USING DEEP LEARNING TECHNIQUE FOR PRICE ACTION ANALYSIS IN GOLD TRADING

Mr. Kwan Boon-long

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

การใช้เทคนิคการเรียนรู้เชิงลึกเพื่อการวิเคราะห์เชิงพฤติกรรมราคาสำหรับการซื้อขายทองคำ

นายขวัญ บุญ-หลง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2564

Thesis Title	USING DEEP LEARNING TECHNIQUE FOR PRICE ACTION ANALYSIS IN GOLD TRADING
By	Mr. Kwan Boon-long
Field of Study	Computer Science
Thesis Advisor	PITTIPOL KANTAVAT, Ph.D.

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

-- Dean of the FACULTY OF ENGINEERING

(Professor SUPOT TEACHAVORASINSKUN, Ph.D.)

THESIS COMMITTEE

-- Chairman

(Professor BOONSERM KIJSIRIKUL, Ph.D.)

-- Thesis Advisor

(PITTIPOL KANTAVAT, Ph.D.)

-- External Examiner

(Assistant Professor Kridsda Nimmanunta, Ph.D.)

ขวัญ บุญ-หลง : การใช้เทคนิคการเรียนรู้เชิงลึกเพื่อการวิเคราะห์เชิงพฤติกรรมราคาสำหรับการซื้อขายทองคำ. (USING DEEP LEARNING TECHNIQUE FOR PRICE ACTION ANALYSIS IN GOLD TRADING) อ.ที่ปรึกษาหลัก : อ. ดร.พิตติพล คันธวัฒน์

ในวิทยานิพนธ์นี้นิสิตได้ใช้เทคนิคการเรียนรู้เชิงลึกเพื่อจุดประสงค์หลักในการเรียนรู้รูปแบบพฤติกรรมราคาของกราฟราคาทองคำ ซึ่งเป็นวิธีที่นักวิเคราะห์นิยมใช้ในการซื้อขายทองคำ การวิเคราะห์เชิงพฤติกรรมราคานั้น เป็นการวิเคราะห์ และทำนาย โดยดูจากแบบแผน (pattern) ที่เปลี่ยนแปลง และส่งผลต่อการปรับเปลี่ยนของราคา เพื่อทำนายราคาของทองคำในอนาคต โดยวิธีที่นิสิตเสนอนั้นเป็นการนำโครงข่ายประสาทแบบคอนโวลูชัน ซึ่งมีความสามารถในการเรียนรู้แบบแผนผสมผสานกับโครงข่ายประสาทเทียมแบบวนกลับชนิดพิเศษ Long Short-Term Memory (LSTM) ซึ่งมีความสามารถในการเรียนรู้ลำดับขั้น โดยพบว่าผลการทดสอบการซื้อขายย้อนหลังด้วยวิธีที่เสนอสามารถสร้างผลตอบแทนได้ดีกว่าวิธีการซื้อขายทองคำแบบดั้งเดิม เนื่องจากโครงสร้างการเรียนรู้เชิงลึกแบบผสมผสานโครงข่ายประสาทแบบคอนโวลูชัน และโครงข่ายประสาทเทียมแบบ LSTM มีความสามารถในการเรียนรู้ และจดจำรูปแบบการเปลี่ยนแปลงของแบบแผนพฤติกรรมของราคาในอดีตเทียบกับปัจจุบัน เพื่อทำนายอนาคตได้อย่างแม่นยำ

สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์	ลายมือชื่อนิสิต ....
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

6370031421 : MAJOR COMPUTER SCIENCE
KEYWORD: Gold Trading, Machine Learning, Deep Learning, Trading Signal, Technical Analysis, Price Action

Kwan Boon-long : USING DEEP LEARNING TECHNIQUE FOR PRICE ACTION ANALYSIS IN GOLD TRADING. Advisor: PITTIPOL KANTAVAT, Ph.D.

