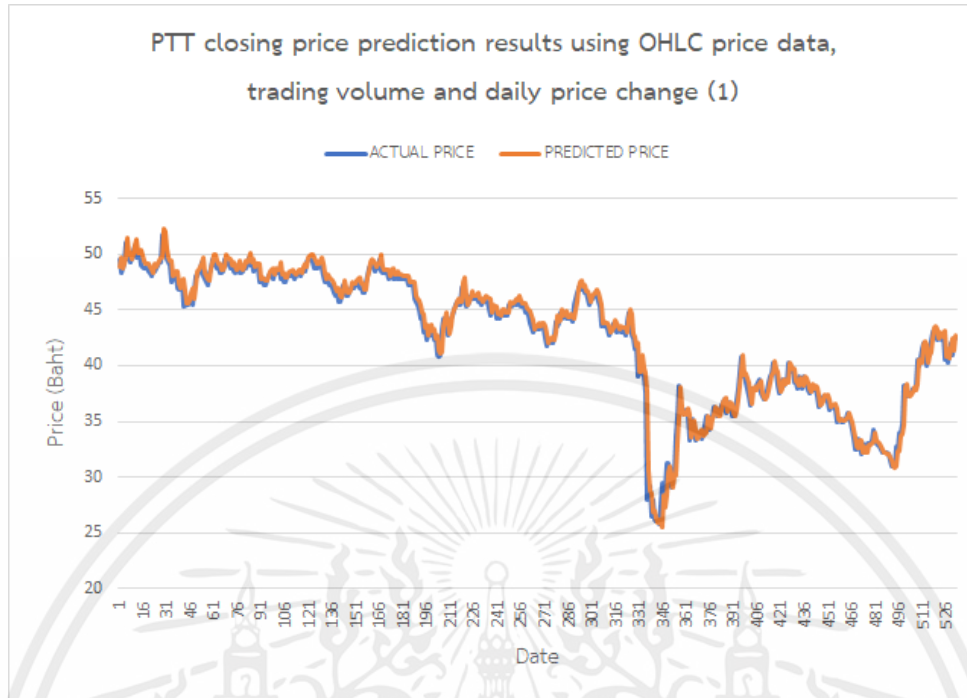
จากการทดสอบการพยากรณ์ราคาปิดวันถัดไปของหุ้นสามัญทั้ง 3 คือ PTT SPCG และ TTW ได้ผลการพยากรณ์ดังนี้

4.1 ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วัน

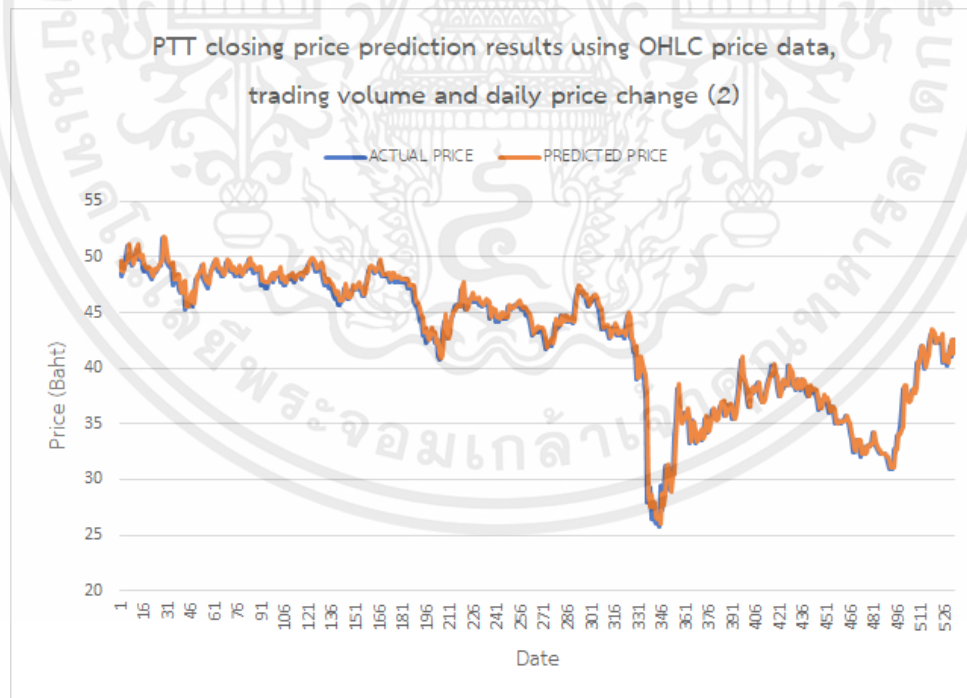
ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย และราคาที่เปลี่ยนแปลง (รายวัน) รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.839608 ดังตารางที่ 4.1 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.1, 4.2 และ 4.3 ตามลำดับ

ตารางที่ 4.1 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคา

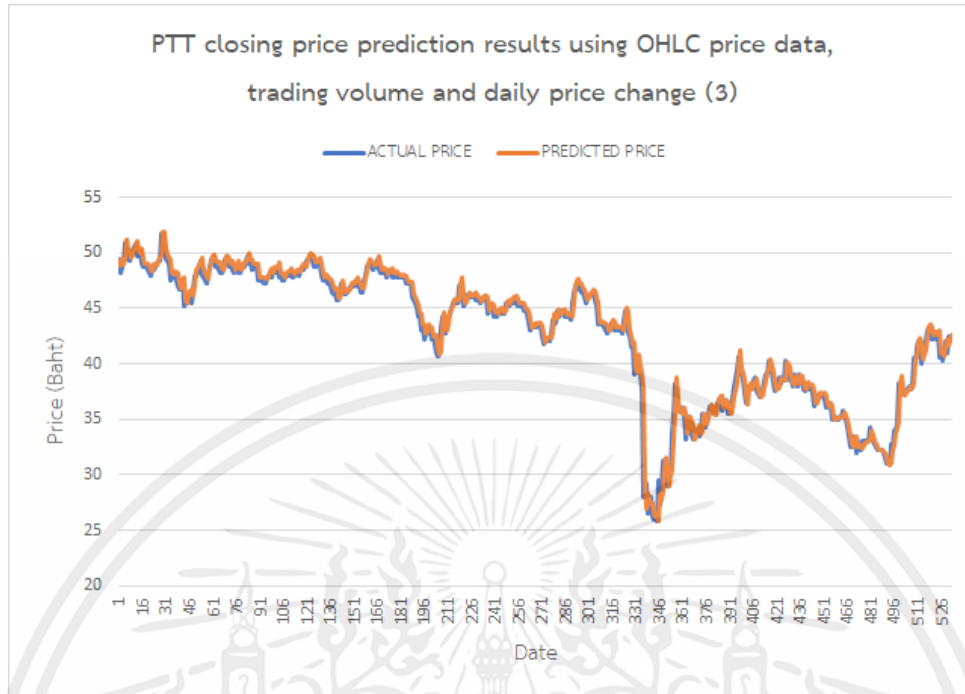
Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
PTT	1	0.248565	0.869657	0.005669
	2	0.202731	0.807739	0.004828
	3	0.227453	0.841429	0.005289
	Avg.	0.22625	0.839608	0.005262



รูปที่ 4.1 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 1



รูปที่ 4.2 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 2

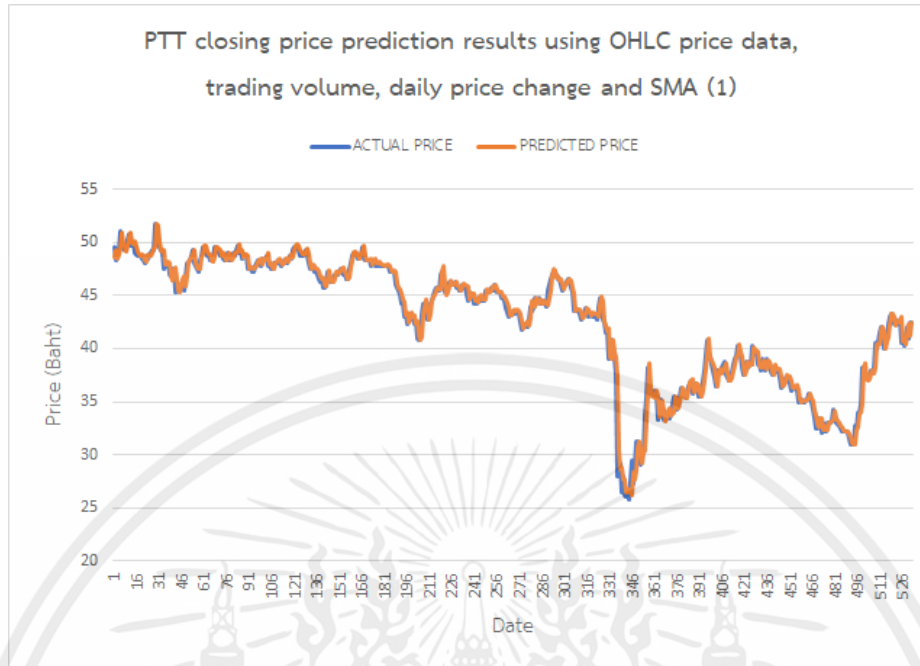


รูปที่ 4.3 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (PTT) รอบที่ 3

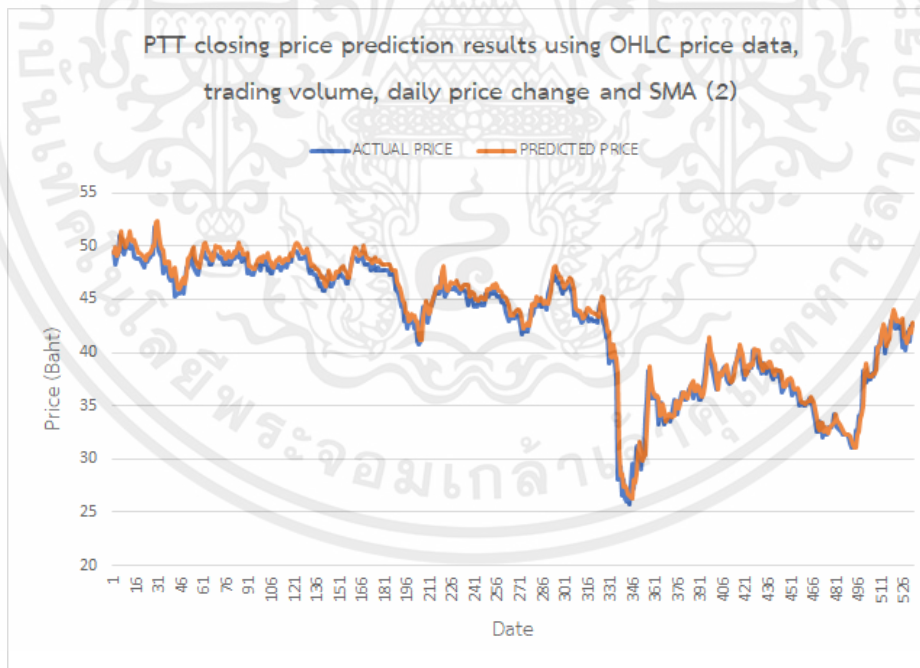
ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ SMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.994671 ดังตารางที่ 4.2 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.4, 4.5 และ 4.6 ตามลำดับ

ตารางที่ 4.2 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ SMA

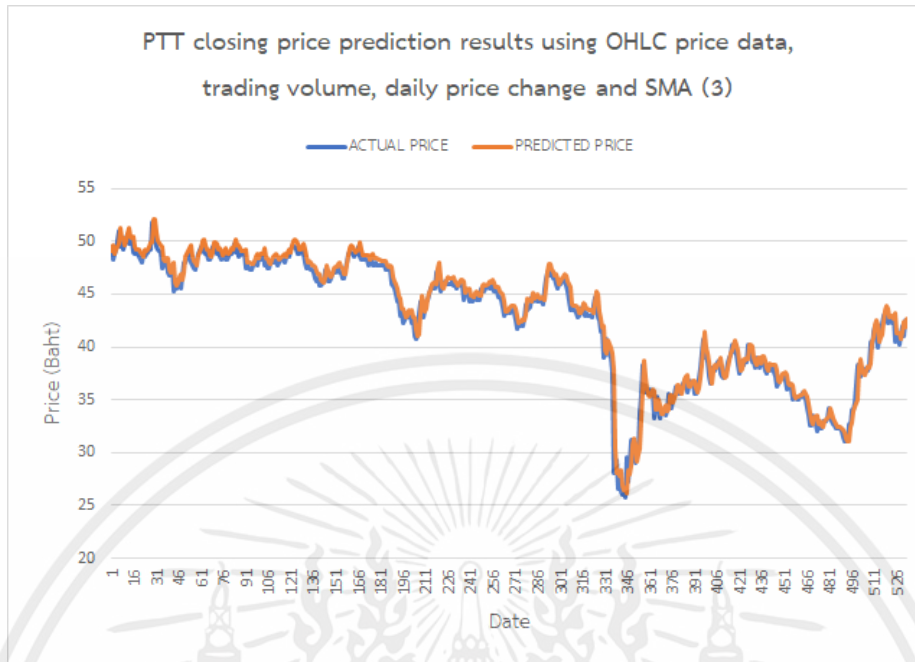
Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
PTT	1	0.096249	0.771108	0.002507
	2	0.483516	1.062159	0.011323
	3	0.394387	0.960927	0.009333
	Avg.	0.324717	0.994671	0.007721



รูปที่ 4.4 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 1



รูปที่ 4.5 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 2

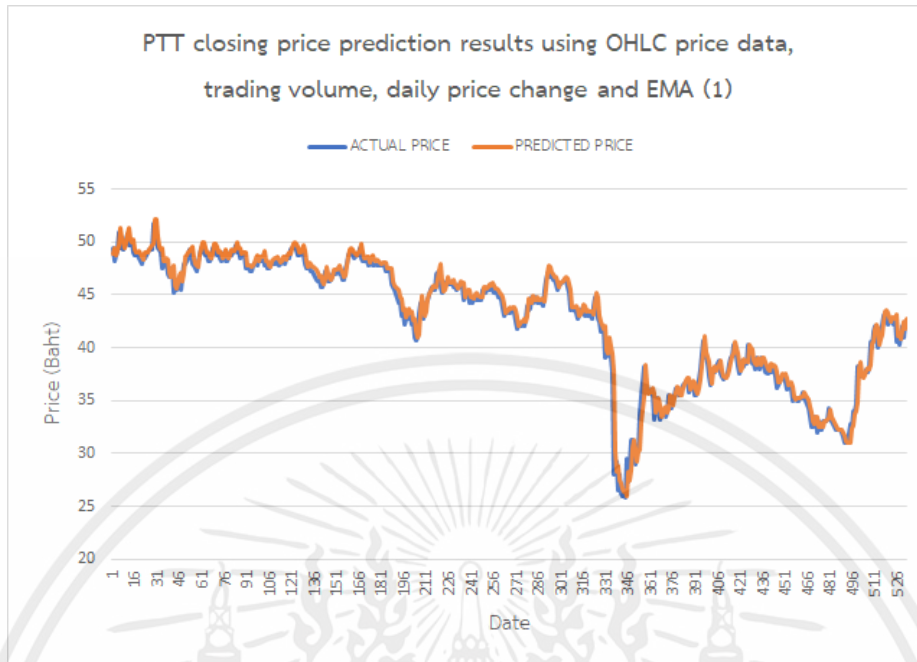


รูปที่ 4.6 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (PTT) รอบที่ 3

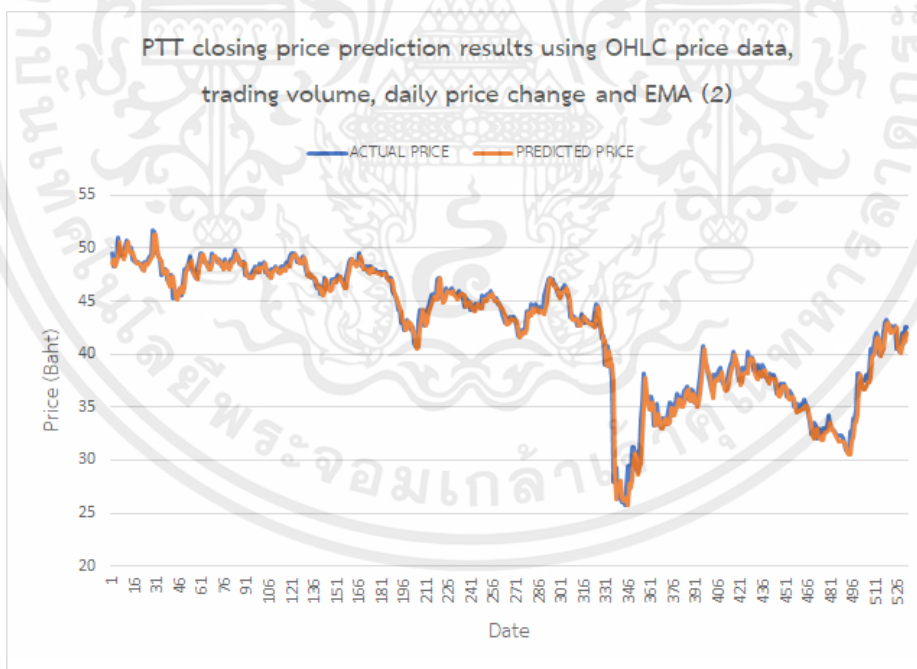
ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ EMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.852602 ดังตารางที่ 4.3 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.7, 4.8 และ 4.9 ตามลำดับ

ตารางที่ 4.3 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ EMA

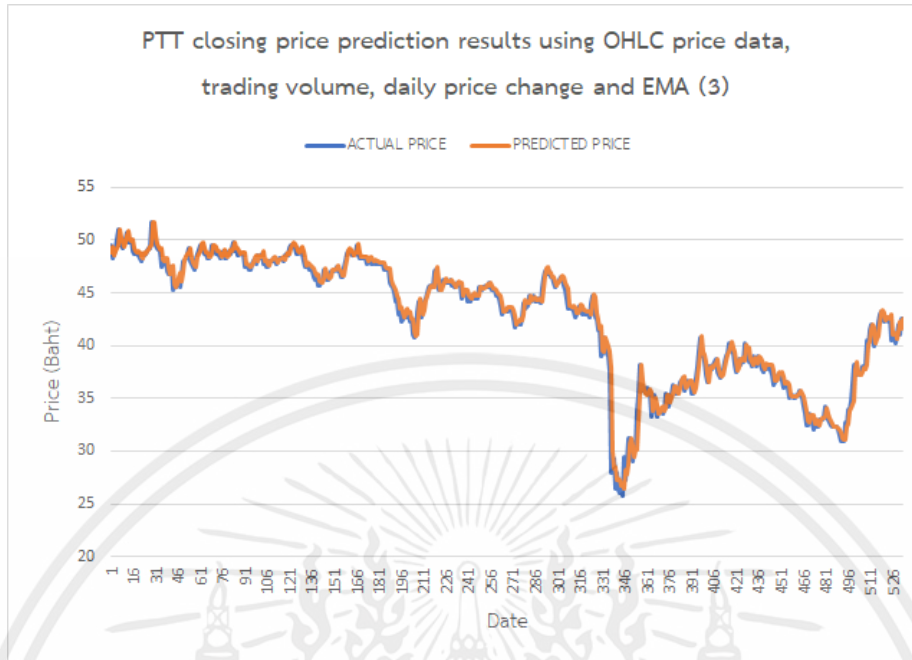
Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
PTT	1	0.297938	0.885805	0.0071
	2	-0.25084	0.86947	-0.00607
	3	0.165836	0.802532	0.004178
	Avg.	0.070978	0.852602	0.001736



