การทำนายแนวโน้มของตลาดหุ้นด้วยแบบจำลองการเรียนรู้เชิงลึกที่เพิ่มประสิทธิภาพร่วมกับข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษร

นายพิศุทธ อ่อนเจริญ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Stock Market Movement Prediction Using Enhanced Deep Learning Model with Numerical and Textual Information

Mr. Pisut Oncharoen

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

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 2018

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การทำนายแนวโน้มของตลาดหุ้นด้วยแบบจำลองการเรียนรู้เชิงลึกที่เพิ่มประสิทธิภาพร่วมกับข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษร
โดย	นายพิศุทธ อ่อนเจริญ
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกูล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

---------- คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

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

---------- อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกูล)

---------- กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.พงศ์ศักดิ์ เหลืองอร่าม)

---------- กรรมการภายนอกมหาวิทยาลัย
(ดร.ธนภัทร ฆังคะจิตร)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

พิศุทธ อ่อนเจริญ : การทำนายแนวโน้มของตลาดหุ้นด้วยแบบจำลองการเรียนรู้เชิงลึกที่เพิ่มประสิทธิภาพร่วมกับข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษร. (Stock Market Movement Prediction Using Enhanced Deep Learning Model with Numerical and Textual Information) อ.ที่ปรึกษาหลัก : ผศ. ดร.พีรพล เวทีกูล

การทำนายแนวโน้มของตลาดหุ้นเป็นสิ่งที่ยากเนื่องจากตลาดหุ้นมีความผันผวนสูงและได้รับอิทธิพลจากปัจจัยภายนอกอื่น ๆ ในปัจจุบันเริ่มมีการนำเทคนิคการเรียนรู้เชิงลึก (Deep Learning) มาใช้ในการทำนายแนวโน้มของตลาดหุ้น โดยที่ข้อมูลรับเข้าของแบบจำลองสามารถแบ่งออกได้เป็น 2 ประเภทคือ 1) ข้อมูลเชิงตัวเลข เช่น ราคาในอดีตและตัวชี้วัดทางเทคนิค และ 2) ข้อมูลเชิงตัวอักษร ซึ่งได้แก่หัวข้อข่าวและเนื้อข่าว เป็นต้น แต่อย่างไรก็ตามงานวิจัยส่วนใหญ่มักจะมุ่งเน้นไปที่สร้างแบบจำลองโดยใช้ข้อมูลประเภทใดประเภทหนึ่งเท่านั้น ในขณะที่นักลงทุนส่วนใหญ่ทำการวิเคราะห์พฤติกรรมของตลาดโดยพิจารณาจากข้อมูลหลากหลายประเภท งานวิจัยนี้ได้นำเสนอแบบจำลองการเรียนรู้เชิงลึกที่สามารถพิจารณาข้อมูลทั้งสองประเภทเพื่อทำนายแนวโน้มของตลาดหุ้น ซึ่งแบบจำลองนี้ประกอบไปด้วยนิวรอลเน็ตเวิร์กแบบคอนโวลูชัน (Convolutional Neural Network) และหน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory) โดยใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัวซึ่งสกัดได้จากหัวข้อข่าว ราคาในอดีตและตัวชี้วัดทางเทคนิคซึ่งสร้างจากข้อมูลของราคาในอดีต รวมทั้งได้ทำการนำเสนอฟังก์ชันวัตถุประสงค์ชนิดใหม่ที่สามารถช่วยเพิ่มประสิทธิภาพในด้านผลตอบแทนเฉลี่ยต่อปีที่ได้จากการจำลองการซื้อขายโดยการนำเอาค่าชาร์ปเรโชซึ่งเป็นตัวชี้วัดผลตอบแทนเมื่อเทียบกับความเสี่ยงมาใช้ร่วมกับค่าครอสเอนโทรปี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อนิสิต .....

ปีการศึกษา 2561 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

5971017621 : MAJOR COMPUTER SCIENCE

KEYWORD: Deep Learning, Convolutional Neural Network, Event Embedding, Stock Market Prediction

Pisut Oncharoen : Stock Market Movement Prediction Using Enhanced Deep Learning Model with Numerical and Textual Information. Advisor: Asst. Prof. Peerapon Vateekul, Ph.D.

Stock market prediction is difficult because markets are volatile and influenced by many factors. Recently, many studies attempt to predict stock market trend using deep learning approach. These prediction models employed two types of input as (1) numerical information of historical prices and technical indicators, and (2) textual information including news contents or headlines. However, most of the studies focused on prediction model development based on a single input type, while investors analyzed market behavior based on a variety of information. In this work, we proposed a deep neural network for stock market prediction, which can analyze both types of inputs. The proposed model consists of convolutional neural network and long-short term memory and takes event embedding vectors extracted from news headlines, historical price data, and a set of technical indicators as input. Moreover, we also introduced a new objective function that can improve annualized return based on trading simulations by using Sharpe ratio, which is a measure of return relative to risk, and Cross-entropy.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Field of Study: Computer Science Student's Signature .....

Academic Year: 2018 Advisor's Signature .....

กิตติกรรมประกาศ

การที่วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีนั้น นอกจากการทำงานของตัวผู้วิจัยแล้ว ยังมีบุคคลท่านอื่นที่เป็นส่วนสำคัญที่ได้ให้ความช่วยเหลือในการจัดทำวิทยานิพนธ์ฉบับนี้ขึ้นมา ผู้วิจัยรู้สึกทราบซึ้งในความกรุณาเหล่านี้เป็นอย่างมากจึงใคร่ขอใช้เนื้อหาในส่วนกิตติกรรมประกาศของวิทยานิพนธ์ฉบับนี้แสดงความขอบพระคุณเป็นอย่างสูงมา ณ ที่นี้

ขอขอบพระคุณอาจารย์ที่ปรึกษา ผศ. ดร. พีรพล เวทีกูล ผู้ที่คอยให้ความช่วยเหลือและให้คำปรึกษาอย่างเต็มที่ รวมทั้งผลักดันให้งานวิจัยและวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ซึ่งประกอบไปด้วย ศ. ดร. บุญเสริม กิจสิริกุล ผศ. ดร. พงศ์ศักดิ์ เหลืองอร่าม และ ดร. ธนภัทร ฆังคะจิตร ที่ได้กรุณาให้เกียรติเป็นคณะกรรมการ รวมทั้งให้คำปรึกษาและข้อเสนอแนะอันเป็นประโยชน์อย่างมากต่อการทำวิจัยและวิทยานิพนธ์ฉบับนี้

ขอขอบคุณเพื่อน ๆ พี่ ๆ น้อง ๆ สมาชิก “Data Mining Group, MIND Lab” ทุกท่าน สำหรับกำลังใจ และคำแนะนำต่าง ๆ เพื่อนำมาปรับใช้ในวิทยานิพนธ์ฉบับนี้

สุดท้ายนี้ขอขอบพระคุณครอบครัวและบุคคลอันเป็นที่รักของผู้วิจัยที่ให้การสนับสนุนในทุก ๆ ด้าน และคอยให้กำลังใจตลอดระยะเวลาในการดำเนินการทำงานวิจัยนี้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

พิศุทธ อ่อนเจริญ

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ค

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญรูปภาพ ..... ฎ

สารบัญตาราง ..... ฐ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญ 1

1.2 วัตถุประสงค์ของงานวิจัย 2

1.3 ขอบเขตการวิจัย 3

1.4 ประโยชน์ที่ได้รับ 3

1.5 วิธีดำเนินการวิจัย 3

1.6 ผลงานวิจัยที่ตีพิมพ์ 4

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 5

2.1 การแทนข้อความ (Text Representation) 5

2.1.1 ถุงคำ (Bag of words) 5

2.1.2 ทีเอฟไอดีเอฟ (Term Frequency Inverse Document Frequency หรือ TF-IDF) 5

2.1.3 เวกเตอร์วันฮอท (One-hot Vector) 6

2.1.4 คำฝังตัว (Word Embedding) 6

2.2 นิวรอลเน็ตเวิร์ก (Neural Network) 7

2.2.1 เพอร์เซ็ปตรอน (Perceptron) 7

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2.2 นิวรอลเน็ตเวิร์กแบบป้อนไปข้างหน้า (Feedforward Neural Network).....8
2.2.3 ฟังก์ชันกระตุ้น (Activation Function).....9
2.2.4 ฟังก์ชันต้นทุน/วัตถุประสงค์ (Cost function หรือ Objective function).....10
2.2.5 การหาค่าที่เหมาะสมที่สุด (Optimization).....10
2.2.6 การแพร่กระจายย้อยกลับและการเรียนรู้ (Back propagation and Training).....11
2.3 นิวรอลเน็ตเวิร์กเชิงลึก (Deep Neural Network).....12
2.3.1 นิวรอลเน็ตเวิร์กคอนโวลูชัน (Convolutional Neural Network หรือ CNN).....13
2.3.2 นิวรอลเน็ตเวิร์กแบบวกกลับ (Recurrent Neural Network).....16
2.3.3 หน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory หรือ LSTM).....17
2.4 การวัดประสิทธิภาพ (Performance Evaluation).....19
2.4.1 ประสิทธิภาพในด้านการทำนายแนวโน้มของตลาดหุ้น.....19
2.4.1.1 คอนฟิวชันเมทริกซ์ (Confusion Matrix).....19
2.4.1.2 ตัววัดประสิทธิภาพจำแนกตามคลาส.....19
2.4.2 ประสิทธิภาพในด้านผลตอบแทน.....20
2.4.2.1 ผลตอบแทนเฉลี่ยต่อปี (Annualized Return).....20
2.4.2.2 ชาร์ปเรโช (Sharpe Ratio).....20
บทที่ 3 งานวิจัยที่เกี่ยวข้อง.....22
3.1 แบบจำลองการเรียนรู้เชิงลึกที่ใช้ข้อมูลรับเข้าเป็นข้อมูลเชิงตัวอักษร.....22
3.1.1 งานวิจัยของ Ding และคณะ ที่นำข้อมูลข่าวมาใช้ทำนายพฤติกรรมของราคาหุ้น.....22
3.1.2 งานวิจัยของ Ding และคณะ ซึ่งปรับปรุงวิธีการใช้หัวข้อข่าว.....23
3.2 แบบจำลองการเรียนรู้เชิงลึกที่พิจารณาข้อมูลเชิงตัวอักษรร่วมกับข้อมูลเชิงตัวเลข.....26
3.2.1 งานวิจัยของ Vargas และคณะ ที่นำหัวข้อข่าวมาใช้ร่วมกับตัวชี้วัดทางเทคนิค.....26
3.2.2 งานวิจัยของ Akita และคณะ ที่นำหัวข้อข่าวมาใช้ร่วมกับราคาหุ้นในอดีต.....27
3.3 ประเด็นที่พบจากงานวิจัยก่อนหน้าและสิ่งที่นำมาปรับปรุงในงานวิจัยนี้.....28

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.3.1 การแทนข้อความด้วยเวกเตอร์ 28
3.3.2 ข้อมูลเชิงตัวเลขที่ป้อนเข้าสู่แบบจำลอง 28
3.3.3 การวัดประสิทธิภาพของแบบจำลอง 29
บทที่ 4 แนวคิดและแบบจำลองที่นำเสนอในเบื้องต้นโดยเน้นที่ความแม่นยำในการทำนาย 30
4.1 แนวคิดและวิธีการที่นำเสนอในเบื้องต้น 30
4.1.1 การเตรียมข้อมูล 30
4.1.1.1 ตัวชี้วัดทางเทคนิค (Technical Indicator) 30
4.1.1.2 ตัวแทนของเหตุการณ์ (Event Representation) 31
4.1.1.3 เหตุการณ์ฝังตัว (Event Embedding) 31
4.1.2 แบบจำลองที่นำเสนอในเบื้องต้น 33
4.1.2.1 แบบจำลองที่ใช้ทำนายแนวโน้มของตลาดหุ้น 33
4.2 การทดลองและผลการทดลองเบื้องต้น 34
4.2.1 ชุดข้อมูล 34
4.2.1.1 ข้อมูลเชิงตัวเลข (Numerical Information) 34
4.2.1.2 ข้อมูลเชิงตัวอักษร (Textual information) 35
4.2.2 แบบจำลองอื่น ๆ เพื่อใช้ในการเปรียบเทียบประสิทธิภาพ 35
4.2.2.1 แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว 35
4.2.2.2 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค 36
4.2.2.3 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวเพียงอย่างเดียว 37
4.2.2.4 แบบจำลองที่ใช้ข้อมูลราคาในอดีตและทางเทคนิค 38
4.2.3 การวัดผล 38
4.2.3.1 ความแม่นยำ (Accuracy) 38
4.2.3.2 การจำลองการซื้อขาย (Trading Simulation) 38
4.2.4 ผลการทดลองเบื้องต้น 39

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2.4.1 ความแม่นยำ (Accuracy)....39
4.2.4.2 การจำลองการซื้อขาย (Trading Simulation)....40
บทที่ 5 แนวคิดในการดำเนินงานและวิธีการที่นำเสนอสำหรับแบบจำลองที่เน้นผลตอบแทน....42
5.1 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing)....42
5.1.1 ตัวชี้วัดทางเทคนิค (Technical Indicator)....42
5.1.2 นอร์มัลไลเซชัน (Normalization)....43
5.1.3 เหตุการณ์ฝังตัว (Event Embedding)....43
5.1.4 วิธีการสร้างผลเฉลย (Labeling Method)....44
5.2 แบบจำลองสำหรับการทำนายแนวโน้มของตลาดหุ้น....44
5.2.1 แบบจำลองที่นำเสนอ....44
5.2.2 การจำลองการซื้อขาย (Trading Simulation)....45
5.2.3 ฟังก์ชันวัตถุประสงค์ (Objective function)....46
5.2.4 กระบวนการสอนแบบจำลอง (Training Processes)....47
5.2.5 การเลือกแบบจำลองที่ดีที่สุด....47
บทที่ 6 การทดลองและผลการทดลอง....48
6.1 การสร้างแบบจำลองเพื่อใช้อ้างอิง....48
6.1.1 ประสิทธิภาพของวิธีการสกัดสารสนเทศแบบเปิด....48
6.1.2 ประสิทธิภาพของแบบจำลองที่สร้างเลียนแบบงานวิจัยของ Ding....49
6.2 ผลกระทบของพารามิเตอร์ที่ใช้ในฟังก์ชันวัตถุประสงค์และตัววัดผลที่นำเสนอ....49
6.2.1 อิทธิพลของพารามิเตอร์อัลฟาของฟังก์ชันที่นำเสนอ....50
6.2.2 อิทธิพลของพารามิเตอร์เบต้าของฟังก์ชันที่นำเสนอ....51
6.2.3 อิทธิพลของขนาดการเลื่อนข้อมูลที่ใช้สร้างผลเฉลย....51
6.3 ประสิทธิภาพของแบบจำลองเมื่อทำการเพิ่มตัวชี้วัดทางเทคนิค....52
6.4 ประสิทธิภาพของฟังก์ชันวัตถุประสงค์ที่นำเสนอ....53

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

6.4.1 ชุดข้อมูลและการตั้งค่าการทดลอง 53
6.4.2 ผลการทดลองและการอภิปรายผล.....54
6.5 การทำนายล่วงหน้าหลาย ๆ วัน 55
6.5.1 ชุดข้อมูลและการตั้งค่าการทดลอง 55
6.5.2 ผลการทดลองและการอภิปรายผล.....55
6.6 การทำนายหุ้นรายตัว.....56
6.6.1 ชุดข้อมูลและการตั้งค่าการทดลอง 56
6.6.2 ผลการทดลองและการอภิปรายผล.....56
6.7 การทำนายแนวโน้มของตลาดหุ้นด้วยข้อมูลเชิงตัวเลขร่วมกับข่าวประเภทต่าง ๆ.....57
6.7.1 ชุดข้อมูลและการตั้งค่าการทดลอง 57
6.7.2 ผลการทดลองและการอภิปรายผล.....57
6.8 พฤติกรรมการลงทุนในระหว่างการจำลองการซื้อขาย.....58
บทที่ 7 สรุปผลการวิจัยและแนวทางการวิจัยในขั้นถัดไป 60
6.1 สรุปผลการวิจัย.....60
6.2 แนวทางการวิจัยถัดไป 61
บรรณานุกรม 62
ประวัติผู้เขียน 66

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญรูปภาพ

รูปที่ 1 โครงสร้างของเพอร์เซ็ปตรอน 7
รูปที่ 2 โครงสร้างของนิวรอลเน็ตเวิร์กแบบป้อนไปข้างหน้า 8
รูปที่ 3 โครงสร้างของนิวรอลเน็ตเวิร์กคอนโวลูชัน 13
รูปที่ 4 ตัวอย่างการทำคอนโวลูชัน 13
รูปที่ 5 การทำคอนโวลูชันแบบกว้างและการเสริมเติม 14
รูปที่ 6 การทำคอนโวลูชันโดยมีข้อมูลรับเข้าขนาด 5x5 ตัวกรองขนาด 3x3 15
รูปที่ 7 การทำคอนโวลูชันโดยมีจำนวนตัวกล่องเท่ากับ3 15
รูปที่ 8 ตัวอย่างชั้นการรวมโดยค่าที่มากที่สุดและค่าเฉลี่ย 16
รูปที่ 9 ชั้นการเชื่อมโยงเต็มรูปแบบ 16
รูปที่ 10 โครงสร้างของนิวรอลเน็ตเวิร์กแบบวกกลับ 17
รูปที่ 11 ตัวอย่างโครงสร้างของหน่วยความระยะสั้นแบบยาว 18
รูปที่ 12 ตัวอย่างข่าว 22
รูปที่ 13 โครงสร้างของแบบจำลองนิวรอลเน็ตเวิร์กเชิงลึก 23
รูปที่ 14 อิทธิพลของข่าวในช่วงระยะเวลาต่าง ๆ 24
รูปที่ 15 โครงสร้างของนิวรอลเท็นเซอเน็ตเวิร์ก 24
รูปที่ 16 อัลกอริทึมที่ใช้สร้างเหตุการณ์ฝังตัว 25
รูปที่ 17 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก 25
รูปที่ 18 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก 26
รูปที่ 19 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก 27
รูปที่ 20 ตัวชี้วัดทางเทคนิคที่ใช้งานวิจัย 31
รูปที่ 21 โครงสร้างของนิวรอลเท็นเซอเน็ตเวิร์ก 31
รูปที่ 22 อัลกอริทึมที่ใช้ฝึกสอนแบบจำลองสำหรับการสร้างเหตุการณ์ฝังตัว 32
รูปที่ 23 แบบจำลองที่นำเสนอ 34

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รูปที่ 24 แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว 36
รูปที่ 25 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค 37
รูปที่ 26 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าว.....37
รูปที่ 27 แบบจำลองที่ใช้ข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิค.....38
รูปที่ 28 ภาพรวมของวิธีการที่นำเสนอ 42
รูปที่ 29 ตัวชี้วัดทางเทคนิคที่ใช้งานวิจัย 43
รูปที่ 30 ตัวอย่างของเวกเตอร์เหตุการณ์ฝังตัว.....43
รูปที่ 31 แบบจำลองที่นำเสนอ.....45
รูปที่ 32 การแบ่งข้อมูลเพื่อใช้สำหรับการสร้างแบบจำลอง.....47

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

ตารางที่ 1 ตัวอย่าง คอนฟิวชันเมทริกซ์ของการจำแนกแบบ 2 คลาส 19
ตารางที่ 2 ค่าความแม่นยำของแบบจำลองเมื่อทดสอบด้วยชุดข้อมูลทดสอบ 40
ตารางที่ 3 ผลตอบแทนเฉลี่ยต่อปีเมื่อทดสอบด้วยชุดข้อมูลทดสอบ 41
ตารางที่ 4 ตัวอย่างการประเมินประสิทธิภาพของกระบวนการสกัดสารสนเทศแบบเปิด 48
ตารางที่ 5 รายละเอียดของการแบ่งชุดข้อมูลที่ใช้ในการทดลอง 49
ตารางที่ 6 ผลการทดลองเมื่อใช้ค่าพารามิเตอร์อัลฟาที่แตกต่างกัน 50
ตารางที่ 7 ผลการทดลองเมื่อใช้ค่าพารามิเตอร์เบต้าที่แตกต่างกัน 51
ตารางที่ 8 ผลการทดลองเมื่อใช้ช่วงการเลื่อนและฟังก์ชันวัตถุประสงค์ที่แตกต่างกัน 52
ตารางที่ 9 ผลการทดลองเมื่อใช้ช่วงการเลื่อนและฟังก์ชันวัตถุประสงค์ที่แตกต่างกัน 53
ตารางที่ 10 จำนวนพารามิเตอร์น้ำหนักและค่าไบแอสที่ใช้ในแต่ละแบบจำลอง 53
ตารางที่ 11 ผลการทดลองเมื่อนำฟังก์ชันวัตถุประสงค์ที่นำเสนอไปใช้กับแบบจำลองต่าง ๆ 54
ตารางที่ 12 ผลการทดลองเมื่อทำการทำนายล่วงหน้าหลาย ๆ วัน 55
ตารางที่ 13 รายละเอียดจำนวนหัวข้อข่าวที่ถูกนำมาใช้ในการทดลอง 56
ตารางที่ 14 ผลการทดลองเมื่อนำฟังก์ชันวัตถุประสงค์ที่นำเสนอไปใช้กับแบบจำลองต่าง ๆ 57
ตารางที่ 15 ผลการทดลองเมื่อใช้ข้อมูลรับเข้าที่แตกต่างกัน 58
ตารางที่ 16 สัดส่วนพฤติการการลงทุนในระหว่างการจำลองการซื้อขาย 58

Chulalongkorn University

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

เป้าหมายของนักลงทุนส่วนใหญ่ คือ การคาดการณ์พฤติกรรมของตลาดหุ้น โดยมีจุดมุ่งหมายเพื่อใช้ในการตัดสินใจซื้อหรือขายหุ้นโดยมุ่งเน้นให้เกิดกำไรสูงสุด การคาดการณ์พฤติกรรมของตลาดเป็นสิ่งที่ยาก เนื่องจากตลาดหุ้นมีความผันผวนสูงและได้รับอิทธิพลจากปัจจัยภายนอก เช่น นโยบายทางด้านเศรษฐกิจ เศรษฐกิจโลก การเมืองความคาดหวังของนักลงทุน และอื่น ๆ

ทฤษฎีการเดินสุ่ม (Random Walk Theory) [1] ได้กล่าวว่า ในตลาดที่มีประสิทธิภาพ การเปลี่ยนแปลงของราคาหุ้นเป็นกระบวนการสุ่มที่ไม่มีแบบแผน ดังนั้นจึงเป็นไปไม่ได้ที่จะคาดการณ์ราคาในอนาคตได้ แต่อย่างไรก็ตามด้วยความก้าวหน้าในด้านปัญญาประดิษฐ์และขนาดของข้อมูลที่มีมากขึ้นในปัจจุบัน ทำให้นักวิจัยสามารถคาดการณ์พฤติกรรมของตลาดได้ดีขึ้นและมีประสิทธิภาพกว่ากระบวนการแบบสุ่ม ตัวอย่างเช่นงานวิจัยใน [2-7]

งานวิจัยกลุ่มแรก [2, 3] มุ่งเน้นไปที่การทำนายแนวโน้นพฤติกรรมของตลาดหุ้น โดยประยุกต์ใช้ ความรู้ด้านปัญญาประดิษฐ์ร่วมกับการวิเคราะห์ปัจจัยทางเทคนิค (Technical Analysis) ซึ่งเป็นการศึกษารูปแบบพฤติกรรมการเคลื่อนไหวของราคาในตลาดโดยเชื่อว่า ราคาเป็นผลรวมที่สะท้อนปัจจัยต่าง ๆ ในตลาดไว้หมดแล้วและพฤติกรรมของราคาจะเคลื่อนที่ซ้ำรอยเดิม แต่อย่างไรก็ตามวิธีการเหล่านี้มีข้อจำกัดเนื่องจากพฤติกรรมของตลาดหุ้นนั้นยังคงตอบสนองต่อปัจจัยภายนอกอื่น ๆ ซึ่งไม่ได้รวมอยู่ในข้อมูลของราคาหุ้นในอดีตที่นำมาใช้ทำการวิเคราะห์ทางเทคนิค

งานวิจัยในกลุ่มถัดมา [4-7] ได้รับแรงบันดาลใจจากการวิเคราะห์ปัจจัยพื้นฐาน (Fundamental Analysis) ซึ่งเป็นการวิเคราะห์ภาวะทางเศรษฐกิจ การเมืองและการศึกษาผลประกอบการในแต่ละกลุ่มธุรกิจ เพื่อคาดคะเนแนวโน้มในอนาคต งานวิจัยในกลุ่มนี้มุ่งเน้นไปที่การประยุกต์ใช้เทคนิคด้าน การวิเคราะห์ข้อความ (Text Mining) และการเรียนรู้ของเครื่อง (Machine Learning) เพื่อวิเคราะห์ ข้อมูลที่เป็นตัวอักษร (Textual Information) เช่น บทความข่าว รายงานทางการเงิน เพื่อหาความสัมพันธ์และนำมาใช้ในการทำนายพฤติกรรมของตลาดหุ้น

เนื่องจากประสิทธิภาพในการประมวลผลและความสามารถในการจัดการข้อมูลขนาดใหญ่ของคอมพิวเตอร์ในปัจจุบันเพิ่มสูงขึ้น ดังนั้นจึงมีความเป็นไปที่ได้จะประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องที่ซับซ้อนมากขึ้น เช่น แบบจำลองการเรียนรู้เชิงลึก (Deep Learning Model) เพื่อใช้ในการทำนายพฤติกรรมของตลาดหุ้น โดยที่แบบจำลองการเรียนรู้เชิงลึกที่โดนเด่นในปัจจุบันได้แก่ นิวรอลเน็ตเวิร์กแบบคอนโวลูชัน (Convolutional Neural Network หรือ CNN) [8, 9] นิวรอลเน็ตเวิร์กแบบย้อนกลับ (Recurrent Neural Network หรือ RNN) [10, 11] และนิวรอลเน็ตเวิร์กแบบ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

คอนโวลูชันย้อนกลับ (Recurrent Convolutional Neural Network หรือ RCNN) [12, 13] ตัวอย่างงานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้แบบจำลองการเรียนรู้เชิงลึก เพื่อใช้คาดการณ์อนุกรมเวลาทางการเงิน (Financial Time Series) ได้แก่งานวิจัยใน [14, 15] ซึ่งผู้เขียนได้ใช้นิวรอลเน็ตเวิร์กเชิงลึก (Deep Neural Network) และป้อนข้อมูลตัวแทนของเหตุการณ์ (Event Representation) ที่สกัดจากหัวข้อข่าวทางการเงิน (Financial News Articles) เพื่อใช้ในการคาดการณ์ทิศทางของราคาของหุ้น และ ดัชนีเอสแอนด์พี 500 (S&P 500 Index)

งานวิจัยที่อ้างถึงข้างต้นล้วนมุ่งเน้นที่การสร้างแบบจำลองโดยใช้ข้อมูลเชิงตัวเลขหรือใช้ข้อมูลที่เป็นตัวอักษร เพียงอย่างใดอย่างหนึ่งเท่านั้น ในขณะที่นักลงทุนส่วนใหญ่ทำการวิเคราะห์พฤติกรรมของตลาดโดยพิจารณาจากข้อมูลทั้ง 2 ประเภท ดังนั้นแนวทางในการวิจัยในช่วง 2 ปีที่ผ่านมาจึงเริ่มมีการพัฒนาแบบจำลองการเรียนรู้เชิงลึกโดยพิจารณาใช้ข้อมูลเชิงตัวเลขร่วมกับข้อมูลเชิงตัวอักษร ซึ่งได้แก่งานวิจัย [16, 17] ที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีต [16] และตัวชี้วัดทางเทคนิค (Technical Indicator) [17] แต่อย่างไรก็ตาม งานวิจัยดังกล่าวได้ใช้วิธีการพื้นฐานในการสร้างเวกเตอร์เพื่อใช้แทนหัวข้อข่าว ซึ่งเวกเตอร์ที่ได้นั้นอาจจะไม่สามารถสื่อถึงใจความสำคัญของข่าวนั้น ๆ ได้ ส่งผลให้ประสิทธิภาพของแบบจำลองอาจจะไม่ดีนัก

งานวิจัยนี้นำเอาจุดเด่นของงานวิจัย [15] ซึ่งนำเอาเหตุการณ์ฝังตัว (Event Embedding) มาใช้เป็นตัวแทนของเหตุการณ์เพื่อช่วยประสิทธิภาพในการทำนาย และงานวิจัย [17] ซึ่งนำเอาตัวชี้วัดทางเทคนิค มาใช้ร่วมกับหัวข้อข่าวเพื่อใช้ในการทำนายการเปลี่ยนแปลงของตลาดหุ้น โดยที่งานวิจัยนี้จะทำการสร้างแบบจำลองการเรียนรู้เชิงลึกที่ประกอบไปด้วยนิวรอลเน็ตเวิร์กแบบคอนโวลูชันและหน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory) เพื่อทำนายแนวโน้มของตลาดหุ้นโดยใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัวซึ่งสกัดได้จากหัวข้อข่าว ราคาในอดีตและตัวชี้วัดทางเทคนิคซึ่งสร้างจากข้อมูลของราคาในอดีต รวมทั้งได้ทำการนำเสนอฟังก์ชันวัตถุประสงค์ชนิดใหม่ที่สามารถช่วยเพิ่มประสิทธิภาพในด้านผลตอบแทนเฉลี่ยต่อปีที่ได้จากการจำลองการซื้อขาย โดยการนำเอาค่าชาร์ปเรโชซึ่งเป็นตัวชี้วัดผลตอบแทนเมื่อเทียบกับความเสี่ยงมาใช้ร่วมกับค่าครอสเอนโทรปี

