

การผสมผสานการวิเคราะห์เชิงเทคนิคและแบบจำลองการเรียนรู้เชิงลึกสำหรับการซื้อขายน้ำมันดิบ



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

COMBINING TECHNICAL ANALYSIS AND DEEP LEARNING MODELS FOR CRUDE OIL
TRADING



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2021
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การผสมการวิเคราะห์เชิงเทคนิคและแบบจำลองการ เรียนรู้เชิงลึกสำหรับการซื้อขายน้ำมันดิบ
โดย	นายวิศรุต เลิศทวีเดช
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	อาจารย์ ดร.พิตติพล คັນธวัชน์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

----- คณะบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

----- ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ)

----- อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

----- อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(อาจารย์ ดร.พิตติพล คันธวัชน์)

----- กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.กฤษฎา นิมนานันท์)

น้ำมันดิบเป็นสินค้าอุปโภคที่มีความสำคัญในโลก เพราะน้ำมันดิบถือเป็นแหล่งพลังงานหลักของโลก ราคาของน้ำมันดิบนั้นมีส่วนเกี่ยวข้องกับหลาย ๆ อุตสาหกรรม เช่น การขนส่ง, การผลิตพลังงานไฟฟ้า และอุตสาหกรรมปิโตรเคมี ดังนั้นการคาดการณ์ราคาน้ำมันดิบจึงมีความสำคัญสำหรับหลายภาคส่วน แต่ก็เป็นเรื่องที่ทำนายยากเช่นกัน เนื่องจากราคาน้ำมันดิบมีความผันผวนสูง มีหลายงานวิจัยจำนวนมากที่เสนอการใช้การเรียนรู้ของเครื่องเพื่อทำนายราคาน้ำมัน โดยงานวิจัยนี้ได้นำเสนอเทคนิคการใช้โครงข่ายประสาทเทียมแบบผสมผสานกันระหว่างโครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional neural networks - CNN) และ หน่วยความจำระยะสั้นแบบยาว (Long short-term memory - LSTM) เพื่อใช้ทำนายแนวโน้มราคาน้ำมันและส่งสัญญาณการซื้อขายน้ำมันให้ดียิ่งขึ้นเมื่อเทียบกับกลยุทธ์การซื้อขายน้ำมันแบบดั้งเดิม โดยหลักการของแบบจำลองคือ CNN สามารถตรวจจับรูปแบบในตำแหน่งต่าง ๆ ของข้อมูล Time Series ได้ ในขณะที่ LSTM สามารถใช้รักษาความจำทั้งระยะสั้นและระยะยาวสำหรับข้อมูล Time Series ได้ การผสมผสานคุณสมบัติเหล่านี้จึงเพิ่มความสามารถให้แบบจำลองได้ จากการศึกษาที่พบว่าการผสมผสานกันของ CNN และ LSTM สามารถเพิ่มความสามารถในการทำกำไรจากการซื้อขายน้ำมันดิบได้ในระยะยาว

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ปีการศึกษา 2564

ลายมือชื่อนิสิต

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ลายมือชื่อ อ.ที่ปรึกษาร่วม

6370267221 : MAJOR COMPUTER SCIENCE

KEYWORD: crude oil trading, machine learning, deep learning, trading signal,
technical analysis, artificial intelligent

Wisaroot Lertthaweedeche : COMBINING TECHNICAL ANALYSIS AND DEEP
LEARNING MODELS FOR CRUDE OIL TRADING. Advisor: Prof. BOONSERM
KIJSIKUL, D.Eng. Co-advisor: PITTIPOL KANTAVAT, Ph.D.

Crude oil is an important commodity in the world because it is the primary source of global energy production. Crude oil price involves in various industries such as transportation, power generation, and petrochemical. Therefore, the prediction of crude oil price is vital for many sectors, but it is very challenging because of its high volatility. Several research papers proposed different machine learning techniques to forecast crude oil prices. In this study, we propose an artificial neural network (ANN) with various combinations of convolutional neural networks (CNN) and long short-term memory (LSTM) to predict crude oil price trends and provide better trading signals for crude oil compared to traditional trading strategies. The concept of our model is that CNN could detect patterns in different locations of time series data, while LSTM could maintain memory for both short-term and long-term for time series data. The combination of their properties could enhance the performance of the neural network. This study found that the combination of CNN and LSTM could significantly improve trading performance in the long run.

Field of Study: Computer Science

Academic Year: 2021

Student's Signature

Advisor's Signature

Co-advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความกรุณาจาก อาจารย์ ดร. พิตติพล คันธวัฒน์ และ ศาสตราจารย์ ดร. บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ทั้งสองท่านเป็นผู้ที่คอยให้คำปรึกษา ในการทั้งในด้านวิชาการและด้านอื่น ๆ นอกจากนี้ยังเป็นผู้ช่วยตรวจทานให้คำแนะนำในการปรับปรุง วิทยานิพนธ์นี้จนสำเร็จไปได้ด้วยดี ผมจึงขอขอบพระคุณอาจารย์ทั้งสองท่านเป็นอย่างสูงมา ณ ที่นี้ด้วย ครับ

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร. สุกรี สิ้นธุภิญโญ และ ผู้ช่วยศาสตราจารย์ ดร. กฤษฎา นิมนานันท์ ที่ให้เกียรติมาเป็นประธานกรรมการและกรรมการสอบวิทยานิพนธ์ และช่วยชี้แนะแนวทาง ในการปรับปรุงวิทยานิพนธ์ให้มีคุณภาพยิ่งขึ้น

ขอขอบคุณ บริษัท Chevron Thailand Exploration And Production สำหรับการ สนับสนุนทุนการศึกษาการเรียนปริญญาโทตลอดหลักสูตร

สุดท้ายนี้ขอขอบคุณ คุณพ่อ คุณแม่ ครอบครัว และเพื่อน ๆ ที่คอยให้กำลังใจในการเรียน ระดับปริญญาโท และการทำวิทยานิพนธ์เสมอมา จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

วิศรุต เลิศทวีเดช

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญ

	หน้า
.....	ค
บทคัดย่อภาษาไทย.....	ค
.....	ง
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฉ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	3
1.3 ขอบเขตการดำเนินงาน.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1 การเรียนรู้เชิงลึก.....	4
2.1.1 โครงข่ายประสาทเทียม (Artificial Neural Networks หรือ ANN) [17].....	4
2.1.2 โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Networks) [17].....	6
2.1.3 Recurrent Neural Network (RNN).....	6
2.1.4 หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory - LSTM) [19].....	7
2.1.5 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network - CNN) [17]	
.....	10

2.2 การวิเคราะห์เชิงเทคนิค (Technical Analysis) [21]	11
2.2.1 ค่าเฉลี่ยเคลื่อนที่ (Moving Average หรือ MA)	12
2.2.2 ค่าเฉลี่ยเคลื่อนที่ลู่ออก/ลู่อเข้า (Moving Average Convergence/Divergence หรือ MACD).....	13
2.2.3 ดัชนีกำลังสัมพันธ์ (Relative Strength Index หรือ RSI).....	14
2.2.4 เส้นแกว่งสโตคาสติก (Stochastic Oscillators)	15
บทที่ 3 งานวิจัยที่เกี่ยวข้อง	16
บทที่ 4 วิธีดำเนินงาน	19
4.1 การเตรียมข้อมูลก่อนการทดลอง (Data Preprocessing)	19
4.1.1 ข้อมูลดิบและการคำนวณตัวชี้วัดทางเทคนิค.....	19
4.1.2 การแปลงข้อมูล (Data Transformation).....	23
4.1.3 การติดฉลากข้อมูล (Labeling)	28
4.1.4 การแบ่งข้อมูลเพื่อการฝึกแบบจำลอง.....	30
4.2 Deep Learning Model.....	31
4.3 วิธีการวัดผล	32
4.3.1 ค่าความผิดพลาดจากการทำนาย	32
4.3.2 ประสิทธิภาพการซื้อขาย	33
4.3.3 ระบบจำลองการซื้อขาย (Backtest).....	34
4.4 ขั้นตอนการดำเนินงาน	37
บทที่ 5 ผลการทดลอง	38
5.1 ค่าความผิดพลาดจากการทำนาย.....	38
5.2 ประสิทธิภาพการซื้อขาย	40
บทที่ 6 สรุปการทดลอง และข้อเสนอแนะ	44
6.1 สรุปการทดลอง.....	44

6.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต.....	44
บรรณานุกรม	45
ประวัติผู้เขียน	49



สารบัญตาราง

	หน้า
ตารางที่ 1 ตัวแปรและรายละเอียดที่ใช้เพื่อเป็นข้อมูลนำเข้าไปในแบบจำลอง ประกอบด้วย ราคา, ปริมาณการซื้อขาย, และตัวชี้วัดแบบต่าง ๆ [28] [29].....	20
ตารางที่ 2 การแบ่งข้อมูลสำหรับการทดลอง	30
ตารางที่ 3 MSE ของชุดข้อมูลทดสอบในแต่ละปี โดยมีหน่วยเป็น $1 \times 10^{-2} \%^2$ และตัวหนาแสดงถึงค่า MSE ที่น้อยที่สุดของแต่ละกลุ่มข้อมูล	40
ตารางที่ 4 ผลตอบแทนการลงทุน (ROI) ของทุกกลยุทธ์การซื้อขาย โดยตัวหนาแสดงถึงค่าผลตอบแทนการลงทุนที่สูงที่สุดของแต่ละกลุ่มข้อมูล	41
ตารางที่ 5 Sharpe Ratio ของทุกกลยุทธ์การซื้อขาย โดยตัวหนาแสดงถึงค่าผลตอบแทนการลงทุนที่สูงที่สุดของแต่ละกลุ่มข้อมูล	42

สารบัญรูปภาพ

หน้า

รูปที่ 1 ภาพจำลองแสดงโครงข่ายประสาททางชีววิทยา [17].....	4
รูปที่ 2 สถาปัตยกรรมพื้นฐานของโครงข่ายประสาทเทียม [17].....	5
รูปที่ 3 ตัวอย่างสถาปัตยกรรมโครงข่ายประสาทเทียมที่ประกอบไปด้วย 1 ชั้นนำเข้า, 2 ชั้นที่ถูกซ่อน, และ 1 ชั้นนำออก ต่อกันแบบอนุกรมตามลำดับ [17].....	5
รูปที่ 4 ตัวอย่างสถาปัตยกรรมของโครงข่ายประสาทเทียมเชิงลึกที่มีความซับซ้อนมากกว่าปกติ [17]	6
รูปที่ 5 สถาปัตยกรรมแบบย่อของ RNN และสถาปัตยกรรมแบบแสดงชั้นลำดับของ RNN [17].....	7
รูปที่ 6 สถาปัตยกรรมของ LSTM [19].....	8
รูปที่ 7 สถาปัตยกรรมของ LSTM แบบแสดงรายละเอียดของ 4 เกต [19].....	8
รูปที่ 8 ตัวอย่างการคอนโวลูชันจากข้อมูลนำเข้าขนาด $7 \times 7 \times 1$ และ Filter ขนาด $3 \times 3 \times 1$	10
รูปที่ 9 ตัวอย่างกราฟแสดงราคาและเส้นค่าเฉลี่ยเคลื่อนที่ 20 วันและ 60 วัน [21].....	12
รูปที่ 10 ตัวอย่างกราฟแสดงเส้น MACD และ เส้น Signal [21].....	13
รูปที่ 11 ตัวอย่างกราฟแสดงเส้นดัชนีกำลังสัมพันธ์ (RSI) [21].....	14
รูปที่ 12 ตัวอย่างกราฟแสดงเส้นแกว่งสโตคาสติก [21].....	15
รูปที่ 13 โครงสร้างแบบจำลอง CNN-LSTM ที่ถูกนำเสนอโดย Rezaei และคณะ [25].....	17
รูปที่ 14 สัดส่วนงานวิจัยที่ใช้ตัวแปรต่าง ๆ กันเป็นข้อมูลนำเข้าจากผลการสำรวจในงานทบทวนงานวิจัยการทำนายราคาหลักทรัพย์ของ Kumar และคณะ [11].....	18
รูปที่ 15 ความแตกต่างของข้อมูลราคาเปิดหลังจากแปลงโดยใช้ร้อยละที่แตกต่างเทียบกับราคาปิด ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง.....	24
รูปที่ 16 ความแตกต่างของข้อมูลราคาปิดหลังจากการแปลงโดยใช้ร้อยละที่แตกต่างเทียบราคาปิดในช่วงเวลาก่อนหน้า ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง.....	25
รูปที่ 17 ความแตกต่างของข้อมูล OBV หลังจากการแปลงโดย Quantile Transformation ประกอบด้วย (1) ข้อมูลก่อนการแปลง, (2) การกระจายตัวของข้อมูลก่อนการแปลง, (3) ข้อมูลหลังการแปลง และ (4) การกระจายตัวของข้อมูลหลังการแปลง.....	26

รูปที่ 18 ความแตกต่างของราคาปิดหลังจากการแปลงข้อมูลโดย Data Normalization ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง27

รูปที่ 19 ตัวอย่างการทำเส้นข้อมูลเฉลี่ยโดยใช้ Savitzky-Golay [32] แบบ Polynomial degree ที่ 3 และข้อมูลรอบข้าง 9 จุด.....29

รูปที่ 20 (1) ราคาปิดน้ำมัน WTI และราคาเฉลี่ยโดย Savitzky-Golay filter และ (2) ร้อยละการเปลี่ยนแปลงจากช่วงเวลาก่อนหน้าของราคาเฉลี่ย29

รูปที่ 21 แผนภาพโครงข่ายประสาทเทียม CNN-LSTM ในงานวิจัยนี้ซึ่งมีลำดับของชั้น CNN มาก่อน LSTM.....31

รูปที่ 22 แผนภาพโครงข่ายประสาทเทียม LSTM-CNN ในงานวิจัยนี้ซึ่งมีลำดับของชั้น LSTM มาก่อน CNN.....32

รูปที่ 23 ร้อยละการเปลี่ยนแปลงของราคาเฉลี่ย และผลการทำนายจากแบบจำลอง34

รูปที่ 24 ผลการจำลองการซื้อขายจากกลยุทธ์ที่จากผลการทำนายร้อยละการเปลี่ยนแปลงรายวันของเส้นราคาเฉลี่ยของแบบจำลอง CNN-LSTM ในปี ค.ศ. 201736

รูปที่ 25 ผล MSE ของชุดข้อมูลสำหรับฝึกและชุดข้อมูลตรวจสอบของกลุ่มข้อมูลปี 201739

รูปที่ 26 ผลตอบแทนการลงทุนเฉลี่ยจากทุกกลุ่มข้อมูลในช่วง ค.ศ. 2015 – 2020 ของทุกกลยุทธ์การซื้อขาย41

รูปที่ 27 Sharpe Ratio เฉลี่ยจากทุกกลุ่มข้อมูลในช่วง ค.ศ. 2015 – 2020 ของทุกกลยุทธ์การซื้อขาย43

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

น้ำมันดิบเป็นสินค้าอุปโภคที่มีความสำคัญต่อเศรษฐกิจในปัจจุบันเป็นอย่างมาก เนื่องจากอุปทานพลังงานทั่วโลกกว่า 30% มาจากน้ำมันดิบ [1] และถูกใช้ในกิจกรรมทางเศรษฐกิจต่าง ๆ มากมาย เช่น การเดินทางโดยรถยนต์ เรือ และเครื่องบิน, การผลิตไฟฟ้า, และวัตถุดิบสำหรับอุตสาหกรรมปิโตรเคมี ราคาน้ำมันดิบจึงมีความสำคัญอย่างยิ่งต่ออุตสาหกรรมต่าง ๆ ที่เกี่ยวข้องกับกิจกรรมทางเศรษฐกิจต่าง ๆ เหล่านี้ นอกจากนี้ภาคเอกชนแล้วราคาน้ำมันดิบยังมีความสำคัญกับภาครัฐโดยเฉพาะธนาคารกลาง เนื่องจากน้ำมันดิบส่งผลโดยตรงต่อการเติบโตทางเศรษฐกิจของประเทศทำให้มีผลต่อการกำหนดนโยบายทางการเงินต่าง ๆ จากภาครัฐและธนาคารกลาง

แต่อย่างไรก็ตามราคาน้ำมันดิบมีความผันผวนสูงมาก อันเนื่องมาจากการเปลี่ยนแปลงอุปสงค์และอุปทานต่อน้ำมันดิบในแต่ละช่วงเวลา ซึ่งทั้งอุปสงค์และอุปทานของน้ำมันดิบนั้นขึ้นกับหลาย ๆ ปัจจัย ทั้งกิจกรรมทางเศรษฐกิจ รวมถึงกำลังการผลิตของพลังงานทางเลือกอื่น ๆ เช่น ก๊าซธรรมชาติ, ถ่านหิน, พลังงานทดแทน ซึ่งทำให้เกิดผลกระทบแบบทดแทน (Substitution effect) ได้ ปัจจัยอื่น ๆ เช่น การเติบโตทางเศรษฐกิจ การพัฒนาของเทคโนโลยี เหตุการณ์ไม่ปกติ โดยตัวอย่างที่ชัดเจนคือการระบาดของไวรัสโควิด-19 ปัจจัยต่าง ๆ เหล่านี้ล้วนส่งผลกระทบต่อราคาน้ำมัน ทำให้ราคาน้ำมันมีความผันผวนสูงทั้งสิ้น [2]

มีหลายงานวิจัยที่พยายามศึกษาแบบจำลองการทำนายราคาน้ำมันดิบหลากหลายแบบ เช่น ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine หรือ SVM) [3], [4], Autoregressive Integrated Moving Average (ARIMA) [3], [4], [5], Random walk [3], Genetic algorithm (GA) [6], Generalized autoregressive conditional heteroskedasticity (GARCH) [7], Vector autoregressive (VAR) [8], error correction model (ECM) [9] และด้วยการพัฒนาของเทคโนโลยีการคำนวณที่มีความสามารถมากขึ้นอย่างมาก แบบจำลองที่สามารถเข้าใจความซับซ้อนและไม่เป็นเส้นตรง (nonlinearity) ของอนุกรมเวลาทางการเงิน (Financial time series) ต่าง ๆ ได้ดีขึ้นแต่จำเป็นต้องใช้กำลังการคำนวณสูง อย่างเช่น โครงข่ายประสาทเทียม (Artificial Neural Network) [10] ถูกนำมาใช้ในการทำนายราคาสินทรัพย์ต่าง ๆ อย่างมากมาย

โครงข่ายประสาทเทียมถูกนำมาใช้ในการทำนายราคาของสินทรัพย์ต่าง ๆ เนื่องจากแบบจำลองการเรียนรู้ของเครื่องชนิดนี้สามารถเข้าใจความซับซ้อนและไม่เป็นเส้นตรงของข้อมูลราคา

สินทรัพย์ได้ ทำให้มีประสิทธิภาพในการทำนายราคาสินทรัพย์ต่าง ๆ [11] อย่างเช่นงานวิจัยของ Gudelek และคณะ [12] ใช้ โครงข่ายประสาทเทียมคอนโวลูชัน (CNN) ทำนายราคาหลักทรัพย์ได้ ความแม่นยำสูง สามารถใช้สร้างกลยุทธ์การซื้อขายที่ได้ทำไรมากกว่าการซื้อขายและถือ Tsantekidis และคณะ [13] ยังพบว่าแบบจำลอง สามารถทำนายราคาหลักทรัพย์ได้แม่นยำกว่า SVM อีกด้วย นอกจาก CNN แล้ว LSTM ยังเป็นอีกโครงข่ายประสาทเทียมที่นิยมนำมาใช้ทำนายราคาสินทรัพย์ ต่างๆ อย่างแพร่หลาย อย่างเช่น Li และคณะ [14] พบว่า LSTM มีความถูกต้องในการทำนายราคา หลักทรัพย์มากกว่า SVM ในกรณีที่มีการแกว่งตัวของราคาน้อยกว่า Cheng และคณะ [15] ใช้ Attention-based LSTM เพื่อทำนายประเภทการเคลื่อนไหวของราคาสินทรัพย์จากข้อมูลข่าว ใน วิจัยนี้จึงประยุกต์ใช้ CNN ควบคู่กับ LSTM เพื่อทำนายแนวโน้มราคาน้ำมัน

ในงานวิจัยนี้เราจึงนำเสนอการผสมผสานกันระหว่าง CNN และ LSTM และศึกษาหาลำดับ โครงสร้างที่เหมาะสมของ CNN และ LSTM ในโครงข่ายประสาทเทียม เพื่อใช้ในสร้างสัญญาณซื้อ ขายน้ำมันดิบที่ให้ผลตอบแทนเฉลี่ยดีกว่าผลตอบแทนที่ได้จากการซื้อขายน้ำมันดิบตามสัญญาณบ่งชี้ แบบดั้งเดิมในระยะยาว นอกจากนี้งานวิจัยนี้ยังนำเสนอแนวทางการทำนายแนวโน้มของราคา โดยใช้ ราคาเฉลี่ย (Smoothed price) จาก Savitzky-Golay filter [16] เพื่อเพิ่มประสิทธิภาพการส่ง สัญญาณซื้อขายของแบบจำลอง อีกทั้งงานวิจัยนี้วัดประสิทธิภาพของแบบจำลองด้วยผลตอบแทนการ ลงทุน (Return on investment, ROI) และ Sharpe ratio ของพอร์ตการลงทุนที่สร้างจากการ จำลองการซื้อขายน้ำมันดิบกับข้อมูลราคาในอดีต