รูปที่ 4.7 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 1



รูปที่ 4.8 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 2

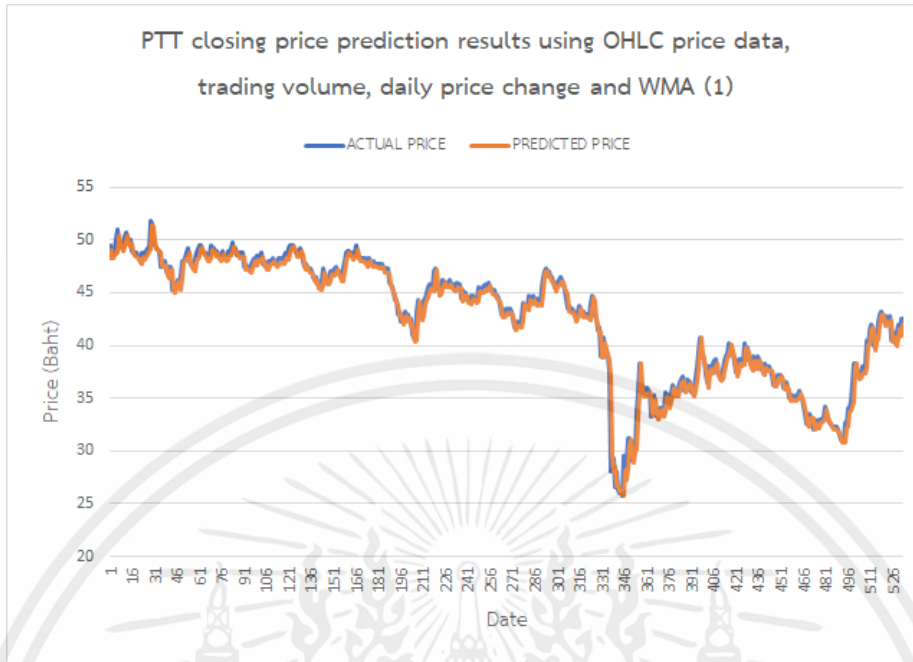


รูปที่ 4.9 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (PTT) รอบที่ 3

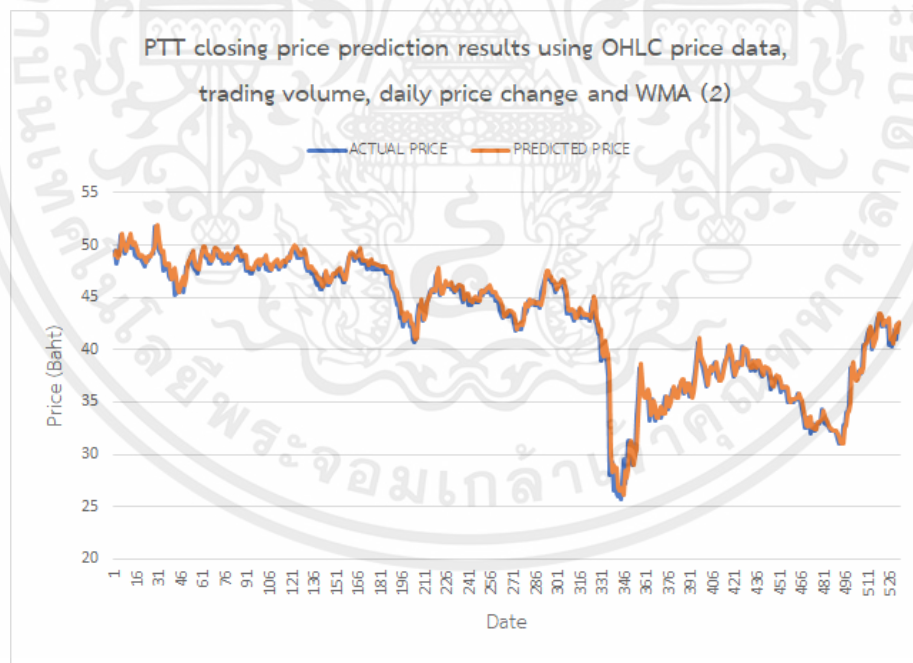
ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ WMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.854980 ดังตารางที่ 4.4 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.10, 4.11 และ 4.12 ตามลำดับ

ตารางที่ 4.4 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ WMA

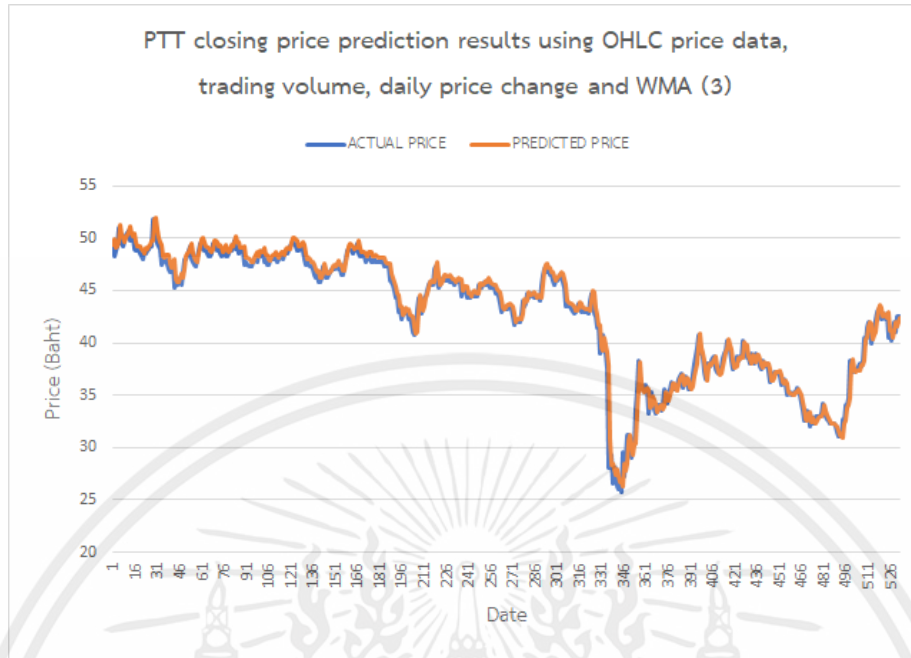
Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
PTT	1	-0.31386	0.839916	-0.0072
	2	0.252742	0.851058	0.006158
	3	0.248892	0.873967	0.005737
	Avg.	0.062591	0.854980	0.001565



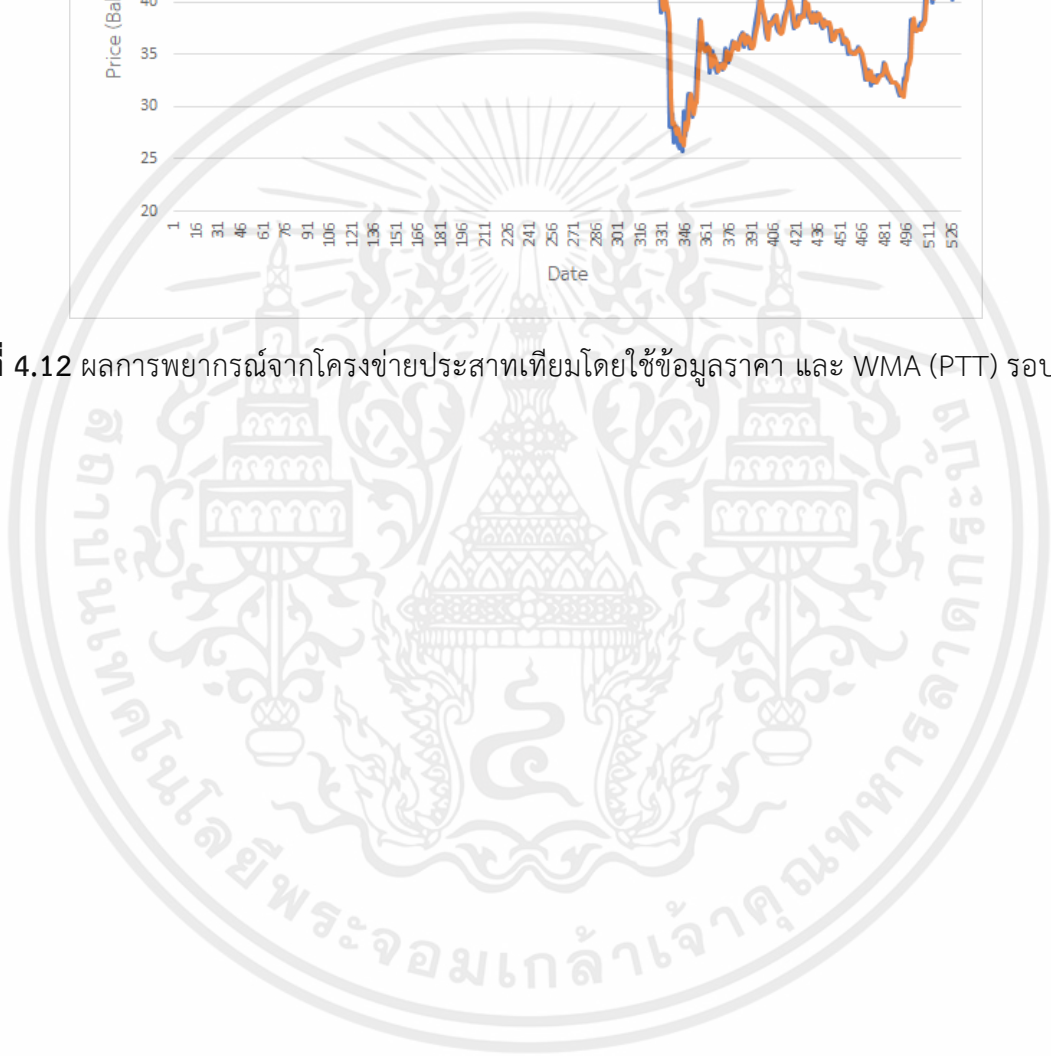
รูปที่ 4.10 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 1



รูปที่ 4.11 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 2



รูปที่ 4.12 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (PTT) รอบที่ 3

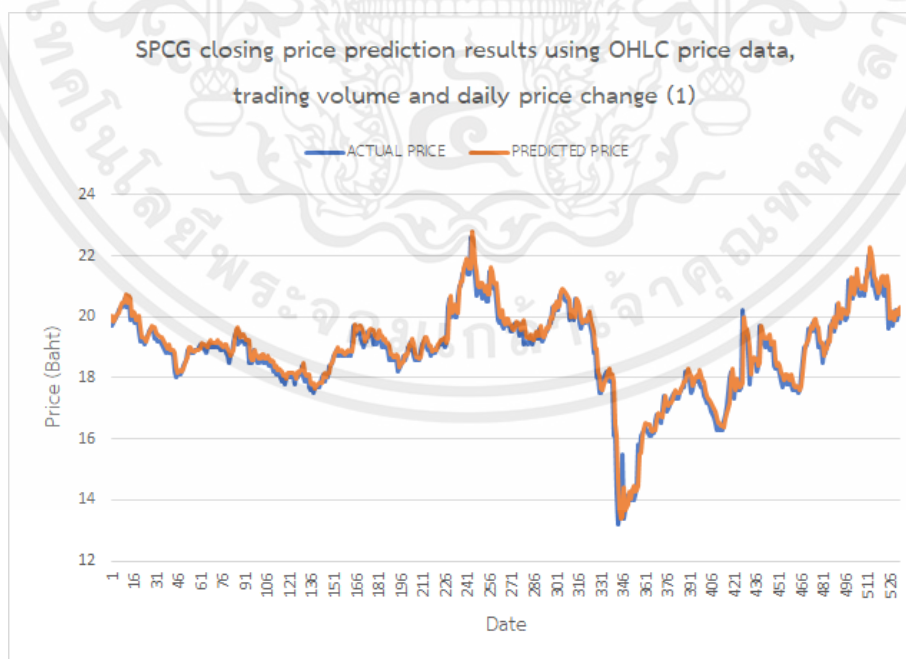


4.2 ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วัน

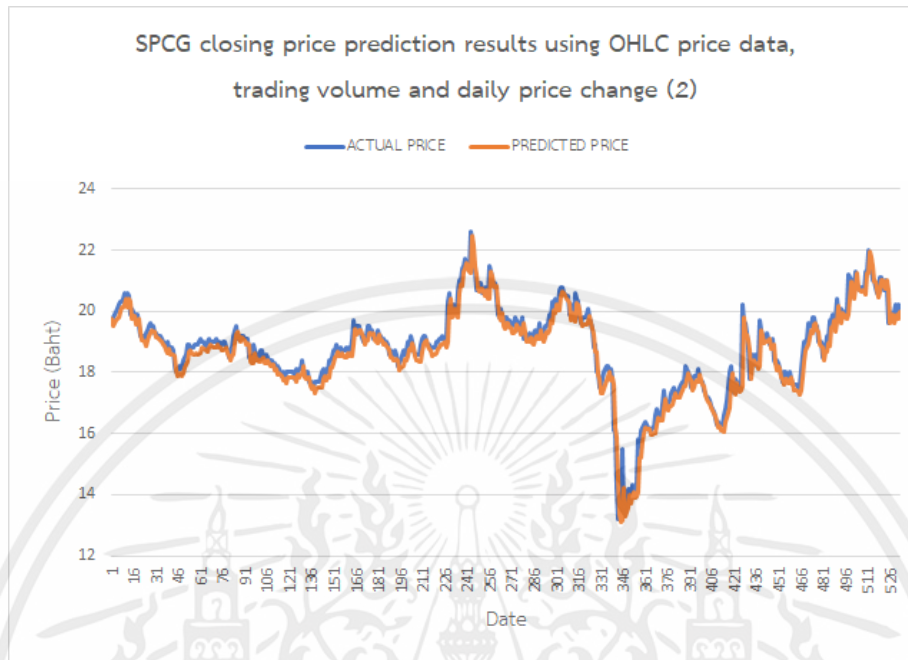
ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย และราคาที่เปลี่ยนแปลง (รายวัน) รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.150718 ดังตารางที่ 4.5 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.13, 4.14 และ 4.15 ตามลำดับ

ตารางที่ 4.5 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคา

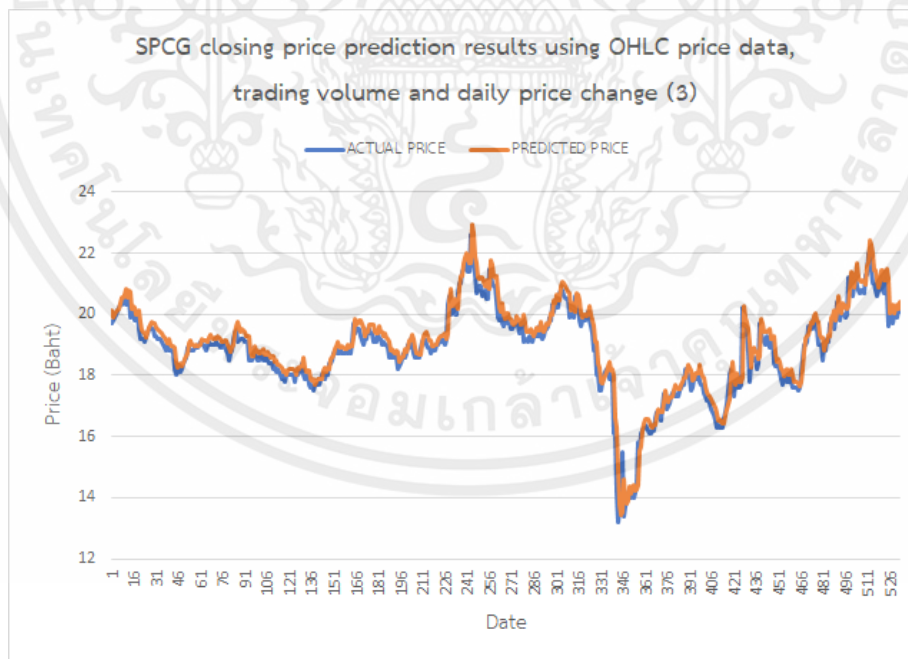
Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
SPCG	1	0.128508	0.132908	0.006961
	2	-0.18088	0.150369	-0.00958
	3	0.213522	0.168878	0.01146
	Avg.	0.053717	0.150718	0.002947



รูปที่ 4.13 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 1



รูปที่ 4.14 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 2

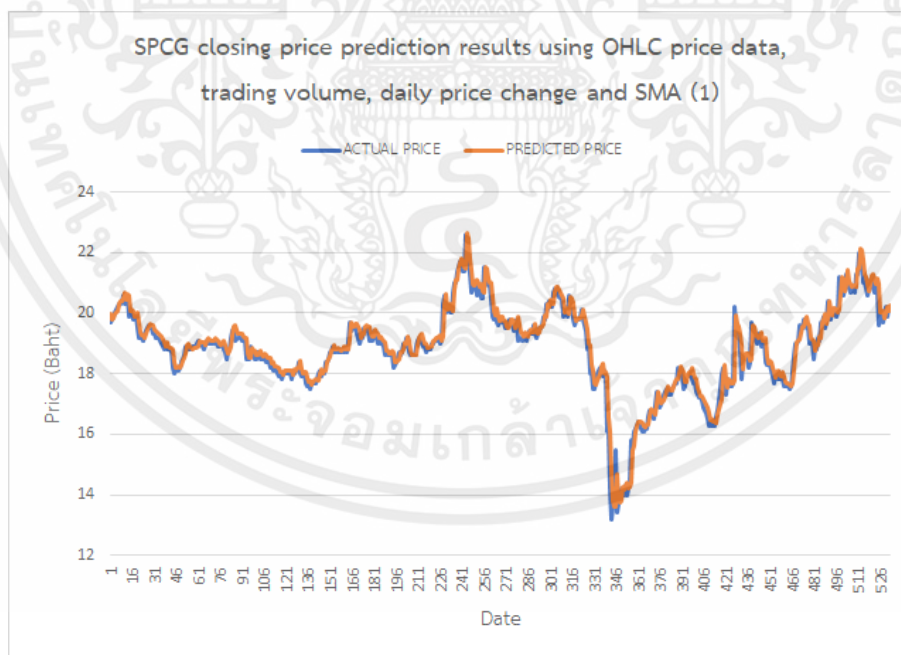


รูปที่ 4.15 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (SPCG) รอบที่ 3

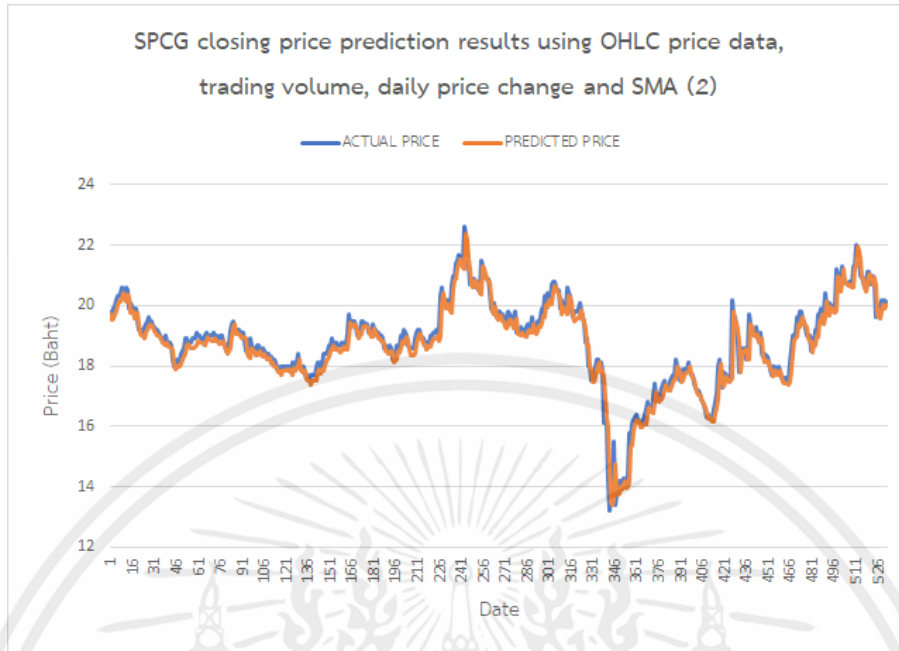
ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ SMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.127071 ดังตารางที่ 4.6 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.16, 4.17 และ 4.18 ตามลำดับ