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อนำเสนอวิธีการทำนายแนวโน้มของตลาดหุ้นโดยใช้แบบจำลองการเรียนรู้เชิงลึกที่เพิ่มประสิทธิภาพร่วมกับการใช้ข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษรโดยมุ่งเน้นที่การปรับปรุงแบบจำลองการเรียนรู้เชิงลึกให้สามารถวิเคราะห์ข้อมูลทั้งสองประเภทที่มีลักษณะแตกต่างกัน ร่วมกันได้

1.3 ขอบเขตการวิจัย

1. ข้อมูลเชิงตัวอักษรที่ใช้ในงานวิจัยนี้ คือ หัวข้อข่าวภาษาอังกฤษจากสามแหล่งข้อมูลได้แก่ รอยเตอร์ (Reuters) เรดดิต (Reddit) และ อินทรินิโอ (Intrinio)
2. ข้อมูลเชิงตัวเลขที่ใช้ในงานวิจัยนี้ คือ ราคาในอดีต (เปิด/ปิด/สูงสุด/ต่ำสุด) ตัวชี้วัดทางเทคนิคจำนวน 15 ชนิดซึ่งอ้างอิงจากงานวิจัย [18] และข้อมูลของตัวชี้วัดทางเทคนิคจะถูกสร้างโดยข้อมูลของราคาหุ้นในอดีตโดยใช้ข้อมูลจากยาฮู! ไฟแนนซ์ (Yahoo! Finance)
3. ทำการวัดประสิทธิภาพของแบบจำลองโดยการพิจารณาค่าความแม่นยำในการทำนายร่วมกับผลตอบแทนที่ได้จากการจำลองการซื้อ-ขาย ตามผลลัพธ์จากแบบจำลอง
4. เปรียบเทียบประสิทธิภาพของแบบจำลอง โดยการเปรียบผลลัพธ์ของการใช้ข้อมูลแต่ละประเภทเพียงอย่างเดียว การใช้ข้อมูลทั้งสองประเภทร่วมกัน

1.4 ประโยชน์ที่ได้รับ

1. สามารถเพิ่มประสิทธิภาพในการทำนายพฤติกรรมของตลาดหุ้นโดยใช้สร้างแบบจำลองการเรียนรู้เชิงลึก ร่วมกับข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษร
2. สามารถปรับปรุงวิธีการนำข้อมูลสองประเภทที่มีลักษณะต่างกัน มาพิจารณาร่วมกันโดยใช้แบบจำลองการเรียนรู้เชิงลึก
3. สามารถนำกรอบงานวิจัยนี้ไปประยุกต์ใช้กับข้อมูลหัวข้อข่าวที่เป็นภาษาอื่นได้
4. สามารถหาความสัมพันธ์ของหัวข้อข่าวและพฤติกรรมการเปลี่ยนแปลงของราคาหุ้นได้

1.5 วิธีดำเนินการวิจัย

1. ศึกษาวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับหัวข้อที่วิจัย
2. ศึกษาชุดข้อมูลและแบบจำลองเบื้องต้นที่ใช้ในการทำนายแนวโน้มของตลาดหุ้น
3. ทดสอบแนวทางการวิจัยในเบื้องต้น
4. วิเคราะห์ผลการทดลอง เพื่อแนวคิดที่ได้มาพัฒนาแบบจำลอง
5. นำผลการสรุปมาปรับปรุงแก้ไขแบบจำลอง
6. วิเคราะห์และสรุปผลการทดลองสุดท้าย
7. จัดทำวิทยานิพนธ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.6 ผลงานวิจัยที่ตีพิมพ์

“Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators” โดย พิศุทธ อ่อนเจริญ และ พีรพล เวทีกูล ในงานประชุมวิชาการ “2018 - 5th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA 2018)” ซึ่งจัดขึ้น ณ โรงแรมบียอน รีสอร์ท จังหวัดกระบี่ ประเทศไทย ระหว่างวันที่ 14-17 สิงหาคม 2561

“Deep Learning Using Risk-Reward Function for Stock Market Prediction” โดย พิศุทธ อ่อนเจริญ และ พีรพล เวทีกูล ในงานประชุมวิชาการ “2018 - 2nd International Conference on Computer Science and Artificial Intelligence (CSAI 2018)” ซึ่งจัดขึ้น ณ โรงแรมโนโวเทลเซินเจิ้นวอเตอร์เกท เขตเซินเจิ้น มณฑลกวางตุ้ง ประเทศจีน ระหว่างวันที่ 8-10 ธันวาคม 2561

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยชิ้นนี้แบ่งออกได้เป็น 4 หัวข้อได้แก่ การแทนข้อความ นิวรอลเน็ตเวิร์ก นิวรอลเน็ตเวิร์กเชิงลึก และการวัดประสิทธิภาพ

2.1 การแทนข้อความ (Text Representation)

การแทนข้อความเป็นกระบวนการหนึ่งที่สำคัญสำหรับการนำข้อมูลเชิงตัวอักษรมาวิเคราะห์ เนื่องจากคอมพิวเตอร์ไม่สามารถประมวลผลข้อมูลที่เป็นตัวอักษรได้โดยตรง ดังนั้นจึงต้องทำการแปลงข้อมูลเหล่านี้ให้เป็นตัวเลขเพื่อให้คอมพิวเตอร์สามารถเข้าใจข้อมูลเหล่านี้ หัวข้อถัดไปจะอธิบายวิธีการแทนข้อความโดยใช้วิธีการต่าง ๆ โดยใช้ข้อความตัวอย่างดังต่อไปนี้ 1) “ฉันชอบไปโรงเรียน โรงเรียนของฉันน่าอยู่” และ 2) “โรงเรียนของฉันสวย”

2.1.1 ถุงคำ (Bag of words)

เป็นวิธีการแทนข้อความให้อยู่ในรูปของเวกเตอร์ที่มีขนาดเท่ากับจำนวนคำทั้งหมดในพจนานุกรมของชุดข้อมูลนั้น ๆ จากข้อมูลตัวอย่างข้างต้นสามารถสร้างเป็นพจนานุกรรมของคำได้ดังนี้ [“ฉัน”, “ชอบ”, “ไป”, “โรงเรียน”,“ของ”,“น่าอยู่”,“สวย”] วิธีการถุงคำจะไม่คำนึงถึงรูปแบบไวยากรณ์และลำดับของคำ โดยจะแทนข้อความด้วยความถี่ของคำที่ปรากฏ ตัวอย่างเช่น

1) “ฉันชอบไปโรงเรียน โรงเรียนของฉันน่าอยู่” สามารถแทนด้วย [2 1 1 2 1 1 0]

2) “โรงเรียนของฉันสวย” สามารถแทนด้วย [1 0 0 1 1 0 1]

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.2 ทีเอฟไอดีเอฟ (Term Frequency Inverse Document Frequency หรือ TF-IDF)

เป็นวิธีการแทนข้อความด้วยถุงคำอย่างหนึ่ง โดยแทนค่าข้อมูลในเวกเตอร์ด้วยความถี่ของคำ (Term Frequency หรือ tf) คูณกับค่าผกผันของความถี่ของคำนั้น ๆ เมื่อเทียบกับทั้งชุดข้อมูล (Inverse Document Frequency หรือ idf) สามารถคำนวณค่า TF-IDF ได้โดยใช้สมการ (1) และ (2)

$$tfidf = tf\ x\ idf \tag{1}$$

$$idf = \log(\frac{N}{n_t}) \tag{2}$$

โดยที่ N คือ จำนวนข้อความทั้งหมดในชุดข้อมูล และ n_t คือ จำนวนของข้อความในชุดข้อมูลที่มีคำนั้น ๆ

จากข้อความตัวอย่างข้างต้น สามารถคำนวณค่าผกผันของความถี่ของคำนั้น ๆ เมื่อเทียบกับทั้งชุดข้อมูลได้เป็น [0 0.3 0.3 0 0 0 0.3 0.3] และสามารถแทนข้อความได้ดังนี้

1) “ฉันชอบไปโรงเรียน โรงเรียนของฉันน่าอยู่” สามารถแทนด้วย [0 0.3 0.3 0 0 0.3 0]

2) “โรงเรียนของฉันสวย” สามารถแทนด้วย [0 0 0 0 0 0 0.3]

2.1.3 เวกเตอร์วันฮอท (One-hot Vector)

เป็นการแทนข้อความด้วยกลุ่มเวกเตอร์ของคำที่มีขนาดเท่ากับจำนวนคำทั้งหมดในพจนานุกรมของชุดข้อมูลนั้น ๆ โดยที่เวกเตอร์เหล่านี้จะเรียงลำดับตามตำแหน่งของคำที่ปรากฏในข้อความ ค่าภายในเวกเตอร์จะมีค่าเป็น 1 เพียงตำแหน่งเดียว ส่วนตำแหน่งอื่น ๆ ในเวกเตอร์นี้จะมีค่าเป็น 0 จากตัวอย่างจะได้ว่าพจนานุกรมของคำคือ [“ฉัน”, “ชอบ”, “ไป”, “โรงเรียน”,“ของ”,“น่าอยู่”,“สวย”] สามารถแสดงการแทนคำด้วยเวกเตอร์วันฮอทได้ดังตัวอย่างต่อไปนี้

“ฉัน” แทนด้วย $\begin{bmatrix}1\\0\\0\\0\\0\\0\\0\end{bmatrix}$, “ชอบ” แทนด้วย $\begin{bmatrix}0\\1\\0\\0\\0\\0\\0\end{bmatrix}$, “ไป” แทนด้วย $\begin{bmatrix}0\\0\\1\\0\\0\\0\\0\end{bmatrix}$ และ “โรงเรียน” แทนด้วย $\begin{bmatrix}0\\0\\0\\1\\0\\0\\0\end{bmatrix}$

สำหรับการแทนข้อความจะนำเวกเตอร์วันฮอทมาเรียงต่อกันตามลำดับของคำที่ปรากฏในข้อความนั้น ๆ ตัวอย่างเช่น

“ฉันชอบไปโรงเรียน โรงเรียนของฉันน่าอยู่” สามารถแทนด้วย $\begin{bmatrix}1&0&0&0&0&0&1&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&0&0\end{bmatrix}$

“โรงเรียนของฉันสวย” สามารถแทนด้วย $\begin{bmatrix}0&0&1&0\\0&0&0&0\\0&0&0&0\\1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สำหรับการแทนคำด้วยวิธีถุงคำและเวกเตอร์วันฮอท เมื่อคำในชุดข้อมูลมีจำนวนมากจะส่งผลให้เวกเตอร์ที่ใช้แทนคำเหล่านั้นมีขนาดใหญ่ ซึ่งอาจจะส่งผลต่อความเร็วและประสิทธิภาพในการประมวลผลข้อมูลได้

2.1.4 คำฝังตัว (Word Embedding)

เป็นวิธีการแทนที่ข้อความด้วยเวกเตอร์คำ (Word Vector) ที่มีขนาดเล็ก เรียงลำดับตามตำแหน่งของคำที่ปรากฏในข้อความนั้น ๆ ทั้งนี้ เวกเตอร์คำนิยมสร้างโดยการวิเคราะห์ชุดข้อมูลทั้งหมด แล้วสร้างเวกเตอร์คำที่มีขนาดเล็ก ๆ ตามที่ต้องการโดยออกแบบให้กลุ่มของคำที่มีความหมายใกล้เคียงกันจะต้องมีระยะห่างระหว่างเวกเตอร์ใกล้เคียงกัน วิธีการสร้างเวกเตอร์คำที่นิยม

ได้แก่ เวิร์ดทูเวก (word2vec) [19] และ โกลฟ (GloVe) [20] กำหนดให้คำจากตัวอย่างข้างต้นสามารถแทนที่ด้วยเวกเตอร์คำดังต่อไปนี้

“ฉัน” แทนด้วย $\begin{bmatrix}0.82\\0.11\\0.56\end{bmatrix}$, “ชอบ” แทนด้วย $\begin{bmatrix}0.77\\0.35\\0.12\end{bmatrix}$, “ไป” แทนด้วย $\begin{bmatrix}0.33\\0.66\\0.11\end{bmatrix}$, “โรงเรียน” แทนด้วย $\begin{bmatrix}0.23\\0.32\\0.18\end{bmatrix}$, “ของ” แทนด้วย $\begin{bmatrix}0.74\\0.27\\0.13\end{bmatrix}$, “น่าอยู่” แทนด้วย $\begin{bmatrix}0.72\\0.54\\0.98\end{bmatrix}$ และ “สวย” แทนด้วย $\begin{bmatrix}0.42\\0.05\\0.28\end{bmatrix}$

เมื่อนำเวกเตอร์คำมารวมกันจะสามารถแทนที่ข้อความตัวอย่างได้ดังนี้

“ฉันชอบไปโรงเรียน โรงเรียนของฉันน่าอยู่” สามารถแทนด้วย

$$\begin{bmatrix}0.82 & 0.77 & 0.33 & 0.23 & 0.23 & 0.74 & 0.82 & 0.72\\0.11 & 0.35 & 0.66 & 0.32 & 0.32 & 0.27 & 0.11 & 0.54\\0.56 & 0.12 & 0.11 & 0.18 & 0.18 & 0.13 & 0.56 & 0.98\end{bmatrix}$$

“โรงเรียนของฉันสวย” สามารถแทนด้วย

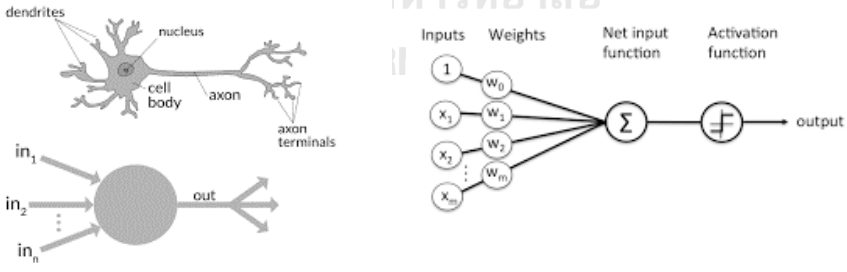
$$\begin{bmatrix}0.23 & 0.74 & 0.82 & 0.42\\0.32 & 0.27 & 0.11 & 0.05\\0.18 & 0.13 & 0.56 & 0.28\end{bmatrix}$$

2.2 นิวรอลเน็ตเวิร์ก (Neural Network)

เป็นแบบจำลองที่ได้รับแรงบรรดาใจมาจากสมองของมนุษย์ โดยสามารถเรียนรู้จากชุดข้อมูลฝึกสอนเพื่อใช้ทำนายข้อมูลทดสอบที่ไม่เคยเห็นขั้นตอนการฝึกสอนได้เป็นอย่างดี ในหัวข้อนี้จะอธิบายตั้งแต่ส่วนที่เล็กที่สุดของโครงข่ายประสาทเทียม ไปจนถึงวิธีการสอนแบบจำลอง

2.2.1 เพอร์เซ็ปตรอน (Perceptron)

เพอร์เซ็ปตอนคือส่วนที่เล็กที่สุดของนิวรอลเน็ตเวิร์ก ซึ่งเปรียบได้กับเซลล์ประสาทหนึ่งเซลล์ที่เรียกว่านิวรอล ลักษณะของเพอร์เซ็ปตอนแสดงดังรูปที่ 1



รูปที่ 1 โครงสร้างของเพอร์เซ็ปตรอน

(อ้างอิงจาก https://appliedgo.net/perceptron/)

เพอร์เซ็ปตรอนสามารถนำมาใช้จำแนกผลลัพธ์เป็นสองกลุ่ม โดยกำหนดให้ฟังก์ชันของเพอร์เซ็ปตรอนแทนด้วย $f(x)$ โดยมีข้อมูลรับเข้าคือ x และข้อมูลส่งออกคือ $\hat{y}$ โดยแสดงการคำนวณข้อมูลที่ส่งออกได้ดังสมการที่ (3)

$$\hat{y} = f(x) = \begin{cases} 1 & if \sum_{i=1}^{m} w_i x_i + b > 0 \\ 0 & otherwise \end{cases} \quad (3)$$

กำหนดให้ w คือน้ำหนัก (weights) b คือค่าไบแอส (bias) และ m คือจำนวนของข้อมูลรับเข้า สำหรับกระบวนการเรียนรู้ของเพอร์เซ็ปตรอน กำหนดให้ชุดข้อมูลตัวอย่างแทนด้วย x และผลลัพธ์จริงนั้นแทนด้วย y สามารถเขียนสมการการเรียนรู้ได้ดังต่อไปนี้

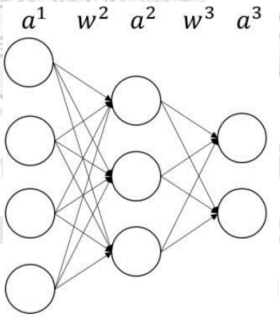
$$w_i \leftarrow w_i + \Delta w_i \quad (4)$$

$$\Delta w_i = \alpha(\hat{y} - y)x_i \quad (5)$$

โดยที่ α คืออัตราการเรียนรู้ (Learning Rate) ซึ่งเป็นค่าที่บ่งบอกว่าการเรียนรู้ในแต่ละรอบจะมีการเปลี่ยนแปลงน้ำหนักด้วยอัตราส่วนของผลต่างของผลลัพธ์ไปมากเท่าใด

2.2.2 นิวรอลเน็ตเวิร์กแบบป้อนไปข้างหน้า (Feedforward Neural Network)

นิวรอลเน็ตเวิร์กแบบป้อนไปข้างหน้ามีลำดับในการคำนวณและส่งผ่านข้อมูลไปในทิศทางเดียว โดยโครงสร้างจะแบ่งออกเป็นลำดับชั้น ซึ่งในแต่ละชั้นจะมีเพอร์เซ็ปตรอนจำนวนหนึ่งซึ่งมีเส้นเชื่อมถึงกันภายในชั้นเดียวกัน แต่จะมีเส้นเชื่อมกับเพอร์เซ็ปตรอนตัวอื่นที่อยู่ในลำดับชั้นที่ติดกันทั้งหมด โดยข้อมูลส่งออกของเพอร์เซ็ปตรอนในชั้นก่อนหน้า จะเป็นข้อมูลรับเข้าของเพอร์เซ็ปตรอนในชั้นปัจจุบัน [21] โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า แสดงได้ดังรูปที่ 2



รูปที่ 2 โครงสร้างของนิวรอลเน็ตเวิร์กแบบป้อนไปข้างหน้า
(อ้างอิงจาก Fig.2 ใน [21])

กำหนดสัญญาลักษณ์แทนการคำนวณไปข้างหน้า (feedforward) โดยให้ a_k^{l-1} แทนผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นที่ *l-1* และ w_{jk}^{l} แทนน้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ *j* ในลำดับชั้น *l* ที่มีเส้นเชื่อมาจากเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นก่อนหน้า และ b_j^l คือ ไบแอส นอกจากนี้ให้ g แทนฟังก์ชันกระตุ้น และให้ n แทนจำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ *l-1* จะสามารถแสดงการคำนวณค่า a_k^l ได้โดยสมการดังต่อไปนี้

$$z_j^l = \sum_{k=1}^{n} w_{jk}^l a_k^{l-1} + b_j^l \quad (6)$$

$$a_j^l = g(z_j^l) \quad (7)$$

2.2.3 ฟังก์ชันกระตุ้น (Activation Function)

สำหรับข้อมูลส่งออกของแต่ละเพอร์เซ็ปตรอน จะมีการใช้ฟังก์ชันกระตุ้น g(z) เพื่อทำให้นิวรอลเน็ตเวิร์กมีความซับซ้อนและสามารถแก้ปัญหาได้หลายหลายมากขึ้น ฟังก์ชันกระตุ้นที่นิยมกันมีดังต่อไปนี้

1) ฟังก์ชันซิกมอยด์ (Sigmoid Function)

เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง 0 ถึง 1 ฟังก์ชันซิกมอยด์สามารถเขียนแทนด้วย σ ซึ่งคำนวณได้จากสมการ (8)

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (8)$$

2) ฟังก์ชันแทนเจต์ไฮเพอร์โบลิก (Hyperbolic Tangent Function)

เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง -1 ถึง 1 ฟังก์ชันแทนเจต์ไฮเพอร์โบลิกสามารถเขียนแทนด้วย *tanh* ซึ่งคำนวณได้จากสมการ (9)

$$tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \quad (9)$$

3) ฟังก์ชันค่าสูงสุดอย่างอ่อน (Softmax Function)

เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง 0 ถึง 1 ซึ่งสามารถเทียบได้กับความน่าจะเป็นของผลลัพธ์ กำหนดให้ชั้นของนิวรอลเน็ตเวิร์กที่ต้องการคำนวณมีผลลัพธ์ทั้งหมด k ตัว ค่าของผลลัพธ์นั้นแทนด้วยสัญลักษณ์ z จะได้ว่า ฟังก์ชันค่าสูงสุดอย่างอ่อนของผลลัพธ์ตัวที่ *j* หรือ แทนด้วยสัญลักษณ์ f_j ซึ่งคำนวณได้จากสมการ (10)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$$f(z)_j = \frac{e^{z_j}}{\sum_{i=1}^{k} e^{z_j}} \quad (10)$$

4) ฟังก์ชันเรคติไฟต์เชิงเส้น (Rectified Linear Unit function หรือ ReLU)

เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาเป็นค่ามากกว่าหรือเท่ากับศูนย์เสมอ สมการฟังก์ชันเรคติไฟต์เชิงเส้นสามารถแทนด้วยสัญลักษณ์ *f* ซึ่งคำนวณได้จากสมการ (11)

$$f(z) = \begin{cases} 0 & if \quad z < 0 \\ z & if \quad z \geq 0 \end{cases} \quad (11)$$

5) ฟังก์ชันขีดแบ่ง (Threshold Function)

เป็นฟังก์ชันรูปทั่วไปของฟังก์ชันเรคติไฟต์เชิงเส้น โดยจะมีค่าขีดแบ่ง t ตามที่กำหนด สมการของฟังก์ชันขีดแบ่ง สามารถแทนด้วยสัญลักษณ์ f ซึ่งคำนวณได้จากสมการ (12)

$$f(z,t)=\begin{cases}0 \ \ if \ \ z<t \\ z \ \ if \ \ z\geq t\end{cases} \tag{12}$$

2.2.4 ฟังก์ชันต้นทุน/วัตถุประสงค์ (Cost function หรือ Objective function)

เป็นฟังก์ชันที่แสดงถึงต้นทุนของนิวรอลเน็ตเวิร์ก กล่าวคือในกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์กนั้น จะทำการปรับค่าน้ำหนักเพื่อที่จะลดค่าของฟังก์ชันต้นทุน ฟังก์ชันต้นทุนที่เป็นที่นิยมมีดังต่อไปนี้ กำหนดให้ J แทนฟังก์ชันต้นทุน n คือ จำนวนข้อมูลทั้งหมดที่ใช้ในการเรียนรู้ y_i แทนผลลัพธ์จริงที่ได้จากชุดข้อมูลที่ i และ $\hat{\mathrm{y}}_i$ แทนผลลัพธ์ที่ทำนายได้จากข้อมูลชุดที่ i

1) ค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Squared Error หรือ MSE)

$$J=\frac{1}{n}\sum_{i=1}^{n}(\hat{\mathrm{y}}_i-\mathrm{y}_i)^2 \tag{13}$$

2) ค่าเฉลี่ยครอสเอนโทรปีแบบทวิภาค (Binary Cross-entropy)

$$J=-\frac{1}{n}\sum_{i=1}^{n}\mathrm{y}_i\log(\hat{\mathrm{y}}_i)+(1-\mathrm{y}_i)\log(1-\mathrm{y}_i) \tag{14}$$

3) ค่าลบลอการิทึมของความเป็นไปได้ (Negative Log-Likelihood)

$$J=-\frac{1}{n}\sum_{i=1}^{n}\mathrm{y}_i\log(\hat{\mathrm{y}}_i) \tag{15}$$

2.2.5 การหาค่าที่เหมาะสมที่สุด (Optimization)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

เป้าหมายของการเรียนรู้ของนิวรอลเน็ตเวิร์กคือการพยายามลดค่าของฟังก์ต้นทุนให้มีค่าน้อยที่สุด โดยใช้วิธีปรับปรุงน้ำหนักของเส้นเชื่อมในนิวรอลเน็ตเวิร์ก วิธีการปรับปรุงน้ำหนักของเส้นเชื่อมที่ได้รับความนิยมมีดังต่อไปนี้ [21]

1) สโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent หรือ SGD)

กำหนดให้ w แทนพารามิเตอร์ซึ่งเป็นน้ำหนักที่ต้องการจะปรับค่า α คือ อัตราการเรียนรู้ $\frac{\partial J_t}{\partial w}$ คือ เกรเดียนของฟังก์ชันต้นทุนเทียบกับ w การเรียนรู้ด้วยสโตแคสติกเกรเดียนเดสเซนท์จะมีการปรับค่าของ w ดังสมการ (16)

$$w_t=w_{t-1}-\alpha\frac{\partial J_t}{\partial w} \tag{16}$$

ปัญหาที่อาจจะเจอในการหว่างการเรียนรู้คือการติดอยู่ในโลคอลออพติมา (Local Optima) ดังนั้นจึงมีการนิยมใช้โมเมนตัม (Momentum) โดยมีจุดประสงค์เพื่อทำให้การเรียนรู้มีการลู่เข้าที่ดีขึ้นและ

หลีกเลี่ยงการติดอยู่ในโลคอลออพติมา กำหนดให้ v แทนค่าความเร็วซึ่งมีการปรับค่าพร้อมกับ w และ γ แทนค่าสัมประสิทธิ์ของโมเมนตัม (Momentum Coefficient) สามารถแสดงสมการในการเรียนรู้ได้ดังสมการ (17) และ (18)

$$v_t = \gamma v_{t-1} + \alpha \frac{\partial J_t}{\partial w} \tag{17}$$

$$w_t = w_{t-1} - v_t \tag{18}$$

2) วิธีเกรเดียนที่ปรับตัวได้ (Adaptive Gradient Method หรือ AdaGrad)

เป็นวิธีที่จะมีการปรับอัตราการเรียนรู้ได้ด้วยตัวเองจากค่าเริ่มต้นที่กำหนด สำหรับการปรับค่าของอัตราการเรียนรู้นั้นจะมีการนำค่าเกรเดียนในอดีตมาใช้ กำหนดให้ g_t แทนเกรเดียนที่เวลา t การเรียนรู้ด้วยวิธีเกรเดียนที่ปรับตัวได้แสดงดังสมการ (19) และ (20)

$$g_t = \frac{\partial J_t}{\partial w} \tag{19}$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\sum_{k=1}^{t} g_i^2}} g_t \tag{20}$$

3) อาร์เอ็มเอสพรอพ (RMSProp)

เป็นวิธีการที่มีการเก็บค่าเกรเดียนของครั้งก่อนหน้าไว้เพื่อที่จะนำมาใช้ในรอบของการเรียนรู้ปัจจุบันโดยการนำไปปรับปรุงอัตราส่วนของอัตราการเรียนรู้ โดยนอกเหนือจากการใช้ g_t แล้วยังมีการใช้ $MeanSquare_t$ สำหรับการเก็บค่าเฉลี่ยของเกรเดียน และให้ γ แทนอัตราการใช้เกรเดียนของอดีตในการเรียนรู้ ซึ่งปกติจะใช้ค่านี้ที่ 0.9 วิธีการอาร์เอ็มเอสพรอพแสดงดังสมการ (21) – (23)

$$g_t = \frac{\partial J_t}{\partial w} \tag{21}$$

$$MeanSquare_t = \gamma MeanSquare_{t-1} + (1-\gamma)\mathrm{g}_t^2 \tag{22}$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{MeanSquare_t}} g_t \tag{23}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2.6 การแพร่กระจายย้อยกลับและการเรียนรู้ (Back propagation and Training)

สำหรับขั้นตอนการป้อนไปข้างข้างหน้า การหาค่าความผิดพลาดของเพอร์เซ็ปตรอนสามารถทำได้โดยการคำนวณค่าเกรเดียนของฟังก์ชันต้นทุนเทียบกับค่าผลลัพธ์ในชั้นสุดท้าย แต่การหาค่าความผิดพลาดของเพอร์เซ็ปตรอนเพื่อใช้สำหรับการเรียนรู้ของลำดับชั้นก่อนหน้านั้นไม่สามารถหาได้โดยตรง จึงต้องอาศัยวิธีการแพร่กระจายย้อนกลับ [21]