1.2 วัตถุประสงค์

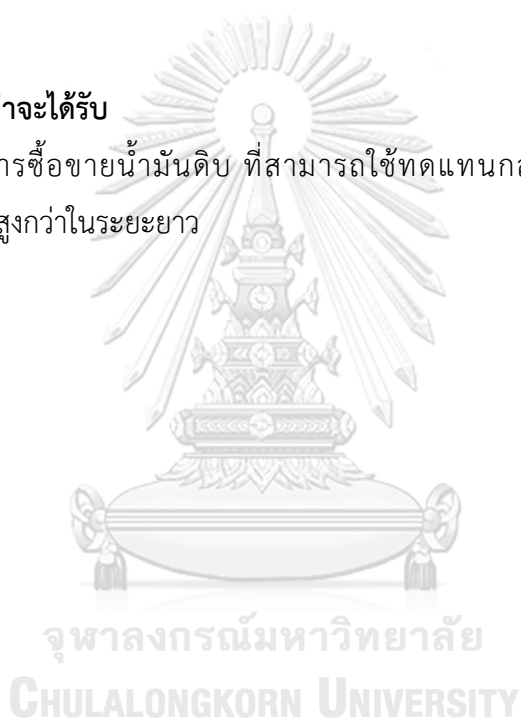
เพื่อสร้างโครงข่ายประสาทเทียมที่สามารถให้สัญญาณการซื้อขายน้ำมันดิบที่ให้ผลตอบแทนเฉลี่ยดีกว่าผลตอบแทนที่ได้จากการซื้อขายน้ำมันดิบตามสัญญาณบ่งชี้แบบดั้งเดิมในระยะยาว

1.3 ขอบเขตการดำเนินงาน

ใช้ข้อมูลราคาน้ำมัน WTI ในช่วงปี ค.ศ. 2005 ถึง ค.ศ. 2020 โดยทดลองซื้อขายจริงในช่วงปี ค.ศ. 2015 ถึง ค.ศ. 2020 และวัดผลโดยใช้ MSE, ผลตอบแทนการลงทุน และ Sharpe Ratio

1.4 ประโยชน์ที่คาดว่าจะได้รับ

ได้กลยุทธ์การซื้อขายน้ำมันดิบ ที่สามารถใช้ทดแทนกลยุทธ์แบบดั้งเดิมได้ และทำผลตอบแทนโดยเฉลี่ยสูงกว่าในระยะยาว



บทที่ 2

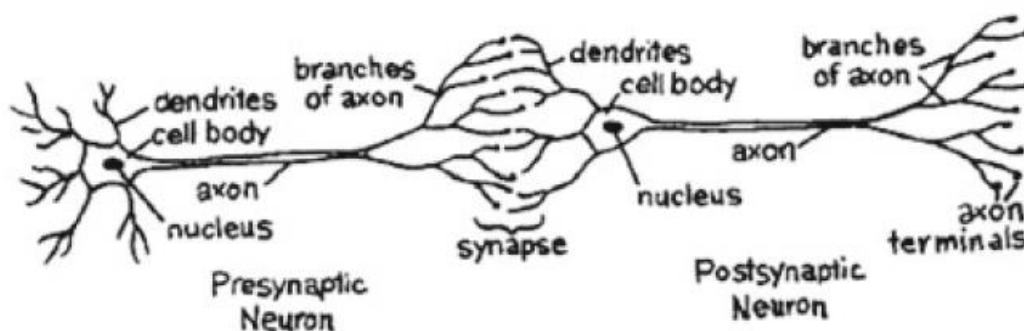
ทฤษฎีที่เกี่ยวข้อง

2.1 การเรียนรู้เชิงลึก

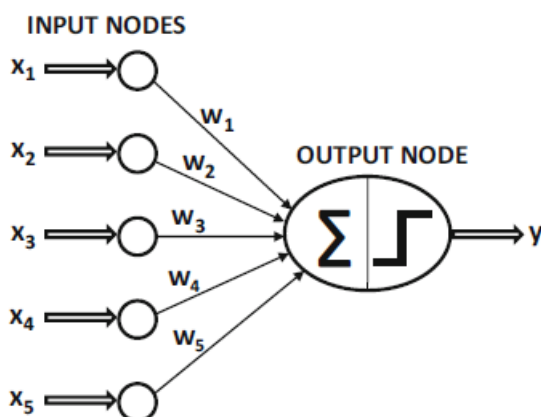
ในส่วนนี้จะอธิบายถึงโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึกที่ใช้เป็นพื้นฐานของแบบจำลองเพื่อการเรียนรู้ของเครื่องในงานวิจัยนี้

2.1.1 โครงข่ายประสาทเทียม (Artificial Neural Networks หรือ ANN) [17]

โครงข่ายประสาทเทียมเป็น วิธีหนึ่งที่ยอมรับใช้เพื่อการเรียนรู้ของเครื่อง (Machine Learning) โครงข่ายประสาทเทียมมาจากแนวคิดเลียนแบบมาจากโครงข่ายประสาทในสมองของมนุษย์ตามภาพจำลองในรูปที่ 1 ซึ่งประกอบไปด้วยเซลล์เรียกว่า เซลล์ประสาท (Neurons) และแต่ละเซลล์ประสาทเชื่อมต่อกันด้วย แกนประสาทนำออก (Axons) และ แกนประสาทนำเข้า (Dendrites) โครงข่ายประสาทเทียมทางชีววิทยานี้ได้ถูกนำมาจำลองเป็นโครงข่ายประสาทเทียมดังแสดงในรูปที่ 2 โครงข่ายประสาทเทียมประกอบด้วย หน่วยคำนวณย่อยเรียกว่า Node เปรียบเสมือนเซลล์ประสาท และระหว่าง Node จะมีเส้นเชื่อมต่อกันคล้ายแกนประสาทนำออก และ แกนประสาทนำเข้า ในโครงข่ายประสาทเทียมทางชีววิทยา เรียกว่า Edge แต่ละ Edge จะบรรจุค่าน้ำหนัก (Weight) โดยเฉพาะตัวของมันอยู่ภายใน เมื่อโครงข่ายประสาทเทียมถูกฝึก ค่าน้ำหนักของแต่ละ Edge ในแบบจำลองจะเกิดการเปลี่ยนแปลง ทำให้แบบจำลองมีการเรียนรู้และจดจำรูปแบบของข้อมูล เกิดเป็นความรู้ของแบบจำลองเพื่อนำไปใช้ประมวลผลข้อมูลอื่น ๆ ในรูปแบบเดียวกันได้

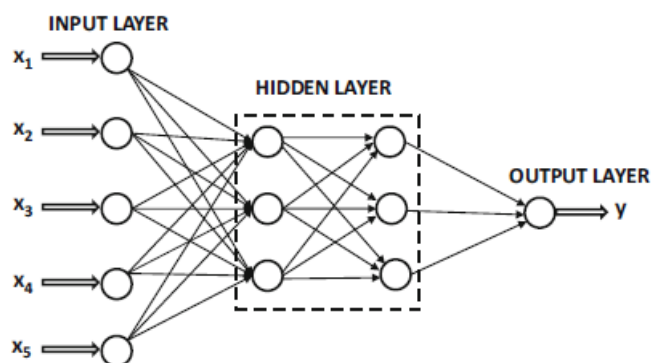


รูปที่ 1 ภาพจำลองแสดงโครงข่ายประสาททางชีววิทยา [17]



รูปที่ 2 สถาปัตยกรรมพื้นฐานของโครงข่ายประสาทเทียม [17]

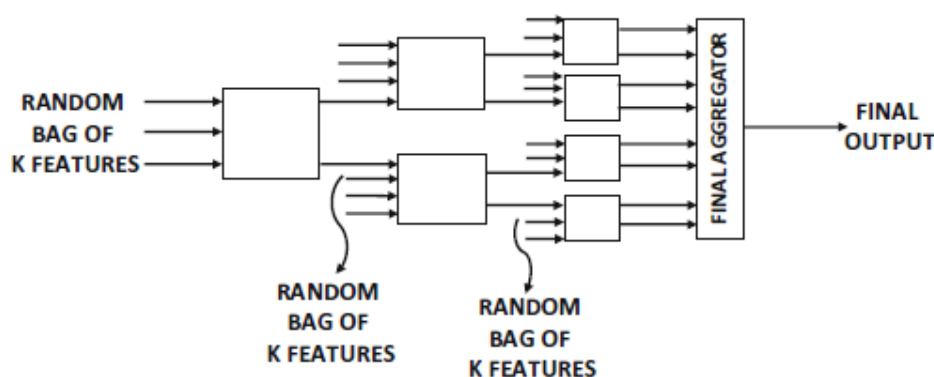
เมื่อ Node เซลล์ประสาทเทียมถูกนำมาเรียงกันแบบขนานจะรวมกันเป็นชั้น (layers) และชั้นสามารถแบ่งประเภทได้เป็น ชั้นนำเข้า (input layers), ชั้นที่ถูกซ่อน (hidden layers), และชั้นนำออก (output layers) ซึ่งจะต่อกันแบบอนุกรมตามลำดับตามตัวอย่างที่แสดงในรูปที่ 3 ชั้นนำเข้าจะส่งข้อมูลไปตามโครงข่าย ทุก ๆ Node ในแต่ละชั้นจะรับจะนำข้อมูลจากชั้นก่อนหน้ามาคำนวณร่วมกับค่าน้ำหนักและฟังก์ชันกระตุ้น (activation function) ก่อนที่จะส่งออกเป็นข้อมูลนำออกให้ชั้นถัดไป จนไปถึงชั้นนำออก ซึ่งข้อมูลนำออกที่ได้จากชั้นนำออกจะเป็นผลการทำนายจากแบบจำลองโครงข่ายก็จะมีทั้งแบบ การจัดแบ่งประเภท (Classification) หรือ การถดถอย (Regression) จะเห็นได้ว่าข้อมูลนำออกที่ได้มาจากโครงข่ายประสาทเทียมจะผ่านกระบวนการคำนวณที่ใช้ค่าน้ำหนักและฟังก์ชันกระตุ้นตามจำนวน Node ในโครงข่าย ทำให้โครงข่ายประสาทเทียมสามารถเรียนรู้ความสัมพันธ์ระหว่างข้อมูลและเป้าหมายที่ซับซ้อนและไม่เป็นเส้นตรงได้



รูปที่ 3 ตัวอย่างสถาปัตยกรรมโครงข่ายประสาทเทียมที่ประกอบไปด้วย 1 ชั้นนำเข้า, 2 ชั้นที่ถูกซ่อน, และ 1 ชั้นนำออก ต่อกันแบบอนุกรมตามลำดับ [17]

2.1.2 โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Networks) [17]

การเรียนรู้เชิงลึก (Deep Learning) คือการใช้โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Networks) ในการเรียนรู้ข้อมูลของเครื่อง ทั้งนี้คำว่า “เชิงลึก” ในที่นี้ไม่ได้หมายถึงความเข้าใจที่ลึกซึ้งขึ้นของแบบจำลอง เพียงแต่หมายถึงจำนวนชั้นที่ถูกซ่อนในโครงข่ายประสาทเทียมมีจำนวนมาก [18] โครงข่ายประสาทเทียมเชิงลึกจึงหมายถึงโครงข่ายประสาทเทียมที่มีชั้นเป็นจำนวนมาก หรือมีโครงข่ายที่ซับซ้อนกว่าปกติตามตัวอย่างในรูปที่ 4 ด้วยความซับซ้อนของแบบจำลองทำให้แบบจำลองสามารถเข้าใจรูปแบบความสัมพันธ์ที่ซับซ้อนมากขึ้นระหว่างข้อมูลนำเข้าและเป้าหมาย แต่อย่างไรก็ตามการเรียนรู้เชิงลึกที่มีประสิทธิภาพต้องใช้กับข้อมูลขนาดใหญ่และคอมพิวเตอร์ที่มีกำลังคำนวณสูง ซึ่งด้วยพัฒนาการอย่างก้าวกระโดดของการเก็บข้อมูล และกำลังการคำนวณของคอมพิวเตอร์ในช่วงทศวรรษที่ผ่านมา ทำให้การเรียนรู้เชิงลึกมีประสิทธิภาพมากขึ้นอย่างก้าวกระโดดและถูกนำมาประยุกต์ใช้กับข้อมูลหลากหลายรูปแบบอย่างแพร่หลายด้วยเช่นกัน

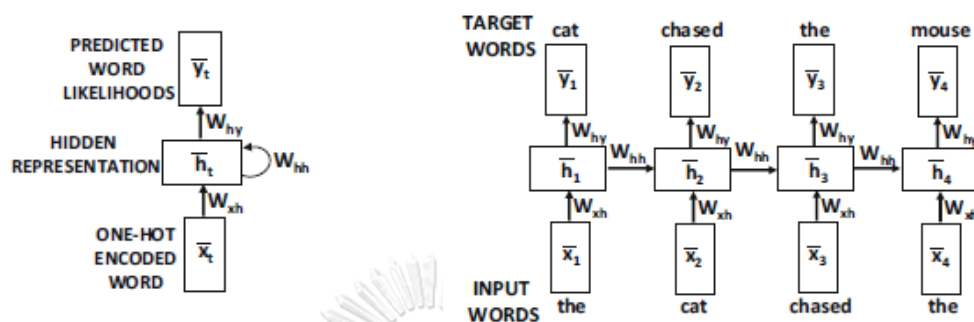


รูปที่ 4 ตัวอย่างสถาปัตยกรรมของโครงข่ายประสาทเทียมเชิงลึกที่มีความซับซ้อนมากกว่าปกติ [17]

2.1.3 Recurrent Neural Network (RNN)

Recurrent Neural Network เป็นโครงข่ายประสาทเทียมประเภทหนึ่งที่มีการวนภายใน Node ทำให้สามารถประมวลผลข้อมูลที่เรียงกันเป็นลำดับหรือมีลักษณะเป็นอนุกรมเวลา (Time-Series) ได้ [18] รูปที่ 5(1) แสดงสถาปัตยกรรมแบบย่อของ RNN ซึ่ง RNN ประมวลผลของอนุกรมข้อมูลนำเข้าโดยการนำข้อมูลนำเข้าของ Node วนกลับเข้าไปใน Node เดิมร่วมกับข้อมูลนำเข้าในลำดับถัดไป ซึ่งการทำเช่นนี้ทำให้ RNN สามารถเก็บและรักษา State ที่บรรจุข้อมูลที่แบบจำลองได้เห็นมาในลำดับก่อนหน้าไว้ได้ ซึ่งการวนรับข้อมูลนำเข้าในลำดับถัดไปในแต่ละครั้ง จะทำให้ State

เปลี่ยนไปและแบบจำลองเรียนรู้ข้อมูลนำเข้าทั้งอนุกรม ซึ่งสามารถดูสถาปัตยกรรมแบบแสดงชั้นลำดับในรูปที่ 5(2) ที่เกิดจากการวางสถาปัตยกรรมแบบย่อในรูปที่ 5(1) ออก

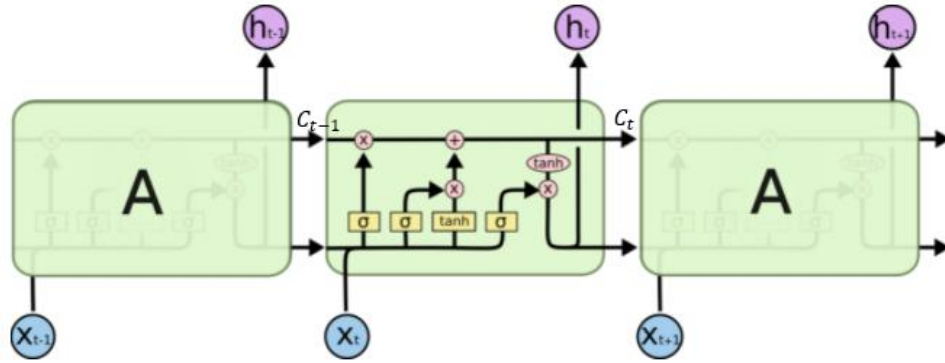


- (1) สถาปัตยกรรมแบบย่อของ RNN (2) สถาปัตยกรรมแบบแสดงชั้นลำดับของ RNN
รูปที่ 5 สถาปัตยกรรมแบบย่อของ RNN และสถาปัตยกรรมแบบแสดงชั้นลำดับของ RNN [17]

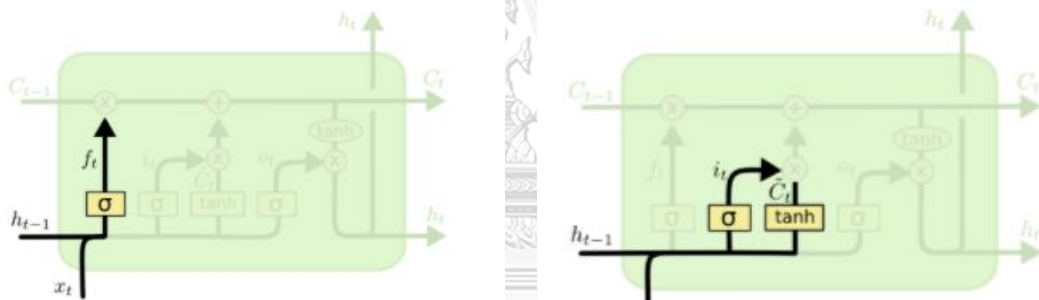
ด้วยรูปแบบการเรียนรู้ใน RNN ที่กล่าวมาข้างต้นทำให้ RNN มีความสามารถในการเรียนรู้ข้อมูลในรูปแบบลำดับหรืออนุกรมเวลาได้ ซึ่งเป็นลักษณะข้อมูลที่พบได้ทั่วไปสำหรับข้อมูลทางการเงินและการซื้อขายสินทรัพย์ต่าง ๆ รวมถึงการซื้อขายน้ำมันดิบเช่นกัน ตัวอย่างข้อมูลเช่น ราคา น้ำมันดิบ, ปริมาณการซื้อขาย ล้วนอยู่ในรูปแบบของอนุกรมเวลาทั้งสิ้น โครงข่าย RNN จึงมีความเหมาะสมเป็นอย่างมากที่จะนำมาประยุกต์ใช้สำหรับงานวิจัยนี้ แต่อย่างไรก็ตาม RNN ยังมีการแบ่งประเภทแยกย่อยไปอีกหลายรูปแบบ ซึ่งจะมีความเหมาะสมกับลักษณะข้อมูล และเป้าหมายของการเรียนรู้ของเครื่องที่แตกต่างกัน โดยจะกล่าวถึงรูปแบบย่อยของ RNN ที่ใช้ในงานวิจัยนี้ในส่วนถัดไป

2.1.4 หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory - LSTM) [19]

LSTM ถูกนำเสนอโดย Hochreiter และ Schmidhuber [20] เพื่อแก้ปัญหาใน RNN ที่มักจะมีปัญหาการปรับเปลี่ยนค่าน้ำหนักในขั้นต้น ๆ น้อยเกินไป หรือ Vanishing gradient เพราะอนุกรมของข้อมูลนำเข้ามีความยาวมากเกินไป เป็นผลให้แบบจำลองไม่สามารถเรียนรู้จากข้อมูลในส่วนแรก ๆ ได้ แต่ LSTM สามารถเรียนรู้และมีความจำสำหรับข้อมูลในระยะยาวได้ โดยส่วนสำคัญของ LSTM ก็คือเซลล์ State (C_t) ที่จะเดินทางยาวไปตาม Node ของ LSTM และมีการรับค่าเพื่อเปลี่ยนแปลงเซลล์ State จาก 4 เกตในแต่ละ Node ดังที่แสดงในรูปที่ 6



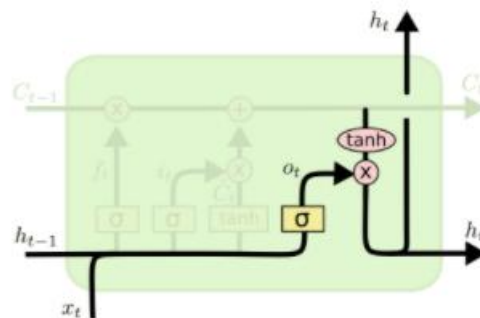
รูปที่ 6 สถาปัตยกรรมของ LSTM [19]



(1) เกตที่หนึ่ง (2) เกตที่สองและสาม

จุฬาลงกรณ์มหาวิทยาลัย

C



(3) เกตที่สี่

รูปที่ 7 สถาปัตยกรรมของ LSTM แบบแสดงรายละเอียดของ 4 เกต [19]

เกตที่ 1 ดังแสดงในรูปที่ 7(1) เป็น เกตซิกมอยด์ หรือสามารถเรียกได้ว่าเป็นเกตฟอรัเกต (Forget gate) คำนำออกจากเกตนี้จะได้ f_t ที่ถูกสร้างขึ้นมาจาก ฟังก์ชันซิกมอยด์ที่รับค่า h_{t-1} และ

x_t ตามสมการ (1) ผลที่ได้จะมีค่าระหว่าง 0 ถึง 1 เพื่อนำมาคูณกับ C_{t-1} โดย 1 หมายถึงเก็บรักษา State ก่อนหน้าไว้ทั้งหมดและ 0 หมายถึงลืม State ก่อนหน้าทั้งหมด