ตารางที่ 4.6 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ SMA

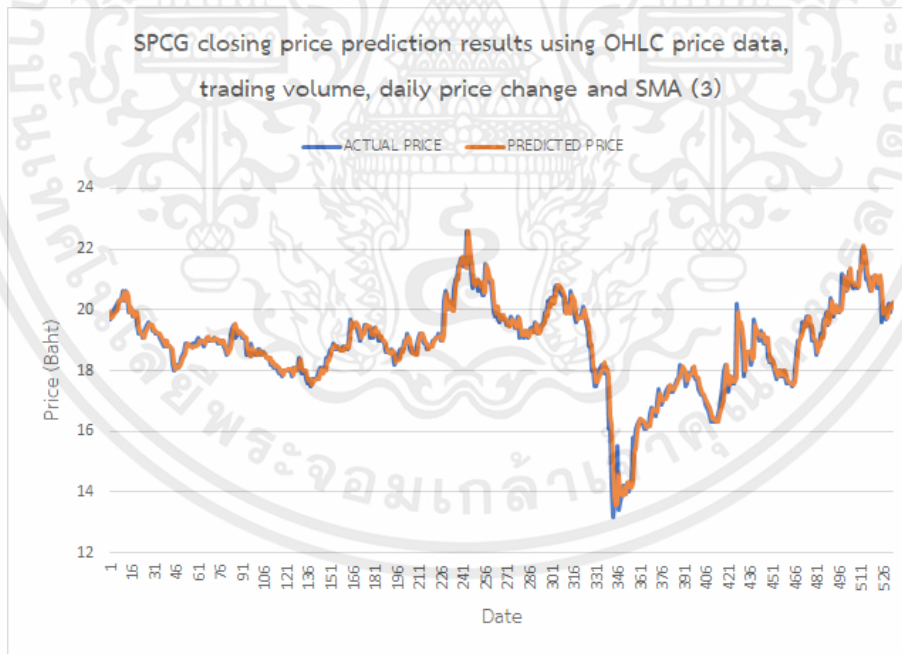
Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.096545	0.12505	0.005337
	2	-0.151640	0.138565	-0.007872
	3	0.014058	0.117596	0.000973
	Avg.	-0.01368	0.127071	-0.00052



รูปที่ 4.16 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 1



รูปที่ 4.17 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 2

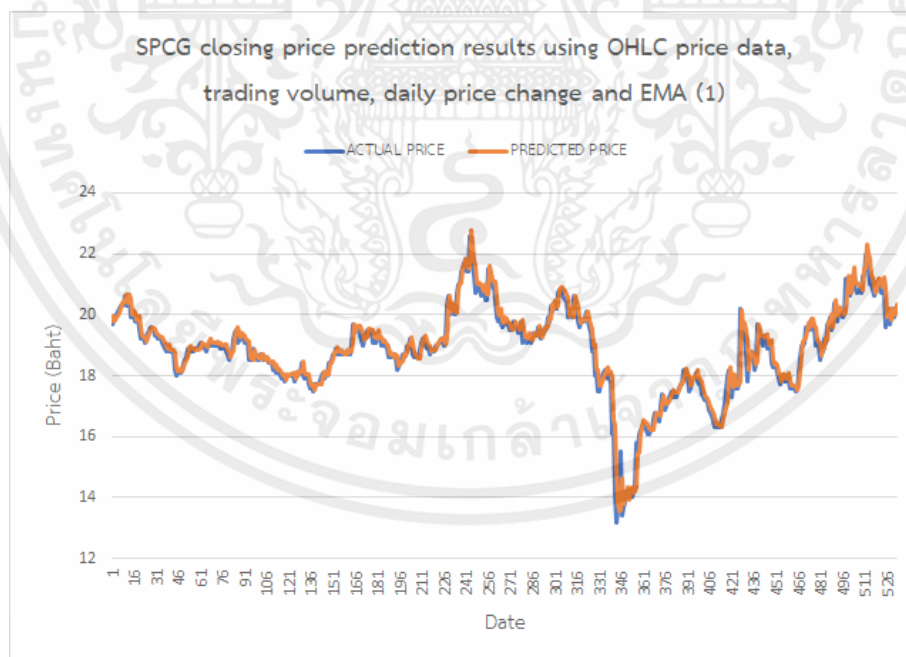


รูปที่ 4.18 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (SPCG) รอบที่ 3

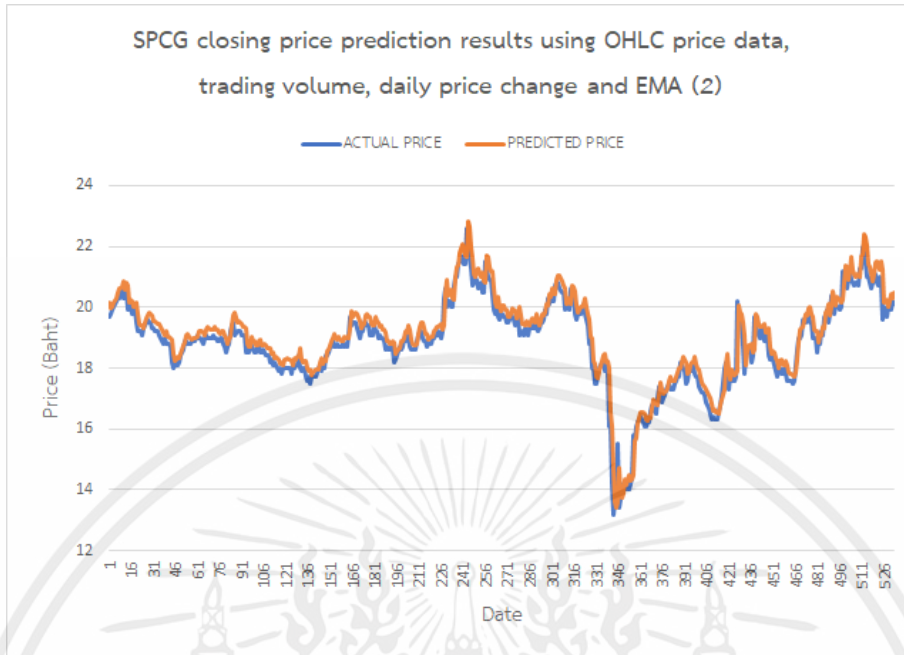
ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ EMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.170201 ดังตารางที่ 4.7 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.19, 4.20 และ 4.21 ตามลำดับ

ตารางที่ 4.7 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ EMA

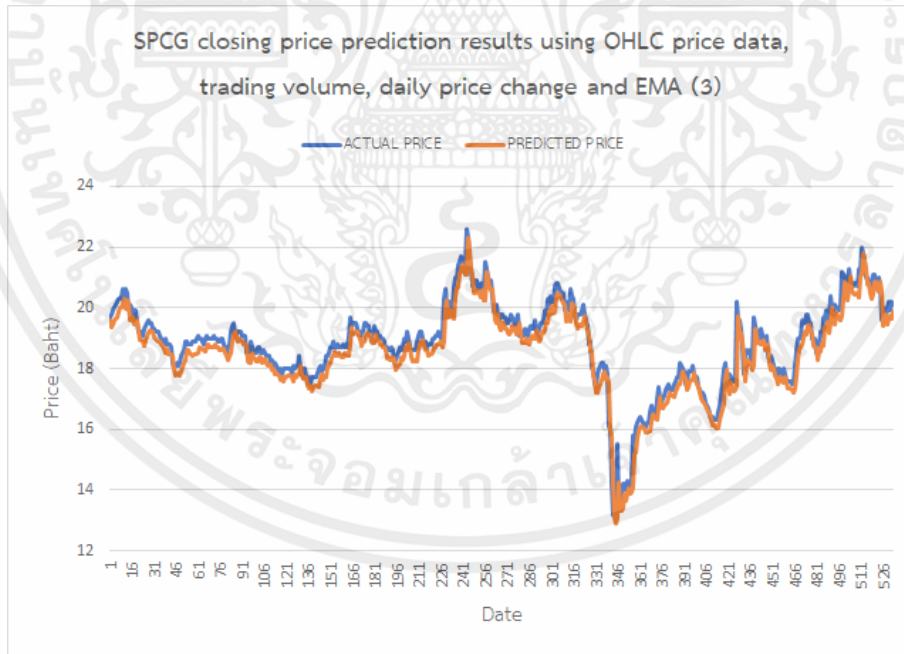
Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.083518	0.12887	0.004616
	2	0.242506	0.177027	0.013
	3	-0.30041	0.204705	-0.01586
	Avg.	0.008538	0.170201	0.000585



รูปที่ 4.19 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 1



รูปที่ 4.20 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 2

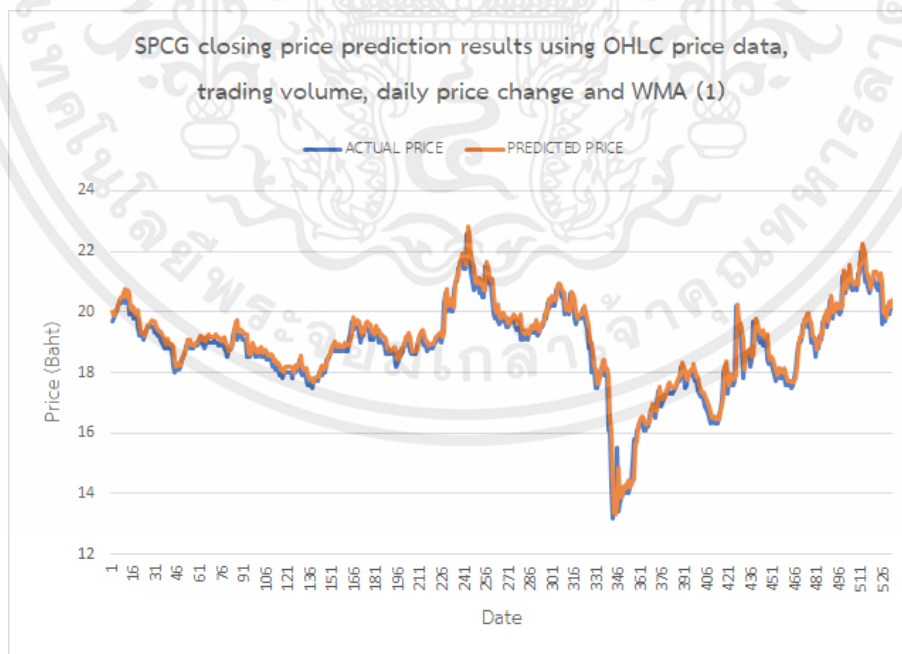


รูปที่ 4.21 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (SPCG) รอบที่ 3

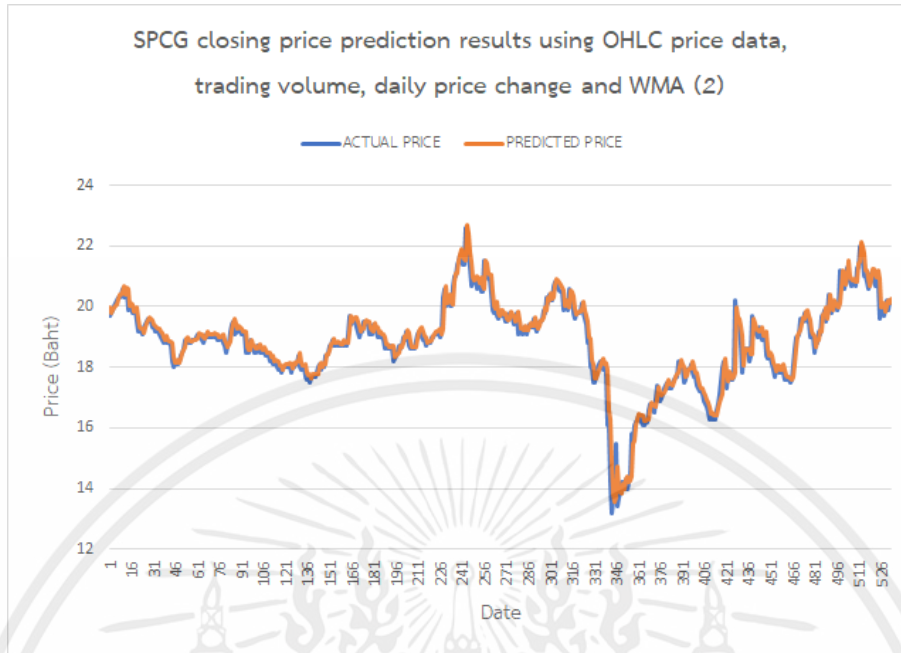
ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ WMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.127516 ดังตารางที่ 4.8 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.22, 4.23 และ 4.24 ตามลำดับ

ตารางที่ 4.8 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาพร้อมกับ WMA

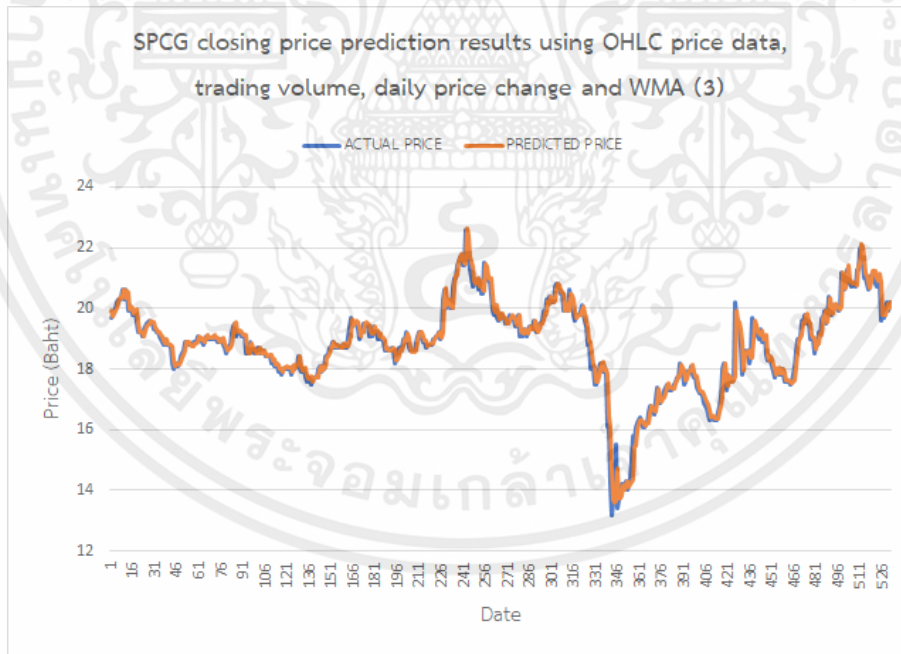
Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.164917	0.141415	0.008913
	2	0.092519	0.125692	0.005144
	3	0.033118	0.11544	0.001933
	Avg.	0.096851	0.127516	0.00533



รูปที่ 4.22 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 1



รูปที่ 4.23 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 2



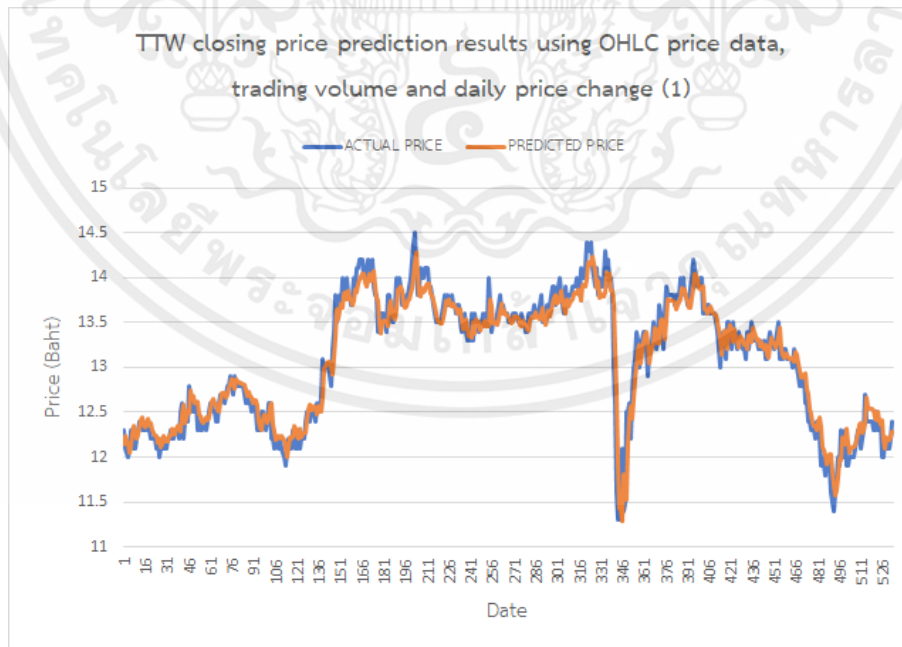
รูปที่ 4.24 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (SPCG) รอบที่ 3

4.3 ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วัน

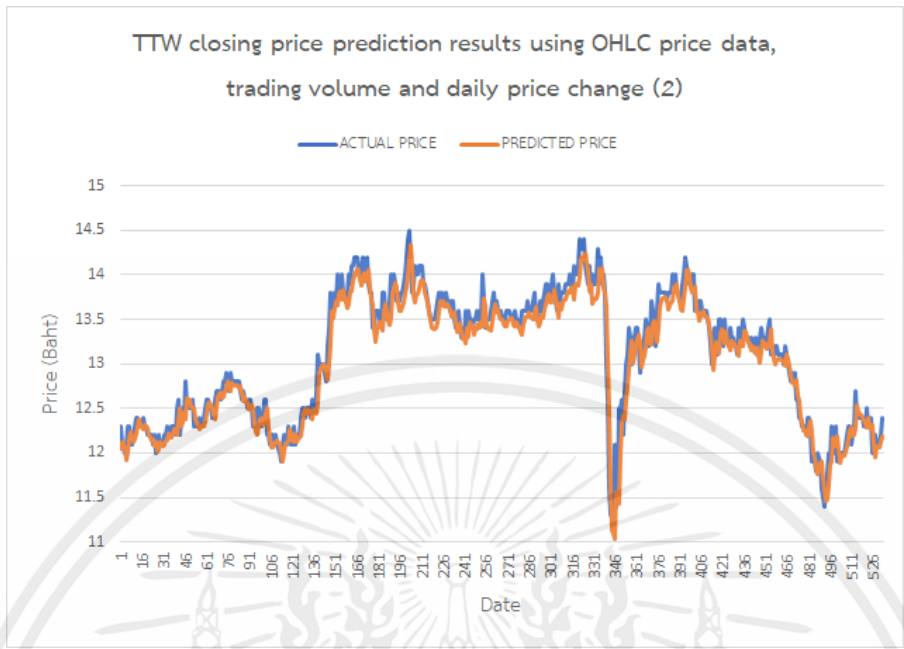
ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย และราคาที่เปลี่ยนแปลง (รายวัน) รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.034332 ดังตารางที่ 4.9 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.25, 4.26 และ 4.27 ตามลำดับ