กำหนดให้ δ_j^l แทนค่าความผิดพลาดของเพอร์เซ็ปตรอนตัวที่ *j* ในลำดับชั้นที่ *l* กำหนดให้ *J* แทนฟังก์ชันต้นทุน กำหนดให้ *z* เป็นค่าที่คำนวณได้ก่อนจะผ่านฟังก์ชันกระตุ้น *g* จะสามารถเขียนสมการของค่าความผิดพลาดได้ดังสมการ (24)

$$\delta_j^l = \frac{\partial J}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l}\frac{\partial a_j^l}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l} g'(z_j^l) \tag{24}$$

สำหรับการหาค่า $\frac{\partial J}{\partial a_j^l}$ นั้น ในลำดับชั้นสุดท้ายสามารถคำนวณหาได้โดยตรงจากฟังก์ชันต้นทุนที่เลือกใช้ ส่วนในลำดับชั้นก่อนหน้า จะต้องหาโดยวิธีการแพร่กระจายย้อนกลับ โดยจะทำคล้ายกับการป้อนไปข้างหน้า เพียงแต่กลับทิศกันเท่านั้น โดยคำนวณได้จากสมการ (25)

$$\frac{\partial J}{\partial a_j^l} = \sum_{k=1}^{m} \frac{\partial J}{\partial z_k^{l+1}}\frac{\partial z_k^{l+1}}{\partial a_j^l} = \sum_{k=1}^{m} \delta_k^{l+1} w_{kj}^{l+1} \tag{25}$$

โดย *m* คือจำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ *l+1* จากนั้น เมื่อคำนวณค่าความผิดพลาดของแต่ละระดับชั้นได้ ก็สามารถหาค่าความผิดพลาดเทียบกับน้ำหนักและค่าไบแอสใด ๆ ได้จากสมการ (26) และ(27)

$$\frac{\partial J}{\partial w_{jk}^l} = \frac{\partial J}{\partial z_j^l}\frac{\partial z_j^l}{w_{jk}^l} = \delta_j^l a_k^{l-1} \tag{26}$$

$$\frac{\partial J}{\partial b_j^l} = \frac{\partial J}{\partial z_j^l}\frac{\partial z_j^l}{b_{jk}^l} = \delta_j^l \tag{27}$$

ในกรณีที่หาค่าที่เหมาะสมโดยวิธีใช้สโตแคสติกเกรเดียนเดสเซนท์ การปรับปรุงค่าน้ำหนัก w_{kj}^l จะสามารถทำได้โดยสมการ (28)

$$w_{kj,t}^l = w_{kj,t-1}^l - \alpha a_{k,t}^{l-1}\delta_{j,t}^l \tag{28}$$

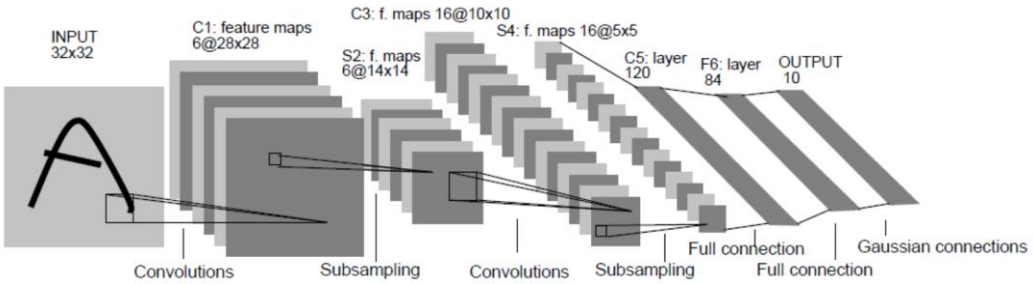
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3 นิวรอลเน็ตเวิร์กเชิงลึก (Deep Neural Network)

คือนิวรอลเน็ตเวิร์กที่มีชั้นซ่อน (Hidden Layer) จำนวนหลาย ๆ ชั้น ตัวอย่างเช่น เน็ตเวิร์กความเชื่อเชิงลึก (Deep Belief Network หรือ DBN) นิวรอลเน็ตเวิร์กแบบวกกลับ (Recurrent Neural Network หรือ RNN) หน่วยความระยะสั้นแบบยาว (Long-Short Term Memory หรือ LSTM) นิวรอลเน็ตเวิร์กคอนโวลูชัน (Convolutional Neural Network หรือ CNN) จุดเด่นของนิวรอลเน็ตเวิร์กเชิงลึกคือความสามารถในการเรียนรู้ฟีเจอร์ที่ซ่อนอยู่ในข้อมูลรับเข้าได้ดี ต่างกับนิวรอลเน็ตเวิร์กทั่ว ๆ ไปที่ต้องเลือกข้อมูลรับเข้าให้เหมาะสม (Feature Selection) หัวข้อถัดไปจะอธิบายถึงรายละเอียดของนิวรอลเน็ตเวิร์กเชิงลึกที่ใช้ในงานวิจัยนี้

2.3.1 นิวรอลเน็ตเวิร์กคอนโวลูชัน (Convolutional Neural Network หรือ CNN)

เป็นนิวรอลเน็ตเวิร์กเชิงลึกที่มีจุดเริ่มต้นจากงานวิจัยด้านการจำแนกภาพตัวอักษร [22] โดยใช้ตัวกรอง (Filter) เพื่อสร้างเป็นฟีเจอร์ใหม่ (Feature Map) เพื่อนำไปใช้เป็นข้อมูลรับเข้าของชั้นถัดไป โครงสร้างของนิวรอลเน็ตเวิร์กคอนโวลูชันแสดงได้ดังรูปที่ 3 ซึ่งเกิดจากนำชั้นหลาย ๆ ประเภทดังต่อไปนี้มาประกอบเข้าด้วยกัน ดังนี้

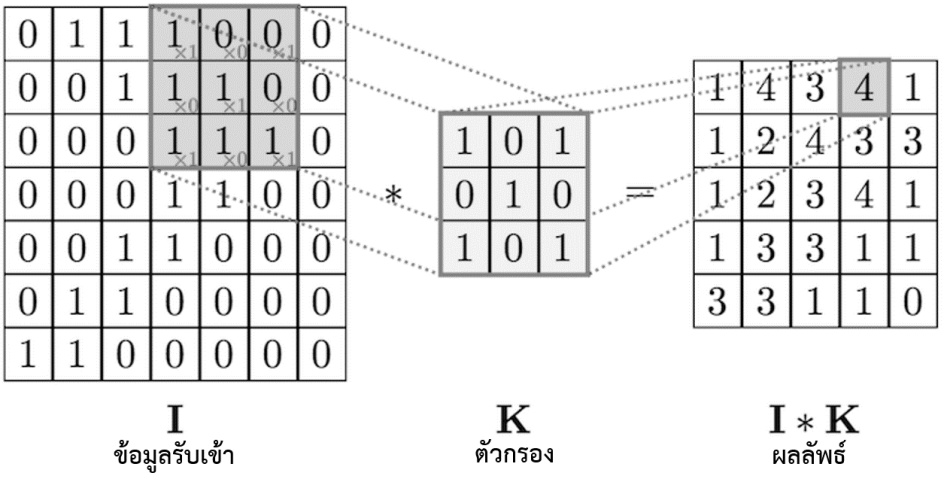


รูปที่ 3 โครงสร้างของนิวรอลเน็ตเวิร์กคอนโวลูชัน (อ้างอิงจาก Fig.2 ใน [22])

1) ชั้นคอนโวลูชัน (Convolutional Layer)

เป็นชั้นที่ทำการหาฟีเจอร์จากลุ่มของข้อมูลรับเข้าที่อยู่ใกล้ ๆ กันโดยใช้วิธีการดอทเมทริกซ์กับตัวกรองโดยที่น้ำหนักของตัวกรองจะใช้ร่วมกันในทุก ๆ การทำคอนโวลูชันของข้อมูลรับเข้า กำหนดให้ข้อมูลรับเข้าแทนด้วยเมทริกซ์ I และตัวกรองแทนด้วยเมทริกซ์ K ซึ่งมีขนาด $h \times w$ ผลลัพธ์ของการทำคอนโวลูชัน สามารถคำนวณได้จากสมการ (29)

$$\left(I * K\right)_{xy} = \sum_{i=1}^{h}\sum_{j=1}^{w} K_{ij} \cdot I_{x+i-1,y+j-1} \tag{29}$$



รูปที่ 4 ตัวอย่างการทำคอนโวลูชัน

(อ้างอิงจาก http://cs231n.github.io/convolutional-networks/)

ในชั้นคอนโวลูชัน มีองค์ประกอบที่ต้องพิจารณาดังต่อไปนี้

- ขนาดของตัวกรอง (Filter Size)

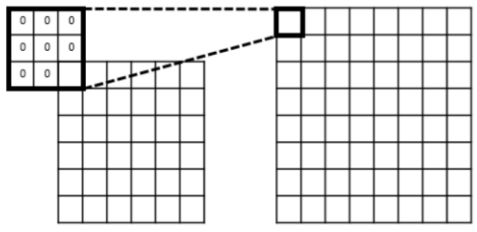
คือความกว้างและความสูงของตัวกรองที่จะนำมาใช้ในการทำคอนโวลูชัน (ค่า w และ h ในสมการที่ 29) ตัวอย่างการทำคอนโวลูชันในรูปที่ 4 นั้นใช้ตัวกรองที่มีขนาด 3x3

- ชนิดของการทำคอนโวลูชัน (Convolution Type)
    - คอนโวลูชันแบบแคบ (Narrow Convolution)

    โดยทั่วไปการทำคอนโวลูชันมักจะเป็นแบบแคบ ซึ่งในการทำคอนโวลูชัน ตัวกรองที่นำมาทำการดอทเมตริส์นั้นจะไม่มีการกระทำเลยขอบของเมตริกซ์ข้อมูลรับเข้า ส่งผลให้ผลลัพธ์ที่ได้จากการทำคอนโวลูชันของข้อมูลรับเข้าที่มีขนาด NxN กับตัวกรองที่มีขนาด MxM จะได้เมตริกซ์ขนาด (N-M+1)x(N-M+1) ตัวอย่างการทำคอนโวลูชันแบบแคบแสดงดังรูปที่ 4

    - คอนโวลูชันแบบกว้าง (Wide Convolution)

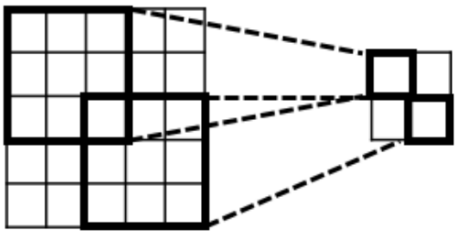
    เป็นการทำคอนโวลูชันที่มีการกระทำเลยขอบของเมตริกซ์ข้อมูลรับเข้าออกไป โดยที่พื้นที่ส่วนที่เกินออกไปนั้นจะมีการแทนค่าของข้อมูล ณ ช่องนั้น ๆ ด้วย 0 ซึ่งเรียกว่า การเสริมเติม (Padding) ผลลัพธ์ที่ได้จากการทำคอนโวลูชันของข้อมูลรับเข้าที่มีขนาด NxN กับตัวกรองที่มีขนาด MxM จะได้เมตริกซ์ขนาด (N+M-1)x(N+M-1) ทั้งนี้การทำคอนโวลูชันแบบกว้างมีขึ้นเพื่อป้องกันการสูญเสียข้อมูลตรงบริเวณขอบของข้อมูลรับเข้า ตัวอย่างการทำคอนโวลูชันแบบกว้างแสดงดังรูปที่ 5



รูปที่ 5 การทำคอนโวลูชันแบบกว้างและการเสริมเติม (อ้างอิงจาก Fig. 2.6 ใน [21])

- ขนาดของการก้าวข้าม (Stride Size)

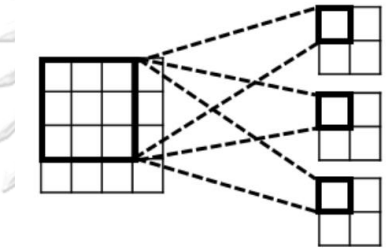
คือจำนวนช่องของข้อมูลรับเข้า ที่จะทำการเลื่อนไปเมื่อทำการหาผลลัพธ์ของการคอนโวลูชันในแต่ละช่อง โดยทั่วไปมักจะใช้ขนาดของการก้าวข้ามเป็น 1 ตัวอย่างการทำคอนโวลูชันที่มีขนาดของการก้าวข้ามเป็น 1 แสดงในรูปที่ 4 และรูปที่ 6 แสดงลักษณะของการทำคอนโวลูชันที่มีขนาดของการก้าวข้ามเป็น 2



รูปที่ 6 การทำคอนโวลูชันโดยมีข้อมูลรับเข้าขนาด 5x5 ตัวกรองขนาด 3x3 และมีขนาดของการก้าวข้ามเป็น 2 (อ้างอิงจาก Fig. 2.7 ใน [21])

- จำนวนตัวกรอง (Number of Filters)

การทำคอนโวลูชันในแต่ละชั้นสามารถมีตัวกรองได้มากกว่าหนึ่งตัว และน้ำหนักของแต่ละตัวกรองอาจต่างกันได้ ซึ่งการกำหนดจำนวนตัวกรองในชั้นคอนโวลูชันใด ๆ จะเป็นการกำหนดจำนวนช่องสัญญาณ (Channel) ของข้อมูลรับเข้าสำหรับชั้นถัดไป ดังรูปที่ 7



รูปที่ 7 การทำคอนโวลูชันโดยมีจำนวนตัวกล่องเท่ากับ3 (อ้างอิงจาก Fig. 2.8 ใน [21])

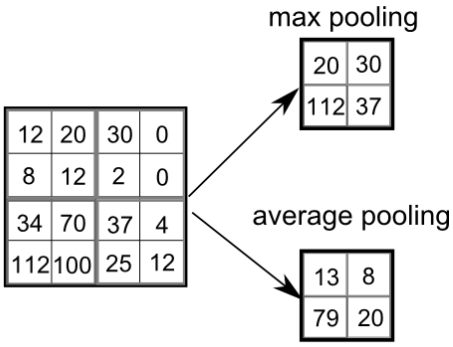
- จำนวนช่องสัญญาณ (Channel)

จำนวนช่องสัญญาณ หรือเรียกได้อีกอย่างว่า ความลึกของข้อมูลรับเข้า อาจจะมีค่ามากกว่าหนึ่งค่าได้ ตัวอย่างเช่น ในการวิจัยทางด้านรูปภาพที่มีการใช้ช่องสัญญาณทั้งหมด 3 ช่องแทนค่าของแม่สี หรือเกิดจากจำนวนของตัวกรองในชั้นคอนโวลูชันก่อนหน้า กำหนดใช้จำนวนช่องสัญญาณมีค่าเป็น k สามารถเขียนสมการในการคำนวณผลลัพธ์ของชั้นคอนโวลูชันได้ดังสมการต่อไปนี้

$$Z_{ij}^{l} = \sum_{c=0}^{k-1}\sum_{a=0}^{m-1}\sum_{b=0}^{m-1} w_{a,b}^{l} a_{c,i+a,j+b}^{l-1} + b^{l} \tag{30}$$

2) ชั้นการรวม (Pooling Layer)

ทำหน้าที่ลดขนาดของข้อมูล เพื่อให้เหลือเฉพาะข้อมูลที่สำคัญ ๆ เท่านั้น ซึ่งมักจะนิยมนำมาต่อกับชั้นคอนโวลูชัน โดยทั่วไปนิยมใช้การเลือกข้อมูลที่มีค่ามากที่สุด (Max Pooling) หรือ ค่าเฉลี่ย (Average Pooling) มาจากแต่ละช่วงของเมตริกซ์เพื่อสร้างเป็นเมทริกซ์ที่มีขนาดเล็ก ดังรูปที่ 8

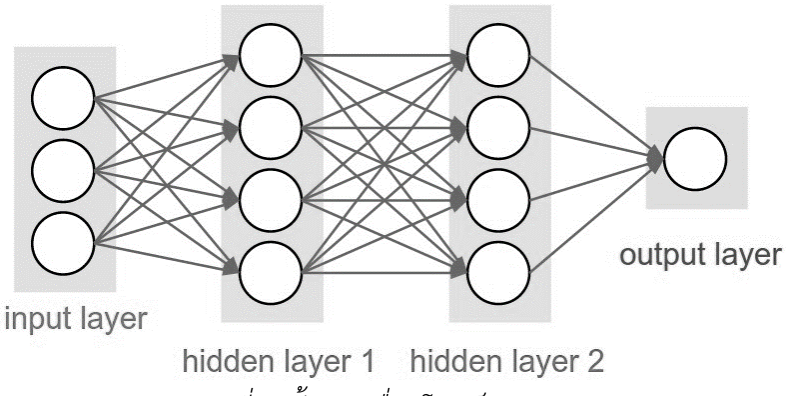


รูปที่ 8 ตัวอย่างชั้นการรวมโดยค่าที่มากที่สุดและค่าเฉลี่ย

(อ้างอิงจาก http://cs231n.github.io/convolutional-networks/)

3) ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer)

หลังจากการประกอบกันของชั้นคอนโวลูชันและชั้นการรวมจำนวนหนึ่งแล้ว ในชั้นสุดท้ายของนิวรอลเน็ตเวิร์กคอนโวลูชันจะเป็นการเชื่อมโยงเต็มรูปแบบ นั่นคือ ในชั้นนี้จะประกอบด้วยชั้นย่อย ๆ ที่มีเพอร์เซ็ปตรอนอยู่จำนวนหนึ่ง โดยที่เพอร์เซ็ปตรอนแต่ละตัว จะมีเส้นเชื่อมกันเพอร์เซ็ปตรอนทุกตัวในชั้นก่อนหน้าและเพอร์เซ็ฟตรอน ทุกตัวในชั้นถัดไป ทำให้สามารถทำการคำนวณการป้อนไปข้างหน้าและการแพร่กระจายย้อนกลับได้ด้วยวิธีการปกติได้ ชั้นการเชื่อมโยงเต็มรูปแบบแสดงดังรูปที่ 9

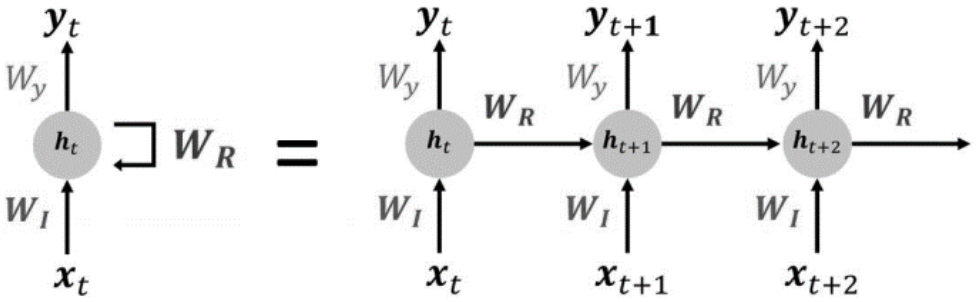


รูปที่ 9 ชั้นการเชื่อมโยงเต็มรูปแบบ

(อ้างอิงจาก http://cs231n.github.io/convolutional-networks/)

2.3.2 นิวรอลเน็ตเวิร์กแบบวกกลับ (Recurrent Neural Network)

เป็นนิวรอลเน็ตเวิร์กที่ถูกออกแบบมาเพื่อใช้สำหรับปัญหาที่เป็นลำดับ (Sequence) โดยอาศัยการเรียนรู้ซึ่งพึ่งพาข้อมูลรับเข้าในอดีตในระยะยาว (Long-term Dependencies) โครงสร้างของนิวรอลเน็ตเวิร์กแบบวกกลับแสดงดังรูปที่ 10



รูปที่ 10 โครงสร้างของนิวรอลเน็ตเวิร์กแบบวกกลับ

(อ้างอิงจาก http://colah.github.io/posts/2015-08-Understanding-LSTMs/)

กำหนดให้ x_t แทนข้อมูลรับเข้า ณ ลำดับที่ t ของชุดข้อมูลใด ๆ W_I แทนค่าน้ำหนักสำหรับข้อมูลรับเข้า W_R แทนค่าน้ำหนักวกกลับ (Recurrent Weight) ซึ่งใช้ร่วมกันในทุก ๆ ลำดับของข้อมูล W_y แทนน้ำหนักของข้อมูลออก h_t แทนสถานะซ่อน (Hidden State) ณ จุดข้อมูลลำดับ t และ y_t แทนข้อมูลออกหลังจากรับข้อมูลรับเข้า ณ ลำดับที่ t

ผลลัพธ์ที่ได้จากข้อมูลรับเข้าในแต่ละลำดับข้อมูลสามารถคำนวณได้ด้วยสมการ (31) และ (32) โดย σ_h แทนฟังก์ชันกระตุ้นในขั้นตอนการคำนวณสถานะซ่อน และ σ_y แทนฟังก์ชันกระตุ้นในขั้นตอนการคำนวณข้อมูลออก ณ ลำดับใด ๆ

$$h_t = \sigma_h(W_I x_t + W_R h_{t-1}) \tag{31}$$

$$y_t = \sigma_y(W_y h_t + b_y) \tag{32}$$

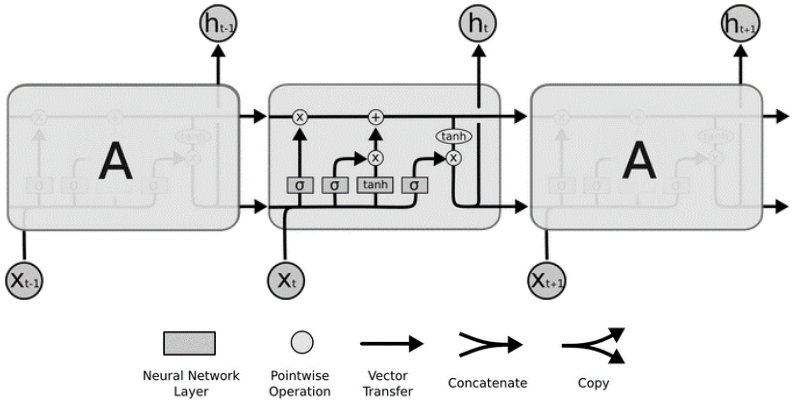
การฝึกสอนนิวรอลเน็ตเวิร์กแบบวกกลับ จะใช้วิธีการแพร่กระจายย้อนกลับตามเวลา (Back Propagation Through Time หรือ BPTT) เพื่อปรับน้ำหนักต่าง ๆ ของนิวรอลเน็ตเวิร์ก ซึ่งวิธีการนี้อาจจะเกิดปัญหา หากความยาวของข้อมูลรับเข้ามีมากเกินไป เนื่องจาก W_R ซึ่งแทนค่าน้ำหนักวกกลับ จะถูกส่งต่อไปยังชั้นสถานะซ่อนถัด ๆ ไป การแพร่กระจายย้อนกลับซึ่งอาศัยกฎลูกโซ่ (Chain Rule) เพื่อใช้ในการปรับน้ำหนัก อาจจะทำให้เกิดปัญหาเนื่องจากเกรเดียนของ W_R ซึ่งเกิดจากการคูณกันของลำดับก่อนหน้า ส่งผลให้ค่าที่ได้มีค่าเป็นศูนย์ (Vanishing Gradient) เมื่อนำหนักอยู่ในระหว่างช่วงศูนย์ถึงหนึ่ง หรืออาจจะเพิ่มมากขึ้นเกินไป (Exploding Gradient) เมื่อนำหนักมีค่ามากกว่าหนึ่ง

2.3.3 หน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory หรือ LSTM)

เป็นนิวรอลเน็ตเวิร์กแบบวกกลับที่ได้รับการปรับปรุงเพื่อแก้ปัญหาการลดหายหรือเพิ่มมากขึ้นของเกรเดียน โดยมีโครงสร้างดังรูปที่ 11 สัญลักษณ์กรอบสีเหลี่ยมผืนผ้าภายในหน่วยความจำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

หมายถึงนิวรอลเน็ตเวิร์กโดยที่สัญลักษณ์ภายในคือฟังก์ชันกระตุ้นที่ใช้ในแต่ละขั้นตอน ส่วนสัญลักษณ์วงกลมสื่อถึงการกระทำเฉพาะข้อมูลตำแหน่งที่ตรงกัน (Pointwise Operation)



รูปที่ 11 ตัวอย่างโครงสร้างของหน่วยความระยะสั้นแบบยาว

(อ้างอิงจาก http://colah.github.io/posts/2015-08-Understanding-LSTMs/)

หน่วยความจำนี้อาศัยสถานะเซลล์ (Cell State) ซึ่งแทนด้วยตัวแปร C_t ในการส่งผ่านข้อมูลเพื่อใช้เรียนรู้ข้อมูลในลำดับเวลาถัดไป เมื่อเซลล์ทำงานกับข้อมูลรับเข้า ณ ลำดับเวลานั้น ๆ เสร็จแล้วจะมีการเพิ่มและการลบข้อมูลออกสถานะเซลล์ผ่านประตูต่าง ๆ ภายในเซลล์ สำหรับการทำงานของเซลล์ในแต่ละรอบ จะรับข้อมูลดังต่อไปนี้ ข้อมูลรับเข้า ณ ลำดับเวลาปัจจุบัน (x_t) สถานะซ่อนในลำดับเวลาก่อนหน้า (h_{t-1}) สถานะเซลล์ในลำดับเวลาก่อนหน้า (C_{t-1}) และเมื่อทำงานเสร็จจะได้ผลลัพธ์เป็น สถานะเซลล์ ณ เวลาปัจจุบัน (C_t) และ สถานะซ่อน ณ เวลาปัจจุบัน (h_t) เพื่อนำมาใช้เป็นข้อมูลออก ณ เวลาปัจจุบัน โดยสามารถอธิบายการทำงานได้ด้วยสมการ (33) ถึง (38)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{33}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{34}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{35}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{36}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{37}$$

$$h_t = o_t * \tanh(C_t) \tag{38}$$

สมการที่ (33) เปรียบเสมือนประตูลืม (Forget Gate) ซึ่งผลลัพธ์ที่ได้จะบ่งชี้ว่าควรเก็บค่าของสถานะเซลล์ จากลำดับเวลาก่อนหน้าหรือไม่ สมการที่ (34) เปรียบเสมือนประตูรับข้า (Input Gate) ซึ่งทำหน้าที่ตัดสินใจว่าจะทำการปรับปรุงค่า ณ ตำแหน่งใดบ้าง โดยจะปรับค่าตามผลลัพธ์จากสมการที่ (36) นอกจากนี้ สมการที่ (37) เปรียบเสมือนประตูข้อมูลออก (Output Gate) ซึ่งทำหน้าที่ควบคุมปริมาณของข้อมูลที่ส่งต่อไปยังการทำงานในลำดับขั้นเวลาถัดไป

2.4 การวัดประสิทธิภาพ (Performance Evaluation)

2.4.1 ประสิทธิภาพในด้านการทำนายแนวโน้มของตลาดหุ้น

คือ การวัดประสิทธิภาพการจำแนกแบบสองคลาส (Binary-Class Classification) ซึ่งสามารถแสดงได้ดังนี้

2.4.1.1 คอนฟิวชันเมทริกซ์ (Confusion Matrix)

คือ เมทริกซ์ที่แสดงผลของการจำแนกโดยแจกแจงจำนวนที่จำแนกได้ตามคลาส ดังตัวอย่างใน ตารางที่ 1 ซึ่งแสดงการจำแนกข้อมูลเป็น 2 คลาส โดยค่าแต่ละแถวแสดงจำนวนข้อมูลที่มีคลาสนั้นเป็นคำตอบที่ถูกต้อง ส่วนค่าในแต่ละหลักแสดงจำนวนข้อมูลที่ทำนายได้คลาสนั้น กำหนดให้สำหรับคลาสใด ๆ

TP คือ จำนวนข้อมูลที่ทำนายได้คลาสที่หนึ่งและทำนายถูก (True Positive)

FP คือ จำนวนข้อมูลที่ทำนายได้คลาสที่หนึ่งและทำนายผิด (False Positive)

TN คือ จำนวนข้อมูลที่ทำนายได้คลาสที่สองและทำนายถูก (True Negative)

FN คือ จำนวนข้อมูลที่ทำนายได้คลาสที่สองและทำนายผิด (True Negative)

ตารางที่ 1 ตัวอย่าง คอนฟิวชันเมทริกซ์ของการจำแนกแบบ 2 คลาส

		คลาสที่ทำนาย	
		แนวโน้มขาขึ้น	**แนวโน้มขาลง**
คลาสจริง	**แนวโน้มขาขึ้น**	TP	FN
	แนวโน้มขาลง	FP	TN

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.4.1.2 ตัววัดประสิทธิภาพจำแนกตามคลาส

โดยทั่วไปตัววัดประสิทธิภาพที่นิยมใช้กันในงานวิจัยมีอยู่ 4 ค่า ดังนี้

- ค่าความเที่ยง (Precision) เป็นการวัดความแม่นยำของแบบจำลองโดยการพิจารณาแยกทีละคลาส ตัวอย่างเช่น การวัดว่าแบบจำลองทำนายว่าคำตอบที่เป็นบวกถูกต้องเท่าไหร่จากผลการทำนายคลาสบวกทั้งหมดเท่าไหร่

$$Precision = \frac{TP}{TP + FP} \tag{39}$$

- ค่าความระลึก (Recall) เป็นการวัดความถูกต้องของแบบจำลองโดยการพิจารณาแยกทีละคลาส ตัวอย่างเช่น การวัดว่าผลการทำนายคลาสบวกความถูกต้องเท่าไหร่เมื่อเทียบกับคลาสบวกจริงทั้งหมด

$$Recall = \frac{TP}{TP + FN} \tag{40}$$

- ค่าเอฟวัน (F1) เป็นการวัดความเที่ยงและความระลึกของแบบจำลองไปพร้อม ๆ กันโดยคำนวณได้จากสมการดังต่อไปนี้