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

ในส่วนถัดไปตามรูปที่ 7(2) จะเป็นการตัดสินใจการใช้หรือจำข้อมูลใหม่เพื่อนำเข้ามาในเซลล์ State ในส่วนนี้จะมี 2 เกต เกตแรกเป็นเกตซิกมอยด์ และเกตที่สองเป็นเกตแทนเอช (tanh gate) ซึ่งจะถูกรวมกันแล้วใช้ในการปรับเปลี่ยนเซลล์ State ตามสมการที่ (2) และ (3)

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

ส่วนสุดท้ายตามรูปที่ 7(3) คือส่วนที่ตัดสินใจการปล่อยค่านำออกของ Node ในส่วนนี้จะเริ่มจากการนำ h_{t-1} และ x_t มาเข้าฟังก์ชันซิกมอยด์ได้เป็น o_t ตามสมการที่ (4) จากนั้นนำ C_t มาเข้าฟังก์ชันแทนเอชและนำไปคูณกับ o_t ตามสมการที่ (5) เพื่อให้ค่านำออกจาก Node มีเพียงส่วนที่ถูกเลือกแล้วเท่านั้น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

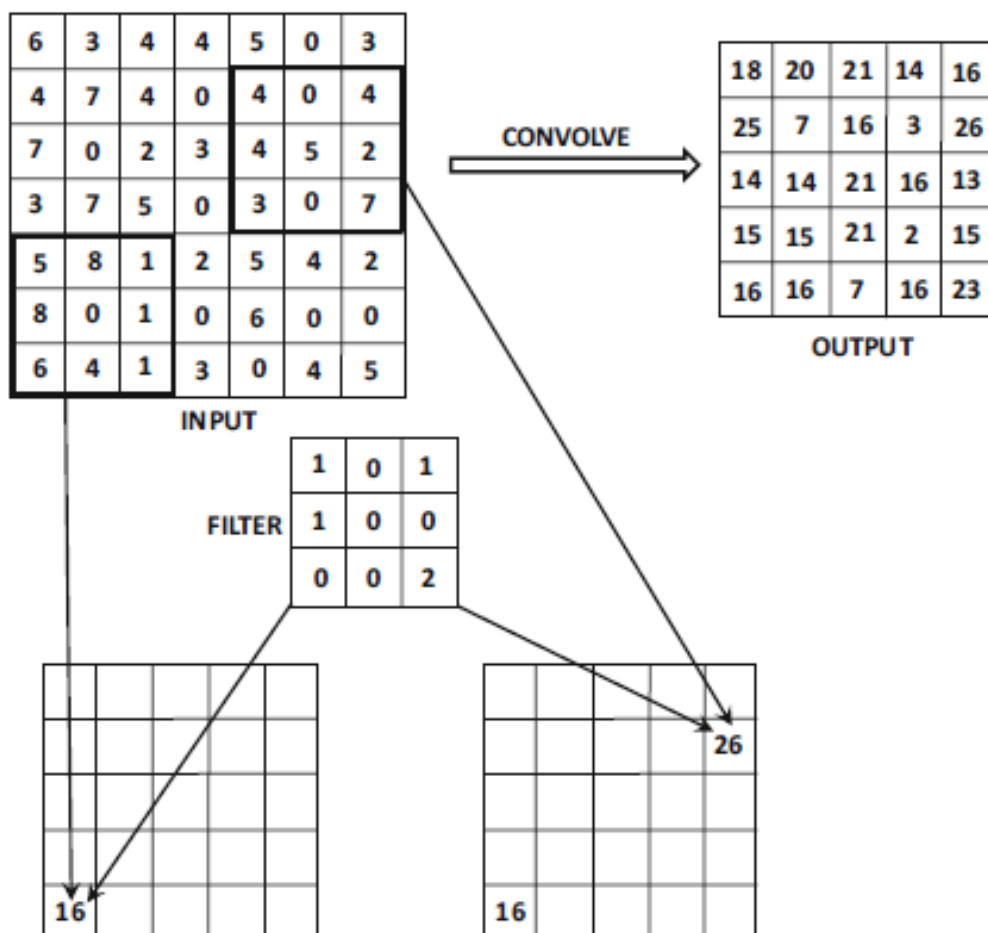
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t * \tanh(C_t) \quad (5)$$

ค่านำออกที่ได้จะเป็น State ของ Node และจะถูกใช้เป็นตัวนำเข้าไปใน Node ถัดไป ร่วมกับเซลล์ State เพื่อช่วยในการจำข้อมูลก่อนหน้านี้ ด้วยรูปแบบของเกตและเซลล์ State ใน LSTM ทำให้แบบจำลองประเภทนี้สามารถแก้ปัญหา Vanishing gradient ได้

2.1.5 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network - CNN) [17]

CNN เป็นแบบจำลองโครงข่ายประสาทเทียมที่นิยมนำมาใช้ในการเรียนรู้ข้อมูลรูปภาพ หรือ Computer Vision โดยเป็นรูปแบบอย่างง่ายของโครงข่ายประสาทเทียมทั่วไป ที่เป็น Fully Connected หรือมี Edges เชื่อมข้อมูลระหว่างชั้นก่อนหน้าและชั้นโครงข่ายประสาทเทียมที่สนใจครบทุกคู่ Node แต่สำหรับ CNN จะเป็นการลดจำนวน Edge ระหว่างสองชั้นนี้ลง แต่ละ Node ของชั้น CNN จะมี Edge เชื่อมไปแค่บาง Node ของชั้นก่อนหน้าเท่านั้น ซึ่ง Edge เหล่านี้เรียกว่า Filter หรือ Kernel ดังแสดงในรูปที่ 8 และค่าน้ำหนักใน Filter จะถูกเรียนรู้ได้ในตอนฝึกแบบจำลอง นอกจากนี้ค่าน้ำหนักยังถูกใช้ซ้ำ ๆ กันสำหรับแต่ละ Node ของชั้น CNN อีกด้วย ทั้งนี้เพื่อลดความสำคัญในตำแหน่งของข้อมูลลงได้ CNN จึงสามารถนำมาใช้เรียนรู้ข้อมูล Time Series ที่อาจจะมีรูปแบบข้อมูลซ้ำ ๆ กันแต่อาจจะมีการเปลี่ยนตำแหน่งไปได้



รูปที่ 8 ตัวอย่างการคอนโวลูชันจากข้อมูลนำเข้าขนาด 7x7x1 และ Filter ขนาด 3x3x1

2.2 การวิเคราะห์เชิงเทคนิค (Technical Analysis) [21]

การวิเคราะห์เชิงเทคนิค คือการใช้ข้อมูลราคาหลักทรัพย์ และปริมาณการซื้อขายหลักทรัพย์ในอดีตมาทำนายแนวโน้มราคาหลักทรัพย์ในอนาคต ซึ่งแตกต่างจากการวิเคราะห์หลักทรัพย์เชิงพื้นฐาน (Fundamental Analysis) ที่ใช้ข้อมูลทางบัญชีหรือข้อมูลอื่น ๆ ในการวิเคราะห์หามูลค่าที่แท้จริงของหลักทรัพย์ ซึ่งต้องใช้สมมติฐานบางอย่างในการคำนวณหรือข้อมูลทางบัญชีที่ไม่มีในบางหลักทรัพย์ เช่น ทอง, เงินตรา, น้ำมัน, บิตคอยน์

การวิเคราะห์เชิงเทคนิคมีข้อดีคือ ใช้ข้อมูลเพียงราคาหลักทรัพย์ และปริมาณการซื้อขายในอดีตของหลักทรัพย์ที่สามารถหาได้โดยง่าย, สามารถใช้กับหลักทรัพย์ที่ไม่สามารถสร้างกระแสเงินสดได้อย่างเช่นสินค้าโภคภัณฑ์ (Commodity), ไม่ได้รับผลกระทบในกรณีเกิดการฉ้อโกงหรือปรับแต่งทางบัญชี แต่การวิเคราะห์เชิงเทคนิคจะไม่เหมาะกับตลาดที่มีสภาพคล่องน้อย และมีการบงการ (Manipulation) มูลค่าหลักทรัพย์อยู่เบื้องหลัง เช่น เงินตราที่มีการบงการมูลค่าโดยธนาคารกลางของแต่ละประเทศ

ทั้งนี้การวิเคราะห์เชิงเทคนิคมีสมมติฐานที่สำคัญคือ ราคาเป็นผลมาจากพฤติกรรมโดยรวมของนักลงทุนที่ทั้งมีเหตุผลและไม่มีเหตุผล ซึ่งพฤติกรรมเหล่านี้ทำให้เกิดเป็นแนวโน้มของราคาและมีรูปแบบซึ่งมีแนวโน้มที่จะเกิดซ้ำและสามารถตรวจจับเพื่อทำนายราคาหลักทรัพย์ในอนาคตได้ ซึ่งสมมติฐานนี้มีนัยยะว่าตลาดไม่มีประสิทธิภาพ (Market Inefficiency) นั่นเอง

ตลาดที่มีประสิทธิภาพ [22] หมายถึง การที่ราคาของหลักทรัพย์สะท้อนข้อมูลทั้งหมดที่มีอยู่อย่างเต็มที่ รวดเร็ว และมีเหตุผล ซึ่งวัดได้จากระยะเวลาระหว่างข้อมูลใหม่จนถึงเวลาที่ราคาของหลักทรัพย์เข้าสู่มูลค่าที่แท้จริงตามข้อมูลใหม่นั้น ซึ่งในความเป็นจริงหลาย ๆ ตลาดไม่ได้มีประสิทธิภาพเต็มที่ เช่น การที่มีข่าวร้ายแล้วมีการขายหลักทรัพย์จนราคาต่ำกว่าความเป็นจริงไปมากซึ่งสะท้อนความไม่มีเหตุผลของนักลงทุน, มีค่าธรรมเนียมการซื้อขายที่เป็นอุปสรรคต่อประสิทธิภาพของตลาด, กฎการห้ามใช้ข้อมูลวงในในการซื้อขาย (Insider Trading) ซึ่งล้นแล้วแต่ลดประสิทธิภาพของตลาดทั้งสิ้น ด้วยเหตุผลข้างต้นการวิเคราะห์เชิงเทคนิคจึงเป็นวิธีที่นิยมใช้กันอย่างแพร่หลายในปัจจุบัน

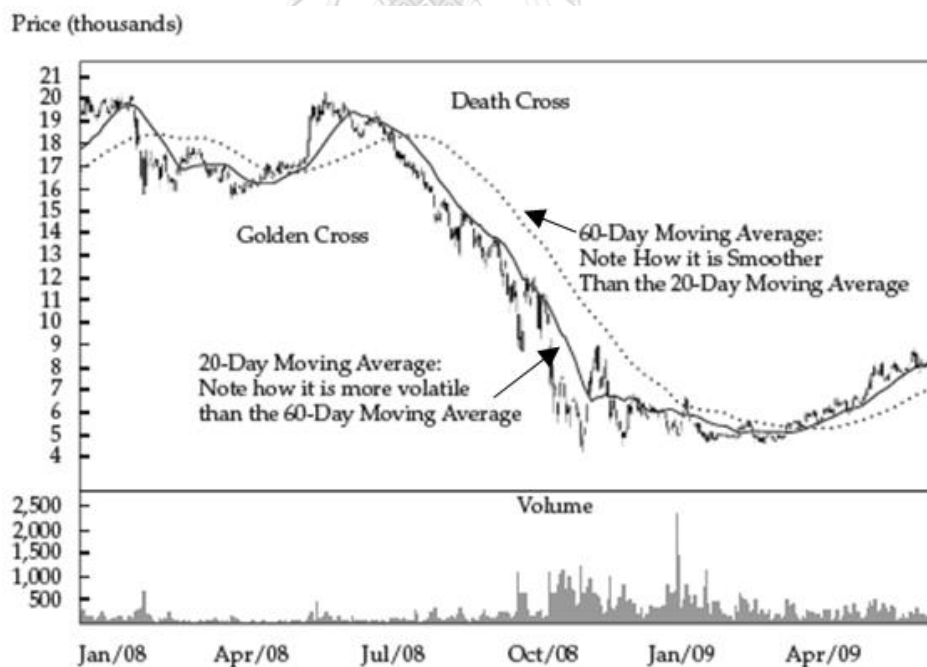
มีเครื่องมือช่วยสำหรับการวิเคราะห์เชิงเทคนิคอยู่หลายประเภท เช่น ประเภทที่หนึ่ง คือ การสร้างเส้นแนวโน้มราคา, เส้นแนวรับ, เส้นแนวต้าน เพื่อหาราคาที่เหมาะสมสำหรับเข้าซื้อ และขาย, ประเภทที่สอง คือ การวิเคราะห์รูปแบบราคาในอดีต เช่น กราฟรูปแบบหัวและไหล่, กราฟ Double Top, กราฟ Double Bottom, กราฟสามเหลี่ยม และประเภทที่สาม คือ การนำข้อมูลราคาหรือปริมาณการซื้อขายมาผ่านกระบวนการคำนวณเพื่อสร้างตัวบ่งชี้ทางเทคนิค (Technical Indicator)

ซึ่งจะเป็นวิธีที่ใช้ในงานวิจัยนี้ จึงขอก้าวถึงเครื่องมือสำหรับการวิเคราะห์เชิงเทคนิคประเภทนี้โดยละเอียดต่อไป

ตัวบ่งชี้ทางเทคนิคมีหลายชนิด แต่ในที่นี้จะอธิบายถึงชนิดที่นิยมใช้กันอย่างแพร่หลาย และเป็นชนิดที่นำมาวิเคราะห์เปรียบเทียบกับแบบจำลองการเรียนรู้ของเครื่องในงานวิจัยนี้ ดังนี้

2.2.1 ค่าเฉลี่ยเคลื่อนที่ (Moving Average หรือ MA)

ค่าเฉลี่ยเคลื่อนที่ เป็นตัวบ่งชี้ทางเทคนิคที่ใช้เพื่อทำให้ราคาของสินทรัพย์ที่มีความผันผวนราบเรียบขึ้น ค่าเฉลี่ยเคลื่อนที่มีหลักการคำนวณง่าย ๆ คือ ค่าเฉลี่ยของราคาปิดล่าสุดจำนวน N ช่วงเวลา ยิง N ที่ค่ามากเส้นค่าเฉลี่ยเคลื่อนที่ที่ได้ยิ่งมีความราบเรียบมากขึ้น ในแนวโน้มขาขึ้นราคาจะอยู่เหนือเส้นค่าเฉลี่ยเคลื่อนที่ ในแนวโน้มขาลง ราคาจะอยู่ใต้เส้นค่าเฉลี่ยเคลื่อนที่ ทั้งนี้เส้นค่าเฉลี่ยเคลื่อนที่มีจำนวนช่วงเวลาที่แตกต่างกันสามารถนำมาใช้ร่วมกันเพราะวิเคราะห์สัญญาณซื้อและขายได้ ตามตัวอย่างที่แสดงในรูปที่ 9 ซึ่งแสดงราคาและเส้นค่าเฉลี่ยเคลื่อนที่ 20 วันและ 60 วัน

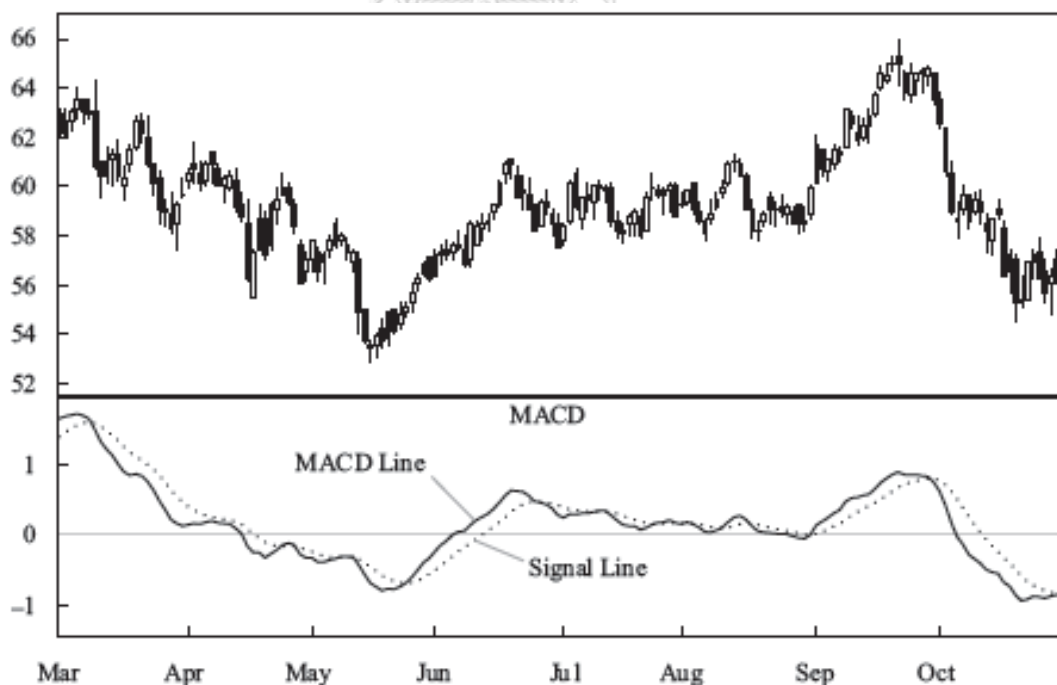


รูปที่ 9 ตัวอย่างกราฟแสดงราคาและเส้นค่าเฉลี่ยเคลื่อนที่ 20 วันและ 60 วัน [21]

เมื่อเส้นค่าเฉลี่ยเคลื่อนที่ระยะสั้นตัดเส้นค่าเฉลี่ยเคลื่อนที่ระยะยาวขึ้นไปอยู่ข้างบน จะเรียกว่า Golden Cross ซึ่งสามารถใช้เป็นสัญญาณบ่งบอกความเป็นขาขึ้นหรือจุดเข้าซื้อที่แน่นอน ในทางตรงกันข้ามเมื่อเส้นค่าเฉลี่ยเคลื่อนที่ระยะสั้นตัดเส้นค่าเฉลี่ยเคลื่อนที่ระยะยาวลงไปอยู่ด้านล่าง จะเรียกว่า Dead Cross ซึ่งสามารถใช้เป็นสัญญาณบ่งบอกความเป็นขาลงหรือจุดขาย

2.2.2 ค่าเฉลี่ยเคลื่อนที่ลู่เข้าลู่ออก (Moving Average Convergence/Divergence หรือ MACD)

ค่าเฉลี่ยเคลื่อนที่ลู่เข้าลู่ออก สามารถสร้างได้จากเส้นค่าเฉลี่ยเคลื่อนที่แบบยกกำลัง ประกอบด้วย 2 เส้น ตามรูปที่ 10 เส้น MACD เป็นค่าความแตกต่างระหว่างเส้นค่าเฉลี่ยเคลื่อนที่แบบยกกำลัง 2 เส้นที่มีจำนวนช่วงเวลาแตกต่างกัน โดยปกติจะใช้ 12 และ 26 วัน เส้น Signal เป็นค่าเฉลี่ยเคลื่อนที่แบบยกกำลัง ที่โดยปกติจะใช้ที่ 9 วัน เมื่อเส้น MACD ตัดเส้น Signal ขึ้นไปอยู่ข้างบนจะเป็นสัญญาณซื้อ และในทางตรงกันข้าม เมื่อเส้น MACD ตัดเส้น Signal ลงไปอยู่ข้างล่างจะเป็นสัญญาณขาย



รูปที่ 10 ตัวอย่างกราฟแสดงเส้น MACD และ เส้น Signal [21]

2.2.3 ดัชนีกำลังสัมพัทธ์ (Relative Strength Index หรือ RSI)

ดัชนีกำลังสัมพัทธ์ มีการคำนวณมาจากอัตราส่วนระหว่างราคาที่เปลี่ยนแปลงเพิ่มขึ้นและราคาที่เปลี่ยนแปลงลดลงในช่วงเวลาที่สนใจ ซึ่งโดยส่วนมากจะใช้เป็น 14 วัน ตามสมการที่ (6) ซึ่งดัชนีกำลังสัมพัทธ์นั้นจะมีการแกว่งตัวอยู่ในช่วง 0 ถึง 100 โดยที่ค่าที่สูง (ตามปกติจะหมายถึงค่าที่มากกว่า 70) บ่งบอกภาวะการซื้อมากเกินไป (overbought) และในทางกลับกัน ค่าที่ต่ำ (ตามปกติจะหมายถึงค่าที่น้อยกว่า 30) บ่งบอกภาวะการขายมากเกินไป (oversold) ตามตัวอย่างกราฟที่มีเส้น RSI ในรูปที่ 11

$$RSI = 100 - \frac{100}{1 + \frac{U_N}{D_N}} \quad (6)$$

U_N คือ ค่าเฉลี่ยของราคาที่เปลี่ยนแปลงเพิ่มขึ้น N ช่วงเวลาล่าสุด

D_N คือ ค่าเฉลี่ยของราคาที่เปลี่ยนแปลงลดลง N ช่วงเวลาล่าสุด

N คือ จำนวนช่วงเวลาที่สนใจ

วิธีการวิเคราะห์สัญญาณซื้อขายจาก RSI มีหลายวิธี แต่วิธีที่ง่ายและเป็นที่ยอมรับกันมากที่สุดคือการซื้อเมื่อเกิดภาวะขายมากเกินไปหรือสัญญาณซื้อจะเกิดเมื่อ RSI ตัด 30 ขึ้นไป และขายเมื่อเกิดภาวะซื้อมากเกินไป หรือสัญญาณขายเกิดเมื่อ RSI ตัด 70 ลงมา