ตารางที่ 4.9 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคา

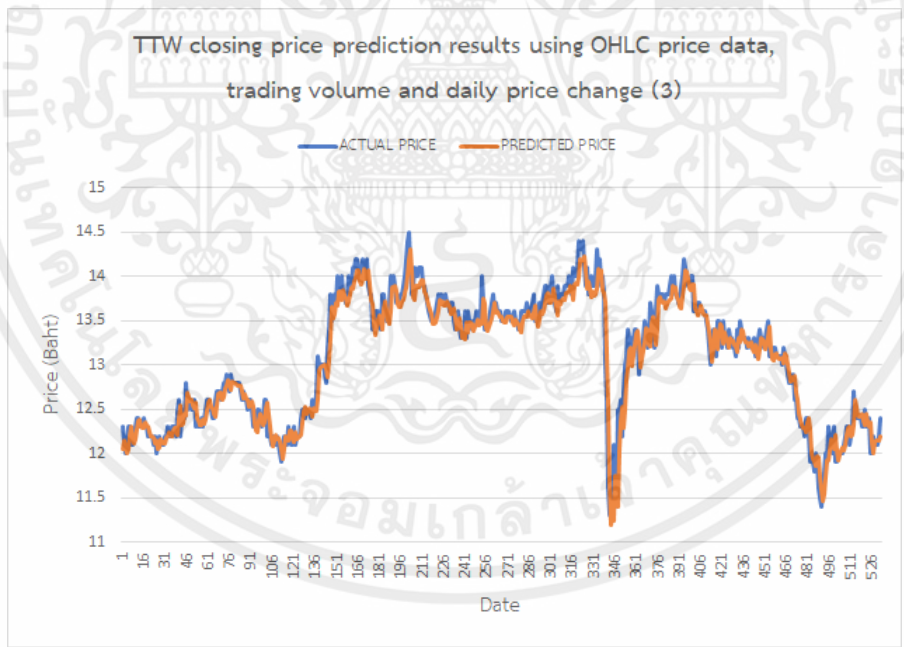
Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
TTW	1	-0.00122	0.032864	0.000257
	2	-0.07565	0.037104	-0.00558
	3	-0.0417	0.033027	-0.00297
	Avg.	-0.03952	0.034332	-0.00276



รูปที่ 4.25 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 1



รูปที่ 4.26 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 2

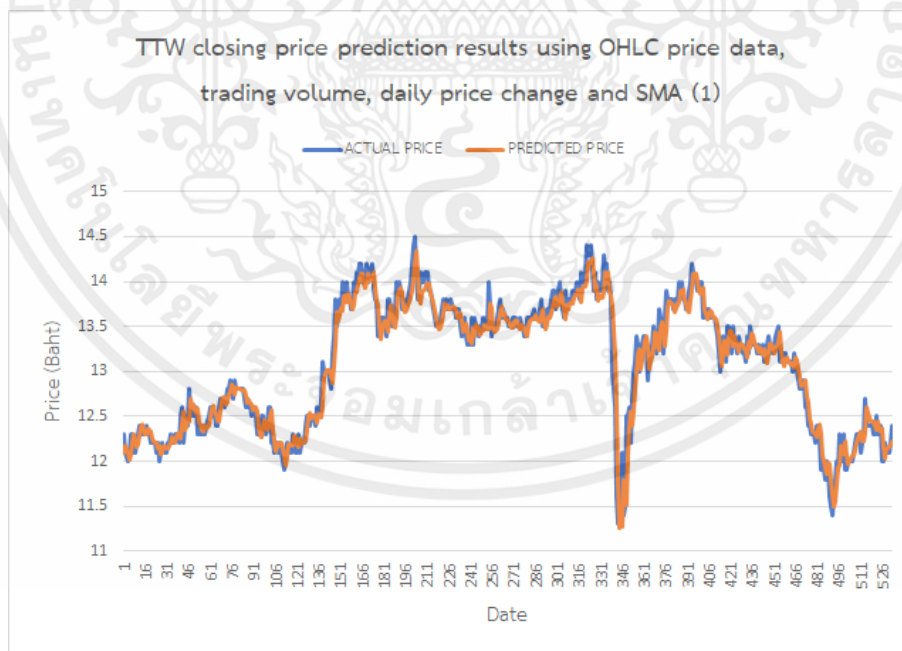


รูปที่ 4.27 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา (TTW) รอบที่ 3

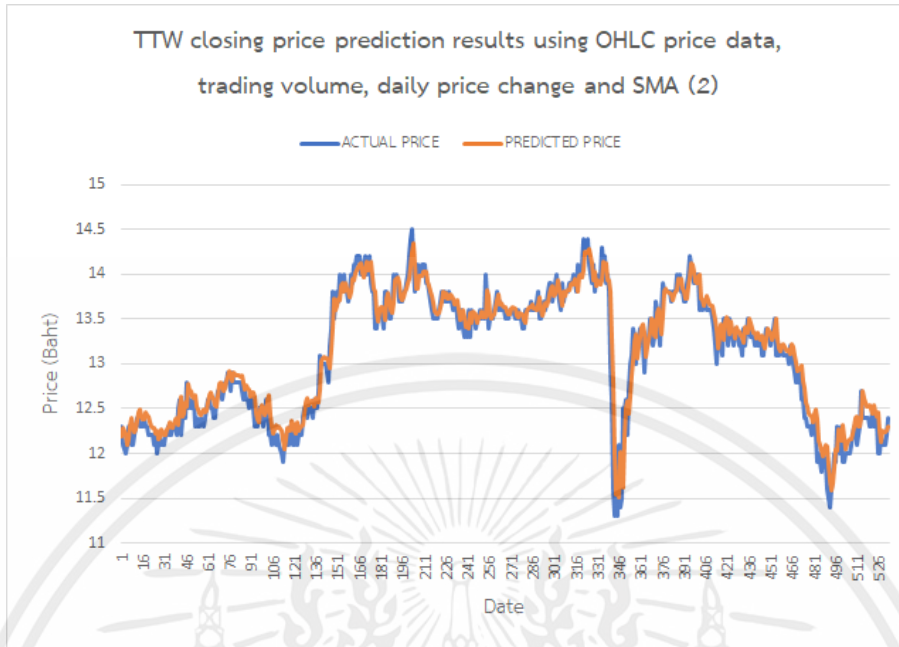
ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ SMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.033400 ดังตารางที่ 4.10 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.28, 4.29 และ 4.30 ตามลำดับ

ตารางที่ 4.10 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ SMA

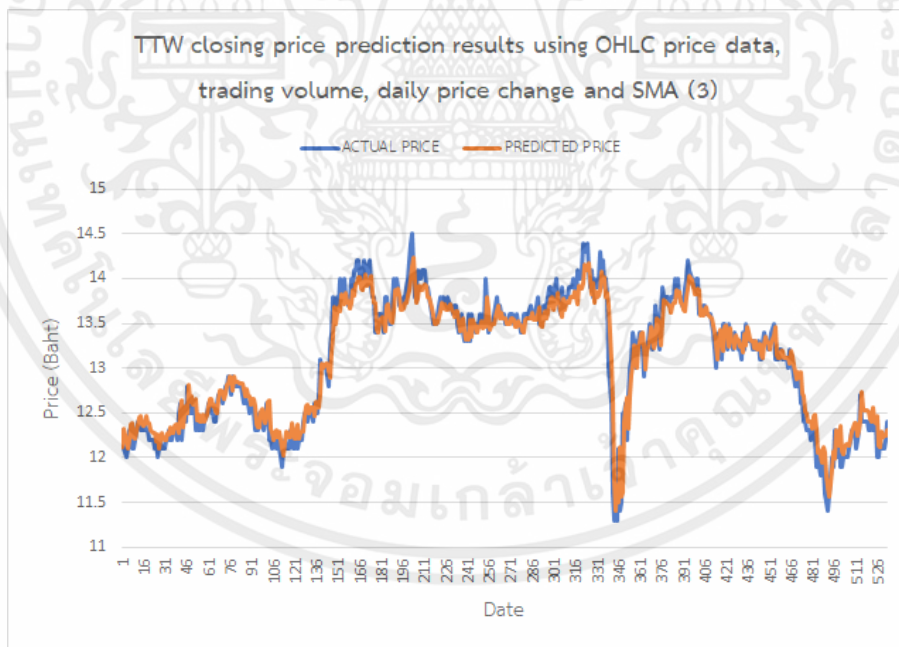
Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.01868	0.031612	-0.0012
	2	0.053199	0.035205	0.00439
	3	-0.000077	0.033384	0.000415
	Avg.	0.011481	0.033400	0.001202



รูปที่ 4.28 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 1



รูปที่ 4.29 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 2

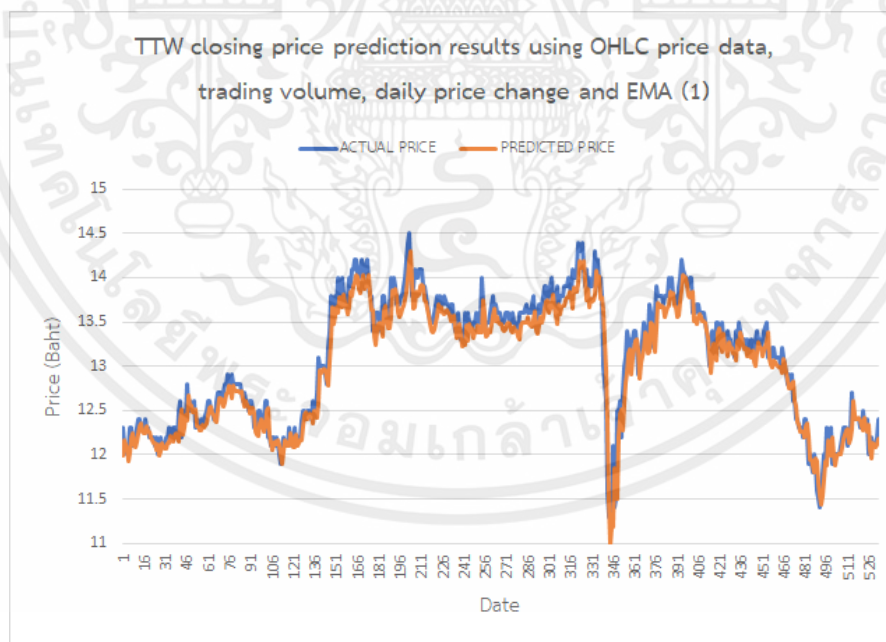


รูปที่ 4.30 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ SMA (TTW) รอบที่ 3

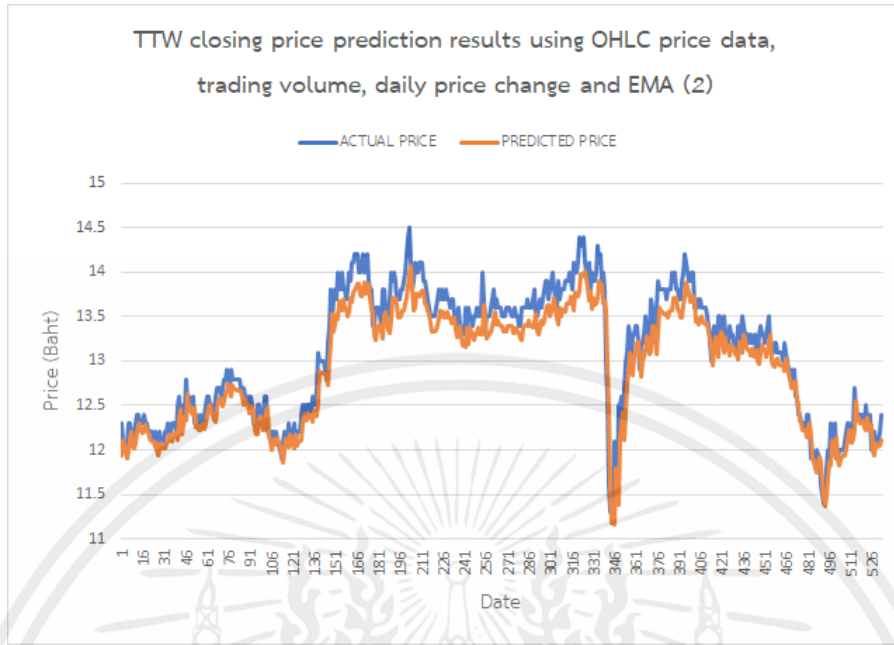
ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ EMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.058776 ดังตารางที่ 4.11 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.31, 4.32 และ 4.33 ตามลำดับ

ตารางที่ 4.11 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ EMA

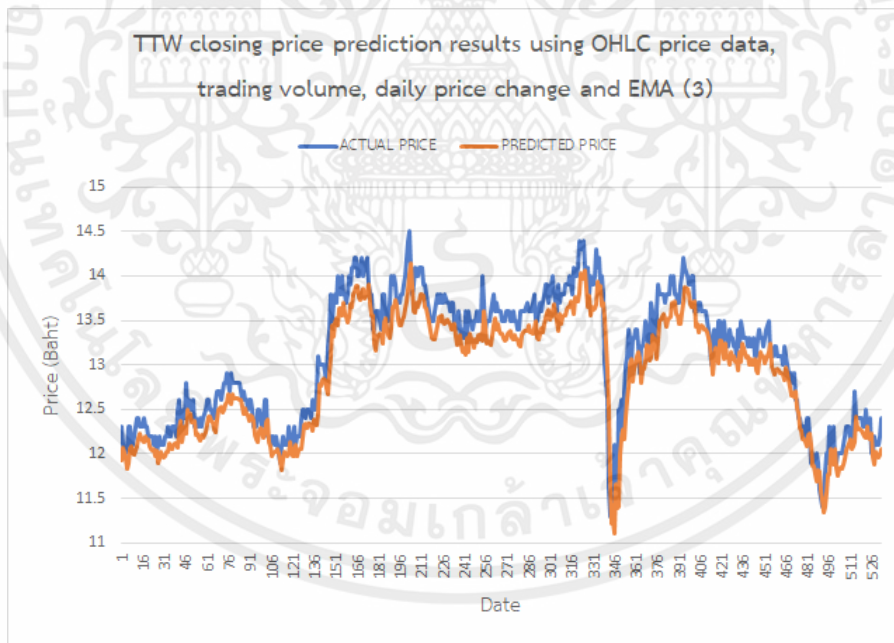
Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.09468	0.039306	-0.00699
	2	-0.16786	0.061295	-0.0125
	3	-0.21088	0.075727	-0.01586
	Avg.	-0.15781	0.058776	-0.01178



รูปที่ 4.31 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 1



รูปที่ 4.32 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 2

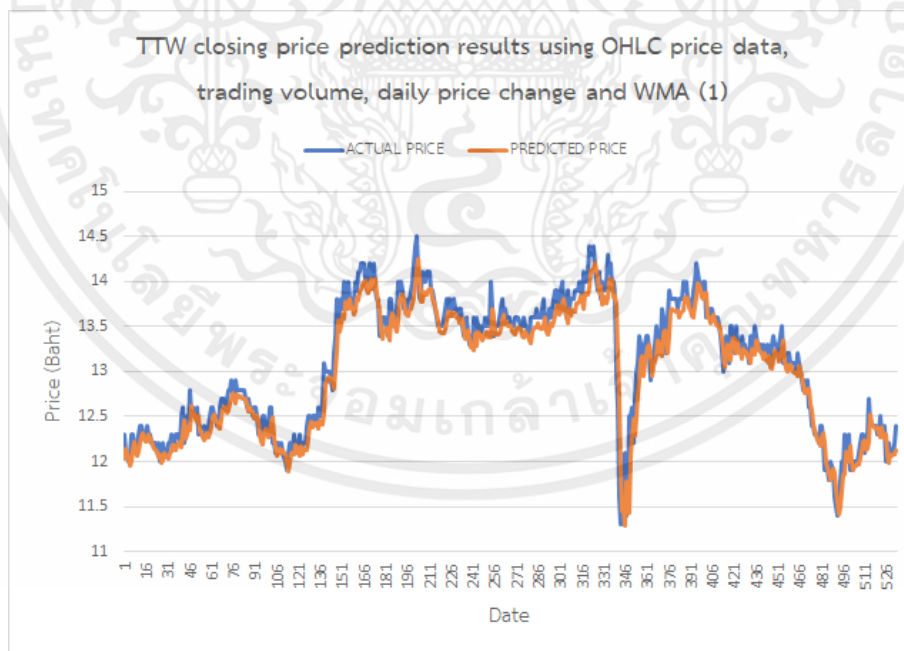


รูปที่ 4.33 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ EMA (TTW) รอบที่ 3

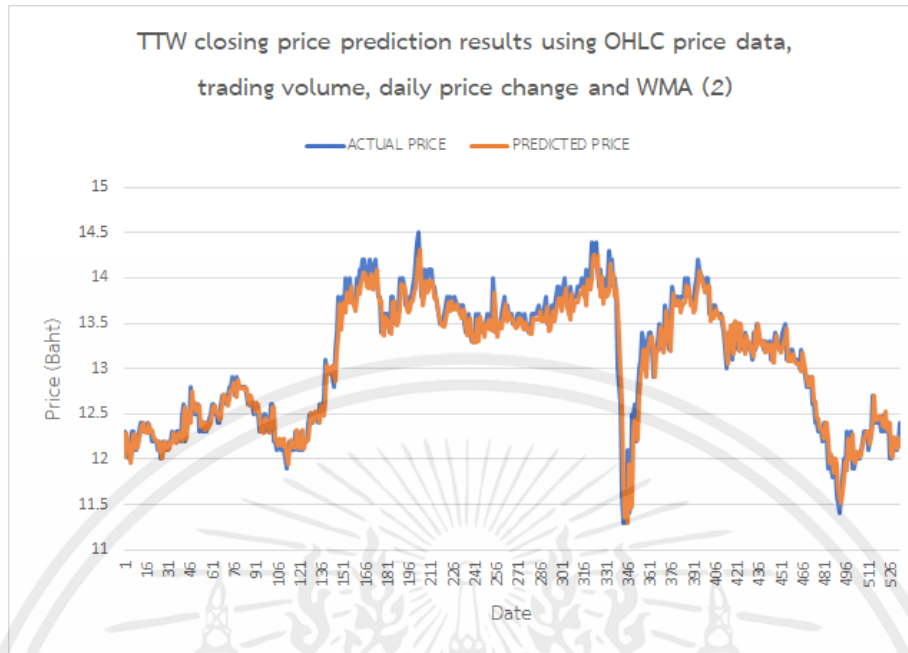
ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคา ปริมาณการซื้อขาย ราคาที่เปลี่ยนแปลง (รายวัน) และค่าเฉลี่ยเคลื่อนที่แบบ WMA รอบที่ 1, 2 และ 3 ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 0.044283 ดังตารางที่ 4.12 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.34, 4.35 และ 4.36 ตามลำดับ

ตารางที่ 4.12 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ WMA

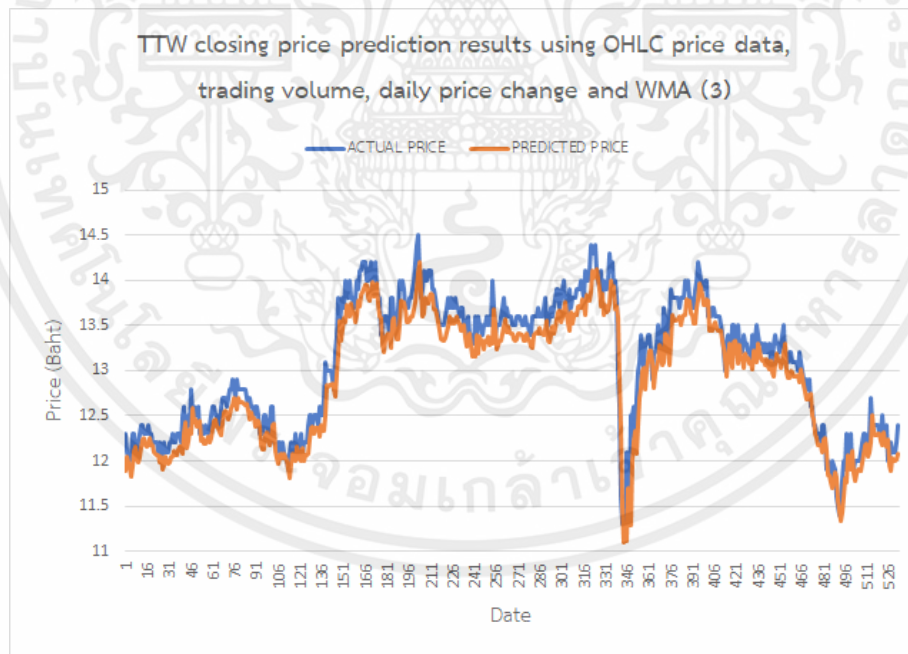
Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.09321	0.040473	-0.00691
	2	-0.03459	0.032351	-0.00238
	3	-0.17229	0.060024	-0.01298
	Avg.	-0.10003	0.044283	-0.00742