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{41}$$

- ค่าความแม่นยำ (Accuracy) เป็นการวัดความแม่นยำของแบบจำลองโดยรวม กล่าวคือ แบบจำลองทำนายถูกกี่ครั้งจากจำนวนการทำนายทั้งหมด

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{42}$$

ในเบื้องต้นงานวิจัยนี้ใช้ค่าความแม่นยำในการวัดประสิทธิภาพของแบบจำลอง แต่อย่างไรก็ตามสำหรับแบบจำลองในขั้นตอนสุดท้าย จะเลือกใช้ค่าเอฟวันในการประเมินประสิทธิภาพของแบบจำลองในด้านความสามารถในการทำนายแนวโน้มของตลาดหุ้น เนื่องจากต้องการที่จะพิจารณาทั้งค่าความเที่ยงและค่าความระลึกไปพร้อม ๆ กัน

2.4.2 ประสิทธิภาพในด้านผลตอบแทน

ผลลัพธ์ที่ได้จากการทำนายแนวโน้มของตลาดหุ้น จะถูกนำมาทำการจำลองการซื้อ-ขาย (Trading Simulation) เพื่อวัดประสิทธิภาพในด้านผลตอบแทน สำหรับขั้นตอนการจำลองการซื้อ-ขายที่ใช้ในงานวิจัยนี้นั้น จะถูกอธิบายในหัวข้อ 5.2.2

2.4.2.1 ผลตอบแทนเฉลี่ยต่อปี (Annualized Return)

คือ การนำผลกำไร/ขาดทุน ที่ได้จากการจำลองการซื้อ-ขาย มาปรับให้เป็นค่าเฉลี่ยต่อปี เพื่อให้ง่ายต่อการเปรียบเทียบ ซึ่งสามารถคำนวณได้จากสมาการ (43) เมื่อ n คือ จำนวนวันทั้งหมดที่ทำการจำลองการซื้อขาย

$$\text{Annualized Return} = \left(\frac{\text{Final Balance}}{\text{Initial Balance}}\right)^{\frac{365}{n}} - 1 \tag{43}$$

2.4.2.2 ชาร์ปเรโช (Sharpe Ratio)

คือ การวัดผลตอบแทนเมื่อเทียบกับความเสี่ยง ซึ่งเป็นอัตราส่วนที่นิยมใช้มายาวนานและเป็นที่รู้จักกันมากที่สุด ชาร์ปเรโชได้ถูกนำเสนอในปี 1966 (ปรับปรุงในปี 1994) โดย William Forsyth Sharp [23] นักเศรษฐศาสตร์ที่ได้รับรางวัลโนเบล ดังแสดงในสมการที่ (44)

$$Sharpe\ Ratio = \frac{E[R - R_f]}{Var[R - R_f]} \tag{44}$$

โดยที่ R คืออัตราผลตอบแทนจากการลงทุนในเวลาช่วงหนึ่ง

R_f คืออัตราผลตอบแทนแบบไร้ความเสี่ยง (risk-free rate)

นอกจากชาร์ปเรโชจะถูกนำมาใช้เปรียบเทียบประสิทธิภาพของแบบจำลองแล้ว ในงานวิจัยนี้ยังได้นำชาร์เรโชที่ได้จากการทำการจำลองการซื้อ-ขาย มาใช้เป็นส่วนหนึ่งของวัตถุประสงค์ โดยมีเป้าหมายเพื่อเพิ่มประสิทธิภาพในด้านผลตอบแทนให้แก่แบบจำลอง สำหรับรายละเอียดนั้นจะถูกอธิบายหัวข้อ 5.2.2

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 3
งานวิจัยที่เกี่ยวข้อง

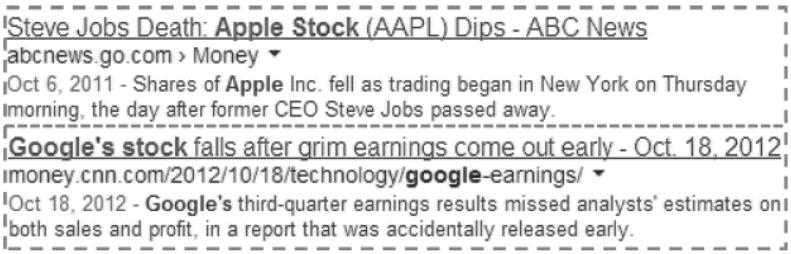
งานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้ คืองานวิจัยที่ทำนายพฤติกรรมของตลาดหุ้นโดยใช้แบบจำลองการเรียนรู้เชิงลึก โดยงานวิจัยแต่ละชิ้นจะมีโครงสร้างนิวรอลเน็ตเวิร์กที่แตกต่างกันรวมทั้งข้อมูลที่รับเข้ามาในเข้าแบบจำลอง ในหัวข้อนี้จะแบ่งงานวิจัยออกเป็น 2 กลุ่มได้แก่ 1) แบบจำลองการเรียนรู้เชิงลึกที่ใช้ข้อมูลรับเข้าเป็นข้อมูลเชิงตัวอักษรซึ่งได้แก่หัวข้อข่าว เป็นต้น และ 2) แบบจำลองการเรียนรู้เชิงลึกที่พิจารณาข้อมูลเชิงตัวอักษรร่วมกับข้อมูลเชิงตัวเลขซึ่งได้แก่ราคาในอดีตหรือตัวชี้วัดทางเทคนิค

3.1 แบบจำลองการเรียนรู้เชิงลึกที่ใช้ข้อมูลรับเข้าเป็นข้อมูลเชิงตัวอักษร

งานวิจัยที่นำข้อมูลเชิงตัวอักษรมาใช้นั้น เป็นงานวิจัยในกลุ่มแรก ๆ ที่เริ่มมีการใช้แบบจำลองการเรียนรู้เชิงลึก โดยงานวิจัยที่เป็นจุดเริ่มต้นมีดังต่อไปนี้

3.1.1 งานวิจัยของ Ding และคณะ ที่นำข้อมูลข่าวมาใช้ทำนายพฤติกรรมของราคาหุ้น

เป็นงานวิจัยแรกที่เริ่มในปี 2014 ซึ่งในงานวิจัยนี้ [14] ได้ตั้งสมมุติฐานว่า ข่าวต่าง ๆ มีผลต่อพฤติกรรมของตลาดหุ้นตัวอย่างเช่น ข่าวการเสียชีวิตของ Steve Job ส่งผลให้ราคาหุ้น Apple (AAPL) ของลดลง หรือ แม้กระทั่งข่าวผลประกอบการของบริษัท Google ที่แย่ในไตรมาสที่สามก็ส่งผลต่อราคาหุ้นของบริษัท



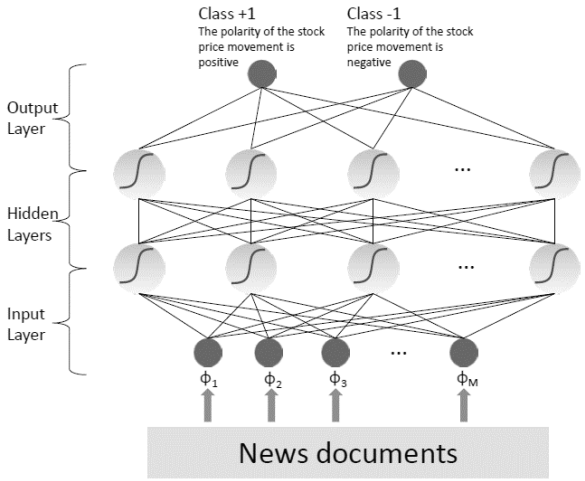
รูปที่ 12 ตัวอย่างข่าว (อ้างอิงจาก Fig.1 ใน [14])

ในงานวิจัยนี้จะพยายามสร้างตัวแทนของเหตุการณ์ (Event Representation) ซึ่งสกัดมาจากข่าวในแต่ละวัน โดยใช้กระบวนการที่ชื่อว่า การสกัดสารสนเทศแบบเปิด (Open Information Extraction) ซึ่งจะแปลงข่าวให้อยู่ในรูปของข้อมูลตัวแทนสามค่าซึ่งได้แก่ แอคเตอร์ (Actor) แอคชัน (Action) และ ออปเจ็ค (Object) ตัวอย่างเช่นข่าวดังต่อไปนี้

“Microsoft agrees to buy Nokia’s mobile phone business for $7.2 billion”

เมื่อนำมาเข้ากระบวนการสกัดสารสนเทศแบบเปิดจะได้ว่า แอคเตอร์ คือ “Microsoft” แอคชัน คือ “buy” และ ออปเจ็ค คือ “Nokia’s mobile phone business”

งานวิจัยนี้นำแบบจำลองนิวรอลเน็ตเวิร์กเชิงลึกมาใช้ในการทำนายราคาหุ้นในช่วงเวลาถัดไป โดยใช้ตัวข้อมูลตัวแทนของเหตุการณ์ที่ได้จากข่าวในแต่ละวัน โครงสร้างของแบบจำลองแสดงดังรูปที่ 13



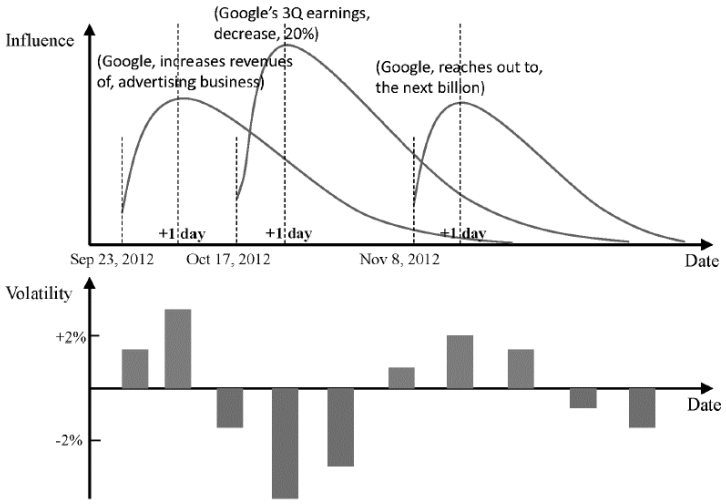
รูปที่ 13 โครงสร้างของแบบจำลองนิวรอลเน็ตเวิร์กเชิงลึก (อ้างอิงจาก Fig.2 ใน [14])

งานวิจัยนี้ได้ทำการทดลองใช้มูลข่าวในแต่ละวัน เพื่อทำนายการเปลี่ยนแปลงของราคาหุ้นในช่วงเวลา 1 วันถัดมา 1 สัปดาห์ถัดมา และ 1 เดือนถัดมา โดยผลการทดลองพบว่าการทำนายการเปลี่ยนแปลงในช่วง 1 วันถัดมาสามารถทำได้ดีที่สุด (ความแม่นยำ 59.6%) เนื่องจากโดยทั่วไปอิทธิพลของข่าวจะมีต่อการเปลี่ยนแปลงของราคาหุ้นในช่วงระยะเวลาสั้น ๆ เท่านั้น นอกจากนี้ในงานวิจัยฉบับนี้ยังได้ทดสอบเปรียบเทียบผลของการใช้หัวข้อข่าวกับการใช้เนื้อข่าว ซึ่งผลการทดลองพบว่าการใช้หัวข้อข่าวเพียงอย่างเดียวนั้นได้ผลลัพธ์ที่ดีกว่าประมาณ 4.95% เนื่องจากหัวข้อข่าวมักจะเป็นข้อความที่สามารถสื่อถึงใจความโดยภาพรวมของเนื้อหาข่าวดังนี้จึงสามารถนำมาสร้างเป็นเหตุการณ์ตัวแทนได้ง่ายกว่า

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

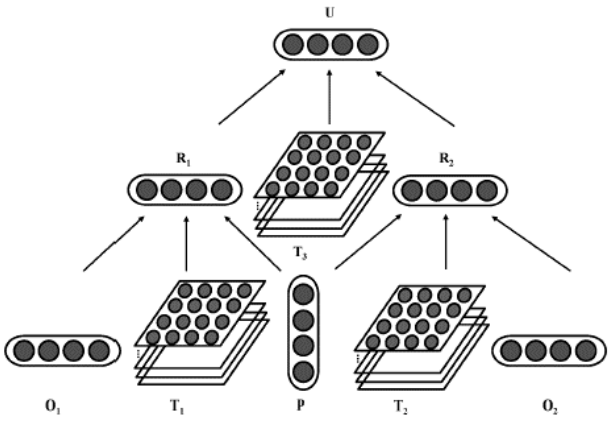
3.1.2 งานวิจัยของ Ding และคณะ ซึ่งปรับปรุงวิธีการใช้หัวข้อข่าว

เป็นงานวิจัย [15] ที่ต่อยอดจากงานวิจัยก่อนหน้า [14] โดยเพิ่มสมมุติฐานเพิ่มเติมว่าข่าวต่าง ๆ จะมีอิทธิพลต่อการเปลี่ยนแปลงของราคาหุ้นเพียงแค่ช่วงเวลาใด เวลาหนึ่งเท่านั้น ดังนั้น การเปลี่ยนของราคาหุ้นในแต่ละวันจึงได้รับอิทธิพลจากข่าวในช่วงเวลาก่อนหน้า ดังแสดงตัวอย่างในรูปที่ 14



รูปที่ 14 อิทธิพลของข่าวในช่วงระยะเวลาต่าง ๆ (อ้างอิงจาก Fig.1 ใน [15])

นอกจากนี้ในงานวิจัยได้ทำการปรับปรุงวิธีการนำหัวข้อข่าวมาใช้ ซึ่งจากเดิมใช้ แอคเตอร์ (Actor, O1) แอคชัน (Action, P) และ ออปเจ็ค (Object, O2) ที่ได้จากกระบวนการสกัดสารสนเทศแบบเปิดมาเป็นวิธีการเหตุการณ์ฝังตัว (Event Embedding) ซึ่งเป็นกระบวนการที่ใช้สร้างเวกเตอร์ของเหตุการณ์โดยใช้นิวรอลเท็นเซอเน็ตเวิร์ก (Neural Tensor Network) ซึ่งมีโครงสร้างดังแสดงในรูปที่ 15



รูปที่ 15 โครงสร้างของนิวรอลเท็นเซอเน็ตเวิร์ก (อ้างอิงจาก Fig.2 ใน [15])

เหตุการณ์ฝังตัวสมารถสร้างโดยใช้ฟังก์ชันต้นทุนที่ออกแบบมาเพื่อเปรียบเทียบการสร้างเหตุการณ์ฝังตัวโดยใช้เหตุการณ์ที่กำหนดกับเหตุการณ์ที่ทำการสุ่มเปลี่ยนแอคเตอร์ โดยมีสมมุติฐานว่าค่าที่ได้จากเหตุการณ์สุ่มจะต้องมีค่าน้อยกว่าค่าที่ได้จากเหตุการณ์จริง อัลกอริทึมที่ใช้แสดงดังรูปที่

Algorithm 1: Event Embedding Training Process

Input: $\mathcal{E} = (E_1, E_2, \cdots, E_n)$ a set of event tuples; the model $EELM$

Output: updated model $EELM'$

random replace the event argument and got the corrupted event tuple

$\mathcal{E}^r \leftarrow (E_1^r, E_2^r, \cdots, E_n^r)$

while $\mathcal{E} \neq [\,]$ **do**

  $loss \leftarrow max(0, 1 - f(E_i) + f(E_i^r) + \lambda \|\mathbf{\Phi}\|_2^2$

  if $loss > 0$ **then**

    $Update(\mathbf{\Phi})$

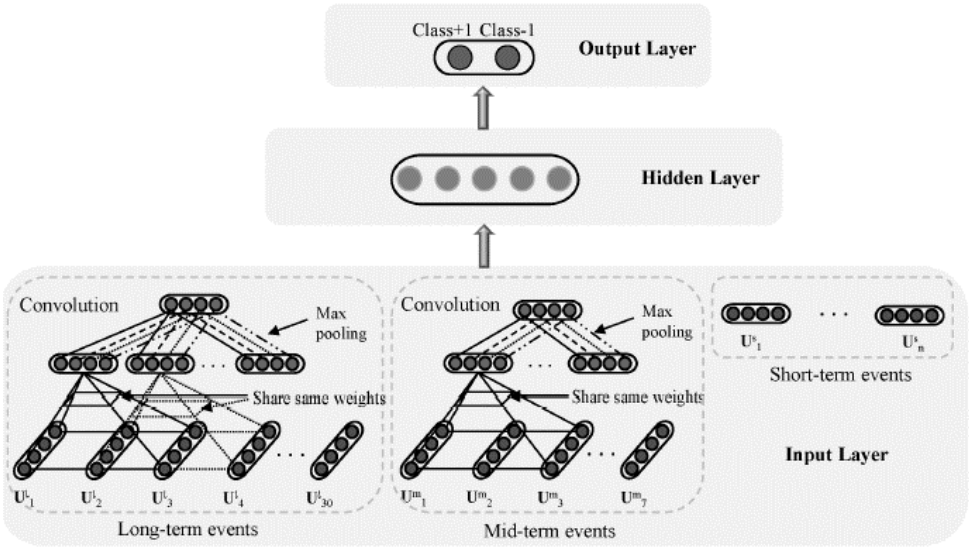
  else

    $\mathcal{E} \leftarrow \mathcal{E}/\{E_i\}$

return $EELM$

รูปที่ 16 อัลกอริทึมที่ใช้สร้างเหตุการณ์ฝังตัว (อ้างอิงจาก [15])

นอกจากนี้ในงานวิจัยนี้ยังทำการปรับปรุงโครงสร้างของแบบจำลอง โดยเลือกใช้นิวรอลเน็ตเวิร์กคอนโวลูชันและนำเหตุการณ์ฝังตัวที่ได้จากหัวข้อข่าวในช่วงระยะเวลา 7 และ 30 วันย้อนหลังมาใช้ร่วมกับเหตุการณ์ฝังตัวที่ได้จากหัวข้อข่าวในแต่วัน เพื่อให้แบบจำลองสามารถเรียนรู้อิทธิพลของข่าวที่เกิดขึ้นจากช่วงเวลาก่อนหน้าได้ โครงสร้างของแบบจำลองการเรียนรู้เชิงลึกที่ใช้แสดงดังรูปที่ 17 ด้วยวิธีการใหม่สามารถเพิ่มความแม่นยำในการทำนายบนชุดข้อมูลทดสอบ จาก 58.83% เป็น 64.21%



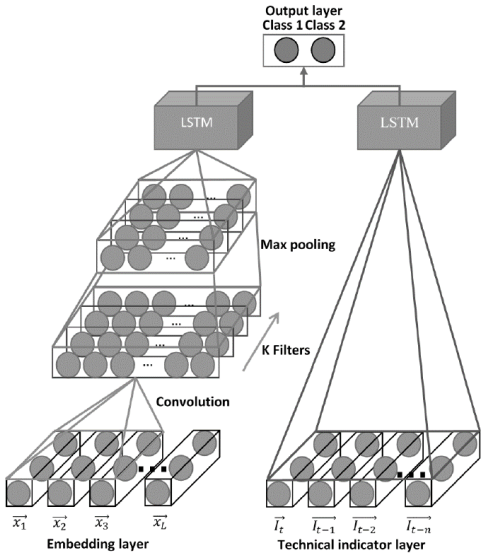
รูปที่ 17 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก (อ้างอิงจาก Fig.2 ใน [15])

3.2 แบบจำลองการเรียนรู้เชิงลึกที่พิจารณาข้อมูลเชิงตัวอักษรร่วมกับข้อมูลเชิงตัวเลข

การใช้แบบจำลองการเรียนรู้เชิงลึกโดยพิจารณาข้อมูลเชิงตัวอักษรร่วมกับข้อมูลเชิงตัวเลขเพื่อใช้ทำนายพฤติกรรมการเปลี่ยนแปลงของราคาหุ้นเป็นแนวทางการวิจัยที่เริ่มได้รับความความนิยมในช่วง 2 ปี ที่ผ่านมา ดังนั้นสำหรับวิทยานิพนธ์ฉบับนี้จึงมุ่งเน้นที่วิธีการนี้ ตัวอย่างงานวิจัยเกี่ยวข้องได้แก่

3.2.1 งานวิจัยของ Vargas และคณะ ที่นำหัวข้อข่าวมาใช้ร่วมกับตัวชี้วัดทางเทคนิค

เป็นงานวิจัย [17] ที่ใช้ข้อมูลหัวข้อข่าวชุดเดียวกับงานวิจัยของ Ding [15] และเพิ่มข้อมูลตัวชี้วัดทางเทคนิคที่สร้างจากข้อมูลของราคาหุ้นในอดีต ซึ่งงานวิจัยนี้ได้แยกข้อมูลนำเข้าเป็นสองส่วน คือ ส่วนที่เป็นหัวข้อข่าวซึ่งได้นำหัวข้อข่าวมาสร้างเป็นเวกเตอร์ของคำ โดยใช้วิธีการเวิร์ดทูเวก (word2vec) หลังจากนั้นจึงใช้นิวรอลเน็ตเวิร์กคอนโวลูชันเพื่อเรียนรู้ความสัมพันธ์ระหว่างข่าวและนำผลลัพธ์ที่ได้มาเรียนรู้ต่อด้วยหน่วยความระยะสั้นแบบยาว เนื่องจากเป็นแบบจำลองที่เหมาะกับข้อมูลที่เป็นอนุกรมเชิงเวลา สำหรับข้อมูลตัวชี้วัดทางเทคนิคนั้นก็จะนำข้อมูลเข้ามาเรียนรู้ด้วยหน่วยความระยะสั้นแบบยาวโดยตรง หลังจากนั้นนำผลลัพธ์ที่ได้จากข้อมูลทั้ง 2 ประเภทมารวมกันและนำมาใช้ทำนายพฤติกรรมการเปลี่ยนแปลงของราคาหุ้น โครงสร้างของแบบจำลองที่ใช้ในงานวิจัยนี้ แสดงดังรูปที่ 18



รูปที่ 18 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก (อ้างอิงจาก Fig.1 ใน [17])

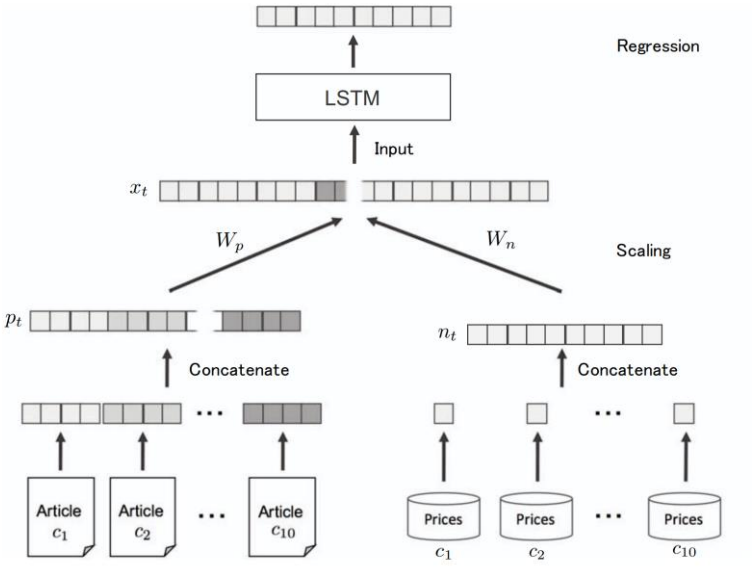
ผลการทดลองแสดงให้เห็นว่าการเพิ่มหน่วยความระยะสั้นแบบยาวและการพิจารณาข้อมูลตัวชี้วัดทางเทคนิคสามารถทำให้ประสิทธิภาพในการทำนายเพิ่มขึ้น แต่อย่างไรก็ตามความแม่นยำที่ได้จากแบบจำลองในงานวิจัยนี้ (62.03%) ยังมีค่าน้อยกว่าค่าที่ได้จากงานวิจัยของ Ding [15] (64.21%)

เนื่องจากนำหัวข้อข่าวมาใช้โดยตรง ซึ่งต่างกับงานวิจัยของ Ding ที่ทำการแปลงหัวข้อข่าวให้เป็นเหตุการณ์ฝังตัว

3.2.2 งานวิจัยของ Akita และคณะ ที่นำหัวข้อข่าวมาใช้ร่วมกับราคาหุ้นในอดีต

งานวิจัยชิ้นนี้ [16] ได้นำหัวข้อข่าวภาษาญี่ปุ่นมาพิจารณาร่วมกับข้อมูลราคาหุ้นในอดีตเพื่อใช้ทำนายพฤติกรรมของราคาหุ้นสำหรับวันถัดไป โดยในงานวิจัยนี้ได้เลือกใช้ วิธีการถดถอย (Regression) เพื่อใช้ทำนายราคาหุ้น ซึ่งแตกต่างกับงานวิจัยที่กล่าวถึงข้างต้น ที่ทำนายพฤติกรรมการเปลี่ยนแปลงโดยวิธีการจำแนกประเภท (Classification)

งานวิจัยนี้ได้ทำการศึกษาข้อมูลหุ้นใน 5 กลุ่มธุรกิจและเลือกใช้ข้อมูลเพียงแค่ 10 บริษัทสำหรับในแต่ละกลุ่มธุรกิจ สำหรับแบบจำลองการเรียนรู้เชิงลึกที่นำเสนอนั้นจะแบ่งข้อมูลเข้าเป็น 2 ส่วนคือ ข้อมูลหัวข้อข่าว ซึ่งนำข้อมูลหัวข้อข่าวในแต่ละวันของทั้ง 10 บริษัทมาทำการแปลงเป็นเวกเตอร์ด้วยวิธีพารากราฟเวกเตอร์ (Paragraph Vector) หลังจากนั้นนำเวกเตอร์ที่ได้มาเชื่อมต่อกัน (Concatenate) อีกส่วน คือ ข้อมูลราคาหุ้นในอดีตซึ่งจะนำข้อมูลของทั้ง 10 บริษัทมาเชื่อมต่อกันโดยตรง หลังจากนำข้อมูลทั้ง 2 ประเภทมาปรับขนาด (Scaling) แล้วนำมาเรียนรู้โดยใช้หน่วยความจำระยะสั้นแบบยาว เพื่อนำมาทำนายราคาสำหรับวันถัดไป โครงสร้างของแบบจำลองแสดงดังรูปที่ 19



รูปที่ 19 โครงสร้างของแบบจำลองการเรียนรู้เชิงลึก (อ้างอิงจาก Fig.1 ใน [16])

งานวิจัยนี้ทำการทดสอบประสิทธิภาพของแบบจำลองด้วยวิธีการจำลองการซื้อ-ขายหุ้นด้วยการจำลองตลาด (Market Simulation) และทำการตัดสินใจซื้อ-ขาย ด้วยค่าของราคาหุ้นที่ได้จากแบบจำลองการเรียนรู้เชิงลึก หลังจากนั้นจึงทำการเปรียบเทียบผลกำไร/ขาดทุนที่ได้

ผลการทดลองแสดงให้เห็นว่าการนำข้อมูลทั้ง 2 ประเภทมาพิจารณาร่วมกันสามารถสร้างผลตอบแทนได้มากกว่าการใช้ข้อมูลประเภทใด ประเภทหนึ่งเพียงอย่างเดียว และการแทนหัวข้อข่าวด้วยวิธีพารากราฟเวกเตอร์ก็ได้ผลลัพธ์ที่ดีกว่าการแทนข้อความด้วยวิธีถุงคำ นอกจากนี้การเรียนรู้ด้วยหน่วยความระยะสั้นแบบยาว ยังใช้ผลลัพธ์ที่ดีกว่าวิธีซัพพอร์ทเวกเตอร์แมชชีน มัลติเลเยอร์เพอร์เซ็ปตรอน และ นิวรอลเน็ตเวิร์กแบบวกกลับ

3.3 ประเด็นที่พบจากงานวิจัยก่อนหน้าและสิ่งที่นำมาปรับปรุงในงานวิจัยนี้

หัวข้อนี้จะกล่าวถึงประเด็นที่พบจากการศึกษางานวิจัยก่อนหน้า รวมทั้งแนวทางในการพัฒนาต่อ โดยมีรายละเอียด ดังต่อไปนี้

3.3.1 การแทนข้อความด้วยเวกเตอร์

สำหรับงานวิจัยที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าว การแทนที่ข้อความด้วยเวกเตอร์เป็นขั้นตอนที่ขาดไม่ได้เนื่องจากแบบจำลองต่าง ๆ ไม่สามารถเรียนรู้จากข้อความได้โดยตรง ดังนั้นจึงต้องมีการแทนข้อความด้วยเวกเตอร์เพื่อให้สามารถป้อนเข้าสู่แบบจำลองได้ วิธีการแทนข้อความที่ใช้ในงานวิจัยต่าง ๆ มีดังนี้ วิธีถุงคำ วิธีคำฝังตัว วิธีพารากราฟเวกเตอร์ และวิธีเหตุการณ์ฝังตัว