รูปที่ 11 ตัวอย่างกราฟแสดงเส้นดัชนีกำลังสัมพัทธ์ (RSI) [21]

2.2.4 เส้นแกว่งสโตคาสติก (Stochastic Oscillators)

เส้นแกว่งสโตคาสติก มาจากข้อสังเกตที่ว่าในช่วงที่ราคามีแนวโน้มเป็นขาขึ้น ราคาปิดมักจะใกล้เคียงกับราคาสูงสุด และในทางกลับกันในช่วงที่ราคามีแนวโน้มเป็นขาลง ราคาปิดมักจะใกล้เคียงกับราคาต่ำสุด ด้วยรูปแบบความสัมพันธ์ระหว่างราคาปิด, ราคาสูงสุด, และราคาต่ำสุดจากข้อสังเกตนี้เอง ได้ถูกนำมาพัฒนาเป็นเส้นแกว่งสโตคาสติก ที่มีการแกว่งตัวในช่วง 0% ถึง 100% และมีช่วงเวลาที่น่าสนใจซึ่งโดยปกติจะใช้ 14 วัน เส้นแกว่งสโตคาสติกนั้นประกอบด้วยสองเส้นได้แก่ %K และ %D ซึ่งสามารถคำนวณได้จากสมการ (7)

$$\%K = 100 \left(\frac{C_t - LL_{t-(14-1)}}{HH_{t-(14-1)} - LL_{t-(14-1)}} \right) \quad (7)$$

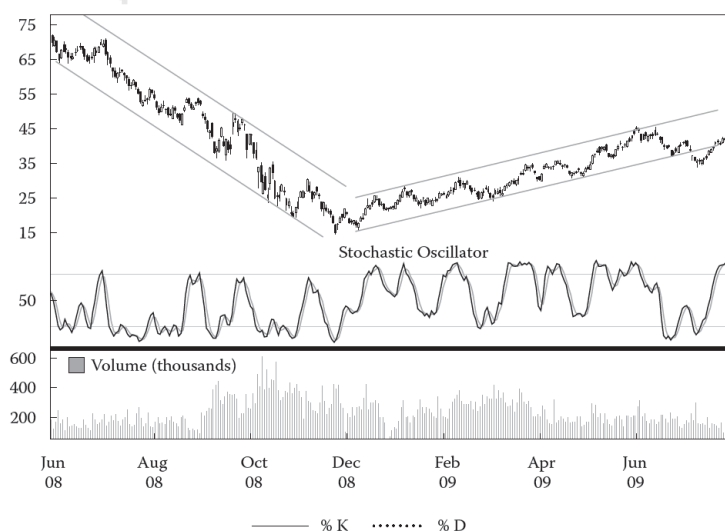
C_t คือ ราคาปิดล่าสุด

$LL_{t-(14-1)}$ คือ ราคาต่ำสุดในช่วงเวลา 14 วัน

$HH_{t-(14-1)}$ คือ ราคาสูงสุดในช่วงเวลา 14 วัน

%D คือ ค่าเฉลี่ยของ %K 3 วันล่าสุด

การวิเคราะห์หาสัญญาณซื้อขายจากเส้นแกว่งสโตคาสติกที่นิยมใช้กันก็คือ เมื่อเส้น %K ตัดเส้น %D ขึ้นในบริเวณขายมากเกินไประดับต่ำกว่า 20% จะเป็นสัญญาณซื้อ และเมื่อเส้น %K ตัดเส้น %D ลงในบริเวณซื้อมากเกินไประดับสูงกว่า 80% จะเป็นสัญญาณขาย ตัวอย่างกราฟที่มีเส้นแกว่งสโตคาสติกแสดงในรูปที่ 12



รูปที่ 12 ตัวอย่างกราฟแสดงเส้นแกว่งสโตคาสติก [21]

บทที่ 3

งานวิจัยที่เกี่ยวข้อง

การเรียนรู้ของเครื่องได้ถูกนำมาใช้ในการวิจัยที่เกี่ยวข้องกับการซื้อขาย หรือทำนายแนวโน้มราคาน้ำมันดิบอย่างมากมาย ด้วยความสามารถของแบบจำลองการเรียนรู้ของเครื่องหลากหลายแบบที่ถูกพัฒนาขึ้นเพื่อใช้กับข้อมูลในรูปแบบอนุกรมเวลา จึงมีหลายงานวิจัยได้นำได้นำแบบจำลองการเรียนรู้ของเครื่องเหล่านี้มาประยุกต์ใช้ในการทำนายแนวโน้มราคาน้ำมันดิบ อย่างเช่น Zhao และ Wang [5] ได้ใช้แบบจำลอง ARIMA ในการทำนายราคาน้ำมันดิบ และพบว่าได้ผลการทำนายในระดับที่น่าพอใจ โดยเฉพาะผลการทำนายในช่วงระยะสั้น ๆ มีความคลาดเคลื่อนจากการทำนายของแบบจำลองโดยเฉลี่ยเพียง 4.1% ต่อมา Yu และคณะ [3] และ Xie และคณะ [4] ทั้งสองงานวิจัยได้เปรียบเทียบประสิทธิภาพแบบจำลอง SVM, โครงข่ายประสาทเทียมแบบ 3 ชั้น, และ ARIMA แล้วพบผลการศึกษาที่สอดคล้องกันว่าแบบจำลอง SVM และโครงข่ายประสาทเทียมแบบ 3 ชั้นมีความแม่นยำในการทำนายราคาน้ำมันดิบได้ดีกว่า ARIMA อย่างมีนัยยะสำคัญ แต่ด้วยปัญหา Overfitting ในโครงข่ายประสาทเทียมทำให้ SVM มีประสิทธิภาพการทำนายได้ดีกว่าอยู่เล็กน้อย ในงานวิจัยของ Ramyar และ Kianfar [8] ได้ศึกษาประสิทธิภาพการทำนายราคาน้ำมันดิบด้วยโครงข่ายประสาทเทียม และ Vector autoregressive (VAR) พบว่าด้วยโครงสร้างโครงข่ายประสาทเทียมที่เหมาะสมและการเลือกใช้ตัวแปรที่เหมาะสมสามารถทำให้โครงข่ายประสาทเทียมทำนายราคาน้ำมันดิบได้เป็นที่น่าพอใจ และมีความแม่นยำกว่า แบบจำลอง Vector autoregressive (VAR)

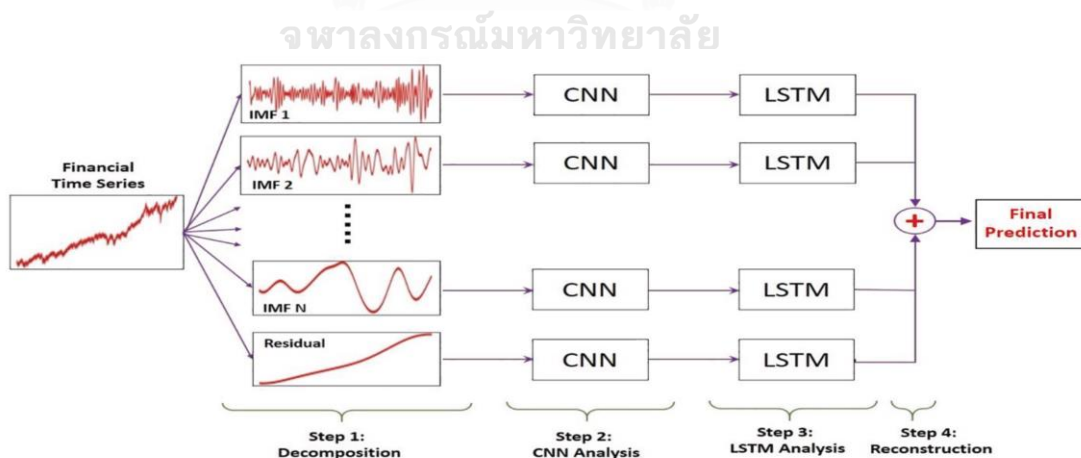
ด้วยการพัฒนาอย่างก้าวกระโดดของการเก็บข้อมูล และกำลังการคำนวณของคอมพิวเตอร์ ในช่วงทศวรรษที่ผ่านมา โครงข่ายประสาทเทียมจึงถูกใช้อย่างแพร่หลายมากขึ้นในงานวิจัยหลาย ๆ ด้าน รวมถึงงานวิจัยด้านอนุกรมเวลาทางการเงินของราคาหลักทรัพย์ซึ่งมีลักษณะข้อมูลคล้ายคลึงราคาน้ำมันดิบในงานวิจัยนี้ด้วยเช่นกัน จากงานทบทวนงานวิจัยเกี่ยวกับการทำนายราคาหลักทรัพย์ของ Kumar และคณะ [11] พบว่าโครงข่ายประสาทเทียมเป็นแบบจำลองที่มีประสิทธิภาพสำหรับการทำนายราคาหลักทรัพย์เนื่องจากสามารถเข้าใจความซับซ้อนและไม่เป็นเส้นตรงของข้อมูลราคาหลักทรัพย์ได้

โครงข่ายประสาทเทียมที่นิยมใช้ในการทำนายราคาหลักทรัพย์ เช่น โครงข่ายประสาทเทียมคอนโวลูชัน เป็นโครงข่ายประสาทเทียมชนิดหนึ่งที่นิยมใช้การทำนายราคาหลักทรัพย์ CNN สามารถช่วยเลือกคุณลักษณะ (Feature Selection) สำหรับแบบจำลองได้แบบอัตโนมัติ [23] CNN ถูกใช้โดย Gudelek และคณะ [12] เพื่อทำนายประเภทการเคลื่อนไหวของราคาจากข้อมูลราคาหลักทรัพย์

และตัวชี้วัดทางเทคนิคในอดีต ซึ่งแบบจำลอง CNN สามารถทำนายได้ความแม่นยำสูง โดย 72.9% ของรายการซื้อขายสามารถทำกำไรได้ และสามารถใช้สร้างกลยุทธ์การซื้อขายที่สร้างผลกำไรได้ดีกว่ากลยุทธ์การซื้อขายและถือ Tsantekidis และคณะ [13] พบว่าแบบจำลอง CNN สามารถทำนายราคาหลักทรัพย์ได้แม่นยำกว่า SVM

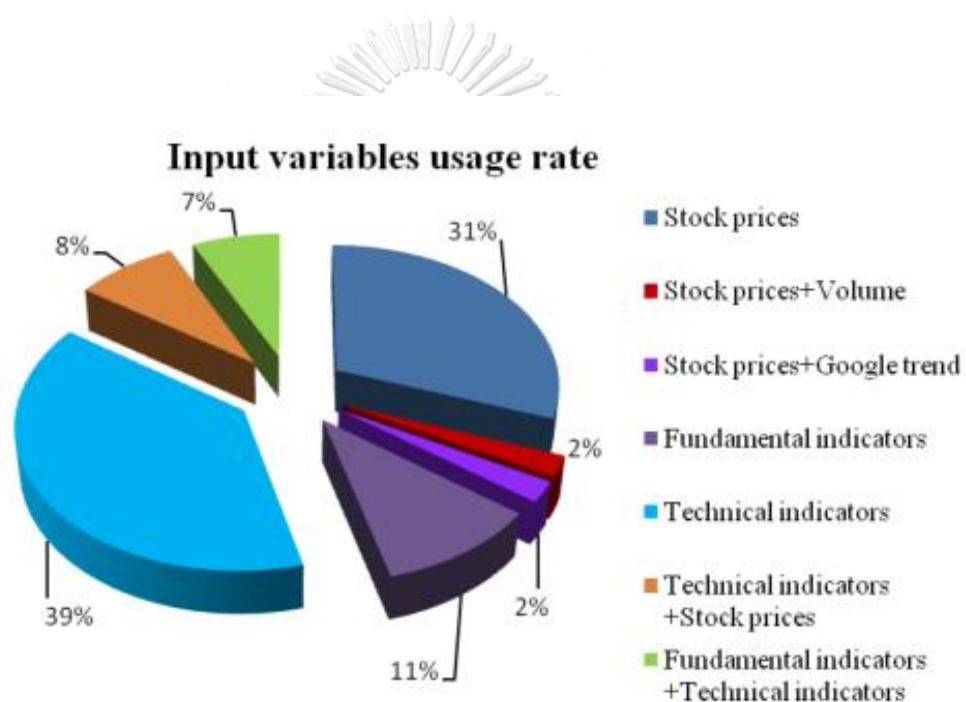
นอกจากนี้ LSTM เป็นโครงข่ายประสาทเทียมอีกประเภทที่ได้รับความนิยมในการใช้เพื่อการทำนายหลักทรัพย์เช่นกัน [24] เพราะ LSTM ถูกออกแบบให้มีความจำภายในโครงข่ายจึงสามารถเรียนรู้ข้อมูลที่มีความยาวมากกว่าโครงข่ายประสาทเทียมทั่วไป [20] Li และคณะ [14] ใช้ LSTM ทำนายราคาหลักทรัพย์ที่มีการแกว่งตัวของราคาในระดับมากขึ้น้อยแตกต่างกันหลายระดับเทียบกับ SVM จากผลการศึกษา LSTM มีความถูกต้องมากกว่า SVM ในหลักทรัพย์ที่มีการแกว่งตัวของราคาน้อยกว่า Cheng และคณะ [15] เสนอแนวทางการทำนายประเภทการเคลื่อนไหวของราคาหลักทรัพย์จากตัวแปรที่สกัดออกมาจากข้อความข่าวโดยใช้ Attention-based LSTM

Rezaei และคณะ [25] ได้เสนอแบบจำลอง CNN-LSTM ตามรูปที่ 13 เพื่อทำนายดัชนีตลาดหุ้น S&P500, Dow Jones, Dax, และ Nikkei225 พบว่ารูปแบบแบบจำลองชนิดนี้มีความเหมาะสมในการวิเคราะห์รูปแบบของข้อมูลอนุกรมเวลาทางการเงินได้อย่างดี โดย CNN-LSTM สามารถทำนายดัชนีตลาดหุ้นได้ค่าความคลาดเคลื่อน RMSE ต่ำกว่าแบบจำลอง LSTM, Support Vector Regression, และ Decision Tree Regression ในทุกๆ ตลาดอย่างชัดเจน ในงานวิจัยนี้จึงนำแบบจำลอง CNN-LSTM เพื่อมาใช้ทำนายแนวโน้มราคาน้ำมัน



รูปที่ 13 โครงสร้างแบบจำลอง CNN-LSTM ที่ถูกนำเสนอโดย Rezaei และคณะ [25]

Kumar และคณะ [11] กล่าวว่านอกจากแบบจำลองแล้วตัวแปรที่ใช้เป็นข้อมูลนำเข้าเองก็มีความสำคัญต่อความแม่นยำของผลการทำนายเช่นกัน จากผลสำรวจ [11] พบว่างานวิจัยด้านการทำนายราคาหลักทรัพย์มากกว่า 50% ใช้ข้อมูลตัวชี้วัดทางเทคนิคเป็นข้อมูลนำเข้าดังแสดงในรูปที่ 14 เนื่องจากเป็นข้อมูลที่หาง่าย ใช้กันอย่างแพร่หลาย และทำให้ผลการทำนายมีประสิทธิภาพมากขึ้น นอกจากนี้ Kumar และคณะ [11] ยังได้เสนอโครงสร้างแนวทางการวิจัยการทำนายราคาหลักทรัพย์ไว้ว่า ข้อมูลที่ใช้นำเข้าไปยังแบบจำลองควรมีการทำให้เป็นมาตรฐานในช่วง $[0, 1]$ และควรใช้มาตรวัดประสิทธิภาพของแบบจำลองหลาย ๆ แบบ ซึ่งจะเป็นแนวทางที่เอามาประยุกต์ใช้กับงานวิจัยนี้เช่นกัน



รูปที่ 14 สัดส่วนงานวิจัยที่ใช้ตัวแปรต่าง ๆ กันเป็นข้อมูลนำเข้าจากผลการสำรวจในงานทบทวนงานวิจัยการทำนายราคาหลักทรัพย์ของ Kumar และคณะ [11]

บทที่ 4

วิธีดำเนินงาน

4.1 การเตรียมข้อมูลก่อนการทดลอง (Data Preprocessing)

ในส่วนนี้จะกล่าวถึงการเตรียมข้อมูลก่อนการทดลองในงานวิจัยนี้ โดยประกอบไปด้วย 4 ส่วนคือ ข้อมูลดิบและการคำนวณตัวชี้วัดทางเทคนิค, การแปลงข้อมูล, การติดฉลากข้อมูล (Labeling), และการแบ่งข้อมูลเพื่อการฝึกแบบจำลอง โดยมีรายละเอียดดังต่อไปนี้

4.1.1 ข้อมูลดิบและการคำนวณตัวชี้วัดทางเทคนิค

ข้อมูลราคาน้ำมันดิบที่ใช้ในงานวิจัยนี้เป็นราคาน้ำมัน WTI ที่นำมาจากเว็บไซต์ Yahoo Finance [26] โดยใช้เป็นราคาน้ำมันดิบ WTI รายวันตั้งแต่ปี ค.ศ. 2005 ถึงปี ค.ศ. 2020 ข้อมูลในแต่ละวันประกอบด้วย ราคาเปิด ราคาสูงสุด ราคาต่ำสุด ราคาปิด และปริมาณการซื้อขาย จากนั้นใช้ TA package [27] เพื่อคำนวณตัวชี้วัดทางเทคนิคตามที่ได้ถูกเสนอในงานวิจัยของ Kumar และคณะ [28] รายละเอียดของตัวชี้วัดทางเทคนิคแสดงในตารางที่ 1 [28] งานวิจัยนี้ใช้ข้อมูลตัวชี้วัดทางเทคนิคหลากหลายประเภทที่นิยมใช้โดยทั่วไป เพื่อให้แบบจำลองสามารถเรียนรู้สามารถเรียนรู้การให้สัญญาณการซื้อขายจากตัวชี้วัดหลาย ๆ แบบ โดยมีการให้น้ำหนักที่แตกต่างกัน ซึ่งตัวชี้วัดที่ใช้ในงานวิจัยนี้แบ่งได้ เป็น 4 ประเภทดังนี้

1. ตัวชี้วัดประเภทแนวโน้มราคา มักจะใช้บ่งบอกทิศทางของราคา ได้แก่ Simple Moving Average (SMA), Moving Average Convergence Divergence (MACD), Average Directional Movement Index (ADX), และ Commodity Channel Index (CCI)
2. ตัวชี้วัดประเภทโมเมนตัม มักจะใช้บ่งบอกความแข็งแกร่งในการเคลื่อนที่ของราคา ได้แก่ Rate of Change (ROC), Relative Strength Index (RSI), True strength index (TSI), Stochastic RSI %K (%K), Stochastic RSI %D (%D), และ Williams %R (%R)
3. ตัวชี้วัดประเภทความแปรปรวน มักจะใช้บ่งบอกความเสี่ยงหรือความแปรปรวนของราคาสินทรัพย์ ได้แก่ Bollinger Bands (BB), Average True Range (ATR), และ Ulcer Index (UI)
4. ตัวชี้วัดประเภทปริมาณการซื้อขาย มักจะใช้บ่งบอกภาวะความกดดันการซื้อขายสินทรัพย์ ได้แก่ Accumulation/Distribution Index (ADI), On-balance volume (OBV), Chaikin

Money Flow (CMF), Force Index (FI), Money Flow Index (MFI), Volume-price trend (VPT), และ Volume Weighted Average Price (VWAP)

ตารางที่ 1 ตัวแปรและรายละเอียดที่ใช้เพื่อเป็นข้อมูลนำเข้าไปในแบบจำลอง ประกอบด้วย ราคา, ปริมาณการซื้อขาย, และตัวชี้วัดแบบต่าง ๆ [28] [29]

ตัวชี้วัด	รายละเอียด
O_t	ราคาเปิด ณ ช่วงเวลา t
H_t	ราคาสูงสุด ณ ช่วงเวลา t
L_t	ราคาต่ำสุด ณ ช่วงเวลา t
C_t	ราคาปิด ณ ช่วงเวลา t
V_t	ปริมาณการซื้อขาย ณ ช่วงเวลา t
$SMA_{fast,t}(n)$	Simple Moving Average เป็นตัวบ่งชี้บ่งบอกแนวโน้มของอนุกรมเวลา โดยลดความแปรปรวนของราคาจากการเฉลี่ย
$SMA_{slow,t}(k)$	
	$SMA_t(n) = \frac{1}{n} \sum_{x=0}^{n-1} C_{t-x}$
$MACD_t(n, k)$	Moving Average Convergence Divergence หรือ MACD เป็นตัวบ่งชี้ประเภทโมเมนตัม (Momentum) คำนวณจากความแตกต่างระหว่างค่าเฉลี่ยเคลื่อนที่เร็ว (n) และ ช้า (k) <p style="text-align: center;"> $MACD_t(n, k) = EMA_t(n) - EMA_t(k)$ $EMA_t(n) = \alpha(C_t - EMA_{t-1}(n)) + EMA_{t-1}(n)$ $\alpha = \frac{2}{n + 1}$ </p>
$ADX_t(n)$	Average Directional Movement Index วัดความแข็งแกร่งของแนวโน้มราคา โดยคำนวณรวมกันจากตัวชี้วัดทิศทางการเคลื่อนไหวของราคา DI^+ และ DI^- <p style="text-align: center;"> $ADX_t(n) = \frac{ DI^+ - DI^- }{ DI^+ + DI^- } \times 100$ </p> <p style="text-align: center;"> $Up = H_t - H_{t-1} \qquad Dn = L_{t-1} - L_t$ </p> <p style="text-align: center;"> <i>if</i> $Up > Dn$ and $Up > 0$: $DM^+ = Up$, else $DM^+ = 0$ <i>if</i> $Dn > Up$ and $Dn > 0$: $DM^- = Dn$, else $DM^- = 0$ </p> <p style="text-align: center;"> $DI^+ = \frac{EMA_t(DM^+, n)}{ATR(n)} \qquad DI^- = \frac{EMA_t(DM^-, n)}{ATR(n)}$ </p>