รูปที่ 4.34 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 1



รูปที่ 4.35 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 2



รูปที่ 4.36 ผลการพยากรณ์จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคา และ WMA (TTW) รอบที่ 3

4.4 การเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์

ตารางที่ 4.13 แสดงผลความคลาดเคลื่อนเฉลี่ยในการพยากรณ์

Ticker	Mean Squared Error (MSE)			
	Without MA	SMA (5)	EMA (5)	WMA (5)
PTT	0.839608	0.994671	0.852602	0.854980
SPCG	0.150719	0.127070	0.170201	0.127515
TTW	0.034332	0.033400	0.058776	0.044283
Average	0.341553	0.385047	0.360526	0.342259

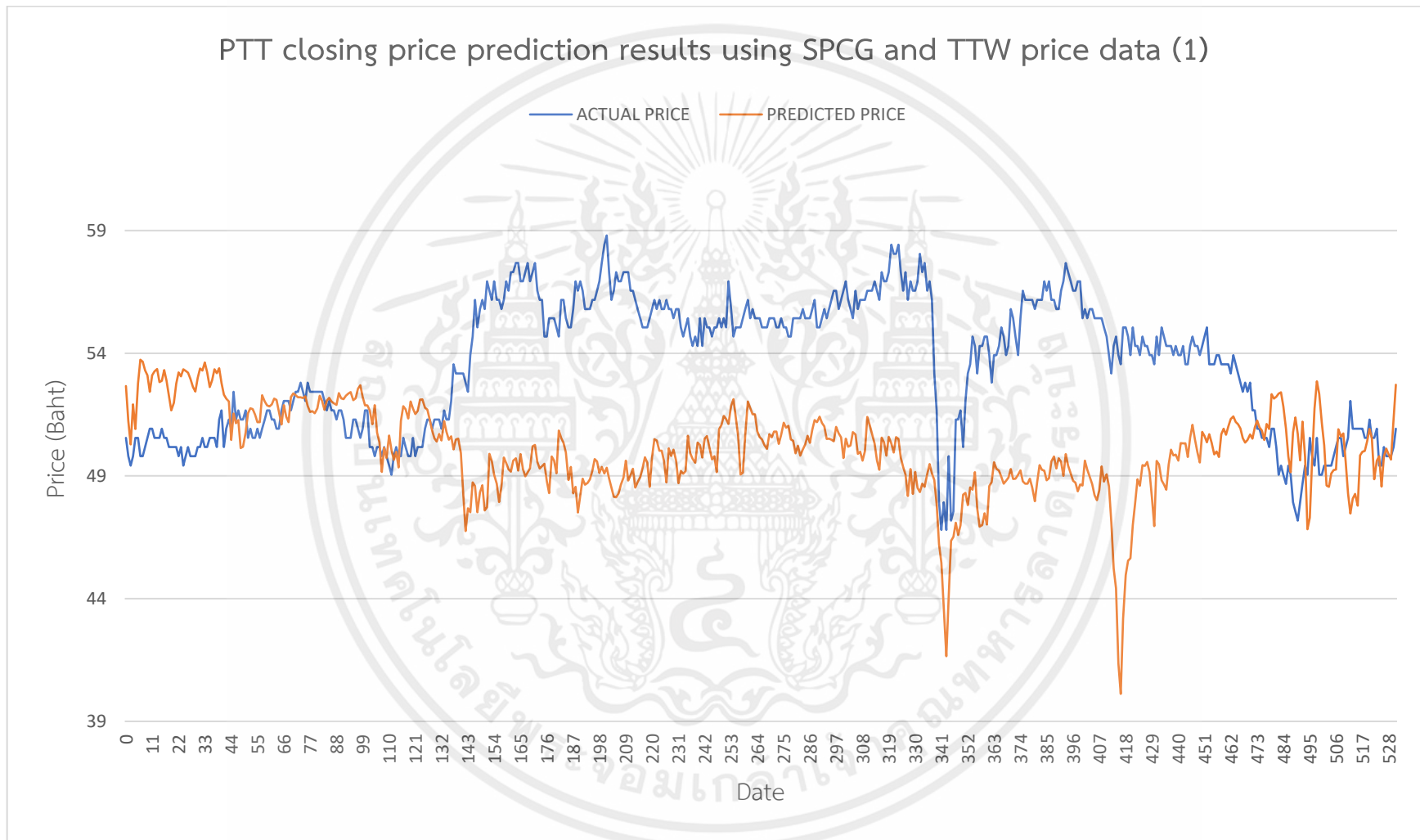
จากตารางที่ 4.13 เห็นได้ว่าสำหรับ PTT ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ Without MA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.839608 สำหรับ SPCG ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ SMA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.127070 และสำหรับ TTW ประสิทธิภาพในการพยากรณ์จากการใช้โครงข่ายประสาทเทียม LSTM ร่วมกับ SMA มีประสิทธิภาพดีที่สุด โดยมีค่า MSE เท่ากับ 0.033400

4.5 ผลการพยากรณ์ราคาปิดหุ้น PTT จากการศึกษาเพิ่มเติม

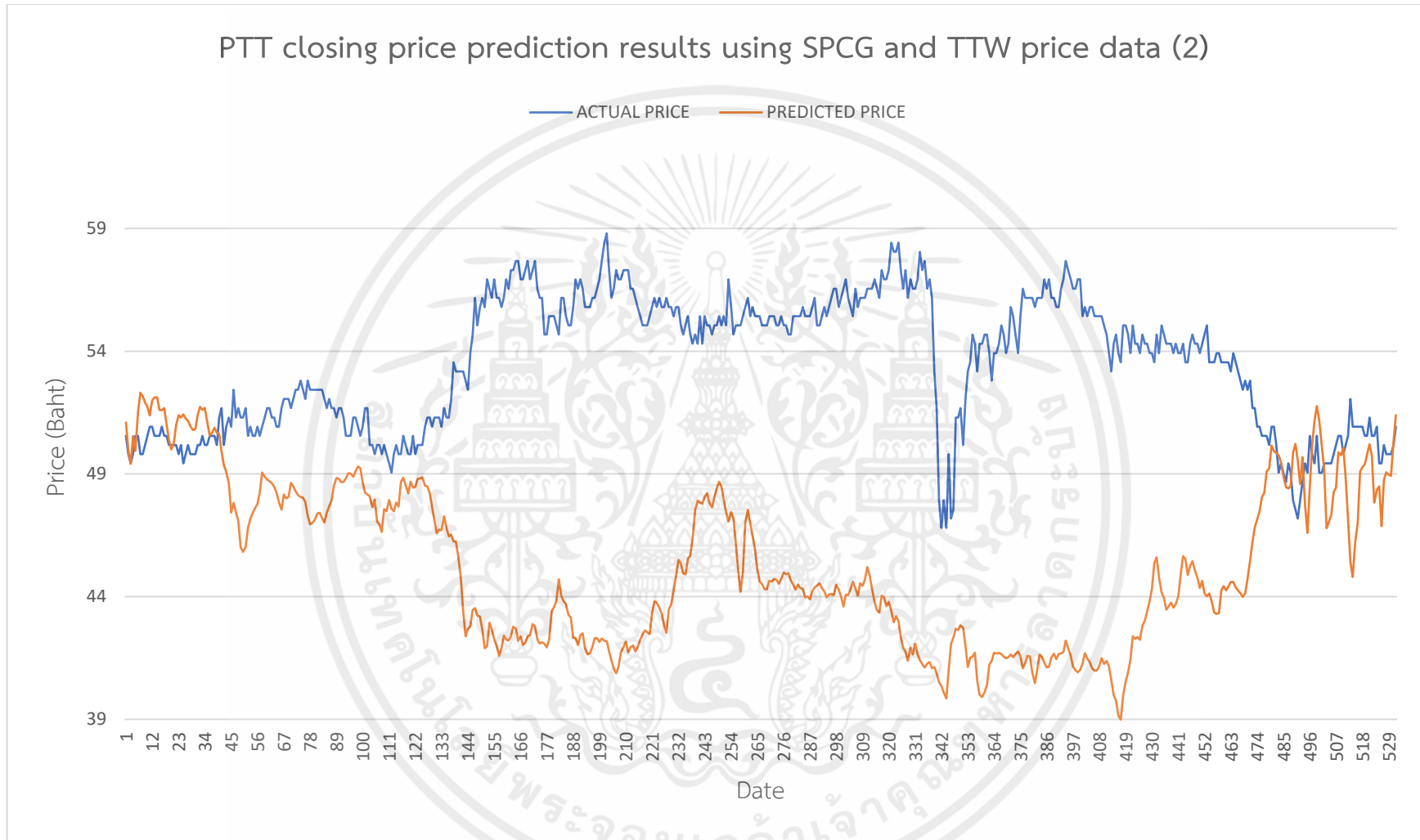
ผลการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น SPCG และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 83.38899 ดังตารางที่ 4.14 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.37, 4.38 และ 4.39

ตารางที่ 4.14 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาของหุ้น SPCG และ TTW

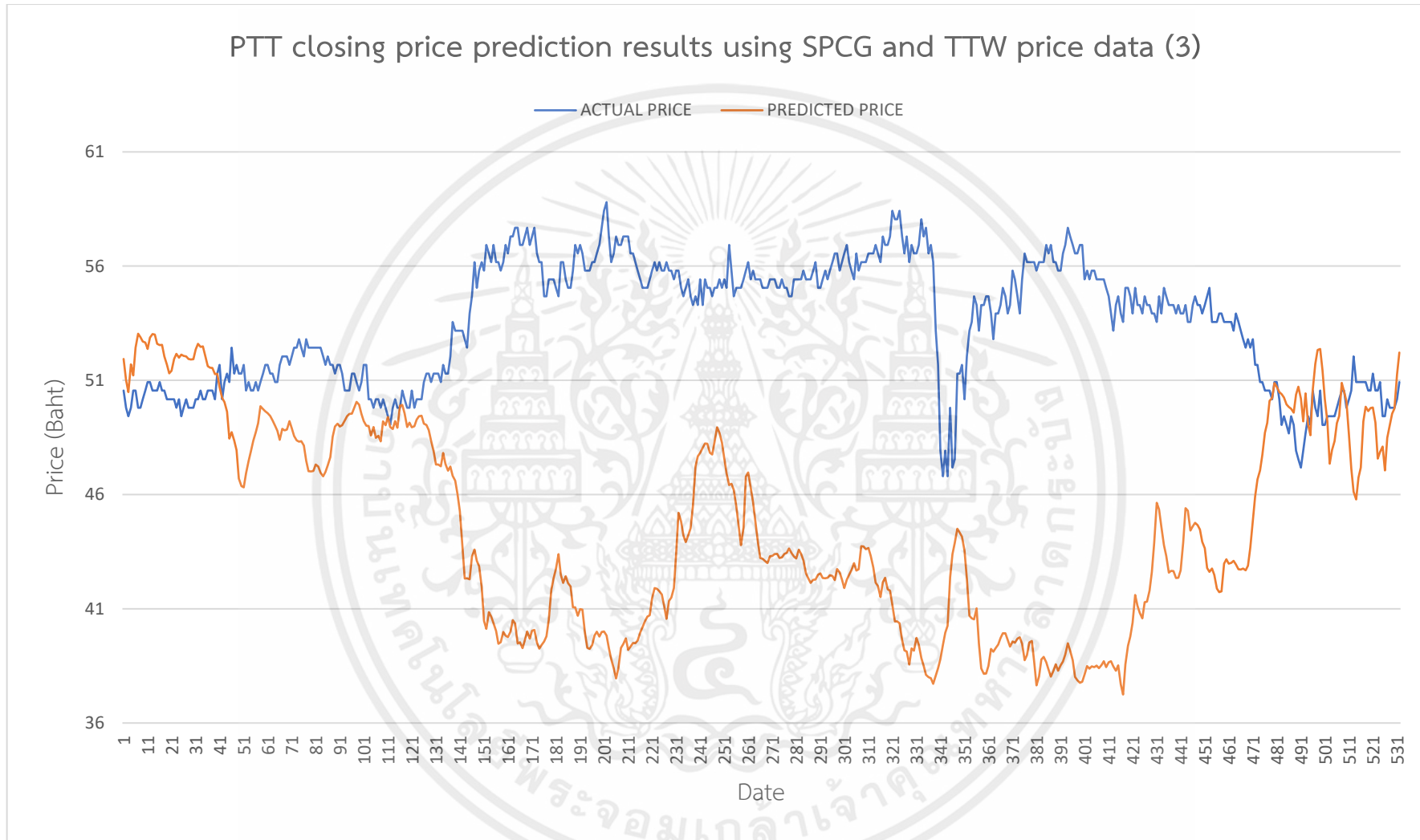
Ticker	Round	Loss Value	MSE	Percent Error
PTT	1	-3.43547	25.42338	-0.06123
	2	-8.22299	98.25544	-0.14923
	3	-8.95946	126.4882	-0.16191
	Avg.	-6.87264	83.38899	-0.12412



รูปที่ 4.37 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (1)



รูปที่ 4.38 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (2)



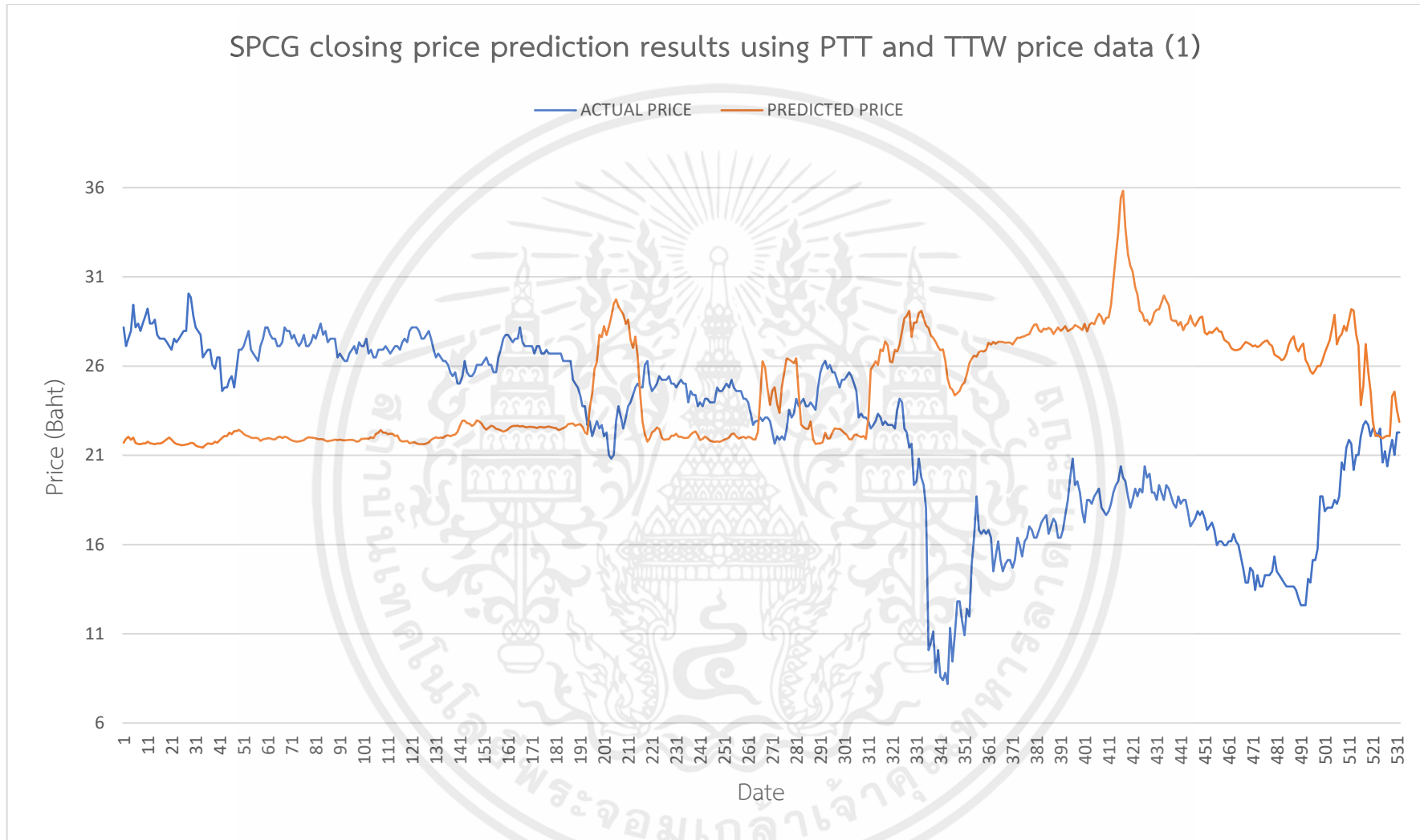
รูปที่ 4.39 ผลการพยากรณ์ราคาปิด PTT จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ SPCG และ TTW (3)

4.6 ผลการพยากรณ์ราคาปิดหุ้น SPCG จากการศึกษาเพิ่มเติม

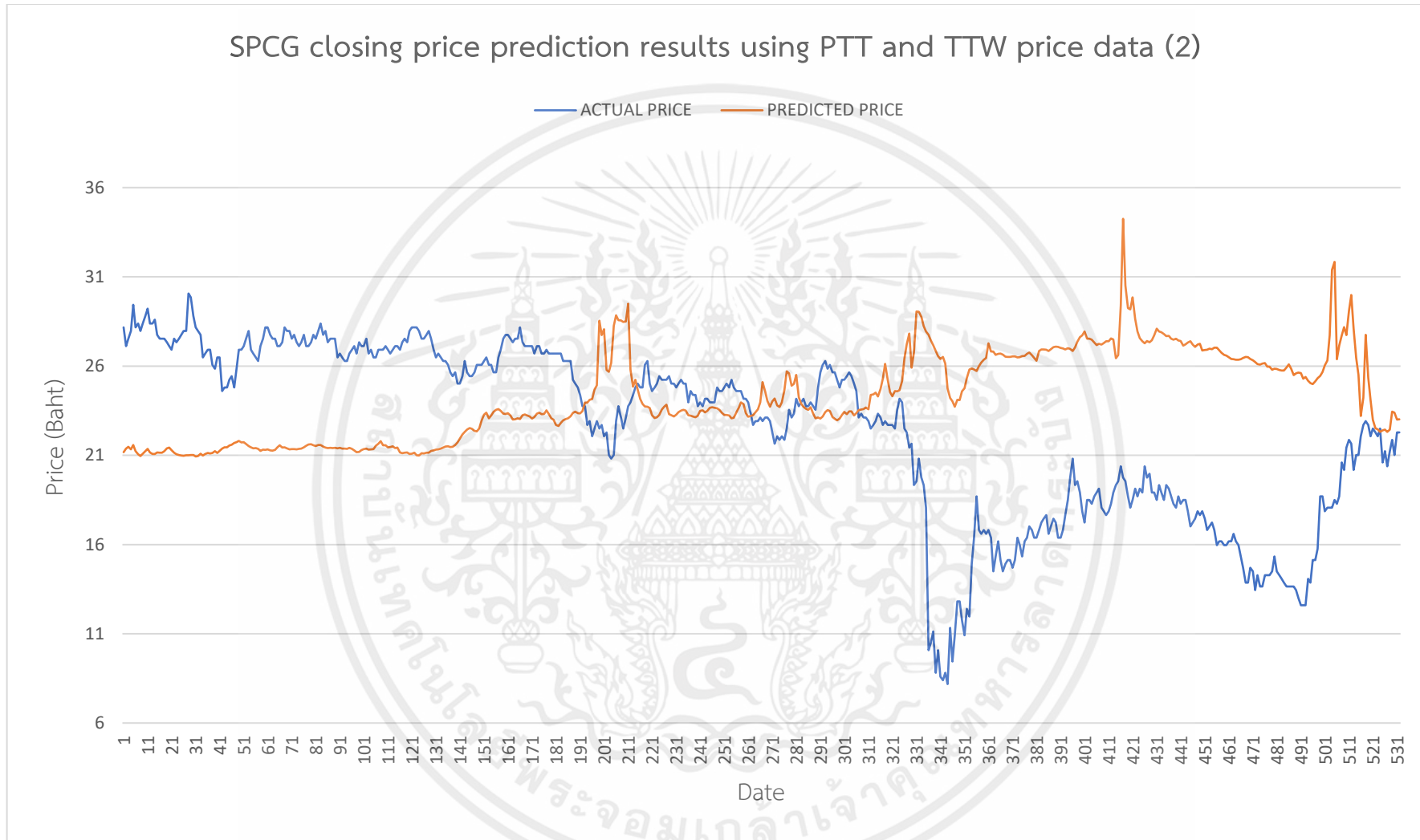
ผลการพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 58.03953 ดังตารางที่ 4.15 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.40, 4.41 และ 4.42