วิธีถุงคำและวิธีคำฝังตัวเป็นวิธีการที่ให้ผลลัพธ์ได้ไม่ดีนัก เนื่องจากวิธีการเหล่านี้พิจารณาเพียงแค่คำในหัวข้อข่าวแล้วทำการแปลงเป็นเวกเตอร์หลังจากนั้นจึงนำเวกเตอร์เหล่านั้นมาเรียงต่อกัน ต่างกับวิธีพารากราฟเวกเตอร์ซึ่งพิจารณาข้อความทั้งหมดเพื่อสร้างเป็นเวกเตอร์ตัวแทนของข้อความนั้น ๆ งานวิจัย [16] แสดงให้เห็นว่าการแทนหัวข้อข่าวด้วยวิธีพารากราฟเวกเตอร์ให้ผลลัพธ์ที่ดีกว่าวิธีถุงคำ สำหรับวิธีเหตุการณ์ฝังตัวที่ถูกนำเสนอในงานวิจัยของ Ding [15] ก็ให้ผลลัพธ์ที่ดีกว่าวิธีถุงคำและคำฝังตัว เนื่องจากวิธีการนี้สามารถสร้างเวกเตอร์ตัวแทนที่สามารถสื่อถึงใจความสำคัญของหัวข้อข่าวที่มีลักษณะคล้าย ๆ กันได้ดีกว่า ดังนั้นงานวิจัยนี้จึงเลือกใช้วิธีการเหตุการณ์ฝังตัวในแทนที่หัวข้อข่าวด้วยเวกเตอร์ของเหตุการณ์

3.3.2 ข้อมูลเชิงตัวเลขที่ป้อนเข้าสู่แบบจำลอง

ข้อมูลเชิงตัวเลขที่ถูกใช้ในงานวิจัยต่าง ๆ คือ ราคาในอดีตและตัวชี้วัดทางเทคนิคซึ่งเป็นฟีเจอร์ใหม่ที่ถูกสร้างขึ้นจากราคาในอดีต งานวิจัย [17] ได้นำเสนอแบบจำลองการเรียนรู้เชิงลึกที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าวและตัวชี้วัดทางเทคนิคจำนวน 7 ชนิดโดยอ้างอิงจากงานวิจัย [24] ซึ่งนำเสนอแบบจำลองเพื่อทำนายแนวโน้มของตลาดหุ้นโดยใช้วิธีซัพพอร์ทเวกเตอร์แมชชีน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ในเบื้องต้นงานวิจัยนี้จะเลือกใช้ตัวชี้วัดทางเทคนิคทั้ง 7 ชนิดเหมือนกับงานวิจัย [17] แต่จะทำการศึกษาเพิ่มเติมเพื่อเลือกใช้ตัวชี้วัดทางเทคนิคอื่น ๆ ตามที่ได้นำเสนอใน [18] เนื่องจากการใช้ตัวชี้วัดทางเทคนิคจำนวน 7 ชนิดนั้นเหมาะสมกับวิธีซัพพอร์ทเวกเตอร์แมชชีนซึ่งเป็นเทคนิคการเรียนรู้ของเครื่องที่จะต้องคำนึงถึงฟีเจอร์ที่จะป้อนเข้าสู่แบบจำลอง แต่สำหรับเทคนิคการเรียนรู้เชิงลึกซึ่งมีจุดเด่นในการเรียนรู้ฟีเจอร์ที่ซ่อนอยู่ในข้อมูลรับเข้า การใช้ตัวชี้วัดทางเทคนิคที่มากขึ้นก็อาจจะส่งผลให้แบบจำลองที่ได้มีประสิทธิภาพสูงขึ้นได้

3.3.3 การวัดประสิทธิภาพของแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองที่ใช้ในงานวิจัยมี 2 วิธีคือ การวัดค่าความแม่นยำในการทำนายของแบบจำลองและการวัดผลตอบแทนที่ได้จากการจำลองการซื้อ-ขายตามผลลัพธ์จากแบบจำลอง งานวิจัยส่วนใหญ่นิยมวัดประสิทธิภาพของแบบจำลองด้วยค่าความแม่นยำในการทำนาย แต่อย่างไรก็ตามแบบจำลองที่เน้นค่าความแม่นยำอาจจะไม่สามารถนำมาใช้ตัดสินใจซื้อ-ขายหุ้นเพื่อให้ได้ผลกำไรอย่างมีประสิทธิภาพ เนื่องจากการเปลี่ยนของราคาหุ้นในแต่ละวันมีค่าแตกต่างกัน การทำนายถูกในวันที่ราคาเปลี่ยนแปลงเล็กน้อยย่อมส่งผลให้ได้ผลตอบแทนน้อย ดังนั้นงานวิจัยนี้จึงมุ่งเน้นที่การใช้วิธีการวัดผลทั้ง 2 ประเภทร่วมกัน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
แนวคิดและแบบจำลองที่นำเสนอในเบื้องต้นโดยเน้นที่ความแม่นยำในการทำนาย

สำหรับในช่วงแรกของงานวิจัยนี้ ได้มีการนำเสนอเทคนิคการสร้างแบบจำลองด้วยวิธีการเรียนรู้เชิงลึกเพื่อทำนายแนวโน้มขอตลาดหุ้น โดยพิจารณาข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษรร่วมกัน โดยที่กระบวนฝึกสอนแบบจำลองจะมุ่งเน้นไปที่ความแม่นยำของแบบจำลองเท่านั้น แต่อย่างไรก็ตามเมื่อนำผลการทำนายไปทำการจำลองการซื้อขาย (Trading Simulation) พบว่าอัตราผลตอบแทนที่ได้ยังคงมีค่าที่ไม่ดีนัก ด้วยเหตุนี้วิธีการที่นำเสนอในขั้นตอนสุดท้ายของงานวิจัยนี้จึงได้เสนอวิธีที่ใช้แก้ไขปัญหาดังกล่าว ซึ่งการที่ใช้จะถูกอธิบายในบทถัดไป สำหรับบทนี้จึงเป็นเพียงแค่การนำเสนอแนวคิดและผลการทดลองในเบื้องต้นเท่านั้น

4.1 แนวคิดและวิธีการที่นำเสนอในเบื้องต้น

งานวิจัยนี้นำเอาจุดเด่นในงานวิจัยของ Ding [15] ซึ่งนำเอาเหตุการณ์ฝังตัว (Event Embedding) มาใช้เป็นตัวแทนของเหตุการณ์เพื่อช่วยประสิทธิภาพในการทำนาย และงานวิจัยของ Vargas [17] ซึ่งนำเอาตัวชี้วัดทางเทคนิค (Technical Indicator) มาใช้ร่วมกับหัวข้อข่าวเพื่อใช้ในการทำนายการเปลี่ยนแปลงของตลาดหุ้น โดยที่งานวิจัยนี้จะทำการสร้างแบบจำลองเพื่อทำนายแนวโน้มของตลาดหุ้นโดยใช้เหตุการณ์ฝังตัวซึ่งสกัดได้จากหัวข้อข่าวร่วมกับตัวชี้วัดทางเทคนิคซึ่งสร้างจากข้อมูลของราคาในอดีต

4.1.1 การเตรียมข้อมูล

หัวข้อนี้จะกล่าวถึงการเตรียมข้อมูลเพื่อนำไปใช้เป็นข้อมูลรับเข้าของแบบจำลองที่ใช้ทำนายแนวโน้มของตลาดหุ้นโดย ซึ่งมีดังต่อไปนี้

4.1.1.1 ตัวชี้วัดทางเทคนิค (Technical Indicator)

คือ ฟีเจอร์ใหม่ที่สร้างจากราคาในอดีต ในเบื้องต้นของงานวิจัยนี้ได้เลือกใช้ ตัวชี้วัดทางเทคนิคจำนวน 7 ชนิดที่ถูกนำเสนอในงานวิจัย [24] รายละเอียดการคำนวณแสดงดังรูปที่ 20

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Feature	Formula	Feature	Formula
Stochastic %K	$\frac{C_t - LL_n}{HH_n - LL_n}$	William's %R	$\frac{H_n - C_t}{H_n - L_n} \times 100$
Stochastic %D	$\frac{\sum_{i=0}^{n-1} \%K_{t-i}}{n}$	A/D Oscillator	$\frac{H_t - C_{t-1}}{H_t - L_t}$
Momentum	$C_t - C_{t-4}$	Disparity 5	$\frac{C_t}{MA_5} \times 100$
Rate of Change	$\frac{C_t}{C_{t-n}} \times 100$		

C_t is the closing price at day t, L_t is the lowest price at day t, H_t is the highest price at day t, MA_n is the moving average of the past n days, LL_n and HH_n is the lowest low and highest high in the past n days, respectively.

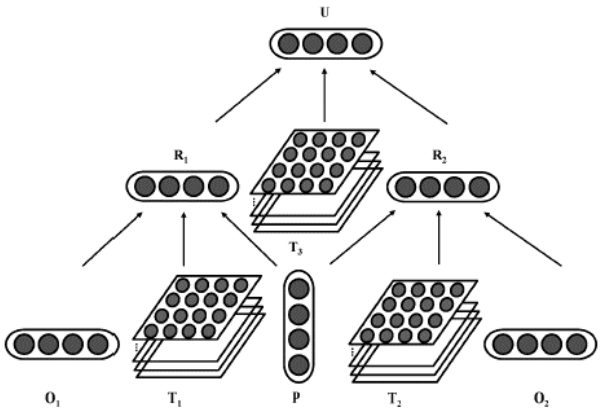
รูปที่ 20 ตัวชี้วัดทางเทคนิคที่ใช้งานวิจัย (อ้างอิงจาก Table 1 ใน [24])

4.1.1.2 ตัวแทนของเหตุการณ์ (Event Representation)

อาศัยแนวคิดของงานวิจัย [14] ที่พยายามสร้างตัวแทนของเหตุการณ์ (Event Representation) ซึ่งสกัดมาจากหัวข้อข่าวในแต่ละวัน โดยใช้กระบวนการที่ชื่อว่า การสกัดสารสนเทศแบบเปิดซึ่งจะแปลงข่าวให้อยู่ในรูปของข้อมูลตัวแทนสามค่าซึ่งได้แก่ แอคเตอร์ แอคชัน และ ออปเจ็ค สำหรับงานวิจัยนี้เลือกใช้ซอฟต์แวร์สำหรับกระบวนการสกัดสารสนเทศแบบเปิดซึ่งถูกพัฒนาโดยมหาวิทยาลัยสแตนฟอร์ด (Stanford Open Information Extraction) [25] ผลลัพธ์ที่ได้จากกระบวนการนี้จะถูกนำไปใช้ในการสร้างเหตุการณ์ฝังตัว

4.1.1.3 เหตุการณ์ฝังตัว (Event Embedding)

เป็นแนวคิดจากงานวิจัย [15] ซึ่งทำการสร้างเวกเตอร์เพื่อใช้แทนเหตุการณ์โดยใช้นิวรอลเท็นเซอเน็ตเวิร์ก ซึ่งมีโครงสร้างดังแสดงในรูปที่ 21



รูปที่ 21 โครงสร้างของนิวรอลเท็นเซอเน็ตเวิร์ก (อ้างอิงจาก Fig.2 ใน [15])

ข้อมูลรับเข้าสำหรับนิวรอลเท็นเซอเน็ตเวิร์ก คือ แอคเตอร์ (O_1) แอคชัน (P) และออปเจ็ค (O_2) จากหัวข้อ 4.1.1.2 แล้วนำไปแปลงเป็นเวกเตอร์คำเฉลี่ย (Average Word Vector) จากรูปที่ 21 เวกเตอร์ R_1 สามารถคำนวณได้จากสมการ (45) เมื่อกำหนดให้ f คือฟังก์ชันแทนเจต์ไฮเพอร์โบลิก k คือขนาดของเวกเตอร์รับเข้า และ W คือ เวกเตอร์น้ำหนักซึ่งมีขนาด k x 2k

$$R_1 = f(O_1^T \cdot T_1^{[1:k]} \cdot P + W \begin{bmatrix} O_1 \\ P \end{bmatrix} + b) \tag{45}$$

ค่า R_2 สามารถคำนวณได้โดยการแทนที่ O_1 ในสมการ (45) ด้วย O_2 สำหรับเวกเตอร์ U หรือเหตุการณ์ฝังตัว นั้นสามารถคำนวณได้จากสมการเดียวกันโดยใช้ R_1 และ R_2 แทน O_1 และ P

การฝึกสอนแบบจำลองที่ใช้สร้างเหตุการณ์ฝังตัวสามารถสร้างโดยใช้ฟังก์ชันต้นทุนที่ออกแบบมาเพื่อเปรียบเทียบการสร้างเหตุการณ์ฝังตัวโดยใช้เหตุการณ์ที่กำหนดกับเหตุการณ์ที่ทำการสุ่มเปลี่ยนแอคเตอร์ โดยมีสมมุติฐานว่าค่าที่ได้จากเหตุการณ์สุ่มจะต้องมีค่าน้อยกว่าค่าที่ได้จากเหตุการณ์จริง สมการของฟังก์ชันต้นทุนซึ่งอ้างอิงมาจากงานวิจัย [15] แสดงดังสมการ (46)

$$loss(E, E^r) = \max(0, 1 - f(E) - f(E^r)) + \lambda \|\Phi\|_2^2) \tag{46}$$

โดยที่ทูเพิลของเหตุการณ์ใด ๆ แทนด้วย E = (O_1, P, O_2) E^r คือ เหตุการณ์ใหม่ที่ถูกสร้างขึ้นด้วยการสุ่มเปลี่ยนแอคเตอร์ E^r = (O_1, P, O_2) และ ø = (T_1, T_2, T_3, W, b) คือ เซ็ตของพารามิเตอร์ทั้งหมด อัลกอริทึมที่ใช้ฝึกสอนแสดงดังรูปที่ 22

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Algorithm 1: Event Embedding Training Process

Input: $\mathcal{E} = (E_1, E_2, \cdots, E_n)$ a set of event tuples; the model $EELM$
Output: updated model $EELM'$

random replace the event argument and got the corrupted event tuple
$\mathcal{E}^r \leftarrow (E_1^r, E_2^r, \cdots, E_n^r)$
while $\mathcal{E} \neq [\,]$ **do**
  $loss \leftarrow max(0, 1 - f(E_i) + f(E_i^r) + \lambda \|\mathbf{\Phi}\|_2^2$
  if $loss > 0$ **then**
    $Update(\mathbf{\Phi})$
  else
    $\mathcal{E} \leftarrow \mathcal{E}/\{E_i\}$
return $EELM$

รูปที่ 22 อัลกอริทึมที่ใช้ฝึกสอนแบบจำลองสำหรับการสร้างเหตุการณ์ฝังตัว (อ้างอิงจาก [15])

4.1.2 แบบจำลองที่นำเสนอในเบื้องต้น

หัวข้อนี้จะกล่าวถึงรายละเอียดของแบบจำลองการเรียนรู้เชิงลึกที่สามารถรับข้อมูลเชิงตัวอักษรและข้อมูลเชิงตัวเลขเพื่อให้ทำนายแนวโน้มของตลาดหุ้น

4.1.2.1 แบบจำลองที่ใช้ทำนายแนวโน้มของตลาดหุ้น

งานวิจัยนี้ได้ทำการปรับปรุงแบบจำลองที่นำเสนอในงานวิจัย [15] โดยการเพิ่มข้อมูลราคาในอดีต (เปิด/ปิด/สูงสุด/ต่ำสุด) และตัวชี้วัดทางเทคนิคเพื่อช่วยเพิ่มประสิทธิภาพในการทำนายพฤติกรรมของตลาดหุ้น โดยที่โครงสร้างของแบบจำลองที่นำเสนอแสดงดังรูปที่ 23 สำหรับขั้นตอนการทำงานของแบบจำลองนั้น สามารถแบ่งออกได้เป็น 3 ส่วนดังนี้

1) การพิจารณาข้อมูลเชิงตัวอักษร

ข้อมูลรับเข้าในส่วนนี้คือ ข้อมูลเหตุการณ์ตัวแทนซึ่งสร้างจากค่าเฉลี่ยของเวกเตอร์เหตุการณ์ฝังตัวในแต่วัน สำหรับข้อมูลเหตุการณ์ตัวแทนจะถูกแบ่งออกเป็น 3 ส่วนคือ เวกเตอร์แทนเหตุการณ์ย้อนหลัง 30 วัน (ขนาด 30 x 100) เวกเตอร์แทนเหตุการณ์ย้อนหลัง 7 วัน (ขนาด 7 x 100) และเวกเตอร์แทนเหตุการณ์ในแต่ละวัน (ขนาด 1 x 100) โดยที่เวกเตอร์ของเหตุการณ์ย้อนหลัง 7 และ 30 วัน จะถูกป้อนเข้านิวรอลเน็ตเวิร์กคอนโวลูชันเพื่อสร้างฟีเจอร์แมพที่ใช้แทนเหตุการณ์ที่สำคัญในอดีต หลังจากนั้นนำเอาฟีเจอร์แมพที่สร้างจากเหตุการณ์ย้อนหลัง 7 และ 30 วัน มาเชื่อมต่อ กับเวกเตอร์ของเหตุการณ์ในแต่ละวัน แล้วป้อนเข้าสู่ชั้นซ่อนที่เชื่อมต่อกับโครงสร้างชั้นถัดไปที่มี เพอร์เซ็ปตรอน 2 ตัว ที่ใช้ฟังก์ชันกระตุ้นเป็นค่าสูงสุดอย่างอ่อน โดยมีจุดประสงค์เพื่อลดความซับซ้อนของผลลัพธ์และสร้างผลลัพธ์ที่ได้จากการใช้มูลเชิงตัวอักษร

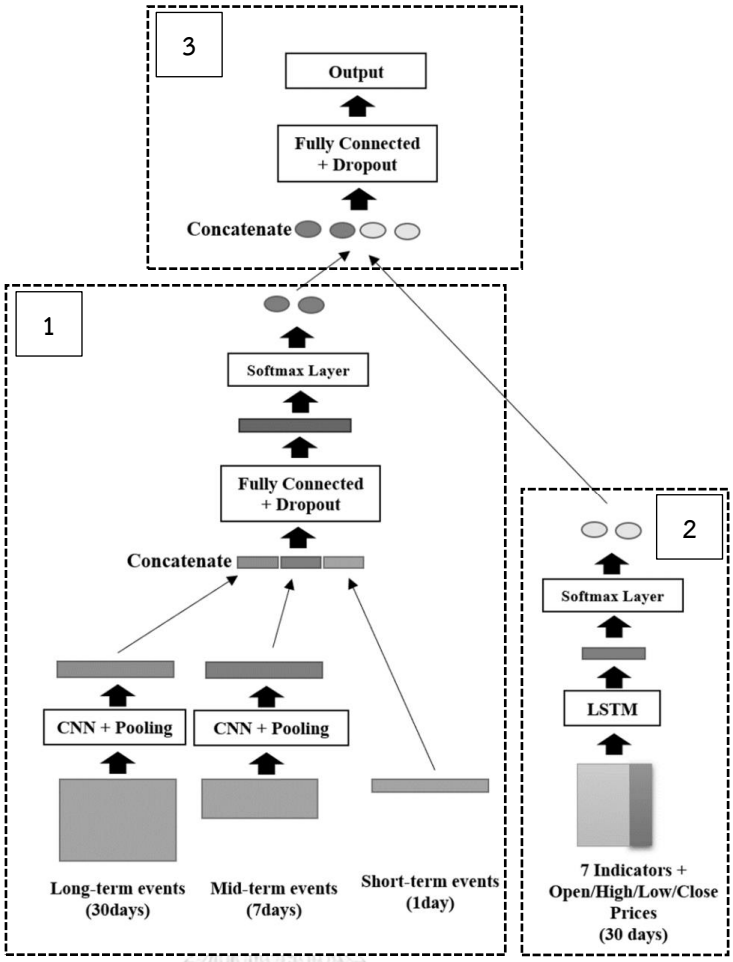
2) การพิจารณาข้อมูลเชิงตัวเลข

สำหรับข้อมูลรับเข้าส่วนที่สองคือ เวกเตอร์แทนราคาในอดีตและ 7 ตัวชี้วัดทางเทคนิคย้อนหลัง 30 วัน (ขนาด 11 x 30) ซึ่งจะถูกป้อนเข้า หน่วยความจำระยะสั้นแบบยาว ซึ่งมีจุดเด่นในการวิเคราะห์ข้อมูลอนุกรมเชิงเวลา หลังจากนั้นนำผลลัพธ์ที่ได้ไปเชื่อมต่อกับเพอร์เซ็ปตรอน 2 ตัว ที่ใช้ฟังก์ชันกระตุ้นเป็นค่าสูงสุดอย่างอ่อน โดยมีจุดประสงค์เพื่อลดความซับซ้อนของผลลัพธ์และสร้างผลลัพธ์จากข้อมูลเชิงตัวเลข

3) การนำผลลัพธ์จากข้อมูลทั้งสองประเภทมาพิจารณาร่วมกัน

คือ การนำผลลัพธ์จากเพอร์เซ็ปตรอนที่ได้จากพิจาราณาข้อมูลแต่ละประเภทมาเชื่อต่อกัน แล้วป้อนเข้าชั้นซ่อนอีกครั้งเพื่อทำนายผลลัพธ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 23 แบบจำลองที่นำเสนอ (ใช้ข้อมูลเหตุการณ์ฝังตัวร่วมกับตัวชี้วัดทางเทคนิค)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2 การทดลองและผลการทดลองเบื้องต้น

หัวข้อนี้จะกล่าวถึงการชุดข้อมูลที่ใช้ในการทดลอง แบบจำลองอื่น ๆ ที่ถูกสร้างขึ้นเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่นำเสนอในเบื้องต้น วิธีการวัดผล และผลการทดลองเบื้องต้น

4.2.1 ชุดข้อมูล

ชุดข้อมูลที่ใช้ในงานวิจัยนี้แบ่งออกเป็น 2 ประเภทดังนี้

4.2.1.1 ข้อมูลเชิงตัวเลข (Numerical Information)

คือ ข้อมูลของราคาหุ้นในอดีตซึ่งนำมาจากยาฮู! ไฟแนนซ์ (Yahoo! Finance) ข้อมูลที่นำมาใช้ได้แก่ ดัชนีเอสแอนด์พี 500 (Standard & Poor's 500 index) ซึ่งเป็นดัชนีที่ใช้เป็นตัวแทนของตลาดหุ้นอเมริกาที่สร้างขึ้นจากบริษัทที่มีมูลค่า (Market Capitalization) สูงสุด 500 อันดับแรก และ ดัชนีอุตสาหกรรมดาวโจนส์ (Dow Jones Industrial Average) ซึ่งเป็นตัวแทนของ 30 บริษัทชั้นนำในตลาดหุ้นอเมริกา

4.2.1.2 ข้อมูลเชิงตัวอักษร (Textual information)

คือ ข้อมูลหัวข้อข่าวในแต่ละวัน โดยงานวิจัยนี้จะนำเอาข้อมูลหัวข้อข่าวมาจาก 3 แหล่งข้อมูลดังต่อไปนี้

1. รอยเตอร์ (Reuters) เป็นข้อมูลหัวข้อข่าวทางด้านการเงินในอดีตตั้งแต่ 20/10/2006 ถึง 19/11/2013 ข้อมูลชุดนี้ได้มาจากงานวิจัย [14]

2. เรดดิต (Reddit) เป็นข้อมูลหัวข้อข่าวจำนวน 25 อันดับแรกที่ได้รับการโหวตสูงสุดจากยูสเซอร์ในเว็บไซต์เรดดิตในแต่ละวัน ตั้งแต่ 8/08/2008 ถึง 1/06/2016 ซึ่งข้อมูลชุดนี้ถูกจัดทำโดยยูสเซอร์ชื่อ Aaaron7sun จากเว็บไซต์แคกเกิ้ล (Kaggle) [26]

3. อินทรินิโอ (Intrinio) เป็นข้อมูลหัวข้อข่าวของแต่ละบริษัทที่ถูกนำมาคำนวณดัชนีอุตสาหกรรมดาวโจนส์ (จำนวน 30 บริษัท) ตั้งแต่ 1/08/2016 ถึง 12/12/2017 โดยที่ข้อมูลชุดนี้ได้มาจากบริษัทอินทรินิโอ ซึ่งเป็นผู้ขายข้อมูลทางด้านการเงิน (Financial Data Vendor)

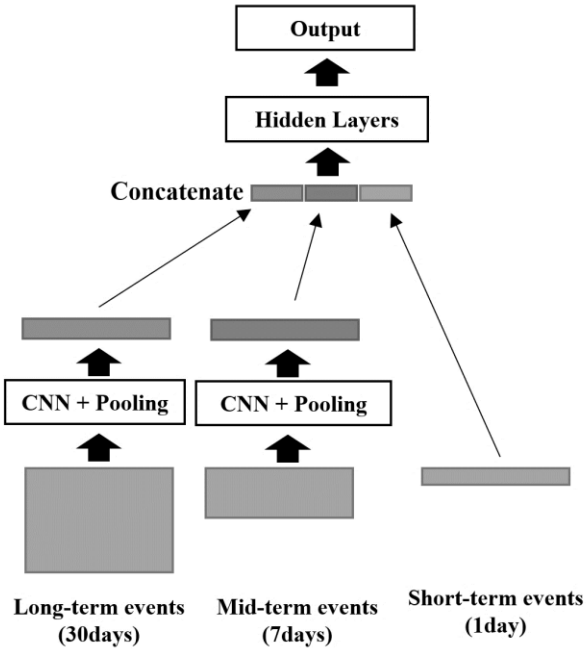
4.2.2 แบบจำลองอื่น ๆ เพื่อใช้ในการเปรียบเทียบประสิทธิภาพ

หัวข้อนี้จะกล่าวถึงแบบจำลองอื่น ๆ ที่ถูกสร้างขึ้นเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่นำเสนอ ซึ่งมีรายละเอียดดังต่อไปนี้

4.2.2.1 แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว

คือ แบบจำลองที่ถูกนำเสนอในงานวิจัย [15] ซึ่งใช้ข้อมูลรับเข้าเป็นข้อมูลเหตุการณ์ตัวแทนซึ่งสร้างจากค่าเฉลี่ยของเวกเตอร์เหตุการณ์ฝังตัวในแต่วัน สำหรับข้อมูลเหตุการณ์ตัวแทนจะถูกแบ่งออกเป็น 3 ส่วนคือ เวกเตอร์แทนเหตุการณ์ย้อนหลัง 30 วัน (ขนาด 30 x 100) เวกเตอร์แทนเหตุการณ์ย้อนหลัง 7 วัน (ขนาด 7 x 100) และเวกเตอร์แทนเหตุการณ์ในแต่ละวัน (ขนาด 1 x 100) โดยที่เวกเตอร์ของเหตุการณ์ย้อนหลัง 7 และ 30 วัน จะถูกป้อนเข้านิวรอลเน็ตเวิร์กคอนโวลูชันเพื่อสร้างฟีเจอร์แมพ ที่ใช้แทนเหตุการณ์ที่สำคัญในอดีต หลังจากนั้นนำเอาฟีเจอร์แมพที่สร้างจากเหตุการณ์ย้อนหลัง 7 และ 30 วัน มาเชื่อมต่อกับเวกเตอร์ของเหตุการณ์ในแต่ละวัน แล้วป้อนเข้าสู่ชั้นซ่อนเพื่อทำนายผลลัพธ์ โครงสร้างของแบบจำลองแสดงดังรูปที่ 24

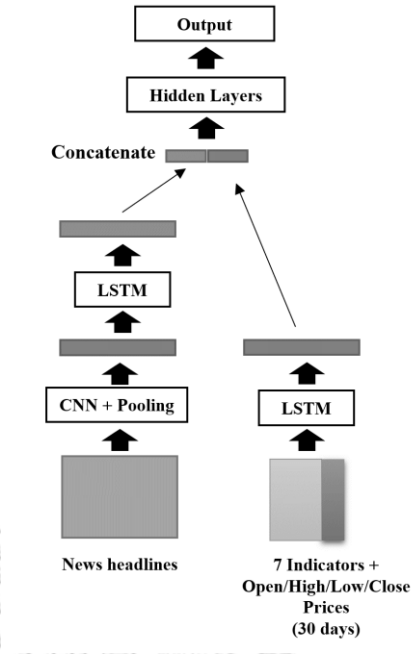
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 24 แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว

4.2.2.2 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค

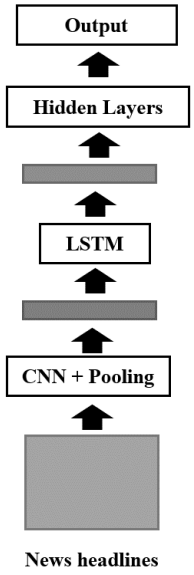
คือ แบบจำลองที่ถูกปรับปรุงจากงานวิจัย [17] โดยการเพิ่มข้อมูลราคาในอดีตเป็นข้อมูลรับเข้า แบบจำลองนี้ใช้ข้อมูลรับเข้า 2 ส่วนคือ เวกเตอร์ตัวแทนของหัวข้อข่าวซึ่งสร้างจากค่าเฉลี่ยของเวกเตอร์คำของแต่หัวข้อข่าว หลังจากนั้นจึงป้อนเข้าสู่นิวรอลเน็ตเวิร์กคอนโวลูชันเพื่อเรียนรู้ความสัมพันธ์ระหว่างข่าวและนำผลลัพธ์ที่ได้มาเรียนรู้ต่อด้วยหน่วยความระยะสั้นแบบยาว เนื่องจากเป็นแบบจำลองที่เหมาะกับข้อมูลที่เป็นอนุกรมเชิงเวลา สำหรับข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิคนั้น จะใช้เวกเตอร์ซึ่งแทนราคาเปิด/ปิด/สูงสุด/ต่ำสุดและ 7 ตัวชี้วัดทางเทคนิคย้อนหลัง 30 วัน (ขนาด 11 x 30) แล้วป้อนเข้าหน่วยความระยะสั้นแบบยาว หลังจากนั้นนำผลลัพธ์ที่ได้จากข้อมูลทั้ง 2 ประเภทมารวมกัน แล้วป้อนเข้าชั้นซ่อนเพื่อทำนายพฤติกรรมการเปลี่ยนแปลงของราคาหุ้น โครงสร้างของแบบจำลองนี้ แสดงดังรูปที่ 25



รูปที่ 25 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค

4.2.2.3 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวเพียงอย่างเดียว

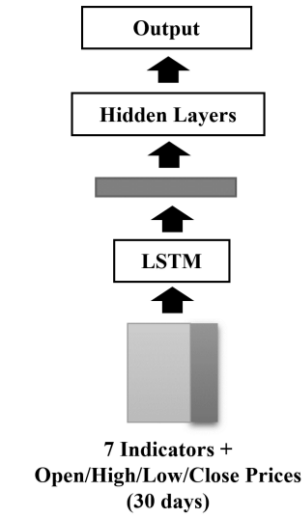
คือ แบบจำลองที่ดัดแปลงจากหัวข้อ 4.2.2.2 โดยทำการตัดข้อมูลรับเข้าที่เป็นราคาในอดีตและตัวชี้วัดทางเทคนิคออกไป โครงสร้างของแบบจำลองแสดงดังรูปที่ 26



รูปที่ 26 แบบจำลองที่ใช้ข้อมูลหัวข้อข่าว

4.2.2.4 แบบจำลองที่ใช้ข้อมูลราคาในอดีตและทางเทคนิค

คือ แบบจำลองที่ดัดแปลงจากหัวข้อ 4.2.2.2 โดยทำการตัดข้อมูลรับเข้าที่เป็นหัวข้อข่าวออกไป โครงสร้างของแบบจำลองแสดงดังรูปที่ 27



รูปที่ 27 แบบจำลองที่ใช้ข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิค

4.2.3 การวัดผล

ในเบื้องต้นของงานวิจัยนี้แบบจำลองจะถูกวัดประสิทธิภาพด้วยวิธีการดังต่อไปนี้

4.2.3.1 ความแม่นยำ (Accuracy)

ทำการวัดว่าแบบจำลองทำนายถูกกี่ครั้งจากจำนวนการทำนายทั้งหมด ค่าความแม่นยำสามารถคำนวณได้จากสมการดังต่อไปนี้

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{47}$$

4.2.3.2 การจำลองการซื้อขาย (Trading Simulation)

นำผลลัพธ์ที่ได้จากแบบจำลองไปทำการจำลองการซื้อขายหุ้นในแต่ละวัน โดยใช้เงื่อนไขดังนี้ เมื่อแบบจำลองทำนายว่าราคาในวันพรุ่งนี้จะเพิ่มขึ้น ก็จะทำการซื้อหุ้นที่ราคาเปิดตลาดเป็นจำนวน k หน่วย แล้วทำการขาย ณ ราคาปิดตลาดของแต่ละวัน หากแบบจำลองทำนายว่าราคาในวันพรุ่งนี้จะลดลง ก็จะทำตรงข้ามกับกรณีก่อนหน้า โดยที่ k สามารถคำนวณได้จากสัดส่วนของเงินที่มีอยู่เทียบราคาเปิดตลาด ดังสมการ (48)

$$k=\max(0, \frac{\text{Balance}}{\text{Open Price}}) \tag{48}$$

ผลกำไร/ขาดทุนจากซื้อขายในแต่ละวันจะถูกสะสมไว้ โดยที่ผลลัพธ์สุดท้ายจะถูกแสดงในรูปของผลตอบแทนเฉลี่ยต่อปี (Annualized Return) ซึ่งสามารถคำนวณได้จากสมาการ (49) เมื่อ n คือ จำนวนวันทั้งหมดที่ทำการจำลองการซื้อขาย

$$\text{Annualized Return} = \left(\frac{\text{Final Balance}}{\text{Initial Balance}}\right)^{\frac{365}{n}} - 1 \quad (49)$$

4.2.4 ผลการทดลองเบื้องต้น

ข้อมูลหัวข้อข่าวที่ใช้ในการทดลองนี้มี 3 ชุด ได้แก่ รอยเตอร์ เรดดิต (Reddit) และอินทรินิโอ สำหรับข้อมูลราคาหุ้นในอดีตนั้นจะใช้ดัชนีเอสแอนด์พี500 และดัชนีอุตสาหกรรมดาว โดยข้อมูลหัวข้อข่าวจากรอยเตอร์จะถูกจับคู่กับดัชนีเอสแอนด์พี500 และข้อมูลหัวข้อข่าวส่วนที่เหลือจะถูกใช้ร่วมกับดัชนีอุตสาหกรรมดาวโจนส์ ข้อมูลทั้ง 3 ชุดจะถูกแบ่งเป็น 3 ส่วน (Training/Validation/Test) และถูกป้อนเข้าแบบจำลองทั้ง 5 ประเภทดังที่ได้กล่าวในหัวข้อก่อนหน้า โดยที่แบบจำลองจะถูกฝึกสอนโดยใช้ข้อมูลสอน (Training Data) หลังจากนั้นจึงเลือกแบบจำลองที่มีความแม่นยำสูงสุดเมื่อทดสอบบนข้อมูลตรวจสอบ (Validation Data) มาใช้แล้วจึงทำการวัดประสิทธิภาพบนข้อมูลทดสอบ (Test Data)

4.2.4.1 ความแม่นยำ (Accuracy)

จากผลการทดลองที่แสดงในตารางที่ 2 เมื่อพิจารณาการใช้ข้อมูลเหตุการณ์ฝังตัวเทียบกับหัวข้อข่าวพบว่าการใช้เหตุการณ์ฝังตัวส่งผลให้แบบจำลองมีความแม่นยำมากขึ้นเมื่อทดสอบบนชุดข้อมูลที่ 1 และ 3 ซึ่งผลลัพธ์ที่ได้ก็สอดคล้องกับงานวิจัย [15] ที่ได้เปรียบเทียบการใช้ข้อมูลหัวข้อข่าวกับการใช้เหตุการณ์ฝังตัว สำหรับผลลัพธ์ในชุดข้อมูลที่ 2 ที่การใช้ข้อมูลเหตุการณ์ฝังตัวได้ผลลัพธ์ที่แย่กว่าการใช้หัวข้อข่าวเพียงอย่างเดียว สาเหตุเกิดจากลักษณะของข้อมูลหัวข้อข่าวที่ใช้คือ หัวข้อข่าวที่ได้รับการโหวตสูงสุดจากยูสเซอร์ในเว็บไซต์เรดดิตในแต่ละวันโดยไม่คำนึงถึงประเภทของหัวข้อข่าว ในขณะที่หัวข้อข่าวในชุดข้อมูลที่ 1 และ 3 เป็นหัวข้อข่าวทางด้านการเงินหรือข่าวที่เกี่ยวข้องกับบริษัทโดยตรง ดังนั้นการสร้างเหตุการณ์ฝังตัวโดยใช้ข้อมูลชุดที่ 2 จึงทำได้ไม่ดีนักและส่งผลให้ความแม่นยำของแบบจำลองมีค่าลดลง

สำหรับการใช้หัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิคพบว่าให้ผลลัพธ์ที่ดีกว่าเมื่อเทียบกับการใช้หัวข้อข่าวเพียงอย่างเดียวในทุก ๆ ชุดข้อมูล ซึ่งผลลัพธ์ที่ได้ก็สอดคล้องกับงานวิจัย [17] ที่ได้เสนอวิธีการนำหัวข้อข่าวมาใช้ร่วมกับตัวชี้วัดทางเทคนิค เมื่อพิจารณาการใช้มูลเหตุการณ์ฝังตัวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค เทียบกับการ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ใช้ข้อมูลเหตุการณ์เพียงอย่างเดียว พบว่าแบบจำลองมีความแม่นยำมากขึ้นในทุก ๆ ชุดข้อมูลที่นำมาทดสอบ

ตารางที่ 2 ค่าความแม่นยำของแบบจำลองเมื่อทดสอบด้วยชุดข้อมูลทดสอบ

ชุดข้อมูล / แบบจำลอง	[1] รอยเตอร์และดัชนีเอสแอนด์พี 500	[2] เรดดิตและดัชนีอุตสาหกรรมดาวโจนส์	[3] อินทรินิโอและดัชนีอุตสาหกรรมดาวโจนส์
1.แบบจำลองที่ใช้ข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิค	56.59%	**52.38%**	63.01%
2. แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวเพียงอย่างเดียว	58.14%	50.53%	63.01%
3. แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว	58.53%	48.94%	64.38%
4. แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค	59.69%	51.85%	65.75%
5. แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค (แบบจำลองที่นำเสนอในเบื้องต้น)	**62.02%**	50.26%	**69.86%**

4.2.4.2 การจำลองการซื้อขาย (Trading Simulation)

จากผลการทดลองที่แสดงในตารางที่ 3 ผลตอบแทนเฉลี่ยต่อปีมีค่าเป็นบวกในชุดข้อมูลที่ 1 ในขณะที่ข้อมูลชุดที่ 2 และ 3 ให้ผลลัพธ์เป็นค่าลบในทุก ๆ แบบจำลอง สาเหตุส่วนหนึ่งเกิดจากแบบจำลองที่นำมาทดสอบถูกเลือกจากแบบจำลองที่มีความแม่นยำมากที่สุดบนชุดข้อมูลตรวจสอบ โดยทั่วไปการเปลี่ยนแปลงของราคาดัชนีต่าง ๆ ในแต่ละวันมักจะมีค่าแตกต่างกัน แม้ว่าแบบจำลองที่ได้จะมีความแม่นยำสูง แต่อาจจะทำนายถูกเฉพาะวันที่ราคาเปลี่ยนแปลงเล็กน้อย หรือ ทายผิดในวันที่ราคาเปลี่ยนแปลงเยอะ ๆ ก็ส่งผลให้ผลตอบแทนเฉลี่ยที่ได้มีค่าติดลบได้ ดังเช่น ผลลัพธ์จากแบบจำลองที่ 5 บนชุดข้อมูลที่ 3 ซึ่งมีความแม่นยำ 69.86% แต่ได้ผลตอบแทนเฉลี่ยต่อปีเป็น -9.75% นอกจากนี้ค่าธรรมเนียมในการซื้อขาย (Transaction Cost) ที่เกิดขึ้น (การทดลองนี้ใช้ 0.1% ของมูลค่าที่ซื้อขายในแต่ละวัน) ก็อาจจะส่งผลให้ผลตอบแทนที่ได้มีค่าเป็นลบได้ เนื่องจากในขั้นตอนการจำลองการซื้อขายได้กำหนดให้ทำการซื้อขายทุก ๆ วัน ซึ่งอาจจะส่งผลให้ผลรวมของค่าธรรมเนียมที่เกิดขึ้นมีค่ามากกว่าผลกำไรที่เกิดขึ้นจากซื้อขาย

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ 3 ผลตอบแทนเฉลี่ยต่อปี (Annualized Return) เมื่อทดสอบด้วยชุดข้อมูลทดสอบ

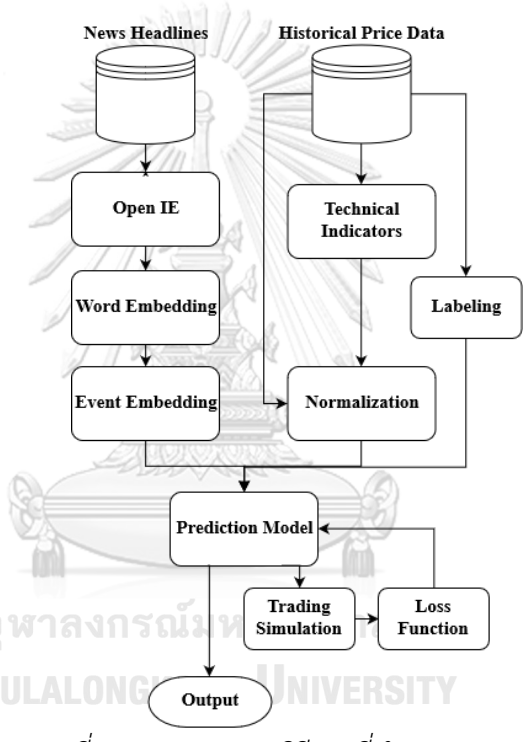
ชุดข้อมูล / แบบจำลอง	[1] รอยเตอร์และดัชนีเอสแอนด์พี 500	[2] เรดดิตและดัชนีอุตสาหกรรมดาวโจนส์	[3] อินทรินิโอและดัชนีอุตสาหกรรมดาวโจนส์
1.แบบจำลองที่ใช้ข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิค	4.04%	-23.78%	-12.07%
2. แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวเพียงอย่างเดียว	1.33%	-24.79%	-13.56%
3. แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวเพียงอย่างเดียว	2.37%	-34.45%	-10.20%
4. แบบจำลองที่ใช้ข้อมูลหัวข้อข่าวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค	10.02%	**-23.12%**	-11.05%
5. แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัวร่วมกับราคาในอดีตและตัวชี้วัดทางเทคนิค (แบบจำลองที่นำเสนอในเบื้องต้น)	**21.44%**	-33.76%	**-9.75%**

การทดลองนี้แสดงให้เห็นว่าการสร้างแบบจำลองโดยพิจารณาเพียงแค่ความแม่นยำของแบบจำลอง อาจจะไม่เหมาะกับการใช้ทำนายแนวโน้มของตลาดแล้วนำไปใช้ตัดสินใจซื้อ-ขายหุ้น ดังนั้นแนวทางในการพัฒนาแบบจำลองต่อไป คือการนำเอาผลตอบแทนที่ได้จากการจำลองการซื้อขายมาพิจารณาร่วมกับความแม่นยำของแบบจำลอง ตัวอย่างเช่น การปรับปรุงฟังก์ชันวัตถุประสงค์ที่ใช้ในกระบวนการเรียนรู้ให้สามารถพิจารณาค่าผลตอบแทนเป็นต้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 5
แนวคิดในการดำเนินงานและวิธีการที่นำเสนอสำหรับแบบจำลองที่เน้นผลตอบแทน

ในบทนี้นำเสนอวิธีการเรียนรู้เชิงลึกเพื่อทำนายแนวโน้มขอตลาดหุ้น โดยพิจารณาข้อมูลเชิงตัวเลขและข้อมูลเชิงตัวอักษรร่วมกัน ซึ่งแบบจำลองที่นำเสนอนั้นจะมุ่งเน้นไปที่ประสิทธิในด้านของผลตอบแทนที่ได้ เมื่อนำผลการทำนายไปทำการจำลองการซื้อขาย ภาพรวมของวิธีการที่นำเสนอแสดงดังรูปที่ 28 ซึ่งแบ่งออกเป็น 2 หัวข้อหลัก ได้แก่ (1) การประมวลผลข้อมูลเบื้องต้น (2) แบบจำลองสำหรับการทำนายแนวโน้มของตลาดหุ้น



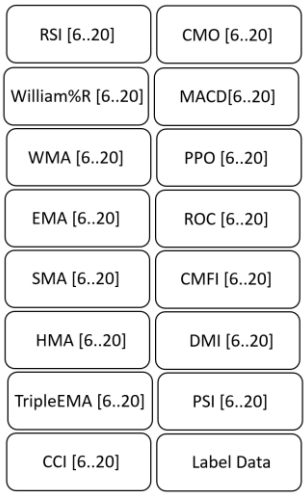
รูปที่ 28 ภาพรวมของวิธีการที่นำเสนอ

5.1 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing)

หัวข้อนี้จะกล่าวถึงการเตรียมข้อมูลเพื่อนำไปใช้เป็นข้อมูลรับเข้าของแบบจำลองที่ใช้ทำนายแนวโน้มของตลาดหุ้นโดย ซึ่งมีรายละเอียดดังต่อไปนี้

5.1.1 ตัวชี้วัดทางเทคนิค (Technical Indicator)

คือ ฟีเจอร์ใหม่ที่สร้างจากราคาในอดีต โดยในงานวิจัยนี้ได้เลือกใช้ ตัวชี้วัดทางเทคนิคจำนวน 15 ตัวและการปรับเปลี่ยนค่าพารามิเตอร์ต่าง ๆ จำนวน 15 รูปแบบ ตามที่ถูกนำเสนอในงานวิจัย [18] ตัวชี้วัดทางเทคนิคที่ใช้แสดงดังรูปที่ 29



รูปที่ 29 ตัวชี้วัดทางเทคนิคที่ใช้งานวิจัย (อ้างอิงจาก Fig.2 ใน [18])

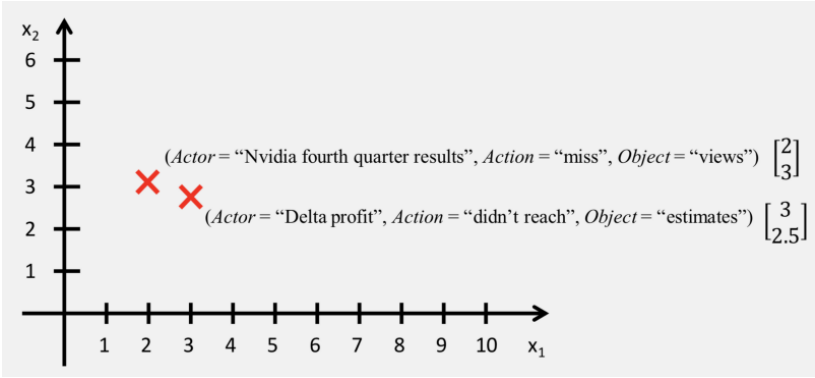
5.1.2 นอร์มัลไลเซชัน (Normalization)

เนื่องจากข้อมูลตัวชี้วัดทางเทคนิคและข้อมูลราคาในอดีตมีค่าอยู่ในช่วงที่แตกต่างกัน ดังนั้นจึงมีการทำนอร์มัลไลเซชันเพื่อให้ชุดข้อมูลอยู่ในช่วงที่ใกล้เคียงกันเพื่อช่วยให้สามารถสอนแบบจำลองได้เร็วขึ้น ในงานวิจัยนี้เลือกใช้การแปลงข้อมูลด้วยวิธีคะแนนมาตรฐาน (z-score) ซึ่งเป็นการแทนข้อมูลด้วยค่าใหม่ โดยการปรับข้อมูลให้มีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนมาตรฐานเป็นหนึ่ง โดยใช้สมการ (50)

$$z(X_i) = \frac{x_i - \mu}{\sigma} \tag{50}$$

5.1.3 เหตุการณ์ฝังตัว (Event Embedding)

เป็นการสร้างเวกเตอร์ที่ใช้แทนเหตุการณ์ต่าง ๆ โดยพยายามให้เหตุการณ์ที่มีลักษณะคล้ายคลึงกันแทนด้วยเวกเตอร์ที่มีลักษณะใกล้เคียงกัน ดังตัวอย่างในรูปที่ 30 ซึ่งวิธีการสร้างเหตุการณ์ฝังนั้นได้ถูกอธิบายบทก่อนหน้า (หัวข้อ 4.1.1.3)



รูปที่ 30 ตัวอย่างของเวกเตอร์เหตุการณ์ฝังตัว

5.1.4 วิธีการสร้างผลเฉลย (Labeling Method)

งานวิจัยชิ้นนี้พยายามที่จะทำนายแนวโน้มของตลาดหุ้นในวันถัดไป ซึ่งแบ่งออกเป็น 3 ประเภท ได้แก่ แนวโน้มแบบขาขึ้น (Upward trend) แนวโน้มแบบขาลง (Downward trend) และแนวโน้มแบบขนานเคลื่อนตัวไปทางด้านข้าง (Sideways tend) ผลเฉลยจะถูกสร้างโดยใช้ข้อมูลของราคาปิดในอดีตโดยใช้วิธีการเลื่อนข้อมูล (Sliding window) เพื่อคำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของราคาในช่วงเวลาดังกล่าว หลังจากนั้นจึงนำมาสร้างเป็นผลเฉลยดังแสดงในสมาการ (51)

$$Trend_i = \begin{cases} Upward\,;\, Close_{i+1} > Avg.\, Close_{window} + S.D.\, Close_{window} \\ Downward\,;\, Close_{i+1} < Avg.\, Close_{window} - S.D.\, Close_{window} \\ Sideways; Otherwise \end{cases} \quad (51)$$

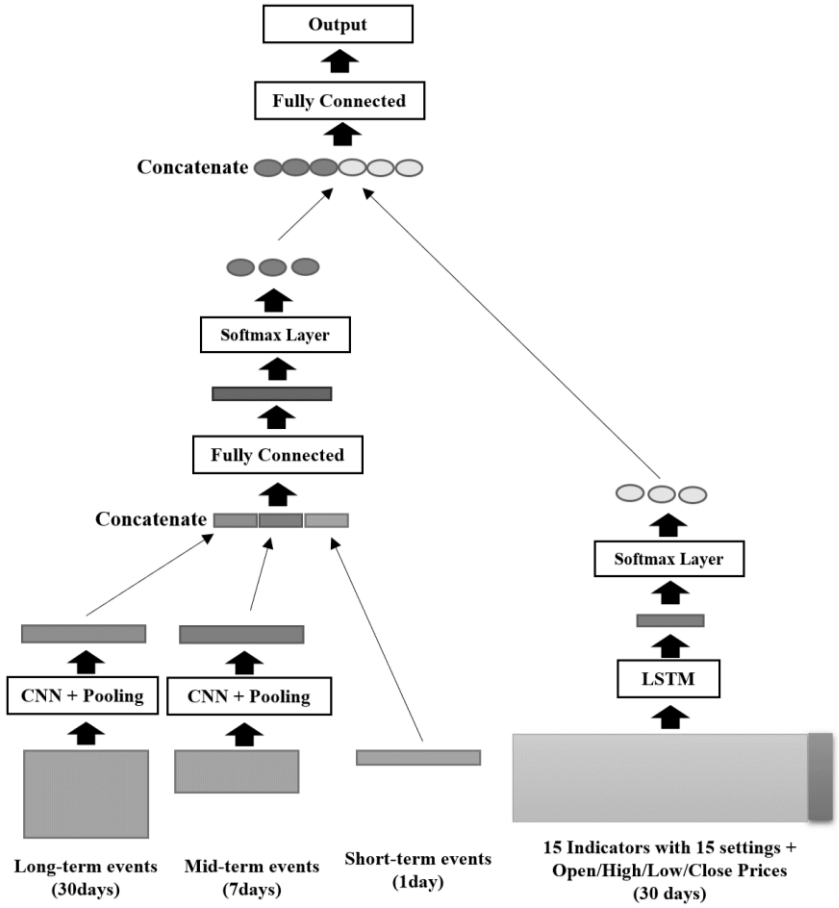
ในเบื้องต้นของงานวิจัยนี้ได้เลือกใช้ขนาดของการเลื่อนข้อมูลเป็น 30 วัน หลังจากนั้นจึงได้มีการทำการทดลองเพื่อทดสอบประสิทธิภาพของแบบจำลองเมื่อใช้ขนาดการเลื่อนเป็น 15, 10 และ 5 วัน สำหรับการสร้างแบบจำลองสุดท้ายนั้นจะเลือกใช้ขนาดการเลื่อนเป็น 5 วัน

5.2 แบบจำลองสำหรับการทำนายแนวโน้มของตลาดหุ้น

5.2.1 แบบจำลองที่นำเสนอ

แบบจำลองที่นำเสนอได้มีการปรับปรุงจากแบบจำลองเบื้องต้นที่ได้กล่าวในบทก่อนหน้าดังนี้ (1) เปลี่ยนโครงสร้างให้สามารถป้อนข้อมูลรับเข้าที่เป็นข้อมูลตัวชี้วัดทางเทคนิคที่เพิ่มขึ้นและราคาในอดีตได้ (15x15 + 4 = 229 ฟีเจอร์) (2) ปรับเปลี่ยนผลลัพธ์ที่ได้จากชั้นฟังก์ชันกระตุ้นค่าสูงสุดอย่างอ่อน (Softmax layer) ซึ่งจากเดิมเป็นเพอร์เซ็ปตรอน 2 ตัว เป็นเพอร์เซ็ปตรอน 3 ตัวเนื่องจากแบบจำลองในขั้นตอนสุดท้ายต้องการที่จะทำนายแนวโน้มของตลาดหุ้นเป็น 3 ประเภท โครงสร้างของแบบจำลองที่นำเสนอแสดงดังรูปที่ 31

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 31 แบบจำลองที่นำเสนอ

5.2.2 การจำลองการซื้อขาย (Trading Simulation)

งานวิจัยนี้ได้นำเอาแนวคิดการจำลองการซื้อหุ้นจาก [18] เพื่อใช้เปรียบเทียบผลตอบแทนที่ได้จากผลการทำนายของแบบจำลอง เงื่อนไขของการจำลองการซื้อขายมีรายละเอียดดังต่อไปนี้ (1) ทำการซื้อหุ้นจำนวน k หน่วย เมื่อแบบจำลองทำนายว่าแนวโน้มในวันถัดไปจะเป็นขาขึ้นและไม่ได้ถือหุ้นไว้ (2) ทำการขายหุ้นเมื่อมีถือหุ้นไว้และแบบจำลองทำนายว่าแนวโน้นในวันถัดไปเป็นขาลง (3) หากไม่ตรงตามเงื่อนไข (1) และ (2) ให้ถือหุ้นไว้ สูตรการคำนวณแสดงดังสมการ (52) โดยที่ k คำนวณได้จากจำนวนเงินที่มีอยู่ ณ เวลา t หารด้วยราคาเปิด ณ เวลา t

$$Action(t) = \begin{cases} Buy\ k\ shares\ of\ stock\ ; Prediciton_{t-1} = Up\ trend\ and\ \#Shares = 0 \\ Sell\ k\ shares\ of\ stock\ ; Prediciton_{t-1} = Down\ trend\ \#Shares > 0 \\ Hold; otherwise \end{cases} \quad (52)$$

ผลกำไร/ขาดทุน จะถูกคำนวณเมื่อเข้าสู่เงื่อนไขการขาย โดยใช้ผลต่างของราคาคูณกับจำนวนหุ้นที่มีอยู่ (k) นอกจากนี้หากมีหุ้นอยู่ ณ วันสุดท้ายที่ทำการจำลองการซื้อขาย (T) ให้ทำการขายหุ้นทั้งหมด ณ ราคาปิดของวันนั้น สมการการคำนวณผลกำไร/ขาดทุน แสดงดังสมการที่ (53)

$$Gain/Loss(t) = \begin{cases} k * (ExitPric\ e - EntryPrice)\ ; Action(t) = Sell \\ k * (ClosePric\ e - EntryPrice)\ ; t = T\ and\ \#Shares > 0 \\ 0\ ; Otherwise \end{cases} \quad (53)$$

ค่าธรรมเนียมการซื้อขายที่ใช้ในงานวิจัยนี้คือ 0.1% ของมูลค่าที่ซื้อขายในแต่ละครั้ง จำนวนเงินที่มีอยู่ ณ เวลา t สามารถคำนวณได้จาก จำนวนเงินจากเวลาก่อนหน้า ผลกำไร/ขาดทุน และค่าธรรมเนียมในการซื้อขาย ดังแสดงในสมการ (54)

$$Balance(t) = Balance(t-1) + Gain/Loss(t) - TransactionCost(t) \quad (54)$$

ผลตอบแทนที่ได้จะถูกคำนวณทุกครั้งที่มีการขายหุ้น และทำการปรับให้เป็นผลตอบแทนเฉลี่ยต่อปี เพื่อให้ง่ายต่อการเปรียบเทียบโดยใช้สมการที่ (55)

$$Return(t) = \left(\frac{Balance(t)}{Balance(t-1)}\right)^{\frac{365}{\#HoldingPeriod}} - 1; if\ Action(t) = Sell \quad (55)$$

งานวิจัยนี้เลือกใช้ชาร์ปเรโช ซึ่งเป็นตัวชี้วัดผลตอบแทนเทียบกับความเสี่ยงที่ใช้กันแพร่หลาย ซึ่งสามารถคำนวณได้ดังสมการ (56)

$$Sharpe\ Ratio = \frac{Mean(Returns)}{S.D.(Returns)} \quad (56)$$

ค่าชาร์ปเรโช ที่ได้จะถูกป้อนกลับเข้าไปในฟังก์ชันวัตถุประสงค์ ซึ่งจะกล่าวในหัวข้อถัดไป