$CCI_t(n)$	<p>Commodity Channel Index วัดทิศทางการเคลื่อนที่ของราคาและความแข็งแกร่งของแนวโน้มจากความแตกต่างระหว่างราคาและค่าเฉลี่ยเคลื่อนที่</p> $CCI_t(n) = \frac{TP_t - SMA_t(TP, n)}{0.015 \sum_{i=1}^n TP_{t-i+1} - SMA_t(TP, n) / n} \times 100$ $TP_t = \frac{(C_t + H_t + L_t)}{3}$
$ROC_t(n)$	<p>Rate of Change เป็นตัวชี้วัดประเภทโมเมนตัมที่วัดร้อยละของการเปลี่ยนแปลงของราคา เพื่อระบุภาวะซื้อมากเกิน หรือภาวะขายมากเกิน</p> $ROC_t(n) = \frac{C_t - C_{t-n}}{C_{t-n}} \times 100$
$RSI_t(n)$	<p>Relative Strength Index เป็นตัวชี้วัดประเภทโมเมนตัมที่วัดขนาดของราคาเพิ่มขึ้นเทียบกับราคาที่ลดลง เพื่อระบุภาวะซื้อมากเกิน หรือภาวะขายมากเกิน</p> $RSI_t(n) = 100 - \frac{100}{1 + \frac{\sum_{i=0}^{n-1} Up_{t-i}/n}{\sum_{i=0}^{n-1} Dn_{t-i}/n}}$ $Up_t = C_t - O_t \text{ if } C_t > O_t \text{ else } 0 \quad Dn_t = O_t - C_t \text{ if } O_t > C_t \text{ else } 0$
$TSI_t(n)$	<p>True strength index เป็นตัวชี้วัดประเภทโมเมนตัมที่ใช้ค่าเฉลี่ยแบบยกกำลังเพื่อระบุภาวะซื้อมากเกิน หรือภาวะขายมากเกิน</p> $TSI_t(n) = \frac{EMA(EMA(PC, n), k)}{EMA(EMA(PC , n), k)}$ $PC_t = C_t - C_{t-1}$
$\%K_t(n)$	<p>Stochastic RSI %K บ่งบอกตำแหน่งของราคาปิดล่าสุดเทียบกับราคาต่ำสุด ($LL_{t-(n-1)}$) และราคาสูงสุด ($HH_{t-(n-1)}$) ในช่วงเวลาที่สนใจ</p> $\%K_t(n) = \frac{C_t - LL_{t-(n-1)}}{HH_{t-(n-1)} - LL_{t-(n-1)}}$
$\%D_t(n)$	<p>Stochastic RSI %D เป็นค่าเฉลี่ยเคลื่อนที่ของ %K ซึ่งทำให้มีความแปรปรวนลดลงและใช้เป็นเส้น Signal คู่กับ %K</p> $\%D_t(n) = \frac{\sum_{i=0}^{n-1} \%K_{t-i}}{n}$
$\%R_t(n)$	<p>Williams %R บ่งบอกตำแหน่งของราคาปิดล่าสุดเทียบกับราคาต่ำสุด ($LL_{t-(n-1)}$) และราคาสูงสุด ($HH_{t-(n-1)}$) ในช่วงเวลาที่สนใจ</p> $\%R_t(n) = \frac{HH_{t-(n-1)} - C_t}{HH_{t-(n-1)} - LL_{t-(n-1)}}$

$BB_{h,t}(n)$	Bollinger Band เป็นตัวชี้วัดบ่งบอกความแปรปรวน นิยามโดย 2 ส่วน เบี่ยงเบนมาตรฐาน จากค่าเฉลี่ยเคลื่อนที่ทั้งทางบวกและทางลบ
$BB_{l,t}(n)$	$BB_{h,t}(n) = SMA_t(n) + 2\sigma(SMA_t(n))$ $BB_{l,t}(n) = SMA_t(n) - 2\sigma(SMA_t(n))$
$ATR_t(n)$	Average True Range วัดความแข็งแกร่งของความแปรปรวน แต่ไม่บอกทิศทางการเคลื่อนที่ $ATR_t(n) = \frac{ATR_{t-1}(n-1) + TR_t}{n}$ $TR_t = \max(H_t - L_t, H_t - C_{t-1} , L_t - C_{t-1})$
$UI_t(n)$	Ulcer Index เป็นตัวชี้วัดประเภทความแปรปรวน โดยวัดความเสี่ยงทางลบของสินทรัพย์ ยิ่ง UI มีค่ามากความเสี่ยงทางลบยิ่งมาก $UI_t(n) = \sqrt{\sum_{i=0}^{n-1} (R_t^i)^2}$ $R_t^i = \frac{C_t - \max(C_{t-n})}{\max(C_{t-n})}$
ADI_t	Accumulation/Distribution Index เป็นตัวชี้วัดประเภทปริมาณการซื้อขายเพื่อใช้บ่งบอกภาวะเก็บสะสม (accumulating) หรือภาวะแจกจ่าย (distribution) สินทรัพย์ เพื่อใช้ทำนายแนวโน้มการกลับตัวของราคา $ADI_t = ADI_{t-1} + CMFV_t$ $CMFV_t = \left(\frac{(C_t - L_t) - (H_t - L_t)}{H_t - L_t} \right) \times V_t$
OBV_t	On-balance volume เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย วัดอารมณ์โดยรวมของผู้ลงทุนเพื่อทำนายทิศทางราคา จากปริมาณการซื้อขาย $OBV_t = OBV_{t-1} + \begin{cases} V_t & \text{if } C_t > C_{t-1} \\ 0 & \text{if } C_t = C_{t-1} \\ -V_t & \text{if } C_t < C_{t-1} \end{cases}$
$CMF_t(n)$	Chaikin Money Flow เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย วัดกระแสเงินผ่านปริมาณการซื้อขายในช่วงเวลาที่สนใจ $MF_t = \frac{(C_t - L_t) - (H_t - C_t)}{H_t - L_t} \times V_t$ $CMF_t(n) = \frac{\sum_{i=0}^n MF_{t-i}}{\sum_{i=0}^n V_{t-i}}$
$FI_t(n)$	Force Index เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย วันความกดดันในการซื้อและขาย ค่าเป็นบวกสูง ๆ หมายถึงราคามีแนวโน้มขึ้นอย่างแข็งแกร่ง ในขณะที่ค่าเป็นลบสูง ๆ หมายถึงราคามีแนวโน้มลงอย่างแข็งแกร่ง $FI_t(1) = (C_t - C_{t-1}) \times V_t$ $FI_t(n) = \sum_{i=0}^n FI_{t-i}(1)$

$MFI_t(n)$	<p>Money Flow Index เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย ที่ใช้ทั้งราคาและปริมาณซื้อขายในการวัดความกดดันในการซื้อและขาย</p> $Raw\ Money\ Flow = TP_t \times V_t$ $Money\ Flow\ Raito_t(n) = \frac{\sum_{i=0}^n Positive\ Money\ Flow_{t-i}}{\sum_{i=0}^n Negative\ Money\ Flow_{t-i}}$ $MFI_t(n) = 100 - \frac{100}{1 + Money\ Flow\ Raito_t(n)}$
VPT_t	<p>Volume-price trend เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย โดยคำนวณมาจากปริมาณการซื้อขายสะสมคูณกับร้อยละการเปลี่ยนแปลงในแต่ละช่วงเวลา</p> $VPT_t = VPT_{t-1} + V_t \times \frac{C_t - C_{t-1}}{C_{t-1}}$
$VWAP_t(n)$	<p>Volume Weighted Average Price เป็นตัวชี้วัดประเภทปริมาณการซื้อขาย ซึ่งมาจากมูลค่าการซื้อขาย ทหารด้วยปริมาณการซื้อขาย</p> $VWAP_t(n) = \frac{\sum_{i=0}^n V_{t-i} \times TP_{t-i}}{\sum_{i=0}^n V_{t-i}}$

หมายเหตุ: ในงานวิจัยนี้ใช้ค่า $n = 10$ และ $k = 15$

4.1.2 การแปลงข้อมูล (Data Transformation)

ด้วยข้อมูลราคาและตัวชี้วัดทางเทคนิคบางตัวจะมีลักษณะเป็น non-Stationary คือมีการเคลื่อนไหวของค่าเฉลี่ย ไม่อยู่ในช่วงใดช่วงหนึ่ง ซึ่งเป็นลักษณะข้อมูลที่ไม่เหมาะสมสำหรับการนำเข้าไปใช้กับโครงข่ายประสาทเทียม [30] ข้อมูลที่ต่างช่วงเวลากัน จะมีค่าที่แตกต่างกันโดยสิ้นเชิง ลักษณะนี้จึงยากสำหรับแบบจำลองที่จะเรียนรู้จากข้อมูลในช่วงเวลาที่ต่าง ๆ กันมาทำนายข้อมูลในอนาคตที่แบบจำลองไม่เคยเรียนรู้มาก่อนได้ ด้วยเหตุนี้ก่อนนำข้อมูลไปใช้กับแบบจำลอง จึงมีความจำเป็นที่จะต้องแปลงข้อมูลจาก non-Stationary Timeseries เป็น Stationary Timeseries

1. คำนวณเป็นร้อยละที่แตกต่างเทียบกับราคาปิด เนื่องจากตัวแปรบางตัวมีการเคลื่อนที่ของค่าเฉลี่ยตามราคาปิด งานวิจัยนี้จึงใช้ร้อยละที่แตกต่างของตัวแปรนั้น ๆ เทียบกับราคาปิดในช่วงเวลาเดียวกัน ตามสมการ (8) โดยตัวแปรที่ถูกแปลงในรูปแบบนี้ได้แก่ ราคาเปิด, ราคาสูงสุด, ราคาต่ำสุด, SMA, BB, ATR, และ VWAP

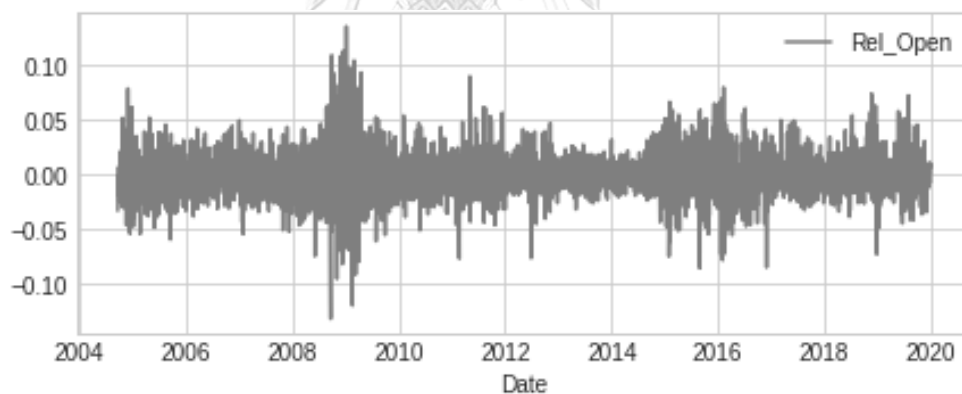
$$Relative\ X_t = \frac{X_t - C_t}{C_t} \quad (8)$$

X_t คือ ข้อมูลอนุกรมเวลาใด ๆ ณ เวลา t

รูปที่ 15 แสดงตัวอย่างสำหรับข้อมูลราคาเปิดก่อนการแปลงในรูปที่ 15(1) ข้อมูลจะมีลักษณะเปลี่ยนแปลงค่าเฉลี่ยอยู่ตลอด ไม่มี Stationary แต่หลังจากแปลงข้อมูลเป็นราคาเปิดสัมพัทธ์ (Relative Open) เทียบกับราคาปิดตามสมการ (8) พบว่าข้อมูลมีความ Stationary มากขึ้นตามที่แสดงในรูปที่ 15(2)



(1) ก่อนการแปลง



(2) หลังการแปลง

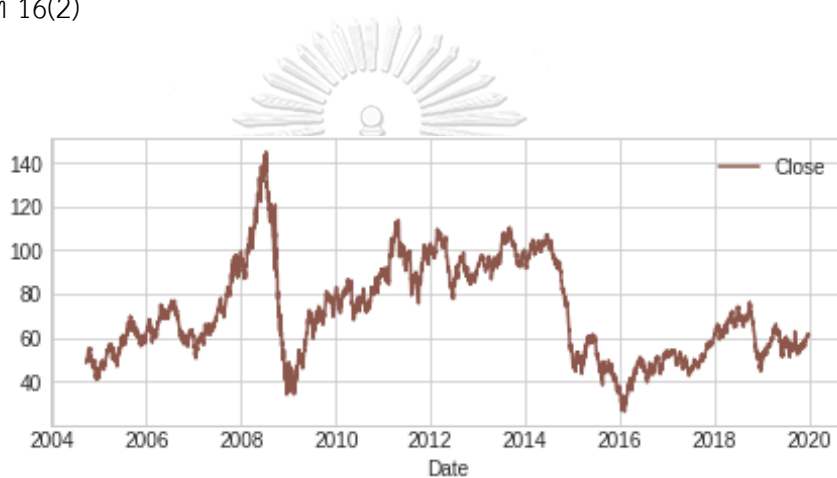
รูปที่ 15 ความแตกต่างของข้อมูลราคาเปิดหลังจากแปลงโดยใช้ร้อยละที่แตกต่างเทียบกับราคาปิด ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง

2. คำนวณเป็นร้อยละเปลี่ยนแปลงจากช่วงเวลาก่อนหน้า ข้อมูลตัวแปรบางชนิดที่ไม่ได้มีการเคลื่อนที่ของค่าเฉลี่ยตามราคาปิด หรือราคาปิดเอง จะถูกทำให้มีความ Stationary โดยแปลงเป็น

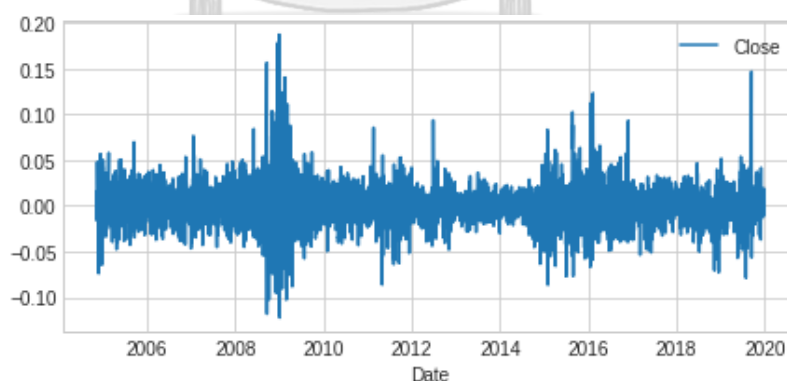
ร้อยละการเปลี่ยนแปลงจากช่วงเวลาก่อนหน้าตามสมการที่ (9) โดยตัวแปรที่ถูกแปลงในรูปแบบนี้ได้แก่ ราคาปิด, ปริมาณการซื้อขาย, ADI, และ OBV

$$\text{Percent Change } X_t = \frac{X_t - X_{t-1}}{X_t} \quad (9)$$

รูปที่ 16 แสดงตัวอย่างสำหรับข้อมูลราคาปิดก่อนการแปลงในรูปที่ 16(1) ข้อมูลจะมีลักษณะเปลี่ยนแปลงค่าเฉลี่ยอยู่ตลอดเวลา ไม่มี Stationary แต่การแปลงข้อมูลเป็นร้อยละการเปลี่ยนแปลงของราคาปิดเทียบกับช่วงเวลาก่อนหน้าตามสมการ (9) พบว่าข้อมูลมีความ Stationary มากขึ้นตามที่แสดงในรูปที่ 16(2)



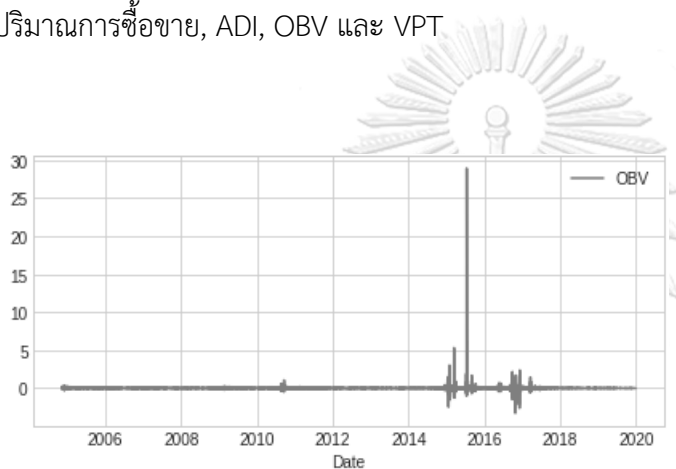
(1) ก่อนการแปลง



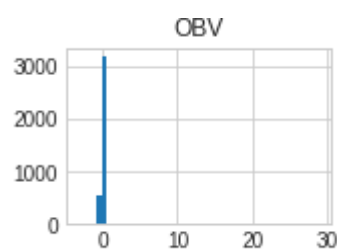
(2) หลังการแปลง

รูปที่ 16 ความแตกต่างของข้อมูลราคาปิดหลังจากการแปลงโดยใช้ร้อยละที่แตกต่างเทียบราคาปิดในช่วงเวลาก่อนหน้า ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง

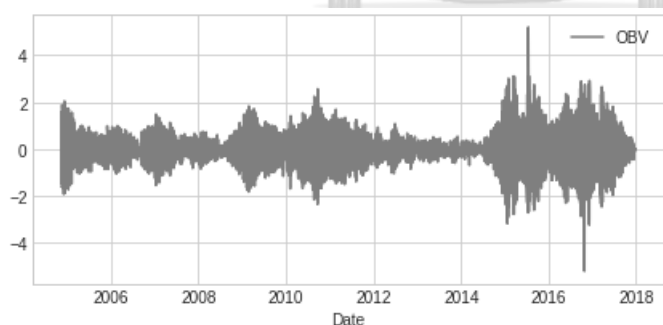
หลังจากนั้นพบว่า การกระจายตัวของข้อมูลนำเข้าบางชนิดมีการกระจายตัวของข้อมูลส่วนใหญ่ที่กระจุกในกรอบแคบ ๆ แต่มีค่าน้อยสุด หรือค่ามากที่สุด ห่างจากค่าเฉลี่ยอย่างเช่นในตัวอย่างตามรูปที่ 17(1) และ รูปที่ 17(2) ซึ่งข้อมูลประเภทนี้ไม่เหมาะสมที่จะนำเข้าไปใช้ฝึกกับโครงข่ายประสาทเทียมเช่นกัน เพราะเมื่อทำ Data Normalization ในขั้นตอนต่อไป ข้อมูลส่วนใหญ่ของตัวแปรเหล่านี้จะถูกบีบให้อยู่ในกรอบแคบ ๆ เนื่องด้วยความแตกต่างของค่าสูงสุดหรือค่าต่ำสุด ทำให้ข้อมูลตัวแปรเหล่านี้จะมีขนาดคณลช่วงเทียบกับข้อมูลตัวแปรอื่น ๆ ดังนั้นในงานวิจัยนี้จึงใช้ Quantile Transformation [31] เพื่อปรับการกระจายตัวของข้อมูลให้มีลักษณะเป็น การกระจายตัวแบบปกติ (Normal Distribution) โดยใช้ Quantile function ตัวแปรที่ถูกแปลงในรูปแบบนี้ได้แก่ ปริมาณการซื้อขาย, ADI, OBV และ VPT



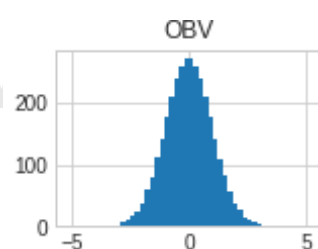
(1) ข้อมูล OBV ก่อน Quantile Transformation