ตารางที่ 4.15 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาของหุ้น PTT และ TTW

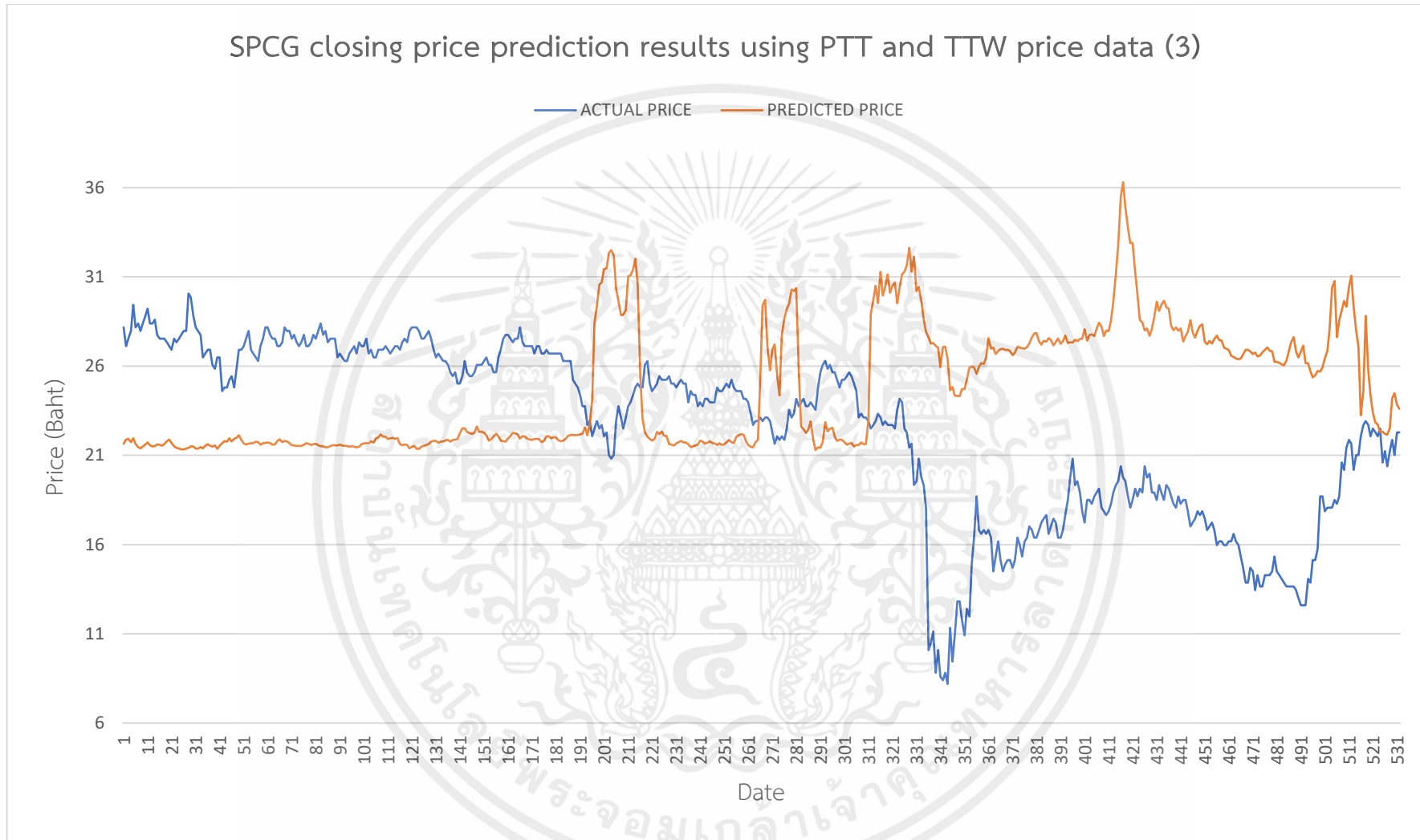
Ticker	Round	Loss Value	MSE	Percent Error
SPCG	1	2.291673	59.64586	0.200275
	2	1.911893	51.88098	0.178004
	3	2.369444	62.59175	0.202814
	Avg.	2.191003	58.03953	0.193698



รูปที่ 4.40 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (1)



รูปที่ 4.41 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (2)



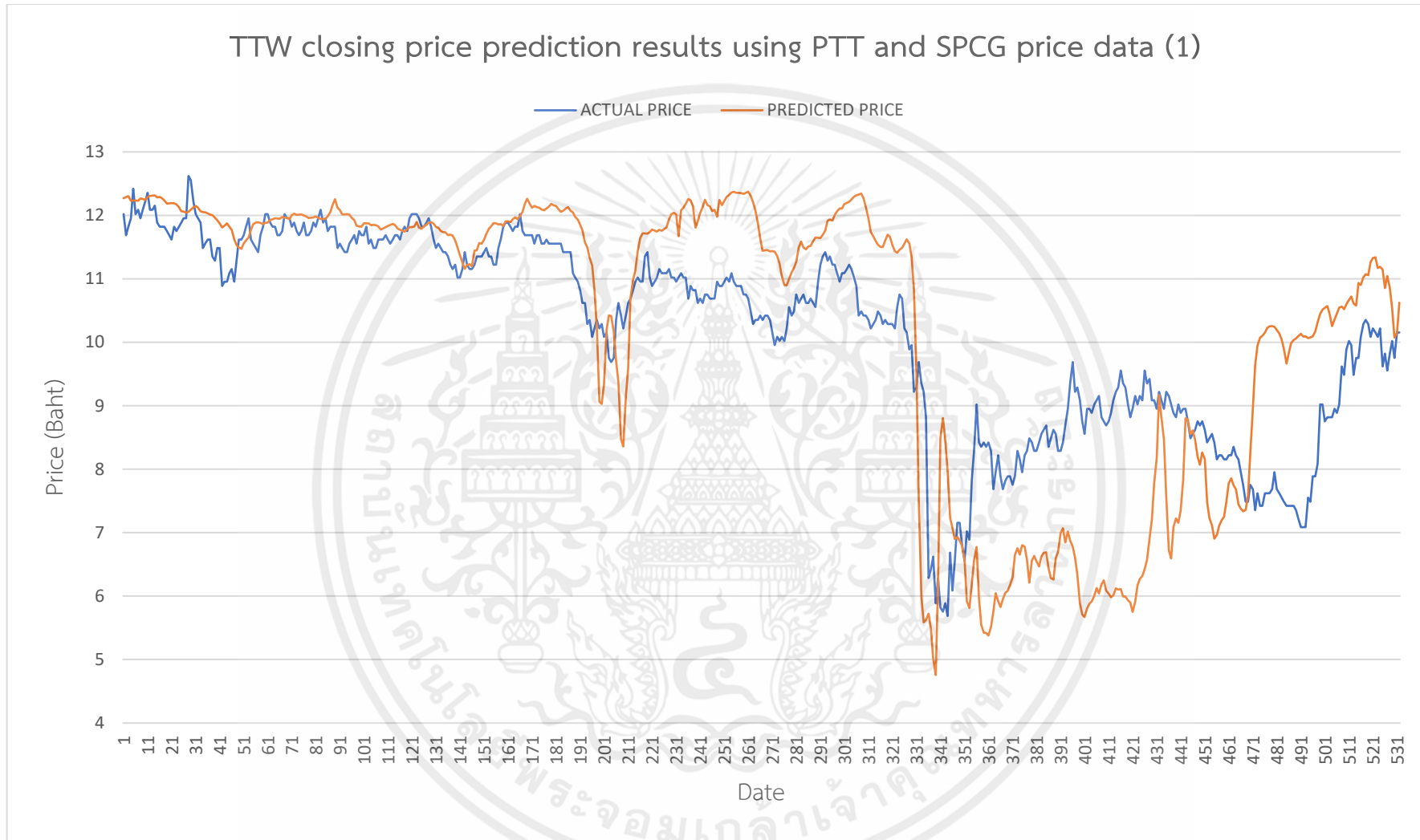
รูปที่ 4.42 ผลการพยากรณ์ราคาปิด SPCG จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ TTW (3)

4.7 ผลการพยากรณ์ราคาปิดหุ้น TTW จากการศึกษาเพิ่มเติม

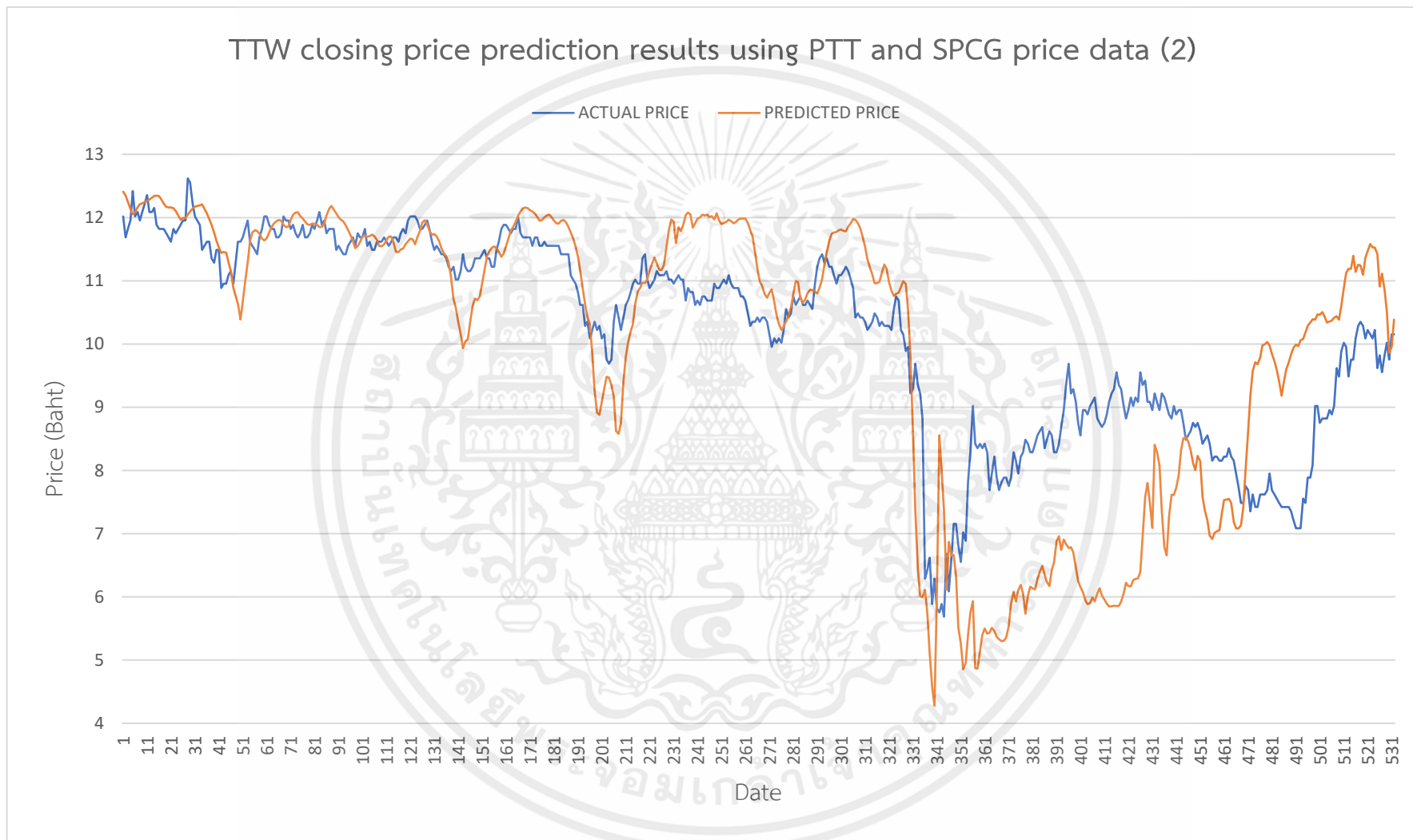
ผลการพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ SPCG ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 1.885817 ดังตารางที่ 4.16 และกราฟเปรียบเทียบระหว่างค่าจริงและค่าจากการพยากรณ์ดังรูปที่ 4.43, 4.44 และ 4.45

ตารางที่ 4.16 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาของหุ้น PTT และ SPCG

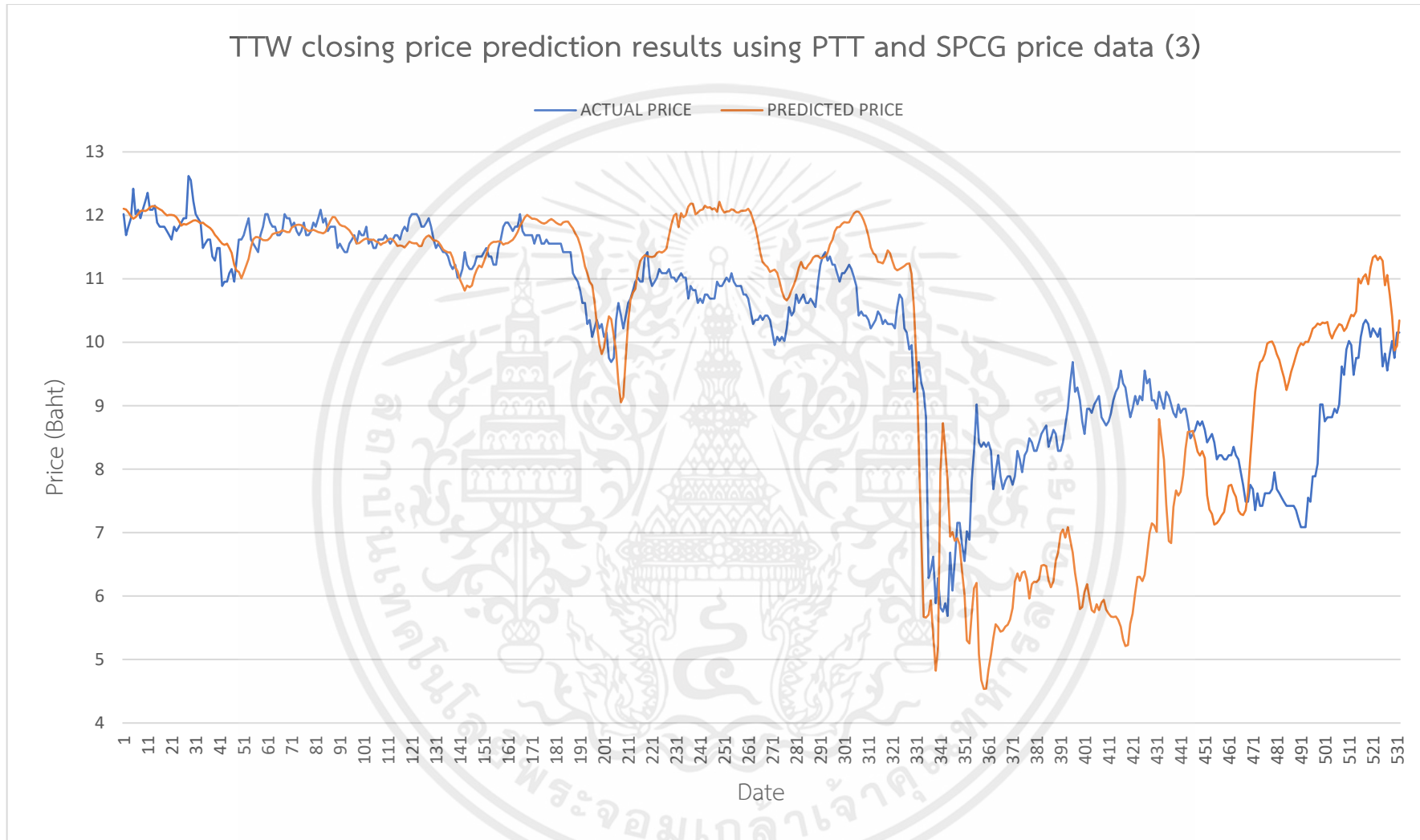
Ticker	Round	Loss Value	MSE	Percent Error
TTW	1	0.117272	1.924176	0.00873
	2	-0.15281	1.844688	-0.01925
	3	-0.0865	1.888587	-0.0117
	Avg.	-0.04068	1.885817	-0.0074



รูปที่ 4.43 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (1)



รูปที่ 4.44 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (2)



รูปที่ 4.45 ผลการพยากรณ์ราคาปิด TTW จากโครงข่ายประสาทเทียมโดยใช้ข้อมูลราคาของ PTT และ SPCG (3)

บทที่ 5 สรุปผลการวิเคราะห์และข้อเสนอแนะ

5.1 ผลสรุปการศึกษา

จากการศึกษาการพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้โครงข่ายประสาทเทียม LSTM ร่วมกับค่าเฉลี่ยเคลื่อนที่ที่ต่างกันประกอบด้วย Simple moving average (SMA), Exponential moving average (EMA) และ Weighted moving average (WMA) โดยใช้ข้อมูลราคาหลักทรัพย์จาก YAHOO FINANCE ทำการรวบรวมข้อมูล ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และปริมาณการซื้อขาย โดยใช้เป็นข้อมูลรายวัน ย้อนหลังเป็นเวลา 10 ปี ตั้งแต่วันที่ 1 มกราคม 2010 ถึง 31 ธันวาคม 2020 มาใช้ในการศึกษาความสามารถพยากรณ์ราคาหลักทรัพย์ในกลุ่มพลังงานคือ PTT SPCG และ TTW ซึ่งได้ทำการแปลงข้อมูลให้อยู่ในช่วง 0 ถึง 1 จากนั้นได้แบ่งข้อมูลสำหรับการดำเนินการในส่วนของโครงข่ายประสาทเทียมออกเป็น 2 ชุด คือ ข้อมูลสำหรับการเรียนรู้ และข้อมูลสำหรับการทดสอบ โดยใช้อัตราส่วน 80 : 20 ตามลำดับ และยังได้กำหนดใช้การวิเคราะห์ Lookback เพื่อให้โครงข่ายประสาทเทียมเกิดการเรียนรู้และทดสอบกระทั่งได้ผลลัพธ์ที่มีประสิทธิภาพซึ่งทำการวัดประสิทธิภาพจากการเปรียบเทียบค่า MSE