We introduce a new structural deep learning model purposing is to be able to learn price action trading features that subjective traders are using to make trading decisions based on the visual recent and actual price movements, rather than relying solely on technical indicators. The model combines convolutional neural networks (CNN) and long short-term memory (LSTM) to improve the trend forecasting of gold prices for better trading signals compared to traditional strategies. As the gold price is a time series data, it is appropriate to apply CNN and LSTM for forecasting. The concept of our model is that CNN could detect price action features or patterns in different locations of time series data; while, LSTM could maintain both short-term and long-term memory as a sequence along with time series data. The collaboration of their abilities could help the neural network model understand complex relationships between recent and actual price movements and trends in gold prices. Our study found that the combining of CNN and LSTM with price action trading features could significantly enhance trading performance in the long run.

Field of Study:	Computer Science	Student's Signature ....
Academic Year:	2021	Advisor's Signature ....

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my thesis advisor, Dr. Pittipol Kantavat for his invaluable help and constant encouragement throughout the course of this research. I am most grateful for his teaching and advice. I would not have achieved this far and this thesis would not have been completed without all the support that I have always received from him. In addition, I would like to thank Prof. Dr. Boonserm Kijsirikul, Asst. Prof. Dr. Kridsda Nimmanunta, and others person for suggestions and all their help. Finally, I most gratefully acknowledge my parents and my friends for all their support throughout the period of this research.

Kwan Boon-long

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

วิทยานิพนธ์

(THESIS)

ชื่อเรื่อง (ภาษาไทย)	การใช้เทคนิคการเรียนรู้เชิงลึกเพื่อการวิเคราะห์เชิงพฤติกรรมราคาสำหรับการซื้อขายทองคำ
ชื่อเรื่อง (ภาษาอังกฤษ)	USING DEEP LEARNING TECHNIQUE FOR PRICE ACTION ANALYSIS IN GOLD TRADING
เสนอโดย	นาย ขวัญ บุญ-หลง
รหัสนิสิต	6370031421
หลักสูตร	วิทยาศาสตร์มหาบัณฑิต
ภาควิชา	วิศวกรรมคอมพิวเตอร์
คณะ	วิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
สถานที่ติดต่อ	ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพ 10330
โทรศัพท์	(+66)2-218-6956, (+66)2-218-6957
อีเมล์	6370031421@student.chula.ac.th
อาจารย์ที่ปรึกษา	อ.ดร. พิตติพล คันธวัฒน์
คำสำคัญ (ภาษาไทย)	การซื้อขายทองคำ, การเรียนรู้ของเครื่อง, การเรียนรู้เชิงลึก, สัญญาณการซื้อขาย, การวิเคราะห์เชิงเทคนิค, พฤติกรรมราคา
คำสำคัญ (ภาษาอังกฤษ)	Gold Trading, Machine Learning, Deep Learning, Trading Signal, Technical Analysis, Price Action

Chapter 1

Introduction

For decades, gold has been used as a precious metal for investment purposes and for manufacturing specific electronic and medical devices. Nowadays, the demand for gold is higher, which results in larger price volatilities. Many factors drive the price of gold, such as Central banks' reserves, the value of USD, Worldwide Jewelry, Industrial Demand, Wealth Protection, and Gold Production [1]. Financial knowledge is indispensable in our lives, and a precise forecast in the financial area will help support the development of capital market and facilitate the development of economy.

Gold traders have used technical analysis for a long time. We can categorize technical analysis trading into two types. First, Mechanical Trading uses technical indicators such as MACD, RSI, and Stochastic Oscillators to generate signals for buying and selling. Second, Price Action Trading is a trading technique that allows a trader to read the market and make subjective trading decisions based on the recent and actual price movement patterns, rather than relying solely on technical indicators. Price Action Trading has room for judgement while no judgement is relied upon in Mechanical Trading.

A number of studies in the past that used traditional statistical techniques to predict the price of the stock, index, currency, and commodities, such as Autoregressive Integrated Moving Average (ARIMA) [2], Random Walk, and Genetic algorithm (GA) [3]. Some researchers also introduced Machine Learning in stock price prediction, Support Vector Machine (SVM) [2] was one of the famous Machine Learning techniques used in stock price prediction. The deep learning technique was introduced in the investment field in recent years. Convolutional neural network (CNN) [4] and Recurrent neural network (RNN) [5] have been the most popular deep learning models for stock price prediction.