(2) การกระจายตัวของ OBV ก่อน Quantile Transformation



(3) ข้อมูล OBV หลัง Quantile Transformation

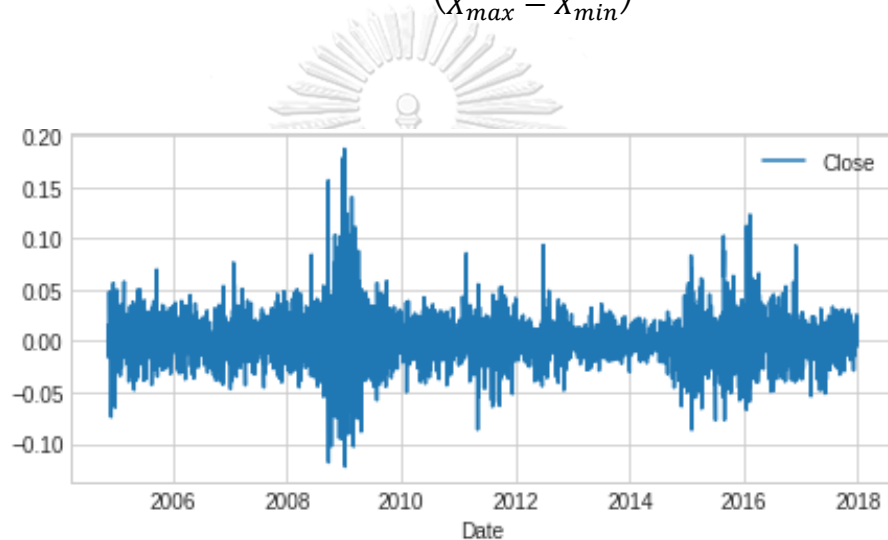


(4) การกระจายตัวของ OBV หลัง Quantile Transformation

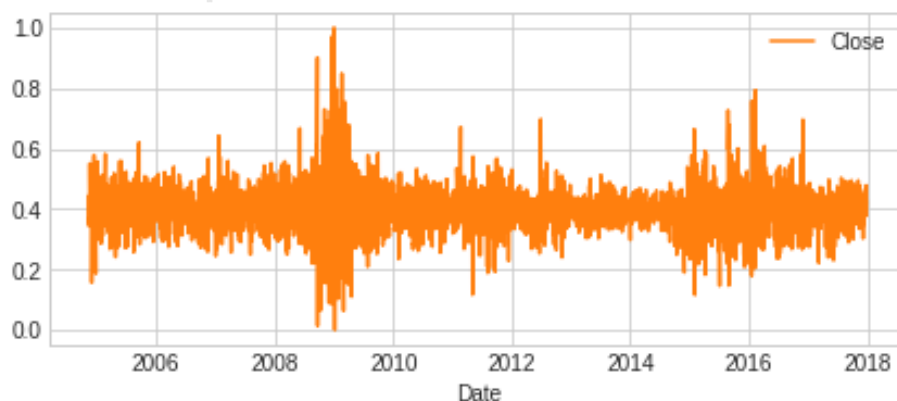
รูปที่ 17 ความแตกต่างของข้อมูล OBV หลังจากการแปลงโดย Quantile Transformation ประกอบด้วย (1) ข้อมูลก่อนการแปลง, (2) การกระจายตัวของข้อมูลก่อนการแปลง, (3) ข้อมูลหลังการแปลง และ (4) การกระจายตัวของข้อมูลหลังการแปลง

รูปที่ 17 แสดงตัวอย่างข้อมูล OBV ก่อนการแปลงข้อมูลโดยใช้ Quantile Transformation ในรูปที่ 17(1) แสดง ค่าสูงสุดของ OBV ค่าเฉลี่ยอย่างมาก และการกระจายตัวของข้อมูลในรูปที่ 17(2) มีลักษณะการกระจายตัวเบ้ขวาอย่างชัดเจน ในขณะที่ข้อมูลหลังจากถูก Quantile Transformation พบว่า ข้อมูลทั้งหมดอยู่ใกล้ค่าเฉลี่ยมากขึ้นตามรูปที่ 17(3) และมีการกระจายตัวแบบปกติตามรูปที่ 17(4)

$$\text{Scaled } X_t = \left(\frac{X_t - X_{min}}{X_{max} - X_{min}} \right) \quad (10)$$



(1) ราคาปิดก่อนการทำ Data Normalization



(2) ราคาปิดหลังการทำ Data Normalization

รูปที่ 18 ความแตกต่างของราคาปิดหลังจากการแปลงข้อมูลโดย Data Normalization ประกอบด้วย (1) ก่อนการแปลง และ (2) หลังการแปลง

ด้วยข้อมูลตัวแปรที่แตกต่างกันแต่ละตัว จะมีช่วงของข้อมูลคนละค่า ซึ่งไม่เหมาะสมจะนำมาใช้ในการฝึกโครงข่ายประสาทเทียม ในงานวิจัยนี้จึงใช้วิธีตามที่เสนอโดย Kumar [11] ที่ได้เสนอไว้ว่าให้มีการทำ Data Normalization ให้มีค่าอยู่ในช่วง $[0, 1]$ ตามสมการที่ (10) ซึ่งข้อมูลนำเข้าทุกอันจะต้องผ่านกระบวนการ Data Normalization ทั้งสิ้น จากตัวอย่างในรูปที่ 18(1) และรูปที่ 18(2) แสดงการแปลงของข้อมูลราคาปิดก่อนและหลังทำ Data Normalization ตามลำดับ จะสังเกตเห็นได้ว่าข้อมูลที่ได้จะมีลักษณะเหมือนข้อมูลก่อนทำทุกประการ แตกต่างเพียงขนาดของข้อมูลในแกนแนวตั้งเท่านั้น

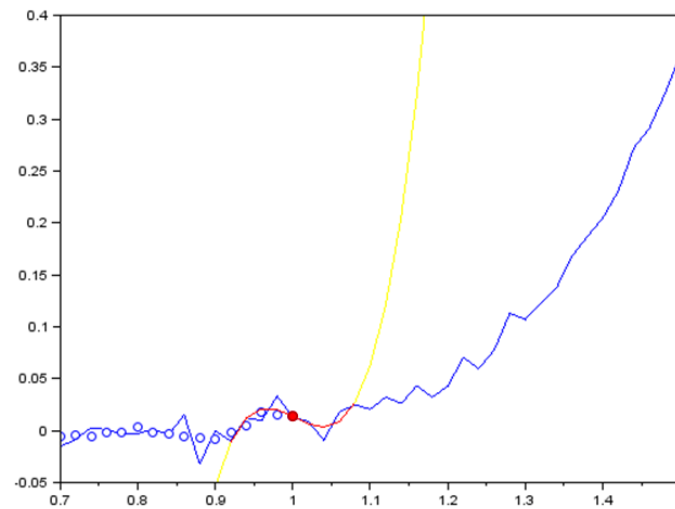
4.1.3 การติดฉลากข้อมูล (Labeling)

ฉลากข้อมูล (Label) ที่ใช้ในงานวิจัยนี้ประกอบด้วยการแปลงข้อมูลราคาปิด 2 ขั้นตอน ขั้นแรกเป็นการสร้างราคาเฉลี่ยจาก Savitzky-Golay filter [16] และขั้นที่สองนำราคาเฉลี่ยที่ได้ไปคำนวณร้อยละการเปลี่ยนแปลงรายวัน

ในขั้นแรกงานวิจัยนี้ใช้ Savitzky-Golay filter จาก SciPy package [16] สร้างเส้นราคาเฉลี่ยเพื่อใช้เป็นฉลากข้อมูล ในการทำนายราคา โดย Savitzky-Golay filter คือการทำ Polynomial Regression เคลื่อนที่ที่สนใจข้อมูลใกล้เคียงรอบข้างตามตัวอย่างรูปที่ 19 เส้นสีแดงคือช่วงของข้อมูลรอบข้างที่นำข้อมูลดิบตามเส้นสีน้ำเงินมาทำ Polynomial Regression และจุดสีแดงคือค่าเฉลี่ยของข้อมูลในช่วงนั้น เมื่อทำ Polynomial Regression กับทุกจุดข้อมูล เส้นตามจุดสีน้ำเงินที่ได้จะเป็นเส้นเฉลี่ยของข้อมูลดิบเส้นสีน้ำเงิน งานวิจัยนี้ใช้ Savitzky-Golay แบบ Polynomial degree ที่ 1 และข้อมูลรอบข้าง 11 จุดมาใช้สร้างเส้นราคาเฉลี่ยดังเส้นสีส้มที่แสดงในรูปที่ 20(1)

เนื่องจากราคาเฉลี่ยยังคงเป็นข้อมูลประเภท non-Stationary เราจึงนำมาแปลงเป็นร้อยละการเปลี่ยนแปลงเทียบกับช่วงเวลาก่อนหน้าตามสมการ (11) จะได้เส้นสีน้ำเงินที่แสดงในรูปที่ 20(2) ซึ่งเป็นฉลากข้อมูล ที่ใช้ในงานวิจัยนี้ การใช้เส้นราคาเฉลี่ยเป็นฉลากข้อมูล จะสามารถใช้แสดงถึงการเปลี่ยนแปลงแนวโน้มของราคา โดยไม่สนใจการแกว่งตัวของราคาในช่วงเวลาสั้น ๆ เพื่อให้แบบจำลองสามารถเรียนรู้ที่จะทำนายแนวโน้มราคาได้

$$y_t = \frac{\text{Smoothed } C_{t+1} - \text{Smoothed } C_t}{\text{Smoothed } C_t} \quad (11)$$



รูปที่ 19 ตัวอย่างการทำเส้นข้อมูลเฉลี่ยโดยใช้ Savitzky-Golay [32] แบบ Polynomial degree ที่ 3 และข้อมูลรอบข้าง 9 จุด



(1) ราคาปิดน้ำมัน WTI



(2) ร้อยละเปลี่ยนแปลงรายวันของราคาน้ำมันเฉลี่ย

รูปที่ 20 (1) ราคาปิดน้ำมัน WTI และราคาเฉลี่ยโดย Savitzky-Golay filter และ (2) ร้อยละการเปลี่ยนแปลงจากช่วงเวลาก่อนหน้าของราคาเฉลี่ย

4.1.4 การแบ่งข้อมูลเพื่อการฝึกแบบจำลอง

ข้อมูลราคาน้ำมันและตัวแปรทั้งหมดที่ได้เตรียมไว้ในขั้นตอน 4.1.1 ถึง 4.1.3 ถูกแบ่งเป็น 6 กลุ่ม แต่ละกลุ่มมีชื่อตามปีของชุดข้อมูลทดสอบเพื่อทดสอบกลยุทธ์การซื้อขายในแต่ละปีที่แตกต่างกันตั้งแต่ปี ค.ศ. 2015 ถึง 2020 โดยแต่ละกลุ่มจะมีข้อมูล 3 ชุด ประกอบด้วย ชุดข้อมูลสำหรับฝึก (9 ปี), ชุดข้อมูลตรวจสอบ (1 ปีหลังจากชุดข้อมูลสำหรับฝึก), และชุดข้อมูลทดสอบ (1 ปีหลังจากชุดข้อมูลตรวจสอบ) แบ่งตามช่วงเวลาแสดงในตารางที่ 2

ชุดข้อมูลสำหรับฝึกใช้เพื่อป้อนข้อมูลให้แบบจำลองเกิดการเรียนรู้และปรับค่าน้ำหนัก ชุดข้อมูลตรวจสอบใช้สำหรับคำนวณค่าความผิดพลาดจากการทำนายเพื่อประเมินเลือกจำนวนรอบการฝึก และ Hyperparameters ที่เหมาะสมที่สุด ชุดข้อมูลทดสอบใช้เพื่อการจำลองสถานการณ์การซื้อขายจริง โดยใช้ข้อมูลที่แบบจำลองไม่เคยเห็นมาก่อน

ตารางที่ 2 การแบ่งข้อมูลสำหรับการทดลอง

กลุ่มข้อมูล	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	
2015	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ					
2016		ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ				
2017			ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ			
2018				ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ		
2019					ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ	
2020						ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ฝึก	ตรวจสอบ	ทดสอบ

ชุดข้อมูลฝึก

ชุดข้อมูลตรวจสอบ

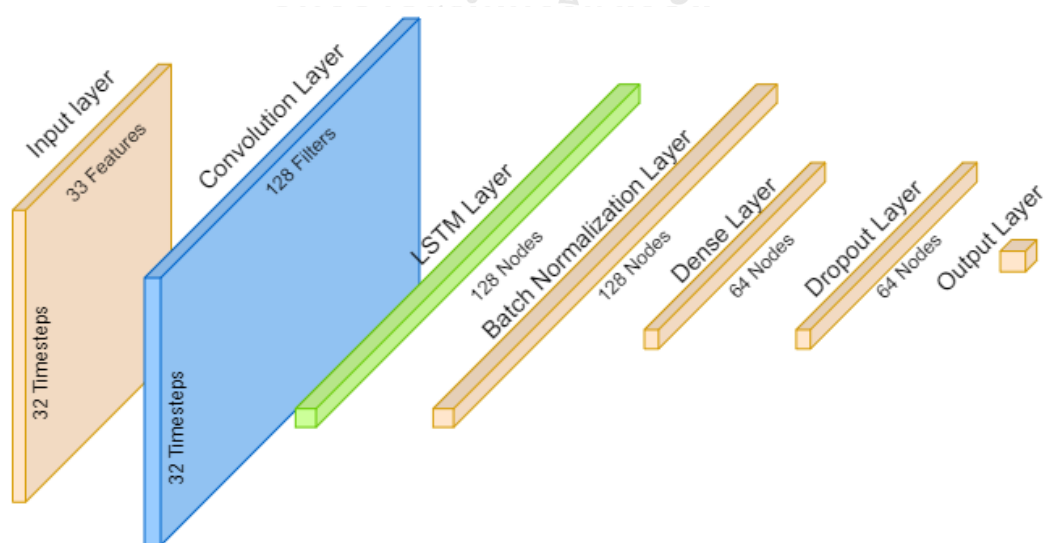
ชุดข้อมูลทดสอบ

4.2 Deep Learning Model

ในงานวิจัยนี้เราได้เลือกแบบจำลองการเรียนรู้ของเครื่องชนิดโครงข่ายประสาทเทียมผสมระหว่าง CNN และ LSTM ซึ่งมีการเรียงลำดับที่แตกต่างกันดังนี้ 1. CNN-LSTM และ 2. LSTM-CNN ตามรูปที่ 21 และรูปที่ 22 ตามลำดับ นอกจากนี้เรายังได้ทำการทดลองกับแบบจำลอง Single CNN และ Single LSTM เปรียบเทียบผลการทดลองกับแบบจำลองแบบผสมอื่นอีกด้วย

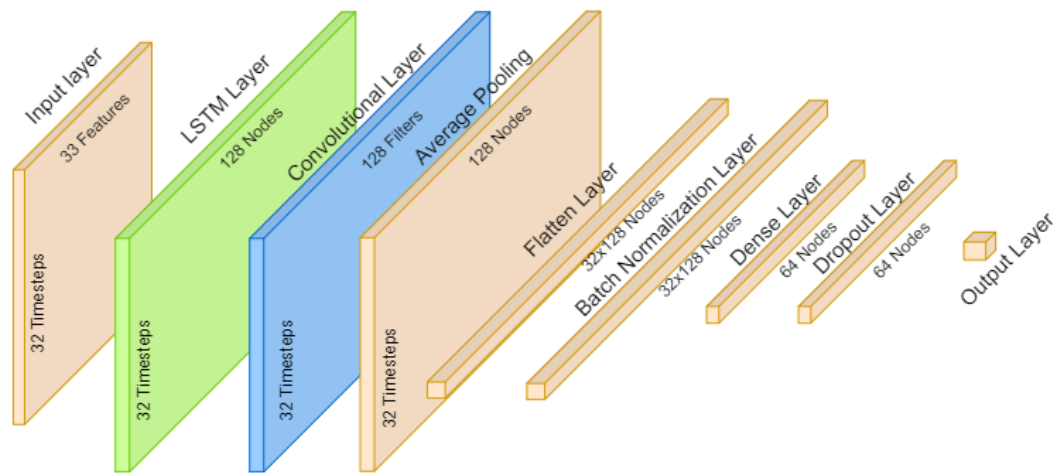
แบบจำลอง CNN-LSTM ดังแสดงในรูปที่ 21 ประกอบไปด้วยชั้นข้อมูลนำเข้า 33 ตัวแปร ตามที่ได้เตรียมข้อมูลไว้ในหัวข้อ 4.1.1 ถึง 4.1.3 โดยแต่ละหนึ่งข้อมูลตัวอย่างจะมีข้อมูลย้อนหลัง 32 ช่วงเวลา ข้อมูลทั้งหมดนี้ถูกส่งต่อไปยังชั้น CNN ซึ่งทำหน้าที่เลือกตัวแปร โดยการเรียนรู้และให้น้ำหนักตัวแปรแต่ละตัวที่แตกต่างกัน เพื่อเรียนรู้แบ่งระดับความสำคัญของตัวแปร ก่อนส่งต่อไปยังชั้น LSTM เพื่อเรียนรู้ความสัมพันธ์ของตัวแปรต่างๆ ในอนุกรมเวลากับฉลากข้อมูล จากนั้นส่งข้อมูลต่อไปยังชั้น Batch Normalization เพื่อลดปัญหา Overfit ข้อมูลทั้งหมดจะถูกประมวลผลในโครงข่ายประสาทเทียม Dense และในชั้นสุดท้ายจะเป็น Dropout layer เพื่อลดปัญหา Overfit ก่อนจะออกเป็นชั้นข้อมูลนำออก ซึ่งจะเป็นผลทำนายที่ได้จากแบบจำลอง

ส่วนแบบจำลอง LSTM-CNN ตามที่แสดงในรูปที่ 22 ประกอบด้วยชั้นข้อมูลนำเข้ารูปแบบเดียวกันกับแบบจำลอง CNN-LSTM แต่มีการสลับลำดับชั้น LSTM มาก่อน CNN จากนั้นมีการเพิ่มชั้น Average Pooling และ ชั้น Flatten เพื่อจัดการข้อมูลที่ออกมาจากชั้น CNN ก่อนจะนำไปเข้าชั้น Batch Normalization, ชั้น Dense, ชั้น Dropout และชั้นข้อมูลนำออกที่เหมือนกับ CNN-LSTM ตามลำดับ



รูปที่ 21 แผนภาพโครงข่ายประสาทเทียม CNN-LSTM ในงานวิจัยนี้ซึ่งมีลำดับของชั้น CNN มาก่อน

LSTM



รูปที่ 22 แผนภาพโครงข่ายประสาทเทียม LSTM-CNN ในงานวิจัยนี้ซึ่งมีลำดับของชั้น LSTM มาก่อน CNN

4.3 วิธีการวัดผล

การวัดผลประสิทธิภาพของแบบจำลองจะพิจารณา 2 แบบ คือ ค่าความผิดพลาดจากการทำนายเพื่อศึกษาความสามารถในการเรียนรู้ของแบบจำลอง และ ประสิทธิภาพการซื้อขายจากการทดลองซื้อขายด้วยข้อมูลย้อนหลัง (Backtest)

4.3.1 ค่าความผิดพลาดจากการทำนาย

ในงานวิจัยนี้ใช้ Mean Square Error (MSE) เพื่อวัดความสามารถในการทำนายของแบบจำลองโครงข่ายประสาทเทียมในขั้นตอนการฝึกแบบจำลอง MSE สามารถคำนวณได้จากสมการ (12) โดยวัดความแตกต่างระหว่างค่าที่แบบจำลองทำนายและฉลากข้อมูล เพื่อใช้วัดความสามารถในการเรียนรู้ของแบบจำลอง MSE จะถูกคำนวณทุกรอบการฝึก เมื่อได้ค่า MSE ในแต่ละรอบการฝึกแบบจำลองจะเรียนรู้และปรับค่าน้ำหนักในแต่ละ Node ในโครงข่าย เพื่อพยายามลดค่า MSE ในรอบถัดไปลง แต่อย่างไรก็ตามค่า MSE ไม่สามารถลดได้เรื่อย ๆ จะมีค่าน้ำหนักในโครงข่ายประสาทเทียมที่เหมาะสมที่ทำให้ได้ค่า MSE ต่ำที่สุด ค่าน้ำหนักที่เหมาะสมนี้จะถูกเลือกจากรอบการฝึกที่ทำให้ได้ MSE ของชุดข้อมูลตรวจสอบต่ำที่สุด

$$MSE = \frac{\sum_{t=0}^n (\hat{y}_t - y_t)^2}{n} \quad (12)$$

\hat{y}_t คือ ผลการทำนายจากแบบจำลอง

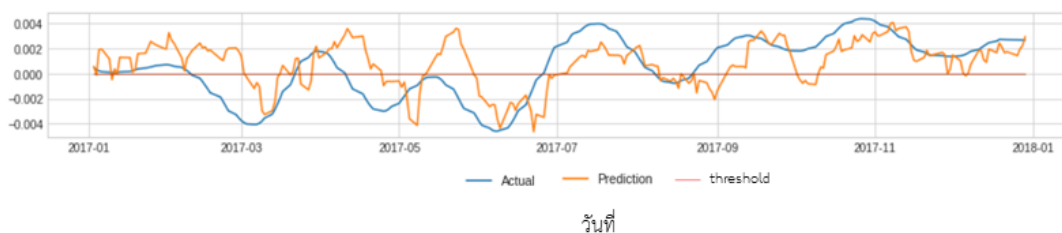
y_t คือ ค่า ราคาสินทรัพย์จริง

ทั้งนี้ในขั้นตอนการฝึกแบบจำลองเราใช้ Adaptive Moment Estimation (Adam) [33] เป็น Optimizer ในการเรียนรู้ของโครงข่ายประสาทเทียมด้วย Learning rate เริ่มต้นที่ 1×10^{-5} และมี Decay Factor เป็น 0.95 ด้วย Patience 10 รอบ กำหนด Batch size เป็น 128 และใช้ Stopping Algorithm เป็นจำนวน 500 รอบหรือ Patience 30 รอบ นอกจากนี้เรายังมีการทำ Grid Search สำหรับจำนวน filters ในชั้น CNN จำนวน 64 และ 128 และจำนวน Node ในชั้น LSTM เป็นจำนวน 64 และ 128 อีกด้วย โดยเลือกจำนวน Filter ในชั้น CNN และจำนวน Node ในชั้น LSTM ที่ทำให้ค่า MSE ของชุดข้อมูลตรวจสอบน้อยที่สุด