จากการศึกษาพบว่า การพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้โครงข่ายประสาทเทียม LSTM ร่วมกับค่าเฉลี่ยเคลื่อนที่แบบ WMA มีประสิทธิภาพสูงสุด เนื่องจากมีค่า MSE ที่ต่ำที่สุดเท่ากับ 0.342259 เมื่อเทียบกับการพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้โครงข่ายประสาทเทียม LSTM ร่วมกับค่าเฉลี่ยเคลื่อนที่ร่วมกับ SMA และ EMA ซึ่งมีค่า MSE เท่ากับ 0.385047 และ 0.360526 ตามลำดับ

นอกจากนี้เรายังพบว่า การพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้โครงข่ายประสาทเทียม LSTM ที่ไม่มีการใช้ร่วมกับค่าเฉลี่ยเคลื่อนที่กลับมีประสิทธิภาพที่มากกว่าการใช้ร่วมกับค่าเฉลี่ยเคลื่อนที่ทั้ง 3 แบบ โดยมีค่า MSE เท่ากับ 0.341553 ซึ่งไม่เป็นไปตามสมมติฐาน

จากการศึกษาเพิ่มเติมพบว่าการพยากรณ์ราคาปิดของหุ้น PTT ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น SPCG และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 83.38899 การพยากรณ์ราคาปิดของหุ้น SPCG ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ TTW ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 58.03953 การพยากรณ์ราคาปิดของหุ้น TTW ล่วงหน้า 1 วันโดยใช้ข้อมูลราคาและปริมาณการซื้อขายของหุ้น PTT และ SPCG ได้ค่าความคลาดเคลื่อนเฉลี่ยกำลังสองเท่ากับ 1.885817

5.2 ข้อสังเกตและข้อเสนอแนะ

จากการศึกษาพบว่า การเลือกใช้เส้นค่าเฉลี่ยช่วยในการพยากรณ์ ส่งผลกับผลการทดลองการพยากรณ์จริง และให้ผลความคลาดเคลื่อนมากขึ้นเมื่อใช้ร่วมกับค่าเส้นและค่าเฉลี่ยแต่ละแบบให้ผลความคลาดเคลื่อนที่ต่างกักรวมไปถึงราคาที่สูงขึ้นให้ผลความคลาดเคลื่อนที่น้อยลง นอกจากนี้ อาจเพิ่มจำนวนวันในการพยากรณ์ล่วงหน้าเพื่อให้เห็นการสะสมของความคลาดเคลื่อนมากขึ้นและอาจส่งผลให้ความแม่นยำจากการพยากรณ์ให้ผลที่แตกต่างกัน เนื่องจากการทำนายราคาล่วงหน้า 1 วันราคาปิดที่ไม่มีการเปลี่ยนแปลงมากนักทำให้ไม่สามารถเห็นผลกระทบจากการเพิ่มข้อมูลค่าเฉลี่ยได้อย่างชัดเจน

จากการทดสอบเพิ่มเติมในการใช้ราคาและปริมาณซื้อขายของหุ้นตัวอื่นร่วมกันเพื่อทำนายราคาปิดของหุ้นอีกตัวให้ผลความคลาดเคลื่อนที่สูงกว่าการใช้ข้อมูลของตัวเองในการทำนายเนื่องจากเป็นข้อมูลของหุ้นคนละตัวกัน แม้จะอยู่ในอุตสาหกรรมเดียวกัน แต่มีทิศทางการเคลื่อนไหวและช่วงราคาที่ไม่เหมือนกัน และข้อมูลที่ถูกนำมาใช้เพื่อฝึกโมเดลก็เป็นข้อมูลอดีตที่มีสภาพแวดล้อมในตลาดจริงต่างกัน

เมื่อได้แบบจำลองโครงข่ายประสาทเทียม LSTM สำหรับการพยากรณ์แล้ว อาจจำลองการสร้างเงื่อนไขสำหรับจำลองการซื้อขายหุ้นเพื่อทดสอบความแม่นยำในการซื้อขายจริง หรือช่วยในการวิเคราะห์ความเสี่ยงจากการเคลื่อนไหวของราคาหุ้นและปรับใช้ให้เหมาะสมเพื่อให้เกิดประโยชน์สูงสุดแก่นักลงทุน

บรรณานุกรม

- [1] ปิยศรีณีย์ ลิขิตาภรณ์ (2563). การพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้ Wavelet Transform, วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมการเงิน บัณฑิตวิทยาลัย มหาวิทยาลัยหอการค้า, 2563, หน้า 27 – 29.
- [2] Chaigusin, S., Chirathamjaree, C., and Claden, J. (2008). The Use of Neural Networks in the Prediction of the Stock Exchange of Thailand (SET) index, Edith Cowan University. Perth. 670-673.
- [3] Vijh, M., Chandola, D. and Kumar, A. (2019). Stock Closing Price Prediction Using Machine Learning Techniques, Jaypee Institute of Information Technology. Noida. 599-606.
- [4] Sanboon, T., Keatruangkamala, K. and Jaiyen, S. (2019). A Deep Learning Model for Predicting Buy and Sell Recommendations in Stock Exchange of Thailand using Long Short-Term Memory, King Mongkut's Institute of Technology Ladkrabang. Bangkok. 757-760.
- [5] Negnevitsky M., (2005). *Artificial Intelligence A Guide to Intelligent Systems*, Great Britain by Biddles Ltd, King's Lynn, England.
- [6] Hilpisch Y., (2020). *Artificial Intelligence in Finance*, O'Reilly Media, United States of America.
- [7] วิทยา พรพิชรพงศ์ (2551). โครงข่ายประสาทเทียม (Artificial Neural Networks - ANN), แหล่งที่มา <https://www.gotoknow.org/posts/163433> , เข้าดูเมื่อวันที่ 08/11/2564.
- [8] อาจารย์อำภา สาระศิริ (2559). เทคนิคการเรียนรู้พื้นฐานโครงข่ายประสาทเทียม, [องค์ประกอบของโครงข่ายประสาทเทียม], แหล่งที่มา <https://www.mut.ac.th/research-detail-92> , เข้าดูเมื่อวันที่ 08/11/2564.
- [9] algoaddict. (2019). หลักการทำงานของ ARIMA ในการทำนายค่าของ Time Series. สืบค้นจาก <https://algoaddict.wordpress.com/2019/06/22/%E0%B8%AB%E0%B8%A5%E0%B8%B1%E0%B8%81%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%97%E0%B8%B3%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B8%82%E0%B8%AD%E0%B8%87-arima-%E0%B9%83%E0%B8%99%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%97/>
- [10] Wikipedia. (2019). Autoregressive integrated moving average. สืบค้นจาก https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average

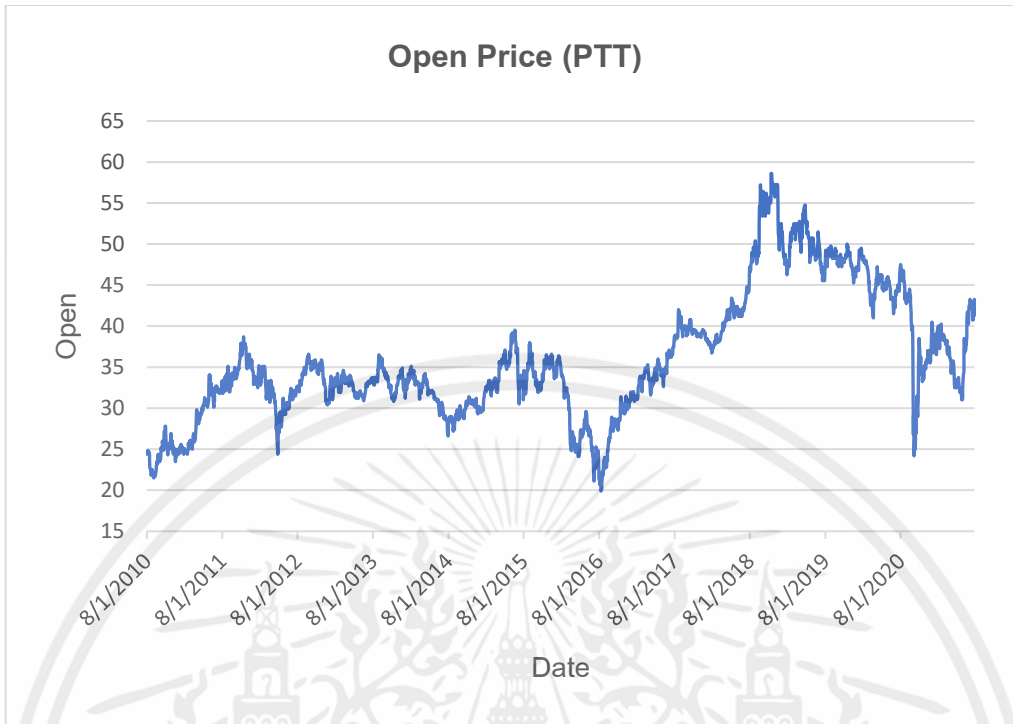
- [11] cway investment. (2018). Volatility Modeling Using GARCH. สืบค้นจาก <https://medium.com/cw-quantlab/volatility-modeling-using-garch-8f3490de6cdd>
- [12] Wikipedia. (2004). Autoregressive conditional heteroskedasticity. สืบค้นจาก https://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity
- [13] Michele D. (1997). *Perceptrons: An Associative Learning Network*. Virginia Tech, Blacksburg.
- [14] นางสาวเมริษา เทพบรรหาร (2555). การประมาณค่าความผันผวนสำหรับอัตราผลตอบแทนของสัญญาฟิวเจอร์สที่อ้างอิงกับดัชนี 50 หลักทรัพย์โดยใช้แบบจำลอง GARCH-X, ปรินญาเศรษฐศาสตรมหาบัณฑิต สาขาวิชาเศรษฐศาสตร์ธุรกิจ คณะวิทยาการจัดการ มหาวิทยาลัยขอนแก่น, 2555, หน้า 17-18.
- [15] Adril M. and Mhamed H. *Stock Market Prediction Using LSTM Recurrent Neural Network*. University Abdelmalek Essaadi, Morocco. 1168-1173.
- [16] สุวพร พรชัย. และ ผศ. ดร. วรณรพี บานชื่นวิจิตร. (2556). ความมีประสิทธิภาพของตลาดหลักทรัพย์แห่งประเทศไทย : กรณีศึกษาบริษัท ไทยรุ่งยูเนี่ยนคาร์ จำกัด (มหาชน). คณะบริหารธุรกิจ มหาวิทยาลัยหอการค้าไทย. หน้า 355-373.
- [17] Dave A. and George M. (1992). *Artificial Neural Networks Technology*. Utica, New York: Kaman Sciences Corporation.
- [18] DOUGLAS C., CHERYL L. and MURAT K. (2015). *Introduction To Time Series Analysis And Forecasting*. Hoboken, New Jersey: John Wiley & Sons.
- [19] John J. (2018). *เทคนิคอลอนาไลซิส* [Technical Analysis of the Financial Markets]. (พิมพ์ครั้งที่ 1). (สุธีร์ ระวีแสงสุรย์). กรุงเทพมหานคร: บริษัท ด่านสุธาการพิมพ์ จำกัด, (1999).
- [20] Jeeva R. (2013). สร้างพอร์ตหุ้นให้รวยมั่งคั่งแบบวอร์เรนบัฟเฟตต์ [Creating a Portfolio like Warren Buffett]. (ศุภิกา กุญชร ณ อยุธยา). นนทบุรี: ห้างหุ้นส่วนจำกัด ภาพพิมพ์, (2012)