Most researchers studied Mechanical Trading, but not many researchers studied Price Action Trading. Many institutional traders have been using Price Action analysis to trade and make profits for decades. Still, they hardly interpret how they trade into defined rules. They are free to make their own decisions for a given scenario to take trading positions, as per their subjective views, experiences, behavioral, and psychological states.

This research studied the Price Action Trading type and combined with deep learning techniques to replace the judgement part in Price Action Trading. We proposed the CNN – LSTM model to learn such visual features movements of a price pattern and moving average. To evaluate our proposed techniques, our results are benchmarked against traditional technical analysis tools and other deep learning techniques.

Chapter 2

RELATED THEORY

There are two related theories 1. Price Action Trading and 2. Deep Learning

2.1 Price Action Trading

Price action describes the characteristics of a security's price movements. This movement is often analyzed with respect to price changes in the recent past. In simple terms, Price Action is a trading technique that allows a trader to read the market and make subjective trading decisions based on recent and actual price movements [6]. A moving average is considered the most common tool for Price Action traders. It represents previous price action for a specific period as a smooth line. For instance, a 50-period simple moving average on the daily chart uses the past 50 days of price action to form a smoothed average.

There are two basic ways that Price Action traders use Moving averages [7]: to determine the market trend and to determine key support and resistance levels.

2.1.1 TO DETERMINE MARKET TREND

The most common usage is comparing the current price with the moving average representing the investor's time horizon. The moving average tends to follow the trend line. For example, many investors use a 200-day moving average. The trend is considered upward if the stock or market average is above its 200-day moving average. Conversely, the trend is considered downward if the stock or market average is below the 200-day moving average. The moving average then becomes a proxy for the trend line and can be used to determine when a trend is potentially changing its direction. In Fig. 1, for example, the later prices have held at both the trend line and the moving average.

Figure 1: Determine Market Trend using Period 200 SMA

2.1.2 TO DETERMINE KEY SUPPORT AND RESISTANCE LEVELS

The moving average often acts as dynamic support or resistance. In Fig. 2, the moving average often acts as support and resistance; therefore, it can be an easy trailing stop mechanism for determining when a position should be liquidated or reduced. In addition, price action tends to reverse when it goes across the curve of moving averages. Most market participants, such as floor traders, institutional investors, use moving averages to time their entries and exits. These may be the reason moving average acts as support and resistance.

Figure 2: Determine Key support and resistance using Period 25 EMA

2.2 Deep Learning Theory

2.2.1 Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) [8] is one of the most powerful Machine Learning techniques; hence, it has been applied in almost the problem domain. It can be categorized into single-layer and multi-layer neural networks. In the single-layer network, Fig. 3, a set of inputs is directly mapped to output using a generalized variation of a linear function. This simple instantiation of a neural network is also referred to as the perceptron. In multi-layer neural networks, Fig. 4, the neurons are arranged in a layered fashion, in which the input and output layers are separated by a group of hidden layers. This layer-wise architecture of the neural network is also referred to as a feed-forward network.