4.3.2 ประสิทธิภาพการซื้อขาย

ประสิทธิภาพการซื้อขายเป็นการวัดประสิทธิภาพของพอร์ตการลงทุน ที่เกิดจากการจำลอง การซื้อขายด้วยราคาในอดีต หรือ Backtest ซึ่งเป็นปัจจัยสำคัญในการพิจารณากลยุทธ์การลงทุนใน สินทรัพย์ต่าง ๆ โดยในการจำลองการซื้อขายจะเป็นการเริ่มลงทุนด้วยเงิน 100 ดอลลาร์สหรัฐ และ ซื้อน้ำมันดิบทั้งหมดที่ราคาปิดของวันที่แบบจำลองส่งสัญญาณการซื้อทั้งหมดตามจำนวนเงินที่มี และ ขายเมื่อเกิดสัญญาณขาย โดยขายน้ำมันดิบทั้งหมดในพอร์ตการลงทุน และมีค่าธรรมเนียมการซื้อขายและ ขาย 0.1% ทั้งนี้เราจะวัดผลเทียบกับกำไรขาดทุนที่ได้จากการซื้อขายกลยุทธ์แบบดั้งเดิมด้วย ได้แก่ กลยุทธ์ซื้อและถือ, กลยุทธ์จากสัญญาณซื้อและขายจาก RSI, กลยุทธ์จากสัญญาณซื้อและขายจาก MACD, กลยุทธ์จากสัญญาณซื้อและขายจาก SMA, กลยุทธ์จากสัญญาณซื้อและขายจาก Stochastics RSI %K และ %D

สัญญาณซื้อและสัญญาณขายถูกแปลผลมาจากผลการทำนายร้อยละการเปลี่ยนแปลงราคา เฉลี่ยรายวันตามตัวอย่างเส้นสีส้มในรูปที่ 23 เมื่อเส้นผลการทำนายเคลื่อนที่จากน้อยกว่าศูนย์ไป มากกว่าศูนย์จะถูกแปลเป็นสัญญาณซื้อ ในทางตรงกันข้ามเมื่อเส้นผลการทำนายเคลื่อนที่จาก มากกว่าศูนย์ไปน้อยกว่าศูนย์จะถูกแปลเป็นสัญญาณขาย



รูปที่ 23 ร้อยละการเปลี่ยนแปลงของราคาเฉลี่ย และผลการทำนายจากแบบจำลอง

ตัวชี้วัดประสิทธิภาพการซื้อขาย ที่ใช้ในงานวิจัยนี้ประกอบไปด้วย 1. ผลตอบแทนการลงทุน (Return on Investment, ROI) และ 2. Shape Ratio โดย ROI จะสามารถใช้วัดผลตอบแทนจากการลงทุนได้ตาม (13) และ Sharpe Ratio เป็นตัวชี้วัดที่ใช้วัดผลตอบแทนโดยคำนึงถึงความเสี่ยง (Risk-adjusted Return) ของพอร์ตการลงทุน ด้วยซึ่งสามารถคำนวณได้จาก (14)

$$ROI = \frac{(Portfolio Value_i - Portfolio Value_{i-1})}{Portfolio Value_{i-1}} \quad (13)$$

$$Sharpe Ratio = \frac{(R_p - R_f)}{\sigma_p} \quad (14)$$

R_p คือ ผลตอบแทนของพอร์ตการลงทุน

R_f คือ ผลตอบแทนจากสินทรัพย์ไร้ความเสี่ยง

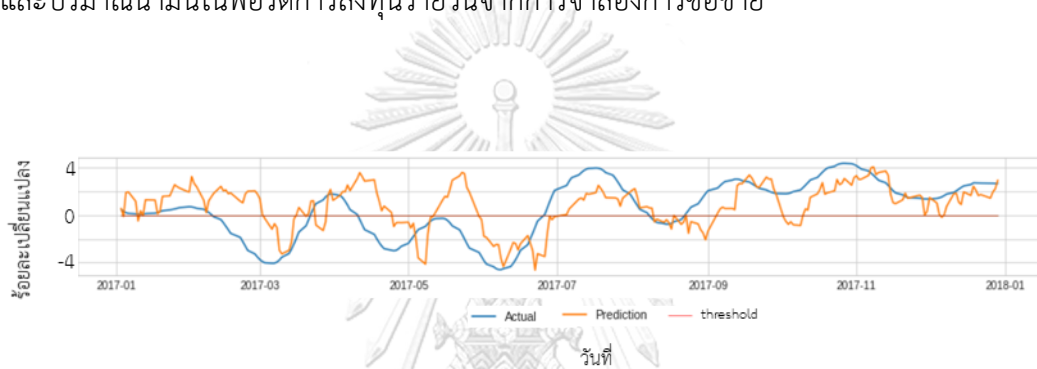
σ_p คือ Standard Deviation ของผลตอบแทนของพอร์ตการลงทุน

4.3.3 ระบบจำลองการซื้อขาย (Backtest)

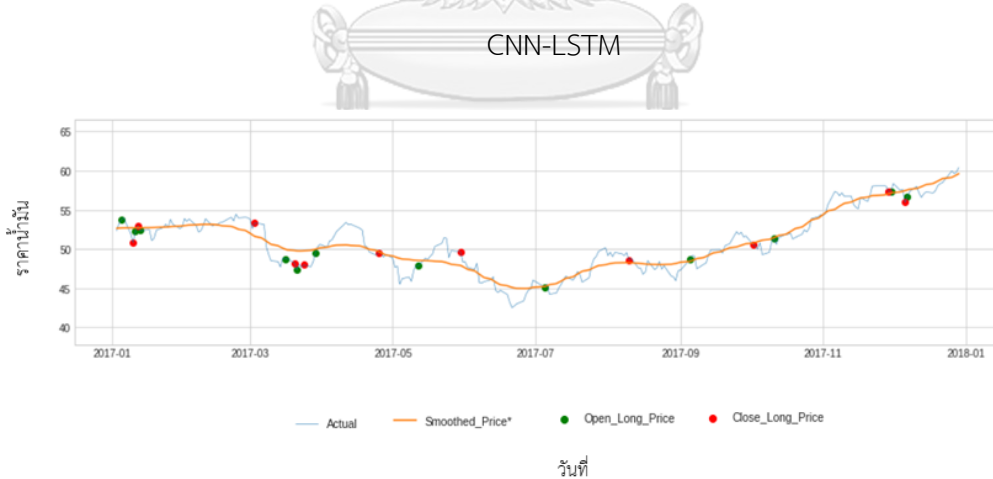
ในการวัดผลด้วยประสิทธิภาพการซื้อขายต้องใช้ระบบจำลองการซื้อขาย โดยเราจะนำแบบจำลองที่ผ่านการฝึกโดยเลือกค่าน้ำหนักและ Hyperparameter ที่เหมาะสมที่สุดมาใช้ในการจำลองการซื้อขาย (Backtest) ในช่วงเวลาทดสอบ ซึ่งเป็นข้อมูลที่แบบจำลองไม่เคยเห็นมาก่อนในช่วงการฝึก

รูปที่ 24 แสดงตัวอย่างผลการจำลองการซื้อขายด้วยกลยุทธ์ที่แปลผลมาจากการทำนายของแบบจำลอง CNN-LSTM ในปี ค.ศ. 2017 โดยรูปที่ 24(1) แสดงค่าร้อยละการเปลี่ยนแปลงรายวันของราคาเฉลี่ยจริงเป็นเส้นสีน้ำเงิน และค่าร้อยละการเปลี่ยนแปลงรายวันของราคาเฉลี่ยจากการทำนายเป็นเส้นสีส้ม จะเห็นได้ว่าทั้งสองเส้นมีลักษณะมีแนวโน้มไปในทิศทางเดียวกัน แต่ค่าที่ได้จากการทำนายจะมีการเปลี่ยนแปลงที่ช้ากว่าค่าจริง เพราะว่าข้อมูลที่แบบจำลองใช้ในการทำนายเป็นข้อมูลในอดีต

ทั้งนี้ในการทดลองเราได้นำเส้นร้อยละการเปลี่ยนแปลงรายวันของราคาเฉลี่ยจากการทำนายมาแปลผลเป็นสัญญาณซื้อขาย โดยเมื่อเส้นดังกล่าวตัดค่าศูนย์ขึ้นไปจะเป็นสัญญาณซื้อ บ่งบอกว่าแบบจำลองทำนายว่าราคาได้ผ่านจุด local minimum แล้วและกำลังจะกลับตัวขึ้นไป ในทางกลับกันเมื่อเส้นดังกล่าวตัดค่าศูนย์ลงมา จะถูกแปลเป็นสัญญาณขาย ซื้อ บ่งบอกว่าแบบจำลองทำนายว่าราคาได้ผ่านจุด local maximum แล้วและกำลังจะกลับตัวลงมา เมื่อได้สัญญาณซื้อและขายจะการแปลผลการทำนายของแบบจำลอง ระบบจำลองการซื้อขายจะซื้อและขายน้ำมันในราคาปิดของวันที่ส่งสัญญาณ ด้วยจำนวนเงินทั้งหมดที่มีในพอร์ตการลงทุน ณ วันนั้น ได้เป็นจุดซื้อ (สีเขียว) และจุดขาย (สีแดง) ดังแสดงในรูปตัวอย่างรูปที่ 24(2) โดยรูปที่ 24(3) แสดงมูลค่าพอร์ตการลงทุนรายวันและปริมาณน้ำมันในพอร์ตการลงทุนรายวันจากการจำลองการซื้อขาย



(1) ค่าร้อยละการเปลี่ยนแปลงรายวันของเส้นราคาเฉลี่ยจริง และผลการทำนายจากแบบจำลอง



(2) ราคาน้ำมันปิดรายวัน, เส้นราคาเฉลี่ย, และจุดซื้อและขายจากกลยุทธ์ที่ได้จากแบบจำลอง CNN-LSTM



(3) มูลค่าพอร์ตการลงทุน และปริมาณน้ำมันในพอร์ตการลงทุน

รูปที่ 24 ผลการจำลองการซื้อและขายจากกลยุทธ์ที่จากผลการทำนายร้อยละการเปลี่ยนแปลงรายวันของเส้นราคาเฉลี่ยของแบบจำลอง CNN-LSTM ในปี ค.ศ. 2017

ในการวัดประสิทธิภาพการซื้อขายจะเปรียบเทียบกับผลที่ได้จากกลยุทธ์การซื้อขายแบบดั้งเดิม [29] จากตัวบ่งชี้ทางเทคนิคที่นิยมใช้กันโดยทั่วไป ดังนี้

1. RSI จะเกิดสัญญาณซื้อเมื่อ RSI มีค่าเพิ่มขึ้นผ่าน 30 ขึ้นไป และในทางตรงกันข้ามจะเกิดสัญญาณขายเมื่อ RSI มีค่าลดลงและต่ำกว่า 70 ลงมา
2. MACD พิจารณาเส้น MACD และ MACD signal โดย MACD signal คือ EMA 9 วันของเส้น MACD ทั้งนี้กลยุทธ์ MACD จะเกิดสัญญาณซื้อเมื่อเส้น MACD ตัดเส้น MACD signal ขึ้นไป และในทางตรงกันข้ามเมื่อเส้น MACD ตัดเส้น MACD signal ลงมาจะเกิดเป็นสัญญาณขาย
3. SMA จะเกิดสัญญาณซื้อเมื่อเส้น SMA 50 วันตัดเส้น SMA 200 วันขึ้นไปและจะเกิดสัญญาณขายเมื่อเส้น SMA 50 วันตัดเส้น SMA 200 ลงมา
4. Stochastics จะพิจารณาเส้น %K และ %D จะเกิดสัญญาณซื้อเมื่อเส้น %K ตัดเส้น %D ขึ้นไป ในขณะที่เส้น %D มีค่าน้อยกว่า 20 และในทางตรงกันข้ามจะเกิดสัญญาณขายเมื่อเส้น %K ตัดเส้น %D ลงมา ในขณะที่เส้น %D มีค่ามากกว่า 80
5. ซื้อและถือ จะเป็นการซื้อที่วันแรกและขายในวันสุดท้ายในช่วงเวลาทดสอบ

4.4 ขั้นตอนการดำเนินงาน

1. นำข้อมูลราคาน้ำมันดิบ WTI รายวันตั้งแต่ปี ค.ศ. 2005 ถึงปี ค.ศ. 2020 จากเว็บไซต์ Yahoo Finance [26]
2. เตรียมข้อมูลโดยคำนวณตัวบ่งชี้ต่างๆ จากข้อมูลดิบ และเปลี่ยนแปลงข้อมูลให้มีความเหมาะสมสำหรับนำไปยังโครงข่ายประสาทเทียม ตามรายละเอียดในหัวข้อ 4.1.1 ถึง 4.1.3
3. แบ่งข้อมูลเป็น 6 ชุด ตามปี ค.ศ. ที่ต้องการทดสอบกลยุทธ์การซื้อขายจากแบบจำลอง ตั้งแต่ปี ค.ศ. 2015 จนถึงปี ค.ศ. 2020 ตามรายละเอียดในหัวข้อ 4.1.4
4. นำข้อมูลไปฝึกค่าน้ำหนักในโครงข่ายประสาทเทียม CNN-LSTM
5. เลือกค่าน้ำหนักโครงข่ายประสาทเทียมที่เหมาะสม จากรอบที่ได้ค่า MSE ของชุดข้อมูลตรวจสอบน้อยที่สุด
6. ทำซ้ำข้อ 4. ถึง 5. โดยทำ Grid Search กับจำนวน Filter ในชั้น CNN เป็น 64 และ 128 และจำนวน Node ในชั้น LSTM เป็น 64 และ 128 โดยเลือกจำนวน Filter ในชั้น CNN และจำนวน Node ในชั้น LSTM ที่ทำให้ค่า MSE ของชุดข้อมูลตรวจสอบน้อยที่สุด
7. นำแบบจำลองมาทำนายกับข้อมูลชุดทดสอบ เพื่อสร้างกลยุทธ์การซื้อขายและจำลองการซื้อขายย้อนหลัง
8. ทำข้อ 4. ถึง 7. ซ้ำสามรอบ ด้วย Random Seed ที่แตกต่างกัน แล้วเฉลี่ยผลการทดลอง
9. ทำซ้ำข้อ 4. ถึง 8. และเปลี่ยนโครงข่ายประสาทเทียมเป็น LSTM-CNN, Single CNN, Single LSTM
10. ทำข้อ 4. ถึง 9. กับข้อมูลทั้ง 6 ชุด
11. เปรียบเทียบและวิเคราะห์ผลการทดลอง

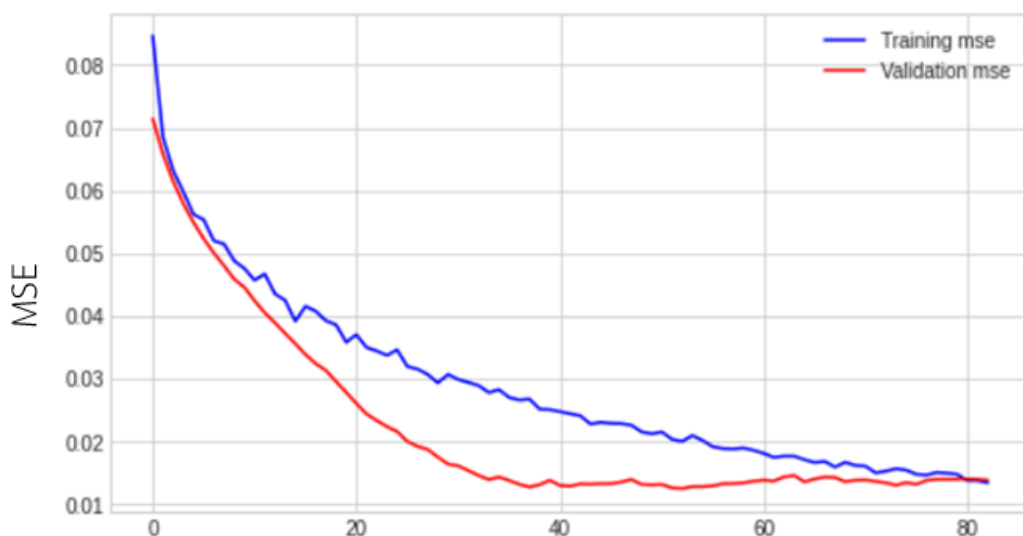
บทที่ 5

ผลการทดลอง

บทนี้จะกล่าวถึงผลการทดลองแบ่งเป็นสองส่วนตามวิธีการวัดผล ประกอบด้วย 1. ค่าความผิดพลาดจากการทำนาย และ 2. ประสิทธิภาพการซื้อขาย โดยค่าความผิดพลาดจากการทำนายจะเป็นการวัดความสามารถในการทำนายของแบบจำลองโครงข่ายประสาทเทียม จะเป็นการเปรียบเทียบความสามารถระหว่างแบบจำลองโครงข่ายประสาทเทียม CNN-LSTM, LSTM-CNN, Single CNN และ Single LSTM ในส่วนประสิทธิภาพการซื้อขายจะเป็นการนำโครงข่ายประสาทเทียมที่ถูกฝึกแล้วมาทดลองแปลผลสัญญาณซื้อขายกับข้อมูลชุดทดสอบ เพื่อวัดความสามารถในการทำกำไรของแบบจำลองแต่ละชนิด นอกจากนี้ยังตัวชี้วัดนี้ยังสามารถใช้ประเมินความสามารถเปรียบเทียบกับกลยุทธ์การซื้อขายแบบดั้งเดิม เช่น กลยุทธ์ซื้อและถือ, กลยุทธ์จากสัญญาณซื้อและขายจาก RSI, กลยุทธ์จากสัญญาณซื้อและขายจาก MACD, กลยุทธ์จากสัญญาณซื้อและขายจาก SMA, กลยุทธ์จากสัญญาณซื้อและขายจาก Stochastics RSI %K และ %D ได้อีกด้วย

5.1 ค่าความผิดพลาดจากการทำนาย

แบบจำลองโครงข่ายประสาทเทียมถูกฝึกด้วยชุดข้อมูลสำหรับฝึกและวัดผลเปรียบเทียบกับกระบวนการฝึกและทำ Grid Search ด้วยชุดข้อมูลตรวจสอบจากตัวอย่างผล MSE จากการฝึกโครงข่ายประสาทเทียม CNN-LSTM กับชุดข้อมูลปี 2017 ในรูปที่ 25 จะแสดงให้เห็นได้ว่าการฝึกในแต่ละรอบโครงข่ายประสาทเทียมจะมีการปรับค่าน้ำหนักจาก Backpropagation process เพื่อให้ได้ค่าน้ำหนักที่เหมาะสมกับชุดข้อมูลสำหรับฝึกมากขึ้น ส่งผลให้ MSE ของชุดข้อมูลสำหรับฝึกในแต่ละรอบการฝึกมีค่าลดลงไปเรื่อย ๆ โดยค่า MSE จะมีค่าสูงสุดในรอบการฝึกแรก และลดลงอย่างรวดเร็วในรอบการฝึกแรก ๆ ทั้งสำหรับชุดข้อมูลสำหรับการฝึกและชุดข้อมูลตรวจสอบจากนั้นเมื่อแบบจำลองผ่านการฝึกไปเรื่อย ๆ อัตราการลดลงของ MSE จะน้อยลงเรื่อย ๆ จนค่อนข้างคงที่ หรือมีการเพิ่มของ MSE เล็กน้อยในรอบการฝึกหลัง ๆ ซึ่งแสดงถึงการเกิด Overfit ของการเรียนรู้ของแบบจำลอง



จำนวนรอบการฝึก (Epochs)

รูปที่ 25 ผล MSE ของชุดข้อมูลสำหรับฝึกและชุดข้อมูลตรวจสอบของกลุ่มข้อมูลปี 2017

ทั้งนี้เรานำค่าน้ำหนักของแบบจำลองโครงข่ายประสาทเทียมจากรอบการฝึกและ Hyperparameter จาก Grid Search ที่ทำให้ค่า MSE ของชุดข้อมูลตรวจสอบน้อยที่สุดไปใช้วัดประสิทธิภาพของแบบจำลองกับชุดข้อมูลสำหรับทดสอบต่อไป ซึ่งเป็นชุดข้อมูลที่แบบจำลองไม่เคยเห็นมาก่อนในช่วงการฝึก จากผลการทดลองได้ค่า MSE เฉลี่ยของแบบจำลองโครงข่ายประสาทเทียม Single CNN, Single LSTM, CNN-LSTM และ LSTM-CNN ดังแสดงในตารางที่ 3 โดยตัวหนา หมายถึงค่า MSE ที่น้อยที่สุดในแต่ละชุดข้อมูล