ภาคผนวก



ภาคผนวก ก.
กราฟดัชนีราคาหุ้น

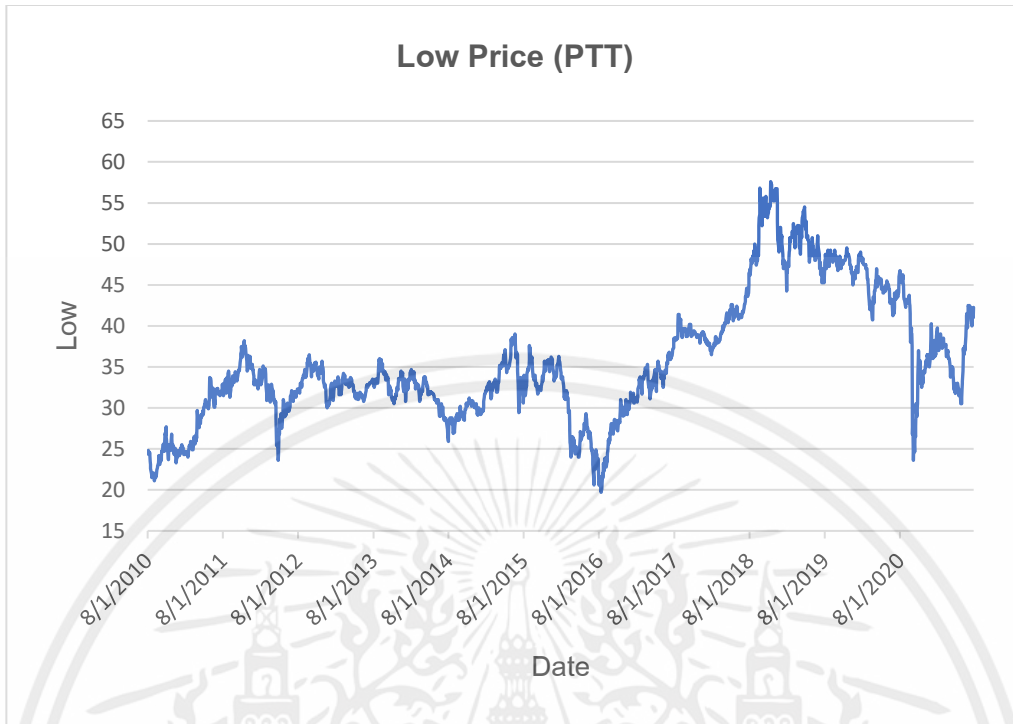




รูปที่ ก.1 ราคาเปิดของ PTT



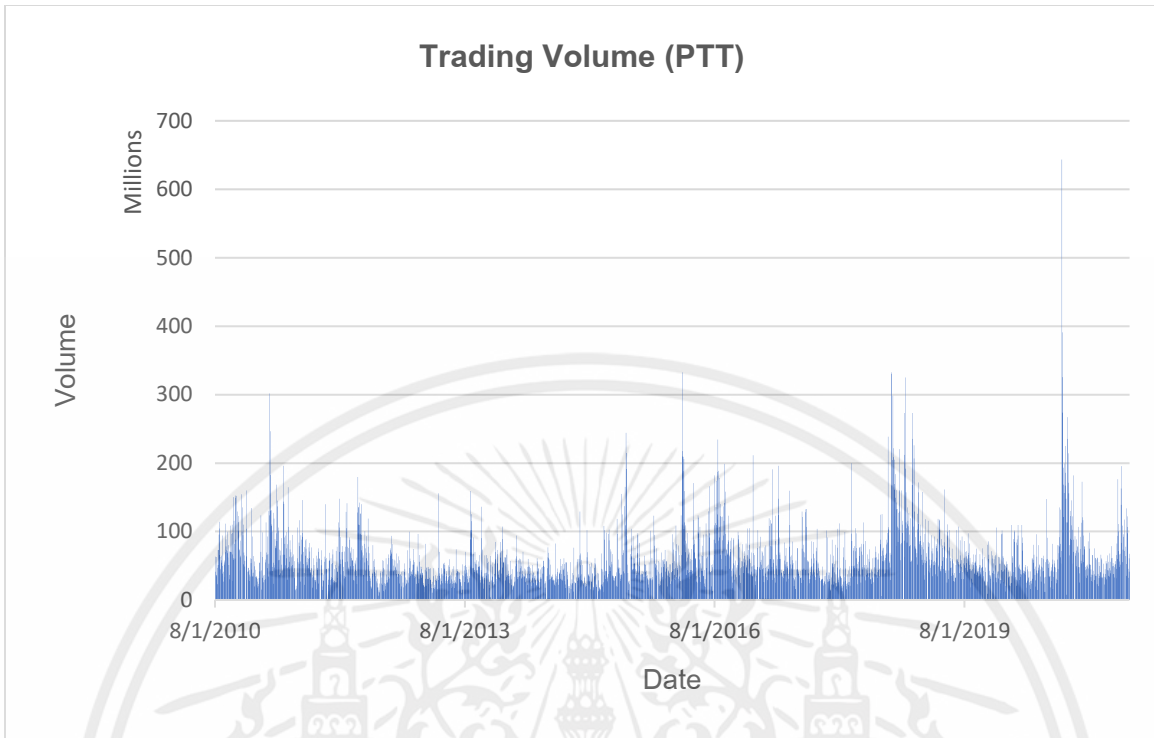
รูปที่ ก.2 ราคาสูงสุดของ PTT



รูปที่ ก.3 ราคาต่ำสุดของ PTT



รูปที่ ก.4 ราคาปิดของ PTT



รูปที่ ก.5 ปริมาณการซื้อขายของ PTT



รูปที่ ก.6 ราคาเปิดของ SPCG



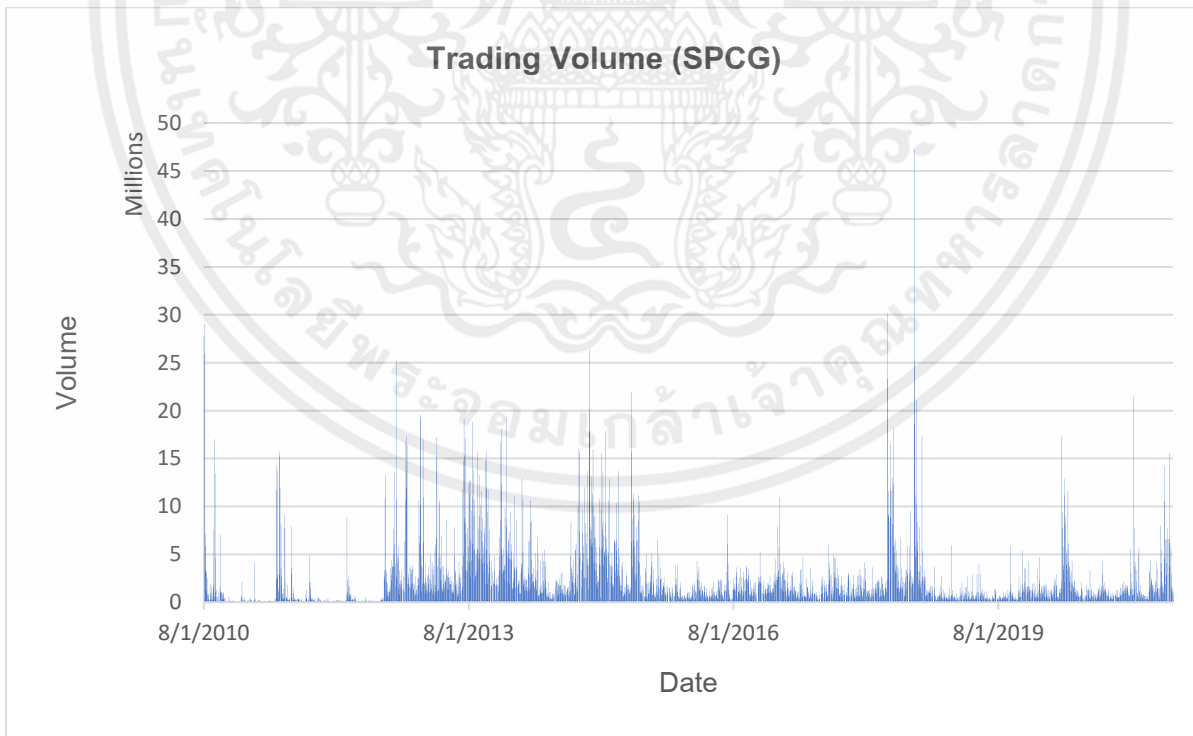
รูปที่ ก.7 ราคาสูงสุดของ SPCG



รูปที่ ก.8 ราคาต่ำสุดของ SPCG



รูปที่ ก.9 ราคาปิดของ SPCG



รูปที่ ก.10 ปริมาณการซื้อขายของ SPCG



รูปที่ ก.11 ราคาเปิดของ TTW



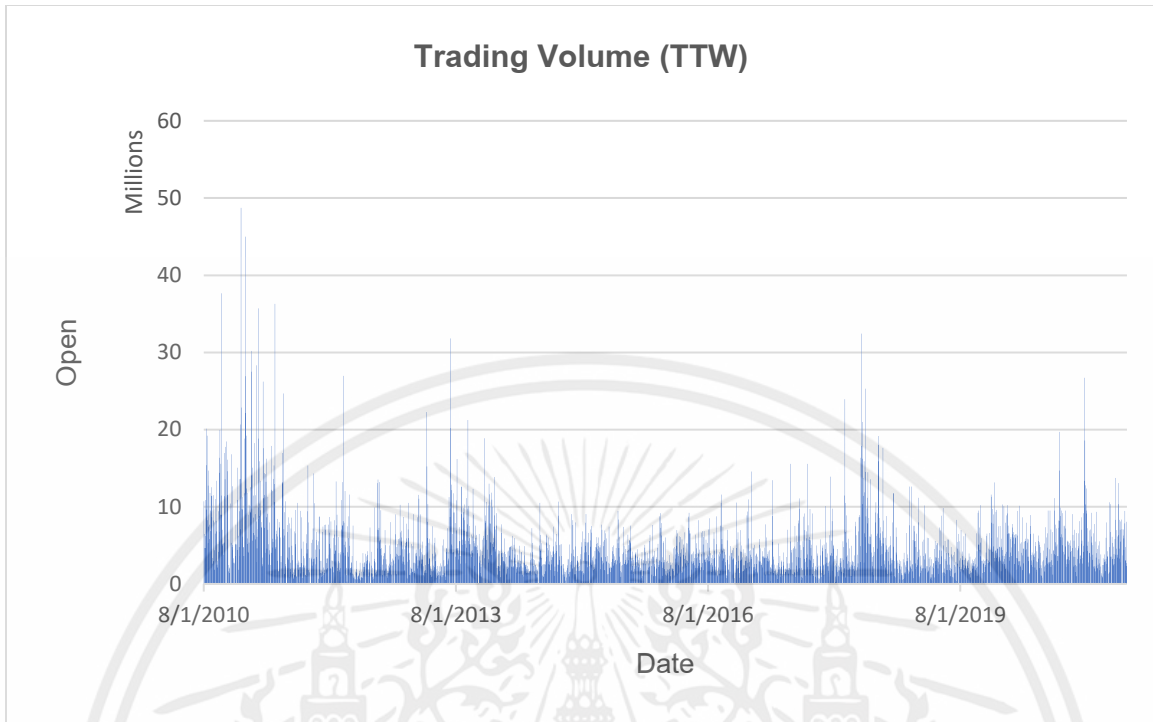
รูปที่ ก.12 ราคาสูงสุดของ TTW



รูปที่ ก.13 ราคาต่ำสุดของ TTW



รูปที่ ก.14 ราคาปิดของ TTW



รูปที่ ก.15 ปริมาณการซื้อขายของ TTW

ภาคผนวก ข.

ตารางแสดงผล



ตารางที่ ข.1 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคา

Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
PTT	1	0.248565	0.869657	0.005669
	2	0.202731	0.807739	0.004828
	3	0.227453	0.841429	0.005289
	Avg.	0.22625	0.839608	0.005262

ตารางที่ ข.2 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาร่วมกับ SMA

Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
PTT	1	0.096249	0.771108	0.002507
	2	0.483516	1.062159	0.011323
	3	0.394387	0.960927	0.009333
	Avg.	0.324717	0.994671	0.007721

ตารางที่ ข.3 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาพร้อมกับ EMA

Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
PTT	1	0.297938	0.885805	0.0071
	2	-0.25084	0.86947	-0.00607
	3	0.165836	0.802532	0.004178
	Avg.	0.070978	0.852602	0.001736

ตารางที่ ข.4 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น PTT โดยใช้ข้อมูลราคาพร้อมกับ WMA

Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
PTT	1	-0.31386	0.839916	-0.0072
	2	0.252742	0.851058	0.006158
	3	0.248892	0.873967	0.005737
	Avg.	0.062591	0.854980	0.001565

ตารางที่ ข.5 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคา

Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
SPCG	1	0.128508	0.132908	0.006961
	2	-0.18088	0.150369	-0.00958
	3	0.213522	0.168878	0.01146
	Avg.	0.053717	0.150718	0.002947

ตารางที่ ข.6 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ SMA

Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.096545	0.12505	0.005337
	2	-0.151640	0.138565	-0.007872
	3	0.014058	0.117596	0.000973
	Avg.	-0.01368	0.127071	-0.00052

ตารางที่ ข.7 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ EMA

Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.083518	0.12887	0.004616
	2	0.242506	0.177027	0.013
	3	-0.30041	0.204705	-0.01586
	Avg.	0.008538	0.170201	0.000585

ตารางที่ ข.8 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น SPCG โดยใช้ข้อมูลราคาร่วมกับ WMA

Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
SPCG	1	0.164917	0.141415	0.008913
	2	0.092519	0.125692	0.005144
	3	0.033118	0.11544	0.001933
	Avg.	0.096851	0.127516	0.00533

ตารางที่ ข.9 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคา

Index	Round	LSTM with daily price change		
		Loss Value	MSE	Percent Error
TTW	1	-0.00122	0.032864	0.000257
	2	-0.07565	0.037104	-0.00558
	3	-0.0417	0.033027	-0.00297
	Avg.	-0.03952	0.034332	-0.00276

ตารางที่ ข.10 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาร่วมกับ SMA

Index	Round	LSTM with SMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.01868	0.031612	-0.0012
	2	0.053199	0.035205	0.00439
	3	-0.000077	0.033384	0.000415
	Avg.	0.011481	0.033400	0.001202

ตารางที่ ข.11 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาพร้อมกับ EMA

Index	Round	LSTM with EMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.09468	0.039306	-0.00699
	2	-0.16786	0.061295	-0.0125
	3	-0.21088	0.075727	-0.01586
	Avg.	-0.15781	0.058776	-0.01178

ตารางที่ ข.12 ผลลัพธ์ของโครงข่ายประสาทเทียม LSTM ของดัชนีราคาตลาดหุ้น TTW โดยใช้ข้อมูลราคาพร้อมกับ WMA

Index	Round	LSTM with WMA		
		Loss Value	MSE	Percent Error
TTW	1	-0.09321	0.040473	-0.00691
	2	-0.03459	0.032351	-0.00238
	3	-0.17229	0.060024	-0.01298
	Avg.	-0.10003	0.044283	-0.00742

ภาคผนวก ค.
Python code



ค.1 ติดตั้ง Libraries ที่จำเป็น

```
#install ta-lib
url = 'https://launchpad.net/~mario-
mariomedina/+archive/ubuntu/talib/+files'
!wget $url/libta-lib0_0.4.0-oneiric1_amd64.deb -qO libta.deb
!wget $url/ta-lib0-dev_0.4.0-oneiric1_amd64.deb -qO ta.deb
!dpkg -i libta.deb ta.deb
!pip install ta-lib
#install ta-lib
!pip install yfinance
```

ค.2 Import libraries ที่จำเป็น

```
#อิมพอร์ตไลบรารีที่จำเป็น
import talib
import yfinance as yf
import numpy as np
import pandas as pd
from pylab import mpl, plt
plt.style.use('seaborn')
mpl.rcParams['font.family'] = 'serif'
%matplotlib inline
```

ค.3 บอกตำแหน่งเข้าถึงไฟล์

```
from google.colab import drive
drive.mount('/content/drive')

#เปลี่ยนไดเรกทอรีเตรียมการสำหรับบันทึกข้อมูล
cd /content/drive/MyDrive/Projectdata
```

ค.4 สร้างฟังก์ชันสำหรับเตรียมข้อมูล

```
#ฟังก์ชันสำหรับเตรียมการข้อมูล
class Data:
    def __init__(self,name):
        self.name = name
        daily = yf.download(self.name+'.BK',start= '2010-1-1', end='2020-12-
31') #ดึงข้อมูลราคา ปริมาณการซื้อขาย ช้อนหลังของหุ้นที่กำหนด
        daily = daily.dropna()
        daily['SMA5'] = talib.SMA(daily['Close'].values,timeperiod=5)
```

```

daily['EMA5'] = talib.EMA(daily['Close'].values,timeperiod=5)
daily['WMA5'] = talib.WMA(daily['Close'].values,timeperiod=5)
daily['Change'] = daily['Close']-daily['Close'].shift(1)
daily = daily.dropna()
self.daily = daily

```

ค.5 ตัวอย่างการพยากรณ์ราคาปิดของหุ้นที่ต้องการ ตัวอย่างหุ้น TTW

```

#บันทึกhistory data
data = Data('TTW') #กำหนดชื่อหุ้นที่ต้องการดึงข้อมูลผ่านฟังก์ชัน
data.daily
data.daily.to_csv('TTWdata.csv')
allhistory = data.daily.copy() #คัดลอกข้อมูล
feature_names = [ 'Close', 'Open', 'High', 'Low', 'Volume', 'Change', 'WMA5' ] #
กำหนดฟีเจอร์ที่ใช้ในการทดสอบตัวอย่างการทดสอบด้วย WMA5

#สเกลข้อมูลให้มีค่าอยู่ระหว่าง 0-1
allhistory[feature_names].head()
from sklearn.preprocessing import MinMaxScaler
sc = MinMaxScaler(feature_range = (0, 1))
allhistory_scaled = sc.fit_transform(allhistory[feature_names])
print(allhistory_scaled.shape)

split = int(len(allhistory_scaled) * 0.8) #แยกข้อมูลระหว่างส่วนtrainกับtest
train = allhistory_scaled[:split].copy()
test = allhistory_scaled[split:].copy()
test_copy = allhistory[split:]

train_features = train[:, :]
train_tragets = train[:,0]
data_len = len(train)
train_tragets = train_tragets.reshape(1,-1).T
features = []
labels = []
lookback = 5
future = 1
for i in range(lookback, data_len-future+1):
    features.append(train[i-lookback:i, :])
    labels.append(train[i+future-1, 0])

features = np.array(features)
labels = np.array(labels) #เตรียมข้อมูลการtrain
labels = labels.reshape(1,-1).T

```

```

import tensorflow as tf
model = tf.keras.models.Sequential([
    tf.keras.layers.LSTM(units = 13, return_sequences = True, input_shape
= (features.shape[1], features.shape[2])),
    tf.keras.layers.LSTM(units = 7, return_sequences = True),
    tf.keras.layers.LSTM(units = 3),
    tf.keras.layers.Dense(units = 1)
]) #กำหนดพารามิเตอร์ของโมเดล
print(model.summary()) #แสดงผลโมเดลอย่างสรุป
model.compile(optimizer = 'adam', loss = 'mean_squared_error') #กำหนดคอมไพเลอร์
และประเภทค่าความผิดพลาด
from time import time
start = time()
history = model.fit(features, labels, epochs = 100, batch_size = 32, valid
ation_data=(features, labels))
end = time()
pic = plt.figure()
plt.figure(figsize=(10,6))
plt.plot(history.history['loss'], label='train')
plt.plot(history.history['val_loss'], label='test')
plt.legend()
#plt.savefig('TTW_WMA_3_train_loss.jpg') แสดงMSEของการtrainและtestแต่ละรอบการทำซ้ำ
test_stock_data_processed = test[:, :]
inputs = test
X_test = []
for i in range(lookback, len(test)-future+1):
    X_test.append(inputs[i-lookback:i, :])
X_test = np.array(X_test)
predicted_stock_price = model.predict(X_test)
prediction_copies = np.repeat(predicted_stock_price, X_test.shape[2], axis
=-1)
realpredict_price = sc.inverse_transform(prediction_copies)[:,0]
realvalue = test[lookback+future-1:,0]
realvalue=realvalue.reshape(1,-1).T
realvaluetransform = np.repeat(realvalue, X_test.shape[2], axis=-1)
realvaluetransform_price = sc.inverse_transform(realvaluetransform)[:,0]
pic = plt.figure()
plt.figure(figsize=(10,6))
plt.plot(realpredict_price, 'r', label='Predicted Price')
plt.plot(realvaluetransform_price, 'b', label='Actual Price')
plt.title('TTW Price Prediction Using Open High Low Close Change And WMA5'
)
plt.xlabel('Date')
plt.ylabel('THB')
plt.legend()

```

```

plt.savefig('TTW_WMA_03.jpg') แสดงผลการเปรียบเทียบการทำนายกับราคาจริง
from sklearn.metrics import mean_squared_error, explained_variance_score, m
ax_error
data = {'ACTUAL PRICE': realvaluetrasform_price, 'PREDICTED PRICE': realpr
edict_price}
result = pd.DataFrame(data=data)
result['SQUARE ERROR'] = np.square((result['ACTUAL PRICE']-
result['PREDICTED PRICE']))
result['Error'] = result['PREDICTED PRICE'] - result['ACTUAL PRICE']
result['PERCENT ERROR'] = result['Error']/result['ACTUAL PRICE']
mean_squared_error(realvaluetrasform_price,realpredict_price, squared=Fals
e)
result['SQUARE ERROR'].mean()
pic = plt.figure()
plt.figure(figsize=(10,6))
result['SQUARE ERROR'].hist(bins=35)
plt.title('Squared Error of TTW Price Prediction Using Open High Low Close
Change And WMA5')
plt.xlabel('Squared Error')
plt.ylabel('Frequency')
plt.legend()
plt.savefig('TTW_WMA_LOSSBIN_03.jpg') แสดงผลค่าสถิติสะสมของความสัมพันธ์
#result.to_csv('TTW_ERRORSTAT_WMA_03.csv') บันทึกค่าที่ทดสอบ

```

ค.6 ตัวอย่าง Python code จากการทดลองเพิ่มเติม

```

TTW = yf.download('TTW'+'.BK',start= '2010-1-1', end='2020-12-31')
TTW=TTW.rename(columns={'Close':'TTWClose','Open':'TTWOpen','High':'TTWHigh',
'Low':'TTWLow','Volume':'TTWVolume'})
PTT = yf.download('PTT'+'.BK',start= '2010-1-1', end='2020-12-31')
PTT=PTT.rename(columns={'Close':'PTTClose','Open':'PTTOpen','High':'PTTHigh',
'Low':'PTTLow','Volume':'PTTVolume'})
SPCG = yf.download('SPCG'+'.BK',start= '2010-1-1', end='2020-12-31')
SPCG=SPCG.rename(columns={'Close':'SPCGClose','Open':'SPCGOpen','High':'SP
CGHigh','Low':'SPCGLow','Volume':'SPCGVolume'})
ALLDATA = pd.concat([TTW,SPCG,PTT],axis=1)
P_TTW = ALLDATA[['TTWClose','PTTClose','PTTOpen','PTTHigh','PTTLow','PTTVo
lume','SPCGClose','SPCGOpen','SPCGHigh','SPCGLow','SPCGVolume']]
P_SPCG = ALLDATA[['SPCGClose','PTTClose','PTTOpen','PTTHigh','PTTLow','PTT
Volume','TTWClose','TTWOpen','TTWHigh','TTWLow','TTWVolume']]
P_PTT = ALLDATA[['PTTClose','TTWClose','TTWOpen','TTWHigh','TTWLow','TTWVo
lume','SPCGClose','SPCGOpen','SPCGHigh','SPCGLow','SPCGVolume']]
from sklearn.preprocessing import MinMaxScaler

```

```

sc = MinMaxScaler(feature_range = (0, 1))
# P_TTW_scaled = sc.fit_transform(P_TTW)
#P_PTT_scaled = sc.fit_transform(P_PTT)
P_SPCG_scaled = sc.fit_transform(P_SPCG)
split = int(len(P_SPCG_scaled) * 0.8) #split data to train and test part
P_SPCG_train = P_SPCG_scaled[:split].copy()
P_SPCG_test = P_SPCG_scaled[split:].copy()
train_features = P_SPCG_train[:, 1:]
train_tragets = P_SPCG_train[:,0]
data_len = len(P_SPCG_train)
train_tragets = train_tragets.reshape(1,-1).T
features = []
labels = []
lookback = 5
future = 1
for i in range(lookback, data_len-future+1):
    features.append(train_features[i-lookback:i, :])
    labels.append(train_tragets[i+future-1, 0])

features = np.array(features)
labels = np.array(labels)#prepare train data
labels = labels.reshape(1,-1).T

test = P_SPCG_test[:, 1:]
X_test = []
for i in range(lookback, len(test)-future+1):
    X_test.append(test[i-lookback:i, :])
X_test = np.array(X_test)
import tensorflow as tf
model = tf.keras.models.Sequential([
    tf.keras.layers.LSTM(units = 13, return_sequences = True, input_shape
= (features.shape[1], features.shape[2])),
    tf.keras.layers.LSTM(units = 7, return_sequences = True),
    tf.keras.layers.LSTM(units = 3),
    tf.keras.layers.Dense(units = 1)
])#define model parameter
print(model.summary())
model.compile(optimizer = 'adam', loss = 'mean_squared_error') #define com
plier and loss type
from time import time
start = time()
history = model.fit(features, labels, epochs = 100, batch_size = 32, valid
ation_data=(features, labels))
end = time()
pic = plt.figure()

```

```

plt.figure(figsize=(10,6))
plt.plot(history.history['loss'], label='train')
plt.plot(history.history['val_loss'], label='test')
plt.legend()
#plt.savefig('TTW_WMA_3_train_loss.jpg') save plot of train loss
predicted_stock_price = model.predict(X_test)
prediction_copies = np.repeat(predicted_stock_price, X_test.shape[2]+1, axis=-1)
realpredict_price = sc.inverse_transform(prediction_copies)[: ,0]
realvalue = test[lookback+future-1: ,0]
realvalue=realvalue.reshape(1,-1).T
realvaluetransform = np.repeat(realvalue, X_test.shape[2]+1, axis=-1)
realvaluetransform_price = sc.inverse_transform(realvaluetransform)[: ,0]
pic = plt.figure()
plt.figure(figsize=(15,6))
plt.plot(realpredict_price, 'r', label='Predicted Price')
plt.plot(realvaluetransform_price, 'b', label='Actual Price')
plt.title('SPCG Price Prediction Using TTW and PTT Price and Volume data 3')
plt.xlabel('Date')
plt.ylabel('THB')
plt.legend()
#plt.savefig('TTW_WMA_03.jpg')save plot of prediction and actual price
from sklearn.metrics import mean_squared_error, explained_variance_score, max_error
data = {'ACTUAL PRICE': realvaluetransform_price, 'PREDICTED PRICE': realpredict_price}
result = pd.DataFrame(data=data)
result['SQUARE ERROR'] = np.square((result['ACTUAL PRICE']-result['PREDICTED PRICE']))
result['Error'] = result['PREDICTED PRICE'] - result['ACTUAL PRICE']
result['PERCENT ERROR'] = result['Error']/result['ACTUAL PRICE']
mean_squared_error(realvaluetransform_price,realpredict_price, squared=False)
result['SQUARE ERROR'].mean()
pic = plt.figure()
plt.figure(figsize=(10,6))
result['SQUARE ERROR'].hist(bins=35)
plt.title('Squared Error of TTW Price Prediction Using Open High Low Close Change And WMA5')
plt.xlabel('Squared Error')
plt.ylabel('Frequency')
plt.legend()
#plt.savefig('TTW_WMA_LOSSBIN_03.jpg') save error histogram
result.to_csv('SPCG_ERRORSTAT_NEWMETHOD_3.csv')#save result

```