5.2.3 ฟังก์ชันวัตถุประสงค์ (Objective function)

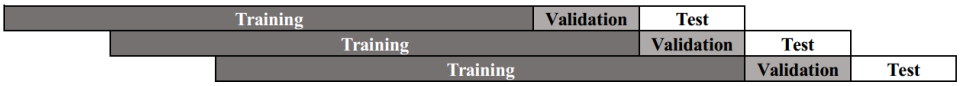
ครอสเอนโทรปี (Cross entropy) เป็นฟังก์ชันวัตถุประสงค์ที่ใช้กันอย่างแพร่หลายในการแก้ปัญหาด้านการจำแนกประเภท (Classification task) ซึ่งฟังก์ชันลักษณะนี้มักจะมุ่งเน้นไปที่ความแม่นยำของการทำนายผลลัพธ์ ในงานวิจัยนี้ได้เสนอฟังก์ชันวัตถุประสงค์แบบใหม่ที่สามารถช่วยเพิ่มประสิทธิภาพของแบบจำลองในด้านของผลตอบแทนเฉลี่ยและชาร์ปเรโช ดังแสดงในสมการที่ (57)

$$Loss = CrossEntropy * (1-\alpha) + \alpha * \log\left(\frac{1}{\max(0.01, SharpeRatio)}\right) \quad (57)$$

ค่าชาร์ปเรโชที่ได้จากการจำลองการซื้อขายจะถูกนำมาคำนวณเป็นฟังก์วัตถุประสงค์ โดยจะถูกปรับค่าด้วยลอการิทึมธรรมชาติ เนื่องจากค่าครอสเอนโทรปีก็ถูกคำนวณจากลอการิทึมธรรมชาติ พารามิเตอร์อัลฟา (α) จะถูกใช้เพื่อกำหนดน้ำหนักไปยังแต่ละส่วนประกอบของฟังก์ชันวัตถุประสงค์ โดยทั่วไปการใช้อัลฟาสูง ๆ ส่งผลให้แบบจำลองที่ได้ ให้ผลลัพธ์ในด้านชาร์ปเรโชที่เพิ่มขึ้น แต่จะให้ค่าเอฟวันที่ต่ำกว่า

5.2.4 กระบวนการสอนแบบจำลอง (Training Processes)

งานวิจัยนี้อาศัยแนวคิดจาก [18] ซึ่งทำการแบ่งข้อมูลเป็นหลาย ๆ ส่วน ตามช่วงเวลาต่าง ๆ แล้วจึงสร้างแบบจำลองสำหรับช่วงเวลานั้น ๆ แล้วทำการวัดประสิทธิภาพของแต่ละแบบจำลองหลังจากนั้นจึงนำผลลัพธ์ที่ได้มาวัดผลร่วมกัน วิธีการแบ่งข้อมูลแสดงดังรูปที่ 32



รูปที่ 32 การแบ่งข้อมูลเพื่อใช้สำหรับการสร้างแบบจำลอง

5.2.5 การเลือกแบบจำลองที่ดีที่สุด

งานวิจัยนี้ได้นำเสนอวิธีการเลือกแบบจำลองโดยการพิจารณาทั้งค่าเอฟวันและค่าชาร์ปเรโชร่วมกันโดยสร้างตัววัดผลขึ้นมาใหม่ที่ชื่อว่า คะแนนเอสเอฟวัน (Sharpe-F1 score, SF1) ดังแสดงในสมการที่ (58)

$$SF1\ score \quad = (1-\beta) * \mathrm{F1}_{norm} + \beta * SharpeRatio_{norm} \tag{58}$$

โดยที่ $\mathrm{F1}_{norm}$ และ $SharpeRatio_{norm}$ ถูกคำนวณด้วยการนอร์มัลไลเซชันด้วยค่าสูงสุดและค่าต่ำสุด (Min-max normalization) โดยใช้ผลลัพธ์ของแบบจำลองต่าง ๆ ที่ทำการวัดผลบนชุดข้อมูลตรวจสอบ (Validation data) สำหรับพารามิเตอร์เบต้า (β) จะถูกใช้เพื่อกำหนดน้ำหนักไปยังแต่ละส่วนประกอบของการคำนวนค่าคะแนนเอสเอฟวัน โดยทั่วไปการใช้เบต้าสูง ๆ ส่งผลให้แบบจำลองที่ได้ ให้ผลลัพธ์ในด้านชาร์ปเรโชที่เพิ่มขึ้น แต่จะให้ค่าเอฟวันที่ต่ำกว่า ในงานวิจัยนี้จะเลือกใช้เบต้า 0.5 เป็นค่ามาตรฐาน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 6
การทดลองและผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองต่าง ๆ ที่จัดทำขึ้นในงานวิจัยนี้ ซึ่งสามารถสรุปได้เป็น 8 หัวข้อ โดยมีรายละเอียดดังต่อไปนี้

6.1 การสร้างแบบจำลองเพื่อใช้อ้างอิง

เนื่องจากงานวิจัยนี้ได้นำเอาแนวคิดของการสร้างเหตุการณ์ฝังตัว ซึ่งนำเสนอโดย Ding [15] ดังนั้นในงานวิจัยนี้จึงได้ทดลองสร้างแบบจำลองตามที่ Ding ได้นำเสนอไว้ เพื่อให้สามารถนำมาเปรียบเทียบประสิทธิภาพกับแบบจำลองที่นำเสนอได้

6.1.1 ประสิทธิภาพของวิธีการสกัดสารสนเทศแบบเปิด

งานวิจัยนี้เลือกใช้ซอฟต์แวร์สำหรับกระบวนการสกัดสารสนเทศแบบเปิด ซึ่งถูกพัฒนาโดยมหาวิทยาลัยสแตนฟอร์ด ดังนั้นจึงต้องจัดทำการทดลองเพื่อวัดประสิทธิภาพของซอฟต์แวร์ก่อน สำหรับการวัดประสิทธิภาพนั้นจะทำโดยการสุ่มเลือกผลลัพธ์จากการป้อนหัวข้อข่าวในแต่ละเดือน จำนวน 2 หัวข่าวต่อเดือน (ณ ต้นเดือน และ ปลายเดือน) หลังจากนั้นผู้วิจัยจะทำการประเมินว่าผลลัพธ์ที่ได้มีความถูกต้องหรือไม่ ตัวอย่างการประเมินแสดงดังตารางที่ 4 จากผลการทดลองพบว่าซอฟต์แวร์ที่เลือกใช้มีความแม่นยำในการสร้างตัวแทนของเหตุการณ์ (Actor, Action, Object) ประมาณ 64%

ตารางที่ 4 ตัวอย่างการประเมินประสิทธิภาพของกระบวนการสกัดสารสนเทศแบบเปิด

หัวข้อข่าวที่ผ่านกระบวนการเตรียมข้อมูล	แอคเตอร์	แอคชัน	ออปเจ็ค	ถูกต้อง?
BlackRock changes managers on energy, growth equity funds	BlackRock	changes	managers	Yes
GM pays $50 million to end lawsuit over 2009 hedge fund deal	GM	pays	$ 50 million	Yes
Oil hits five-year lows in longest losing streak since 2008 crisis	Oil	hits	five-year lows	Yes
Exxon reports big drop in 4Q profit	Exxon	reports	big drop in 4Q profit	Yes
Russian Army Wives Are Protesting For Russia To Come Clean About Where Soldiers Are	Army Wives	Come	Clean	No
Four month old baby pulled alive from Nepal rubble 22 hours after parents lost him during earthquake	baby	old	Four month	No
Will These 2 New Products Boost Apple, Inc.'s Revenue?	Inc.	has	Revenue	No
Former Apple CEO: Trump's 'made in America' focus shouldn't hurt Apple	Apple	made in	America ' focus	No

CHULALONGKORN UNIVERSITY

6.1.2 ประสิทธิภาพของแบบจำลองที่สร้างเลียนแบบงานวิจัยของ Ding

เนื่องจากงานวิจัยนี้ได้นำเอาแนวคิดการสร้างเหตุการณ์ฝังตัวมากจากงานวิจัยของ Ding แต่อย่างไรก็ตามในงานวิจัยได้กล่าวไม่ได้เปิดเผยพารามิเตอร์ทั้งหมดที่ใช้ในการสร้างเหตุการณ์ฝังตัว ดังนั้นจึงต้องมีการเปรียบเทียบผลลัพธ์ที่ได้แบบจำลองที่สร้างเลียนแบบกับผลลัพธ์ที่ถูกนำเสนอโดย Ding ซึ่งผลการทดลองพบว่า แบบจำลองที่สร้างเลียนแบบมีความแม่นยำอยู่ที่ 60.85% ใน ขณะที่ผลลัพธ์ที่ถูกนำเสนอโดย Ding คือค่าความแม่นยำ 64.21% สำหรับสาเหตุของความแตกต่างประมาณ 3.36% นั้น จากตรวจสอบพบว่าเกิดจากความแตกต่างของเวกเตอร์เหตุการณ์ฝังตัวที่ใช้เป็นข้อมูลรับเข้าของแบบจำลองซึ่งเกิดจากความแตกต่างของพารามิเตอร์ที่ใช้ในการสร้างเหตุการณ์ฝังตัวดังที่กล่าวไว้ข้างต้น

6.2 ผลกระทบของพารามิเตอร์ที่ใช้ในฟังก์ชันวัตถุประสงค์และตัววัดผลที่นำเสนอ

การทดลองในหัวข้อนี้จะทำการทดสอบอิทธิพลของแต่ละพารามิเตอร์ที่ใช้ฟังก์ชันวัตถุประสงค์ที่ได้นำเสนอ ซึ่งการทดลองนี้จะใช้แนวคิดและวิธีการที่ได้นำเสนอในบทที่ 5 ยกเว้นส่วนของข้อมูลตัวชี้วัดทางเทคนิคที่ยังคงใช้จำนวน 7 ชนิดตามที่ได้อธิบายในหัวข้อ 4.1.1.1 สำหรับชุดข้อมูลที่ใช้ทดสอบนั้นจะเลือกใช้หัวข้อข่าวที่ได้จากรอยเตอร์และเรดดิตมาสร้างเป็นชุดข้อมูล รวมทั้งทำการแบ่งขั้นตอนการสอนแบบจำลองออกเป็น 3 ช่วงตามที่ได้อธิบายในหัวข้อที่ 5.2.4 รายละเอียดของชุดข้อมูลและช่วงเวลาที่ทำการแบ่งข้อมูลแสดงดังตารางที่ 5

ตารางที่ 5 รายละเอียดของการแบ่งชุดข้อมูลที่ใช้ในการทดลอง

รายละเอียด	ชุดข้อมูลที่ 1	ชุดข้อมูลที่ 2
ข้อมูลราคาในอดีต	ดัชนีเอสแอนด์พี 500	ดัชนีอุตสาหกรรมดาวโจนส์
ข้อมูลหัวข้อข่าว	รอยเตอร์	เรดดิต
จำนวนหัวข้อข่าว	27,158	49,725
ช่วงการแบ่งข้อมูลครั้งที่ 1		
- *ข้อมูลสอน*	20 Oct 06 - 3 May 11	8 Aug 08 - 22 Aug 13
- *ข้อมูลตรวจสอบ*	4 May 11 - 26 Dec 11	23 Aug 13 - 12 May 14
- *ข้อมูลทดสอบ*	27 Dec 11 - 10 Aug 12	13 May 14 - 28 Jan 15
ช่วงการแบ่งข้อมูลครั้งที่ 2		
- *ข้อมูลสอน*	20 Jun 07 - 26 Dec 11	4 May 09 - 12 May 14
- *ข้อมูลตรวจสอบ*	27 Dec 11 - 10 Aug 12	13 May 14 - 28 Jan 15
- *ข้อมูลทดสอบ*	13 Aug 12 - 1 Apr 13	29 Jan 15 - 14 Oct 15
ช่วงการแบ่งข้อมูลครั้งที่ 3		
- *ข้อมูลสอน*	9 Feb 08 - 10 Aug 12	20 Jan 10 - 28 Jan 15
- *ข้อมูลตรวจสอบ*	13 Aug 12 - 1 Apr 13	29 Jan 15 - 14 Oct 15
- *ข้อมูลทดสอบ*	2 Apr 13 - 19 Nov 13	15 Oct 15 - 1 Jul 16

CHULALONGKORN UNIVERSITY

6.2.1 อิทธิพลของพารามิเตอร์อัลฟาของฟังก์ชันที่นำเสนอ

พารามิเตอร์อัลฟา ถูกออกแบบมาเพื่อใช้กำหนดน้ำหนักของแต่ละส่วนประกอบในฟังก์ชันวัตถุประสงค์ที่นำเสนอ ดังแสดงในสมการ (59)

$$Loss = CrossEntropy * (1 - \alpha) + \alpha * \log\left(\frac{1}{\max(0.01, SharpeRatio)}\right) \quad (59)$$

ในการทดลองนี้พารามิเตอร์อัลฟาที่ใช้จะถูกปรับค่าตั้งแต่ 0 – 0.9 หลังจากนั้นจึงทำการวัดประสิทธิภาพของแบบจำลองเพื่อทำการเปรียบเทียบผลลัพธ์ ผลการทดลองที่ได้แสดงในตารางที่ 6

ตารางที่ 6 ผลการทดลองเมื่อใช้ค่าพารามิเตอร์อัลฟาที่แตกต่างกัน

พารามิเตอร์อัลฟา	ชุดข้อมูลที่ 1			ชุดข้อมูลที่ 2		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
0.0	**62.27%**	0.324	19.25%	58.38%	0.063	1.72%
0.1	61.34%	0.323	19.02%	**58.67%**	0.207	1.94%
0.2	60.45%	0.313	17.08%	58.29%	0.284	4.69%
0.3	57.79%	0.490	**21.56%**	54.75%	**0.444**	**6.26%**
0.4	56.73%	0.373	15.35%	56.68%	0.270	5.40%
0.5	56.23%	0.433	14.19%	53.37%	0.359	5.98%
0.6	55.95%	0.535	15.26%	53.56%	0.416	4.88%
0.7	55.40%	0.457	15.97%	54.43%	0.302	5.79%
0.8	55.37%	0.572	13.12%	47.69%	0.256	2.54%
0.9	54.48%	**0.611**	16.43%	51.81%	0.288	1.41%

สำหรับชุดข้อมูลที่ 1 ผลการทดลองแสดงให้เห็นถึงแนวโน้มการลดลงของค่าเอฟวัน เมื่อทำการเพิ่มค่าพารามิเตอร์อัลฟา แต่อย่างไรก็ตาม ค่าชาร์ปเรโชก็จะเพิ่มสูงขึ้นด้วย เนื่องจากการใช้ค่าอัลฟาสูง ๆ จะทำให้ฟังก์ชันวัตถุประสงค์ให้ความสำคัญกับความแม่นยำในการทำนายน้อยลง (ค่าเอฟวัน) แต่เพิ่มความสำคัญไปยังผลลัพธ์ที่ได้จากการจำลองการซื้อขาย (ค่าชาร์ปเรโช) สำหรับชุดข้อมูลที่ 2 แม้ว่าแนวโน้มของค่าเอฟวันและชาร์ปเรโชที่ได้จะมีความผันผวนแต่อย่างไรก็ตามภาพรวมของผลลัพธ์ยังคงสอดคล้องกับชุดข้อมูลที่ 1 นั่นคือ เมื่อใช้พารามิเตอร์อัลฟาที่สูงขึ้นจะส่งผลให้ค่าเอฟวันลดลงและทำให้ชาร์ปเรโชเพิ่มขึ้น

จากผลการทดลองในตารางที่ 6 แสดงให้เห็นว่าค่าพารามิเตอร์อัลฟาที่เหมาะสมสำหรับทั้ง 2 ชุดข้อมูลคือ 0.3 เนื่องจากส่งผลให้ได้ผลตอบแทนเฉลี่ยต่อปีสูงสุดในทั้ง 2 ชุดข้อมูล สำหรับการเลือกใช้พารามิเตอร์อัลฟานั้น ผู้วิจัยแนะนำว่าควรจะใช้ค่าไม่เกิน 0.5 เนื่องจากการใช้ค่าอัลฟาที่มากเกินไปจะส่งผลให้แบบจำลองที่ได้ขึ้นอยู่กับผลลัพธ์จากการทำการจำลองการซื้อขายรวมทั้งวิธีการที่ใช้ในการจำลองการซื้อขายเท่านั้น ซึ่งอาจจะก่อให้เกิดปัญหาการลู่ออก ในกระบวนการฝึกสอนแบบจำลองได้

6.2.2 อิทธิพลของพารามิเตอร์เบต้าของฟังก์ชันที่นำเสนอ

พารามิเตอร์เบต้าจะถูกใช้ในสมการที่ใช้วัดผลคะแนนของแบบจำลองซึ่งถูกสร้างขึ้นเพื่อเลือกแบบจำลองที่ดีที่สุด ดังที่ได้กล่าวในหัวข้อ 5.2.5 โดยที่พารามิเตอร์นี้ได้ถูกออกแบบเพื่อให้ผู้ใช้สามารถเลือกแบบจำลองสุดท้ายตามความเสี่ยงที่ต้องการได้

$$SF1\ score \quad = (1-\beta) * \mathrm{F1}_{norm} + \beta * SharpeRatio_{norm} \tag{60}$$

สำหรับการทดลองนี้ จะเลือกใช้พารามิเตอร์อัลฟาที่ดีที่สุดเท่ากับ 0.3 จากการทดลองก่อนหน้า หลังจากนั้นพารามิเตอร์เบต้าจะถูกปรับค่าตั้งแต่ 0 จนกระทั่ง 1 โดยทำการปรับค่าครั้งละ 0.25 เพื่อเปรียบเทียบอิทธิพลของพารามิเตอร์ดังกล่าว ผลการทดลองแสดงดังตารางที่ 7

ตารางที่ 7 ผลการทดลองเมื่อใช้ค่าพารามิเตอร์เบต้าที่แตกต่างกัน

พารามิเตอร์เบต้า	ชุดข้อมูลที่ 1			ชุดข้อมูลที่ 2		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
0.00	**65.08%**	0.312	21.88%	62.42%	0.211	0.56%
0.25	59.57%	0.354	**23.55%**	**62.43%**	0.125	-0.08%
0.50	57.79%	0.490	21.56%	54.75%	0.444	6.26%
0.75	57.79%	0.490	21.56%	47.92%	0.543	**7.98%**
1.00	63.31%	**0.548**	19.80%	42.63%	**0.553**	4.60%

ผลการทดลองในตารางที่ 7 แสดงให้เห็นถึงแนวโน้มการเปลี่ยนแปลงของค่าเอฟวันและชาร์ปเรโช เมื่อทำการเพิ่มพารามิเตอร์เบต้า ซึ่งผลการทดลองที่ได้ก็สอดคล้องกับวัตถุประสงค์ของสร้างพารามิเตอร์ตัวนี้ กล่าวคือ การเพิ่มค่าพารามิเตอร์เบต้าจะส่งผลให้แบบจำลองสุดท้ายที่ได้มุ่งเน้นไปที่ ค่าชาร์ปเรโชและส่งผลให้ความสามารถในการทำนายผล (ค่าเอฟวัน) ลดลง จากผลการทดลองในตารางข้างต้นพบว่า การเลือกใช้เบต้าเท่ากับ 0.75 ให้ผลลัพธ์ที่เหมาะสมในทั้ง 2 ชุดข้อมูล

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.2.3 อิทธิพลของขนาดการเลื่อนข้อมูลที่ใช้สร้างผลเฉลย

ดังที่ได้กล่าวไปในหัวข้อ 5.1.4 ผลเฉลยจะถูกสร้างโดยใช้ข้อมูลของราคาปิดในอดีตโดยใช้วิธีการเลื่อนข้อมูล (Sliding window) เพื่อคำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของราคาในช่วงเวลาดังกล่าว หลังจากนั้นจึงนำมาสร้างเป็นผลเฉลย สำหรับการทดลองนี้จะเลือกใช้พารามิเตอร์อัลฟาเท่ากับ 0.3 และพารามิเตอร์เบต้าเท่ากับ 0.5 หลังจากนั้นจึงทำการทดสอบแบบจำลองด้วยการใช้ผลเฉลยที่สร้างจากขนาดการเลื่อนที่แตกต่างกัน ผลการทดลองที่ได้แสดงดังตารางที่ 8

ตารางที่ 8 ผลการทดลองเมื่อใช้ช่วงการเลื่อนและฟังก์ชันวัตถุประสงค์ที่แตกต่างกัน

ขนาดการเลื่อน (วัน)	ใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอ			ใช้ฟังก์ชันวัตถุประสงค์แบบเดิม		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
ชุดข้อมูลที่ 1						
30	57.79%	0.490	21.56%	62.27%	0.324	19.25%
15	56.48%	0.436	20.97%	60.86%	0.362	10.02%
10	53.85%	0.569	20.83%	54.07%	0.358	16.42%
5	52.62%	0.516	18.67%	50.78%	0.277	3.88%
ชุดข้อมูลที่ 2						
30	54.75%	0.444	6.26%	58.38%	0.063	1.72%
15	51.49%	0.272	4.48%	55.23%	-0.598	0.17%
10	50.40%	0.193	1.92%	52.46%	-0.226	0.96%
5	44.45%	0.264	1.33%	45.29%	-0.189	-3.28%

จากผลการทดลองข้างต้น พบว่าค่าเอฟวันจะลดลงเมื่อใช้ขนาดการเลื่อนที่เล็กลง เนื่องจากผลเฉลยที่ถูกสร้างจากขนาดการเลื่อนที่เล็กจะมีความฝันผวนของผลเฉลยมากกว่าการใช้ขนาดการเลื่อนที่ใหญ่ ดังนั้นจึงทำให้ทำนายแนวโน้มได้ยากกว่า นอกจากจะแสดงแนวโน้มของผลลัพธ์สำหรับขนาดการเลื่อนต่าง ๆ แล้ว การทดลองนี้ยังแสดงให้เห็นว่า การใช้ฟังก์ชันวัตถุประสงค์ที่ได้นำเสนอนั้นสามารถเพิ่มประสิทธิภาพของแบบจำลองในด้านของผลตอบแทนเฉลี่ยต่อปีและค่าชาร์ปเรโชได้

6.3 ประสิทธิภาพของแบบจำลองเมื่อทำการเพิ่มตัวชี้วัดทางเทคนิค

สำหรับการทดลองในหัวข้อนี้จะทำการเปลี่ยนข้อมูลตัวชี้วัดทางเทคนิคจาก 7 ชนิดเป็น 15 ชนิดและ 15 รูปแบบการตั้งค่า ดังที่ได้อธิบายหัวข้อ 5.1.1 การทดลองนี้จะใช้ชุดข้อมูลเดียวกันกับการทดลองก่อนหน้าโดยที่จะสร้างผลเฉลยโดยใช้ขนาดของการเลื่อนเท่ากัน 5 วัน และเลือกใช้พารามิเตอร์อัลฟาเท่ากับ 0.3 และเบต้าเท่ากับ 0.5 นอกจากนั้นยังทำการเปรียบผลลัพธ์ของการใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอ ซึ่งผลการทดลองได้แสดงในตารางที่ 9

จากผลการทดลองจะเห็นได้ว่าการเปลี่ยนมาใช้ตัวชี้วัดทางเทคนิคที่มากขึ้นส่งผลให้แบบจำลองที่ได้ให้ผลตอบแทนเฉลี่ยนต่อปีและค่าชาร์ปเรโชที่สูงกว่าเดิม แต่ค่าความแม่นยำก็จะลดลงไปแทน นอกการนี้การทดลองนี้ยังแสดงให้เห็นว่าการใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอ สามารถเพิ่มประสิทธิภาพด้านผลตอบแทนเฉลี่ยต่อปีและชาร์ปเรโชได้ แต่ก็จะส่งผลให้ประสิทธิภาพในด้านความแม่นยำในการทำนายลดลงตามไปด้วย

ตารางที่ 9 ผลการทดลองเมื่อใช้ช่วงการเลื่อนและฟังก์ชันวัตถุประสงค์ที่แตกต่างกัน

จำนวนตัวชี้วัดทางเทคนิค	ใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอ			ใช้ฟังก์ชันวัตถุประสงค์แบบเดิม		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
ชุดข้อมูลที่ 1						
7	52.60%	0.516	18.67%	50.78%	0.277	3.88%
15*15	50.08%	0.830	23.50%	52.59%	0.542	17.18%
ชุดข้อมูลที่ 2						
7	44.45%	0.264	1.33%	45.29%	-0.189	-3.28%
15*15	42.09%	0.513	10.87%	42.94%	0.198	8.01%

6.4 ประสิทธิภาพของฟังก์ชันวัตถุประสงค์ที่นำเสนอ

การทดลองในหัวข้อนี้จะนำเอาแบบจำลองอื่น ๆ ที่ใช้โครงสร้างการเรียนรู้เชิงลึกที่ได้กล่าวถึงในหัวข้อ 4.2.2 มาทดสอบสอนแบบจำลองโดยใช้ฟังก์ชันวัตถุประสงค์ที่ได้นำเสนอ

6.4.1 ชุดข้อมูลและการตั้งค่าการทดลอง

ชุดข้อมูลที่ใช้ในการทดลองนี้มี 3 ชุด ซึ่ง 2 ชุดแรกเป็นข้อมูลชุดเดียวกับการทดลองก่อนหน้า สำหรับชุดข้อมูลที่ที่ 3 ได้นำหัวข้อข่าวจากอินทรินิโอ ตั้งแต่ 1 Aug 2016 – 31 Aug 2018 มาใช้ร่วมกับข้อมูลราคาของดัชนีอุตสาหกรรมดาวโจนส์ (DJIA) สำหรับขั้นตอนการเตรียมข้อมูลและกระบวนการจำลองการซื้อขายนั้นจะอ้างอิงตามได้ที่นำเสนอไปในบทที่ 5 โดยเลือกใช้ขนาดการเลื่อนเป็น 5 วันเพื่อสร้างผลเฉลยและแบบจำลองที่นำเสนอนั้นจะเลือกใช้พารามิเตอร์อัลฟาและเบต้าเป็น 0.3 และ 0.5 เหมือนกับการทดลองก่อนหน้า สำหรับจำนวนพารามิเตอร์น้ำหนัก (weights) และค่าไบแอส (bias) ทั้งหมดของแต่ละแบบจำลองแสดงดังตางรางที่ 10

ตารางที่ 10 จำนวนพารามิเตอร์น้ำหนัก (weights) และค่าไบแอส (bias) ที่ใช้ในแต่ละแบบจำลอง

แบบจำลอง	จำนวนพารามิเตอร์น้ำหนักและไบแอส
1) แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าวเพียงอย่างเดียว	197,123
2) แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัวเพียงอย่างเดียว	178,706
3) แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นตัวชี้วัดทางเทคนิคและราคาในอดีต	187,026
4) แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าว ตัวชี้วัดทางเทคนิคและราคาในอดีต	384,929
5) แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัว ตัวชี้วัดทางเทคนิคและราคาในอดีต	363,681

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.4.2 ผลการทดลองและการอภิปรายผล

จากผลการทดลองในตารางที่ 11 แสดงให้เห็นว่าการใช้ข้อมูลทั้ง 2 ประเภทมาป้อนเข้าสู่แบบจำลองสามารถให้ประสิทธิภาพได้ดีกว่าใช้ข้อมูลประเภทใดประเภทหนึ่งเพียงอย่างเดียว สำหรับการเปลี่ยนมาใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอนั้นก็ให้ผลลัพธ์ในด้านผลตอบแทนเฉลี่ยต่อปีและชาร์ปเรโชสูงกว่าการใช้ฟังก์ชันวัตถุประสงค์แบบเดิมแต่อย่างไรก็ตามค่าเอฟวันที่ได้ก็จะมีค่าลดน้อยลง นอกจากนั้นการทดลองนี้ได้แสดงให้เห็นถึงประสิทธิภาพของฟังก์ชันวัตถุประสงค์ที่นำเสนอซึ่งสามารถเพิ่มประสิทธิภาพในด้านผลตอบแทนและชาร์ปเรโชได้ในทุก ๆ แบบจำลองและทุก ๆ ชุดข้อมูล

ตารางที่ 11 ผลการทดลองเมื่อนำฟังก์ชันวัตถุประสงค์ที่นำเสนอไปใช้กับแบบจำลองต่าง ๆ

แบบจำลองที่ใช้	ใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอ			ใช้ฟังก์ชันวัตถุประสงค์แบบเดิม		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
ชุดข้อมูลที่ 1						
แบบจำลอง #1	37.14%	0.362	2.05%	37.28%	0.303	4.40%
แบบจำลอง #2	40.46%	0.163	13.81%	41.64%	0.208	7.69%
แบบจำลอง #3	44.23%	0.576	16.46%	46.73%	0.393	12.26%
แบบจำลอง #4	47.97%	0.333	15.74%	48.31%	0.323	10.05%
แบบจำลอง #5	**50.08%**	**0.830**	**23.50%**	**52.59%**	**0.542**	**17.18%**
ชุดข้อมูลที่ 2						
แบบจำลอง #1	33.17%	0.354	0.67%	33.52%	0.301	-3.51%
แบบจำลอง #2	30.79%	0.375	3.49%	37.22%	**0.333**	3.29%
แบบจำลอง #3	**39.74%**	0.203	9.36%	42.96%	0.170	7.40%
แบบจำลอง #4	37.10%	0.340	8.43%	40.56%	0.257	4.98%
แบบจำลอง #5	42.09%	**0.513**	**10.87%**	**43.52%**	0.198	**8.01%**
ชุดข้อมูลที่ 3						
แบบจำลอง #1	34.80%	0.263	6.15%	35.79%	0.339	4.02%
แบบจำลอง #2	37.87%	0.595	8.19%	41.55%	0.321	4.10%
แบบจำลอง #3	36.49%	**1.140**	6.85%	38.50%	**0.560**	5.67%
แบบจำลอง #4	36.67%	0.556	7.92%	36.85%	0.205	6.33%
แบบจำลอง #5	**43.36%**	0.333	**9.90%**	**43.76%**	0.322	**7.53%**