จากผลการทดลองในตารางที่ 3 แสดงให้เห็นว่าการผสมกันระหว่าง CNN และ LSTM ของทั้งแบบจำลอง CNN-LSTM และ LSTM-CNN สามารถเพิ่มประสิทธิภาพการทำนายของโครงข่ายประสาทเทียมเมื่อเทียบกับ CNN แบบเดี่ยวหรือ LSTM แบบเดี่ยวได้อย่างมีนัยยะสำคัญ เมื่อพิจารณาค่าเฉลี่ยของทุกปีจะพบว่า CNN-LSTM และ LSTM-CNN ให้ค่าเฉลี่ย MSE ใกล้เคียงกันที่ $5.2 \times 10^{-2} \%^2$ ซึ่งมีค่าน้อยกว่าค่าเฉลี่ย MSE จาก CNN แบบเดี่ยวหรือ LSTM แบบเดี่ยวอยู่ถึง 21%

นอกจากนี้ยังสังเกตได้ว่าในปี 2020 แบบจำลองทั้ง 4 แบบมีประสิทธิภาพการทำนายที่แย่กว่าปีทดสอบอื่น ๆ อย่างเห็นได้ชัด ทั้งนี้เกิดจากปี 2020 เป็นปีที่ราคาน้ำมันได้รับผลกระทบจาก

เหตุการณ์ COVID-19 อย่างรุนแรง ซึ่งเป็นเหตุการณ์ที่ไม่เคยเกิดขึ้นมาก่อนในช่วงเวลาในชุดข้อมูล สำหรับฝึกแบบจำลอง ส่งผลให้ค่า MSE ในปี 2020 มีค่าสูงกว่าปีทดสอบอื่น ๆ อย่างมีนัยยะสำคัญ

ตารางที่ 3 MSE ของชุดข้อมูลทดสอบในแต่ละปี โดยมีหน่วยเป็น $1 \times 10^{-2} \%^2$ และตัวหนาแสดงถึงค่า MSE ที่น้อยที่สุดของแต่ละกลุ่มข้อมูล

	CNN	LSTM	CNN-LSTM	LSTM-CNN
2015	2.5	2.2	1.7	1.7
2016	2.3	1.4	1.8	1.6
2017	1.1	0.6	0.7	0.8
2018	1.5	1.3	0.8	1.2
2019	1.2	1.1	0.9	0.8
2020	29.3	31.0	25.1	25.1
Avg	6.3	6.3	5.2	5.2

5.2 ประสิทธิภาพการซื้อขาย

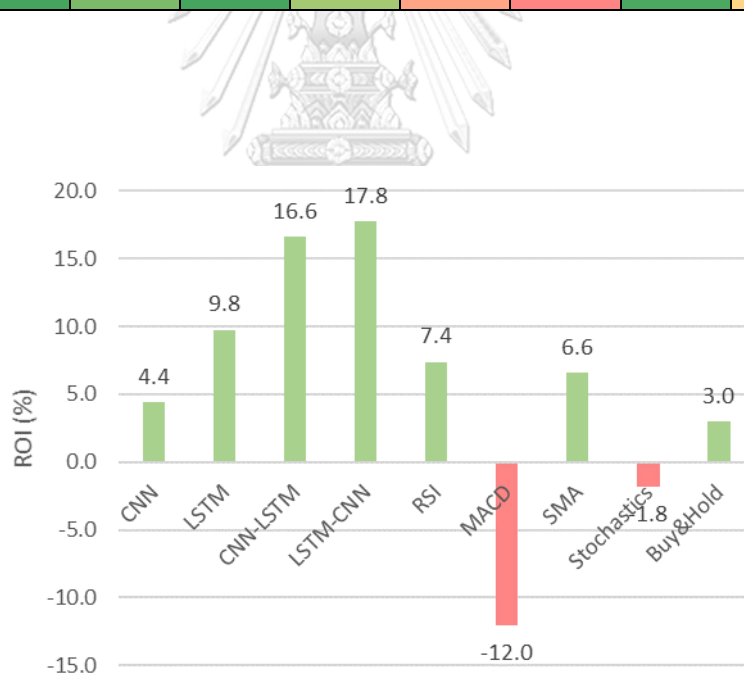
ผลตอบแทนการลงทุนถูกแสดงในตารางที่ 4 โดยตัวหนาแสดงถึงค่าผลตอบแทนการลงทุนที่สูงที่สุดของแต่ละกลุ่มข้อมูล และผลตอบแทนการลงทุนเฉลี่ยในทุกกลุ่มข้อมูลแสดงในรูปที่ 26 ซึ่งเห็นได้ว่าผลตอบแทนการลงทุนเฉลี่ยของแบบจำลอง LSTM-CNN มีค่าสูงที่สุดเทียบกับกลยุทธ์การซื้อขายอื่น ๆ ที่ 17.8% ในขณะที่แบบจำลอง CNN-LSTM มีผลตอบแทนการลงทุนเฉลี่ยที่ต่ำกว่าแบบจำลอง LSTM-CNN อยู่เล็กน้อยที่ 16.6%

ผลตอบแทนการลงทุนของแบบจำลอง CNN แบบเดี่ยวและ LSTM แบบเดี่ยวล้วนให้ผลตอบแทนการลงทุนที่ต่ำกว่าแบบจำลอง LSTM-CNN และแบบจำลอง CNN-LSTM อย่างเห็นได้ชัดที่ 4.4% และ 9.8% ตามลำดับ

ในขณะที่ผลตอบแทนการลงทุนจากกลยุทธ์การลงทุนแบบดั้งเดิมมีค่าต่ำกว่าผลจากแบบจำลอง LSTM-CNN และแบบจำลอง CNN-LSTM อย่างเห็นได้ชัดเช่นกัน โดยอยู่ในช่วง -12.0% – 7.4%

ตารางที่ 4 ผลตอบแทนการลงทุน (ROI) ของทุกกลยุทธ์การซื้อขาย โดยตัวหนาแสดงถึงค่า
ผลตอบแทนการลงทุนที่สูงที่สุดของแต่ละกลุ่มข้อมูล

	CNN	LSTM	CNN-LSTM	LSTM-CNN	RSI	MACD	SMA	Stochastics	Buy & Hold
2015	-15.7	-5.9	0.4	3.8	-20.0	-1.3	-10.7	-30.4	-29.8
2016	7.8	26.7	49.2	51.1	67.5	54.8	21.7	9.8	46.7
2017	3.8	10.9	4.6	25.3	27.8	9.1	10.4	25.0	15.6
2018	-6.9	4.0	4.0	0.9	-18.3	-13.4	-9.4	-12.1	-25.0
2019	-0.4	-3.7	1.0	7.3	1.0	-21.4	-8.6	0.9	31.3
2020	37.8	26.5	40.5	18.3	-13.5	-100.0	36.1	-4.2	-20.8



รูปที่ 26 ผลตอบแทนการลงทุนเฉลี่ยจากทุกกลุ่มข้อมูลในช่วง ค.ศ. 2015 – 2020 ของทุกกลยุทธ์การซื้อขาย

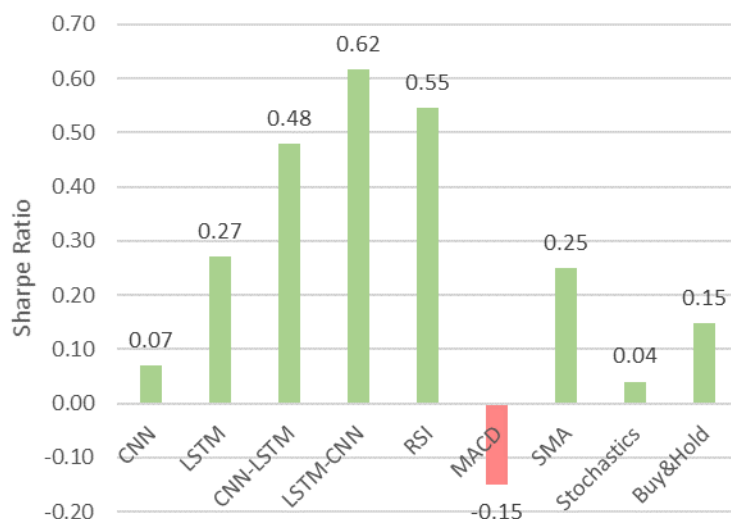
Sharpe Ratio ของแต่ละกลุ่มข้อมูลแสดงในตารางที่ 5 และค่าเฉลี่ยของ Sharpe Ratio ของทุกกลุ่มข้อมูลแสดงในรูปที่ 27 จากผลการทดลองพบว่าแบบจำลอง LSTM-CNN ให้ค่า Sharpe Ratio สูงที่สุดที่ 0.62 ในลำดับถัดไปกลยุทธ์การซื้อขายจากสัญญาณ RSI และแบบจำลอง CNN-LSTM ให้ค่า Sharpe Ratio ต่ำกว่าแบบจำลอง LSTM-CNN อยู่เล็กน้อยที่ 0.55 และ 0.48 ตามลำดับ

ตารางที่ 5 Sharpe Ratio ของทุกกลยุทธ์การซื้อขาย โดยตัวหนาแสดงถึงค่าผลตอบแทนการลงทุนที่สูงที่สุดของแต่ละกลุ่มข้อมูล

	CNN	LSTM	CNN-LSTM	LSTM-CNN	RSI	MACD	SMA	Stochastics	Buy & Hold
2015	-0.24	-0.13	0.06	0.13	-0.47	-0.04	-0.43	-0.95	-0.63
2016	0.31	0.83	1.85	1.61	2.10	1.79	0.71	0.68	0.98
2017	0.19	0.57	0.24	1.40	2.44	0.58	0.66	1.60	0.62
2018	-0.37	0.24	0.27	0.06	-0.71	-0.74	-0.55	-0.53	-0.80
2019	0.01	-0.17	0.02	0.27	0.04	-0.99	-0.41	0.05	0.92
2020	0.51	0.30	0.44	0.23	-0.13	-1.50	1.52	-0.61	-0.20

จากประสิทธิภาพการซื้อขายทั้งในส่วน of ผลตอบแทนการลงทุนและ Sharpe Ratio พบว่าแบบจำลอง LSTM-CNN สามารถให้ค่าทั้งสองสูงที่สุดเทียบกับกลยุทธ์การซื้อขายแบบอื่น ๆ ทั้งหมด ถึงแม้ว่าแบบจำลอง CNN-LSTM และแบบจำลอง LSTM-CNN จะมีความสามารถในการทำนายไม่ต่างกันจากผลของ MSE ในค่าความผิดพลาดจากการทำนาย แต่อย่างไรก็ตามแบบจำลอง LSTM-CNN สามารถทำให้ผลตอบแทนการลงทุนและค่า Sharpe Ratio สูงกว่า CNN-LSTM โดยเฉลี่ย เมื่อเปรียบเทียบในแต่ละปีจะพบว่า แบบจำลอง LSTM-CNN มักจะให้ผลตอบแทนการลงทุนและค่า Sharpe Ratio สูงกว่าหรือใกล้เคียงกับแบบจำลอง CNN-LSTM แต่มีการทดสอบในกลุ่มข้อมูลทดสอบปี ค.ศ. 2020 เพียงปีทดสอบเดียวที่แบบจำลอง CNN-LSTM สามารถให้ค่าผลตอบแทนการลงทุนและค่า Sharpe Ratio สูงกว่าแบบจำลอง LSTM-CNN อย่างเห็นได้ชัด ทั้งนี้เกิดจากเหตุการณ์

COVID-19 ในปี ค.ศ. 2020 ราคาน้ำมันมีแนวโน้มของราคาทั้งขาขึ้นและขาลงที่ชัดเจนในปีเดียวกัน ดังนั้นการส่งสัญญาณซื้อและขายจากแบบจำลองที่คลาดเคลื่อนไปเพียงเล็กน้อย จะส่งผลกระทบต่อผลตอบแทนการลงทุนและค่า Sharpe Ratio ค่อนข้างเยอะ



รูปที่ 27 Sharpe Ratio เฉลี่ยจากทุกกลุ่มข้อมูลในช่วง ค.ศ. 2015 – 2020 ของทุกกลยุทธ์การซื้อขาย

ผลการทดลองค่าความผิดพลาดจากการทำนายให้ผลในลักษณะเดียวกับประสิทธิภาพการซื้อขายในกรณีเปรียบเทียบแบบจำลอง CNN แบบเดี่ยวหรือ LSTM แบบเดี่ยว กับแบบจำลองที่ผสมกันระหว่าง CNN และ LSTM ซึ่งจากผลการทดลองพบว่าแบบจำลองแบบประสานกันระหว่าง CNN และ LSTM โดยไม่คำนึงถึงลำดับการเรียงตัวของชั้น CNN และ LSTM ล้วนแล้วแต่สามารถลด ค่าความผิดพลาดจากการทำนายและเพิ่มประสิทธิภาพการซื้อขายได้เมื่อเทียบกับแบบจำลองแบบเดี่ยวได้ทั้งสิ้น ทั้งนี้เป็นเพราะการผสมกันของแบบจำลอง CNN และ LSTM สามารถเพิ่มความซับซ้อน (Complexity) ในแบบจำลองได้ ทำให้แบบจำลองสามารถเรียนรู้ความสัมพันธ์ที่ซับซ้อนระหว่างตัวแปรนำเข้าและฉลากข้อมูลได้ดียิ่งขึ้น

บทที่ 6

สรุปการทดลอง และข้อเสนอแนะ

6.1 สรุปการทดลอง

ในงานวิจัยนี้เรานำเสนอแบบจำลองโครงข่ายประสาทเทียมหลายรูปแบบจากการผสมกันของแบบจำลอง CNN และแบบจำลอง LSTM เพื่อสร้างสัญญาณการซื้อขายน้ำมันดิบ WTI โดยเราได้สร้างระบบจำลองการซื้อขายเพื่อสร้างพอร์ตการลงทุน ทดสอบความสามารถในการทำผลตอบแทนจากข้อมูลราคาในช่วงปี ค.ศ. 2015 ถึง ค.ศ. 2020 ความสามารถของแบบจำลองถูกวัดผลในสองมิติ ได้แก่ ค่าความผิดพลาดจากการทำนายและประสิทธิภาพการซื้อขาย

จากผลการทดลองพบว่าแบบจำลอง LSTM-CNN มีความสามารถของแบบจำลองที่ดีที่สุดเทียบกับแบบจำลองรูปแบบอื่น ๆ ในงานวิจัยนี้ เนื่องจากแบบจำลองนี้ประกอบไปด้วยชั้น CNN และชั้น LSTM ในลำดับการเรียงตัวที่เหมาะสมที่ทำให้แบบจำลองสามารถเรียนรู้ความสัมพันธ์ที่ซับซ้อนของข้อมูลสำหรับการฝึกได้ดีกว่ากลยุทธ์อื่น ๆ

6.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

1. สามารถเพิ่มความซับซ้อนของแบบจำลองเป็นการผสมกันระหว่างชั้น LSTM และชั้น CNN เพิ่มขึ้นเป็น 3 หรือ 4 ชั้น เพื่อเพิ่ม Complexity ในแบบจำลองเพิ่มขึ้นอีก
2. นอกจากสัญญาณการซื้อขายแล้ว การออกแบบการเรียนรู้ของเครื่องให้สามารถทำนายปริมาณการซื้อขายที่เหมาะสม เพื่อลดความเสี่ยงในการลงทุนก็เป็นปัจจัยที่สำคัญเช่นกัน

บรรณานุกรม

1. International Energy Agency. *Key World Energy Statistics 2020*. 2020 [2021-09-16]; Available from: <https://www.iea.org/reports/key-world-energy-statistics-2020>.
2. Zhao, Y., J. Li, and L. Yu, *A deep learning ensemble approach for crude oil price forecasting*. *Energy Economics*, 2017. **66**: p. 9-16.
3. Yu, L., X. Zhang, and S. Wang, *Assessing Potentiality of Support Vector Machine Method in Crude Oil Price Forecasting*. *EURASIA Journal of Mathematics, Science and Technology Education*, 2017. **13**(12).
4. Xie, W., et al., *New Method for Crude Oil Price Forecasting Based on Support Vector Machines*. *Computational Science*, 2006. **3474**: p. 444-451.
5. Zhao, C. and B. Wang, *Forecasting Crude Oil Price with an Autoregressive Integrated Moving Average (ARIMA) Model*, in *Fuzzy Information & Engineering and Operations Research & Management*. 2014. p. 275-286.
6. Kaboudan, M.A., *Computetric Forecasting of Crude Oil Prices*. *Proceedings of the IEEE Conference on Evolutionary Computation, ICEC*, 2001. **1**: p. 283-287.
7. Hou, A. and S. Suardi, *A nonparametric GARCH model of crude oil price return volatility*. *Energy Economics*, 2012. **34**(2): p. 618-626.
8. Ramyar, S. and F. Kianfar, *Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models*. *Computational Economics*, 2017. **53**(2): p. 743-761.
9. Brigida, M., *The switching relationship between natural gas and crude oil prices*. *Energy Economics*, 2014. **43**: p. 48-55.
10. Sun, S., Y. Wei, and S. Wang, *AdaBoost-LSTM Ensemble Learning for Financial Time Series Forecasting*, in *Computational Science – ICCS 2018*. 2018. p. 590-597.
11. Kumar, G., S. Jain, and U.P. Singh, *Stock Market Forecasting Using Computational Intelligence: A Survey*. *Archives of Computational Methods in Engineering*, 2020. **28**(3): p. 1069-1101.
12. Gudelek, M.U., S.A. Boluk, and A.M. Ozbayoglu, *A Deep Learning based Stock*

- Trading Model with 2-D CNN Trend Detection*. 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2018: p. 1-8.
13. Tsantekidis, A., et al., *Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks*, in *2017 IEEE 19th Conference on Business Informatics (CBI)*. 2017. p. 7-12.
 14. Li, Z. and V. Tam, *A comparative study of a recurrent neural network and support vector machine for predicting price movements of stocks of different volatilities*. 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017: p. 1-8.
 15. Cheng, L.-C., Y.-H. Huang, and M.-E. Wu, *Applied attention-based LSTM neural networks in stock prediction*. 2018 IEEE International Conference on Big Data (Big Data), 2018: p. 4716-4718.
 16. Virtanen, P., et al., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*. Nature Methods, 2020. **17**: p. 261-272.
 17. Aggarwal, C.C., *Neural Networks and Deep Learning*. 2018.
 18. Chollet, F., *Deep Learning with Python*. 2018: Manning Publications Co.
 19. Olah, C. *Understanding LSTM Networks*. [2015-08-22]; Available from: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
 20. Hochreiter, S. and J. Schmidhuber, *Long Short-Term Memory*. Neural Computation, 1997. **9**(8): p. 1735-1780.
 21. CFA Institute, *Alternative Investment and Portfolio Management*. Vol. 6. 2020.
 22. CFA Institute, *Corporate Finance and Equity*. Vol. 4. 2020.
 23. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning: Machine Learning Book*. 2016.
 24. Thakkar, A. and K. Chaudhari, *A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions*. Expert Systems with Applications, 2021. **177**.
 25. Rezaei, H., H. Faaljou, and G. Mansourfar, *Stock price prediction using deep learning and frequency decomposition*. Expert Systems with Applications, 2021. **169**.
 26. NY Mercantile. *Crude Oil (WTI)*. 2021 [2021-10-01]; Available from:

<https://finance.yahoo.com/quote/CL=F>.

27. Padiyal, D.L. *Technical Analysis Library in Python*. 2021 [2021-10-01]; Available from: <https://technical-analysis-library-in-python.readthedocs.io/en/latest/>.
28. Kumar, G., U.P. Singh, and S. Jain, *Hybrid evolutionary intelligent system and hybrid time series econometric model for stock price forecasting*. *International Journal of Intelligent Systems*, 2021. **36**(9): p. 4902-4935.
29. Investopedia. *Financial Term Dictionary*. 2021 [2021-10-03]; Available from: <https://www.investopedia.com/>.
30. Tuychiev, B. *How to Remove Non-Stationarity in Time Series Forecasting*. [2021-09-17]; Available from: <https://towardsdatascience.com/how-to-remove-non-stationarity-in-time-series-forecasting-563c05c4bfc7>.
31. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 2011. **12**: p. 2825-2830.
32. Wikipedia. *Savitzky–Golay filter*. [2021-10-22]; Available from: https://en.wikipedia.org/wiki/Savitzky–Golay_filter.
33. Kingma, D. and J. Ba, *Adam: A Method for Stochastic Optimization*. *International Conference on Learning Representations*, 2014.



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประวัติผู้เขียน

ชื่อ-สกุล	วิศรุต เลิศทวีเดช
วัน เดือน ปี เกิด	14 กรกฎาคม 2533
สถานที่เกิด	กรุงเทพมหานคร, ประเทศไทย
วุฒิการศึกษา	วท.บ. (เกียรตินิยมอันดับหนึ่ง) วิศวกรรมเคมี จุฬาลงกรณ์มหาวิทยาลัย (พ.ศ. 2551 - 2555) วศ.ม. วิศวกรรมปิโตรเลียม จุฬาลงกรณ์มหาวิทยาลัย (พ.ศ. 2557 - 2559)
ที่อยู่ปัจจุบัน	28/194 คอนโด-28-ชิดลม ถนนชิดลม แขวงลุมพินี เขตปทุมวัน กรุงเทพมหานคร 10330