โดยที่ แบบจำลอง #1 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าวเพียงอย่างเดียว

แบบจำลอง #2 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัวเพียงอย่างเดียว

แบบจำลอง #3 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นตัวชี้วัดทางเทคนิคและราคาในอดีต

แบบจำลอง #4 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าว ตัวชี้วัดทางเทคนิคและราคาในอดีต

แบบจำลอง #5 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัว ตัวชี้วัดทางเทคนิคและราคาในอดีต (แบบจำลองที่นำเสนอในงานวิจัยชิ้นนี้)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.5 การทำนายล่วงหน้าหลาย ๆ วัน

ในหัวข้อนี้จะทำการสร้างแบบจำลองเพื่อทำนายแนวโน้มของตลาดหุ้นล่วงหน้าจำนวน 5, 10, 15 และ 30 วัน เพื่อนำมาเปรียบกับการทำนายล่วงหน้า 1 วัน จากการทดลองก่อนหน้า

6.5.1 ชุดข้อมูลและการตั้งค่าการทดลอง

ชุดข้อมูลที่ใช้นั้น ยังคงเป็นชุดข้อมูลเดียวกับการทดลองที่ 6.4 แต่เปลี่ยนเงื่อนไขในการสร้างผลเฉลยซึ่งจากเดิมที่ใช้ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานโดยใช้ขนาดการเลื่อนเป็น 5 วัน แล้วนำมาเปรียบเทียบกับราคาปิดในวันถัดไป โดยจะเปลี่ยนเป็นการเปรียบเทียบกับ 5, 10, 15 และ 30 วันถัดไปแทน รวมทั้งทำการเปลี่ยนเงื่อนไขในการขายหุ้นให้สอดคล้องกับการทำนายผลล่วงหน้า คือ จะทำการขายเมื่อผลการทำนายของช่วงเวลาก่อนหน้าตามจำนวนการทำนายล่วงหน้า เป็นแนวโน้มขาลง

6.5.2 ผลการทดลองและการอภิปรายผล

จากผลการทดลองที่แสดงในตารางที่ 12 พบว่าโดยเฉลี่ยแล้วค่าความแม่นยำ ค่าชาร์ปเรโช และผลตอบแทนเฉลี่ยต่อปี จะมีค่าลดลงเมื่อใช้ระยะเวลาในการทำนายล่วงหน้าที่นานขึ้น ซึ่งแนวโน้มที่ได้นี้ก็สอดคล้องกับงานวิจัยก่อนหน้า [14] ที่ได้ทำการทดลองการทำนายล่วงหน้าหลาย ๆ วัน

ตารางที่ 12 ผลการทดลองเมื่อทำการทำนายล่วงหน้าหลาย ๆ วัน

แบบจำลองที่ทำนายล่วงหน้า	**ประสิทธิภาพของแบบจำลอง**		
	ค่าเอฟวัน	**ชาร์ปเรโช**	**ผลตอบแทนเฉลี่ยต่อปี**
ชุดข้อมูลที่ 1			
1 วัน	**50.08%**	**0.830**	**23.50%**
5 วัน	47.91%	0.242	10.26%
10 วัน	39.42%	0.220	8.26%
15 วัน	39.41%	0.223	5.64%
30 วัน	30.17%	0.173	0.42%
ชุดข้อมูลที่ 2			
1 วัน	**42.09%**	**0.513**	**10.87%**
5 วัน	36.93%	0.313	6.59%
10 วัน	32.97%	0.326	4.29%
15 วัน	31.40%	0.308	-2.91%
30 วัน	27.32%	0.302	-3.75%
ชุดข้อมูลที่ 3			
1 วัน	**43.36%**	0.333	9.90%
5 วัน	36.04%	**0.616**	**10.39%**
10 วัน	32.68%	0.307	-5.24%
15 วัน	22.13%	0.296	-10.72%
30 วัน	24.57%	-0.854	-9.92%

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.6 การทำนายหุ้นรายตัว

การทดลองนี้จะทำเพื่อตัวตรวจสอบประสิทธิภาพของแบบจำลองที่นำเสนอเมื่อนำมาใช้ทำนายแนวโน้มของหุ้นรายตัว

6.6.1 ชุดข้อมูลและการตั้งค่าการทดลอง

การทดลองนี้จะใช้ชุดข้อมูลหัวข้อข่าวจากอินทรินิโอโดยทำการเลือกหุ้นจำนวน 5 ตัวที่มีจำนวนหัวข้อข่าวมากที่สุด หลังจากนั้นจึงทำการสร้าง 5 แบบจำลองโดยการนำข้อมูลหัวข้อข่าวและราคาหุ้น ป้อนเข้าสู่แบบจำลองที่นำเสนอดังที่ได้อธิบายไว้ในบทที่ 5 โดยสร้างผลเฉลยจากขนาดการเลื่อนเท่ากับ 5 วันและเลือกใช้พารามิเตอร์อัลฟาและเบต้าเท่ากับ 0.3 และ 0.5 นอกจากนี้ยังได้ทำการทดลองเปรียบเทียบกับการทำนายแนวโน้มโดยใช้ข้อมูลราคาในอดีตและตัวชี้วัดทางเทคนิคมาใช้สร้างแบบจำลอง

ตารางที่ 13 รายละเอียดจำนวนหัวข้อข่าวที่ถูกนำมาใช้ในการทดลอง

ชื่อย่อ	ชื่อเต็ม	จำนวนหัวข้อข่าว	จำนวนหัวข้อข่าวเฉลี่ยต่อวัน
DJIA	**Dow Jones Industrial Average**	**209,907**	**276**
APPL	Apple Inc.	41,215	54
MSFT	Microsoft Corporation	12,866	17
DIS	The Walt Disney Company	10,458	14
WMT	Walmart Inc.	11,629	15
BA	The Boeing Company	12,094	16

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.6.2 ผลการทดลองและการอภิปรายผล

จากผลการทดลองที่แสดงในตารางที่ 14 พบว่าแนวคิดและแบบจำลองที่นำเสนอสามารถนำมาใช้นายหุ้นรายตัวได้ โดยที่เมื่อพิจารณาประสิทธิภาพของแบบจำลองในด้านเอฟวัน พบว่าแบบจำลองที่ใช้ข้อมูลสองประเภทสามารถทำนายแนวโน้มได้ดีกว่าการใช้ข้อมูลเชิงตัวเลขเพียงอย่างเดียว นอกจากนี้เมื่อพิจารณาประสิทธิภาพในด้านผลตอบเฉลี่ยต่อปี พบว่าการใช้ข้อมูลทั้งสองประเภทเพื่อสร้างแบบจำลองสามารถให้ผลตอบแทนที่ดีกว่าในทุก ๆ ชุดข้อมูล แต่อย่างไรก็ตามเมื่อพิจารณาผลลัพธ์ในด้านชาร์ปเรโช พบว่าการใช้ข้อมูลทั้ง 2 ประเภทนั้นจะให้ค่าที่ต่ำกว่า หรือกล่าวคือ มีค่าความเสี่ยงเมื่อเทียบกับผลตอบแทนที่สูงกว่า

ตารางที่ 14 ผลการทดลองเมื่อนำฟังก์ชันวัตถุประสงค์ที่นำเสนอไปใช้กับแบบจำลองต่าง ๆ

หุ้นที่นำมาทำนาย	แบบจำลองที่ใช้ข้อมูลเหตุการณ์ฝังตัว ตัวชี้วัดทางเทคนิคและราคาในอดีต			แบบจำลองที่ใช้ข้อมูลตัวชี้ทางเทคนิค และราคาในอดีต		
	ค่าเอฟวัน	ชาร์ป เรโช	ผลตอบแทน เฉลี่ยต่อปี	ค่าเอฟวัน	ชาร์ป เรโช	ผลตอบแทน เฉลี่ยต่อปี
DJIA	43.36%	0.333	9.90%	36.49%	1.140	6.85%
APPL	37.27%	0.316	3.13%	36.45%	0.564	2.90%
MSFT	36.31%	0.129	13.85%	34.07%	0.471	7.63%
DIS	36.87%	0.718	3.53%	36.60%	0.515	0.78%
WMT	39.89%	0.333	6.17%	36.75%	0.384	3.97%
BA	44.06%	0.410	11.32%	39.23%	0.504	5.95%

6.7 การทำนายแนวโน้มของตลาดหุ้นด้วยข้อมูลเชิงตัวเลขร่วมกับข่าวประเภทต่าง ๆ

การทดลองนี้จะทำเพื่อตัวตรวจสอบประสิทธิภาพของแบบจำลองที่นำเสนอ เมื่อใช้หัวข่าวประเภทต่าง ๆ เป็นข้อมูลรับเข้า

6.7.1 ชุดข้อมูลและการตั้งค่าการทดลอง

การทดลองนี้ใช้ข้อมูลหัวข้อข่าวจากรอยเตอร์ ซึ่งเป็นหัวข้อข่าวทางด้านเศรษฐกิจและการเงิน หลังจากนั้นจึงทำการแยกข้อมูลหัวข้อข่าวเป็น 2 ประเภท คือ 1) ข้อมูลข่าวที่เฉพาะเจาะจงกับบริษัท คือ หัวข้อข่าวที่มีชื่อบริษัทหรือชื่อย่อปรากฏอยู่ในหัวข้อข่าว 2) ข้อมูลข่าวทั่วไป คือ หัวข้อข่าวอื่น ๆ นอกเหนือจากประเภทที่ 1 ผลที่ได้จากข้อมูลหัวข้อข่าวจากรอยเตอร์ จำนวน 77,243 หัวข้อข่าว สามารถแบ่งได้เป็นข่าวที่เฉพาะเจาะจงจำนวน 27,158 หัวข้อข่าว และ ข่าวทั่วไปจำนวน 50,085 หัวข้อข่าว

6.7.2 ผลการทดลองและการอภิปรายผล

จากผลการทดลองที่แสดงในตารางที่ 15 การนำข้อมูลหัวข้อข่าวมาใช้พิจารณาร่วมกับข้อมูลเชิงตัวเลขสามารถทำนายแนวโน้มได้ดีกว่าการใช้ตัวชี้วัดทางเทคนิคและราคาในอดีตเพียงอย่างเดียว โดยเฉพาะการใช้หัวข้อข่าวที่เฉพาะเจาะจงกับบริษัทจะสามารถให้ผลลัพธ์ที่ดีที่สุด

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ 15 ผลการทดลองเมื่อใช้ข้อมูลรับเข้าที่แตกต่างกัน

แบบจำลองที่ใช้ข้อมูลรับเข้าที่แตกต่างกัน	ประสิทธิภาพของแบบจำลอง		
	ค่าเอฟวัน	ชาร์ปเรโช	ผลตอบแทนเฉลี่ยต่อปี
1) ใช้ข้อมูลตัวชี้วัดทางเทคนิคและราคาในอดีต	44.23%	0.576	16.46%
2) ใช้ข้อมูลหัวข้อข่าวทั่วไป ตัวชี้วัดทางเทคนิคและราคาในอดีต	46.83%	0.583	15.29%
3) ใช้ข้อมูลหัวข้อข่าวเฉพาะเจาะจง ตัวชี้วัดทางเทคนิคและราคาในอดีต	50.08%	0.830	23.50%

6.8 พฤติกรรมการลงทุนในระหว่างการจำลองการซื้อขาย

ในหัวข้อนี้จะนำเอาผลลัพธ์ที่ได้จากการทดลองที่ 6.3 มาทำการวิเคราะห์สัดส่วนพฤติกรรมการลงทุนที่เกิดขึ้นในระหว่างขั้นตอนการจำลองการซื้อขายรวมทั้งทำเปรียบเทียบผลตอบที่ได้จากแบบจำลองต่าง ๆ เทียบกับวิธีการลองทุนแบบซื้อแล้วถือ (Buy and Hold) ผลลัพธ์ที่ได้แสดงดังตารางที่ 16

ตารางที่ 16 สัดส่วนของการทำนายแนวโน้มและพฤติการการลงทุนในระหว่างการจำลองการซื้อขาย

แบบจำลองที่ใช้	สัดส่วนของการทำนายแนวโน้ม			สัดส่วนพฤติกรรมการลงทุน		ผลตอบแทนเฉลี่ยต่อปี
	ขาขึ้น	ขาลง	ไปด้านข้าง	ซื้อ/ขาย	ถือหุ้น	
ชุดข้อมูลที่ 1						
แบบจำลอง #1	39.5%	28.0%	32.4%	31.2%	68.8%	2.05%
แบบจำลอง #2	38.3%	18.8%	42.9%	15.2%	84.8%	13.81%
แบบจำลอง #3	64.2%	19.4%	16.4%	3.6%	96.4%	16.46%
แบบจำลอง #4	48.8%	21.7%	29.5%	7.3%	92.7%	15.74%
แบบจำลอง #5	61.7%	13.3%	25.5%	3.0%	97.0%	**23.50%**
วิธีการซื้อแล้วถือ						20.11%
ชุดข้อมูลที่ 2						
แบบจำลอง #1	37.8%	31.7%	30.6%	20.7%	79.3%	0.67%
แบบจำลอง #2	48.0%	3.0%	49.1%	4.8%	95.2%	3.49%
แบบจำลอง #3	41.9%	45.0%	13.1%	8.5%	91.5%	9.36%
แบบจำลอง #4	68.9%	16.1%	15.0%	5.6%	94.4%	8.43%
แบบจำลอง #5	39.1%	14.8%	46.1%	4.4%	95.6%	**10.87%**
วิธีการซื้อแล้วถือ						3.45%
ชุดข้อมูลที่ 3						
แบบจำลอง #1	63.3%	25.7%	11.0%	30.0%	70.0%	6.15%
แบบจำลอง #2	16.7%	11.4%	71.9%	12.4%	87.6%	8.19%
แบบจำลอง #3	29.5%	44.8%	25.7%	4.8%	95.2%	6.85%
แบบจำลอง #4	59.0%	31.0%	10.0%	9.5%	90.5%	7.92%
แบบจำลอง #5	34.3%	18.1%	47.6%	9.5%	90.5%	**9.90%**
วิธีการซื้อแล้วถือ						-0.66%

โดยที่ แบบจำลอง #1 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าวเพียงอย่างเดียว
แบบจำลอง #2 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัวเพียงอย่างเดียว
แบบจำลอง #3 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นตัวชี้วัดทางเทคนิคและราคาในอดีต
แบบจำลอง #4 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นหัวข้อข่าว ตัวชี้วัดทางเทคนิคและราคาในอดีต
แบบจำลอง #5 คือ แบบจำลองที่ใช้ข้อมูลรับเข้าเป็นเหตุการณ์ฝังตัว ตัวชี้วัดทางเทคนิคและราคาในอดีต
(แบบจำลองที่นำเสนอในงานวิจัยชิ้นนี้)

ผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอนั้น สามารถให้ผลตอบแทนเฉลี่ยต่อปีได้ดีกว่าแบบจำลองการเรียนรู้เชิงลึกในรูปแบบอื่น ๆ ซึ่งผลตอบแทนเฉลี่ยนต่อปีที่ได้นี้ก็มีค่าสูงกว่าวิธีการซื้อและถือ (Buy and hold) ในทุก ๆ ชุดข้อมูล และเมื่อพิจารณาพฤติกรรมการลงทุนในระหว่างการจำลองการซื้อขาย พบว่าแบบจำลองที่สร้างจากการใช้ข้อมูลเชิงตัวเลขหรือข้อมูลเชิงตัวอักษรเพียงอย่างเดียว มักจะมีพฤติกรรมการซื้อขายที่บ่อยกว่าแบบจำลองที่ใช้ข้อมูลทั้งสองประเภทร่วมกัน ซึ่งพฤติกรรมการซื้อขายบ่อย ๆ นี้ก็ส่งผลให้เกิดค่าธรรมเนียมในการซื้อขายที่สูงและส่งผลให้ผลตอบแทนที่ได้มีค่าน้อยกว่า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 7
สรุปผลการวิจัยและแนวทางการวิจัยในขั้นถัดไป

6.1 สรุปผลการวิจัย

วิทยานิพนธ์ชิ้นนี้ ได้นำเสนอแนวคิดและแบบจำลองที่ใช้ทำนายแนวโน้มของตลาดหุ้น โดยใช้ข้อมูลเชิงตัวเลขร่วมกับข้อมูลเชิงตัวอักษร นอกจากนี้ยังได้นำเสนอฟังก์ชันวัตถุประสงค์ที่ใช้ในการฝึกสอนแบบจำลองโดยมีเป้าหมายเพื่อเพิ่มประสิทธิภาพในด้านผลตอบแทนเฉลี่ยต่อปี

สำหรับการเปรียบเทียบประสิทธิภาพของการนำข้อมูลทั้งสองประเภทมาใช้พิจารณาร่วมกันนั้น งานวิจัยนี้ได้เลือกใช้แบบจำลองการรู้เชิงลึกที่ใช้ข้อมูลรับเข้าที่แตกต่างกันมาทดสอบบน 3 ชุดข้อมูลซึ่งผลทดลองพบว่าการนำข้อมูลทั้งสองประเภทมาใช้ร่วมกันสามารถเพิ่มประสิทธิภาพในการทำนายทั้งในด้านความแม่นยำและผลตอบแทนเฉลี่ยต่อปีได้ เนื่องจากแบบจำลองการเรียนรู้เชิงลึกมีจุดเด่นในด้านการเรียนรู้ฟีเจอร์ที่ซ่อนอยู่ในข้อมูล ดังนั้นการนำข้อมูลทั้งสองประเภทมาพิจารณาร่วมกัน ย่อมส่งผลให้แบบจำลองการเรียนรู้เชิงสามารถสกัดฟีเจอร์ได้มากขึ้นและส่งผลให้ได้ผลลัพธ์ที่ดีกว่ากว่านำข้อมูลประเภทใด ประเภทหนึ่ง มาใช้สร้างแบบจำลอง

สำหรับการฝึกสอนแบบจำลองโดยใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอซึ่งนำเอาค่าชาร์ปเรโชที่เป็นตัวชี้วัดผลตอบแทนเมื่อเทียบกับความเสี่ยงมาใช้ร่วมกับค่าครอสเอนโทรปีนั้น ผลการทดลองแสดงให้เห็นว่า วิธีการนี้สามารถเพิ่มประสิทธิภาพในด้านผลตอบแทนเฉลี่ยต่อปีและชาร์ปเรโชได้ ในทุก ๆ ชุดข้อมูล และ ทุก ๆ แบบจำลองที่นำมาใช้อ้างอิง แต่อย่างไรก็ตามประสิทธิภาพในด้านความแม่นยำในการทำนายแนวโน้มก็จะลงลงตามไปด้วยเนื่องจากฟังก์ชันวัตถุประสงค์ที่นำเสนอนั้น ได้ทำการเพิ่มความสำคัญของผลลัพธ์ของการทำการจำลองการซื้อขายในกระบวนฝึกสอนแบบจำลอง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

นอกจากนี้วิทยานิพนธ์ฉบับนี้ยังได้มีการทดลองเพื่อศึกษาอิทธิพลของพารามิเตอร์ต่าง ๆ ที่ถูกใช้ในกระบวนการที่นำเสนอ เพื่อให้สามารถนำเอาค่าพารามิเตอร์ที่ดีสุดที่ได้ ไปใช้เป็นค่าตั้งต้นในการทดลองอื่น ๆ ในอนาคตได้ พร้อมทั้งยังได้ทำการทดลองนำเอาแนวคิดและวิธีการที่นำเสนอไปใช้ทำนายหุ้นรายตัว ทำการทำนายล่วงหน้าหลาย ๆ วันและได้ทำการศึกษาผลกระทบจากการนำเอาหัวข้อข่าวแต่ละประเภทไปใช้ทำนาย ซึ่งผลการทดลองก็แสดงให้เห็นว่าแนวคิดและวิธีการที่นำเสนอสามารถให้ผลลัพธ์ที่ดีกว่าวิธีการทั่วไป

6.2 แนวทางการวิจัยถัดไป

สำหรับแนวทางในการวิจัยในอนาคต สามารถแบ่งออกได้เป็น 3 ส่วนดังนี้

1) การนำข้อมูลเชิงตัวอักษรไปใช้ในแบบจำลอง

เนื่องจากข้อมูลรับเข้าที่เป็นข้อมูลเชิงตัวอักษรนั้น จะต้องถูกนำไปสร้างเหตุการณ์ฝังตัวด้วยวิธีการที่นำเสนอโดยงานวิจัยก่อนหน้า [15] ดังนั้นการปรับปรุงประสิทธิภาพของวิธีการสร้างเหตุการณ์ฝังตัวซึ่งเป็นข้อมูลรับเข้าของแบบจำลองก็มีแนวโน้มที่จะเพิ่มประสิทธิภาพในการทำนายได้ นอกจากนี้ การนำหัวข้อข่าวมาทำการวิเคราะห์อารมณ์และความรู้สึกจากข้อความ (Sentiment Analysis) แล้วนำผลลัพธ์ที่ได้ป้อนเข้าสู่แบบจำลอง ก็เป็นอีกหนึ่งแนวทางที่น่าสนใจ ซึ่งอาจจะช่วยเพิ่มคุณภาพของข้อมูลเชิงตัวอักษรที่นำมาใช้ร่วมกับข้อมูลเชิงตัวเลขได้

นอกจากนี้แบบจำลองที่นำเสนอนั้น สามารถรับข้อมูลหัวข้อข่าวที่เป็นภาษาอังกฤษได้เพียงอย่างเดียว ดังนั้นการพัฒนาแบบจำลองให้สามารถพิจารณาข้อมูลข่าวภาษาไทยได้ ก็เป็นสิ่งที่น่าศึกษาและทำการวิจัยต่อไปในอนาคต

2) กระบวนการฝึกสอนแบบจำลอง

สำหรับกระบวนการฝึกสอนโดยใช้ฟังก์ชันวัตถุประสงค์ที่นำเสนอนั้น มีจุดอ่อนที่ต้องใช้ระยะเวลาค่อนข้างนานในการฝึกสอนแบบจำลอง ดังนั้นหากสามารถลดระยะเวลาในส่วนนี้ได้ ก็อาจจะช่วยให้แบบจำลองสามารถนำไปใช้งานจริงได้ดีขึ้น ตัวอย่างเช่น ในขั้นตอนการฝึกสอนแบบจำลองที่แบ่งชุดข้อมูลเป็นหลาย ๆ ช่วงเวลา แล้วทำการสร้างแบบจำลองที่แตกต่างกัน การนำเอาค่านำหนัก (weights) จากแบบจำลองลองในช่วงเวลาก่อนหน้ามาใช้เป็นค่าเริ่มต้น ก็อาจจะช่วยลดระยะเวลาในการฝึกสอนแบบจำลองได้

3) ฟังก์ชันวัตถุประสงค์ที่ใช้ฝึกสอนแบบจำลอง

ในงานวิจัยนี้ได้เลือกใช้ฟังก์วัตถุประสงค์ที่สามารถฝึกสอบแบบจำลองโดยมุ่งเน้นที่ความแม่นยำในการทำนายและด้านผลตอบแทนเมื่อเทียบกับความเสี่ยง ไปพร้อม ๆ กัน แต่อย่างก็ตาม สิ่งที่นักลงทุนส่วนใหญ่สนใจคือผลแทนที่ได้จากจากลงทุน ดังนั้นการพัฒนาฟังก์ชันวัตถุประสงค์ในรูปแบบใหม่ ที่มุ่งเน้นไปที่ผลตอบแทนและออกแบบให้สามารถเลือกปรับค่าความเสี่ยงตามที่ต้องการได้ ก็เป็นอีกแนวทางที่น่าสนใจและสอดคล้องกับเป้าหมายของนักลงทุนส่วนใหญ่

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บรรณานุกรม

[1] B. G. Malkiel, *A Random Walk Down Wall Street,* 1996.

[2] H. Mizuno, M. Kosaka, H. Yajima, and N. Komoda, "Application of neural network to technical analysis of stock market prediction," *Studies in Informatic and Control,* vol. 7, no. 3, pp. 111-120, 1998.

[3] W. Leigh, R. Purvis, and J. M. Ragusa, "Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support," *Decision Support Systems,* vol. 32, no. 4, pp. 361-377, 2002.

[4] H. Gunduz and Z. Cataltepe, "Borsa Istanbul (BIST) daily prediction using financial news and balanced feature selection," *Expert Systems with Applications,* vol. 42, no. 22, pp. 9001-9011, 2015.

[5] B. Wang, H. Huang, and X. Wang, "A novel text mining approach to financial time series forecasting," *Neurocomput.,* vol. 83, pp. 136-145, 2012.

[6] R. P. Schumaker and H. Chen, "Textual analysis of stock market prediction using breaking financial news: The AZFin text system," *ACM Transactions on Information Systems,* vol. 27, no. 2, 2009, Art. no. a12.

[7] G. Gidofalvi, "Using news articles to predict stock price movements," *Department of Computer Science and Engineering, University of California, San Diego,* 2001.

[8] Y. Kim, "Convolutional neural networks for sentence classification," *Convolutional Neural Networks for Sentence Classification,* 2014.

[9] X. Zhang and Y. LeCun, "Text Understanding from Scratch," *ArXiv e-prints,* Accessed on: February 1, 2015Available: http://adsabs.harvard.edu/abs/2015arXiv150201710Z

[10] T. Mikolov, M. Karafiát, L. Burget, C. Jan, and S. Khudanpur, "Recurrent neural network based language model," in *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2010.

[11] O. Irsoy and C. Cardie, "Opinion mining with deep recurrent neural networks," in

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), 2014, pp. 720-728.

[12] C. Zhou, C. Sun, Z. Liu, and F. C. M. Lau, "A C-LSTM neural network for text classification," *arXiv preprint,*

[13] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent Convolutional Neural Networks for Text Classification," in *Proceedings of the National Conference on Artificial Intelligence*, 2015, vol. 3, pp. 2267-2273.

[14] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Using structured events to predict stock price movement: An empirical investigation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 2014, pp. 1415-1425.

[15] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," presented at the IJCAI International Joint Conference on Artificial Intelligence, 2015.

[16] R. Akita, A. Yoshihara, T. Matsubara, and K. Uehara, "Deep learning for stock prediction using numerical and textual information," in *Proceedings of the 15th International Conference on Computer and Information Science (ICIS)*, 2016, pp. 1-6.

[17] M. R. Vargas, B. S. L. P. De Lima, and A. G. Evsukoff, "Deep learning for stock market prediction from financial news articles," in *Proceedings of the 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 2017, pp. 60-65.

[18] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic Financial Trading with Deep Convolutional Neural Networks: Time Series to Image Conversion Approach," *Applied Soft Computing,* 2018.

[19] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111-3119.

[20] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

natural language processing (EMNLP), 2014, pp. 1532-1543.

[21] K. Thanabhat, "Text Categorization for Thai Corpus Using Character-Level Convolutional Neural Netwok," Master, Department of Computer Engineering, Chulalongkorn University, 2016.

[22] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE,* vol. 86, no. 11, pp. 2278-2324, 1998.

[23] W. F. Sharpe, "The sharpe ratio," *Journal of portfolio management,* vol. 21, no. 1, pp. 49-58, 1994.

[24] Y. Zhai, A. Hsu, and S. K. Halgamuge, "Combining news and technical indicators in daily stock price trends prediction," in *International symposium on neural networks*, 2007, pp. 1087-1096.

[25] *Stanford Open Information Extraction*. Available: https://nlp.stanford.edu/software/openie.html

[26] *Daily News for Stock Market Prediction*. Available: https://www.kaggle.com/aaron7sun/stocknews/

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน

ชื่อ-สกุล	พิศุทธ อ่อนเจริญ
วัน เดือน ปี เกิด	20 กรกฎาคม 2533
สถานที่เกิด	จังหวัดสงขลา
วุฒิการศึกษา	วศ.บ. (เกียรตินิยมอันดับหนึ่ง) วิศวกรรมเครื่องกล มหาวิทยาลัยสงขลานครินทร์ (พ.ศ. 2551 - 2554)
ที่อยู่ปัจจุบัน	545 ซอยสวนพลู 1 ถนนสาทร 3, ทุ่งมหาเมฆ, สาทร, กทม. 10120
ผลงานตีพิมพ์	P. Oncharoen and P. Vateekul, "Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators", 2018 - 5th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA 2018) P. Oncharoen and P. Vateekul, “Deep Learning Using Risk-Reward Function for Stock Market Prediction”, 2018 - 2nd International Conference on Computer Science and Artificial Intelligence (CSAI 2018)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University