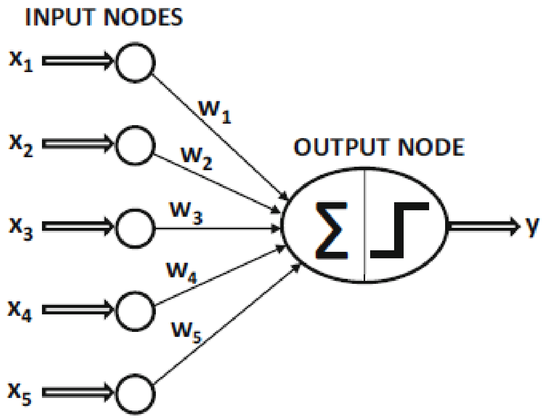

Figure 3: Single layer neural networks

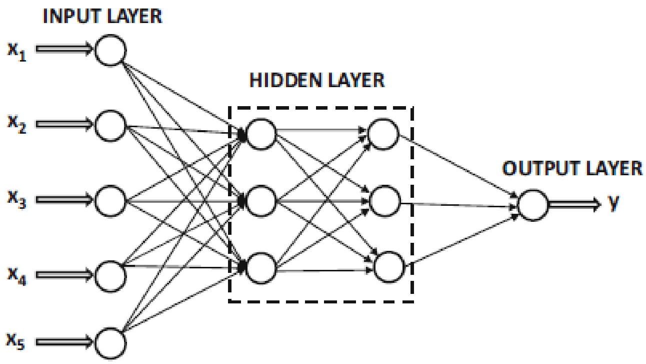

Figure 4: multi-layer neural networks

2.2.2 Convolutional Neural Network (CNN)

The Convolutional neural network (CNN) is a neural network that is widely used for computer vision tasks. In contrast to the original neural network that fully connects all pairs of nodes between layers, CNN decreases the number of nodes between the previous layer and its layer to simplify the network by having a sequence of convolution operations and subsampling operations, as shown in Fig. 5. The fundamental difference between a densely connected layer and a convolution layer is Dense layers learn global patterns in their input feature space, whereas convolution layers learn local patterns, see Fig. 6, in the case of images, patterns found in small 2D windows of the inputs, these filters were all 5 × 5.

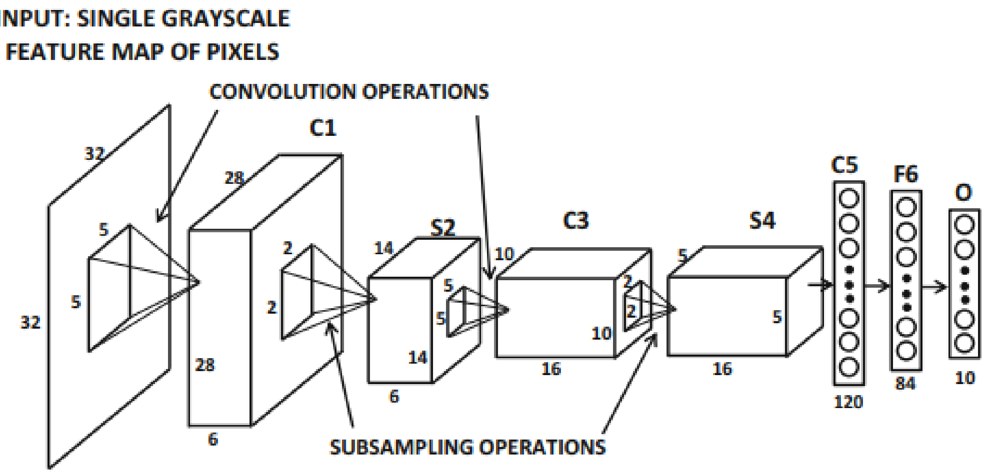

Figure 5: CNN structure

This fundamental characteristic gives CNN two interesting properties. First, the patterns they learn are translation invariant. After learning a certain pattern in the lower-right corner of a picture, a convolution can recognize it anywhere: for example, in the upper-left corner. A densely connected network would have to learn the pattern anew if it appeared at a new location. This mechanism makes convolution data efficient when processing images. Second, It can learn spatial hierarchies of patterns (see Fig. 7). A first convolution layer will learn small local patterns such as edges. A second convolution layer will learn larger patterns made of the features of the first layers, and so on. This allows convolution to efficiently learn increasingly complex and abstract visual concepts.

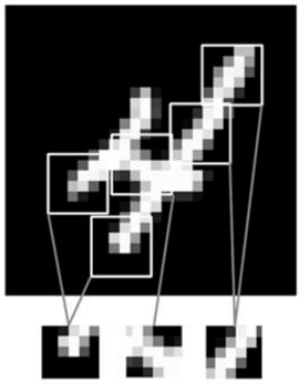

Figure 6: Local patterns

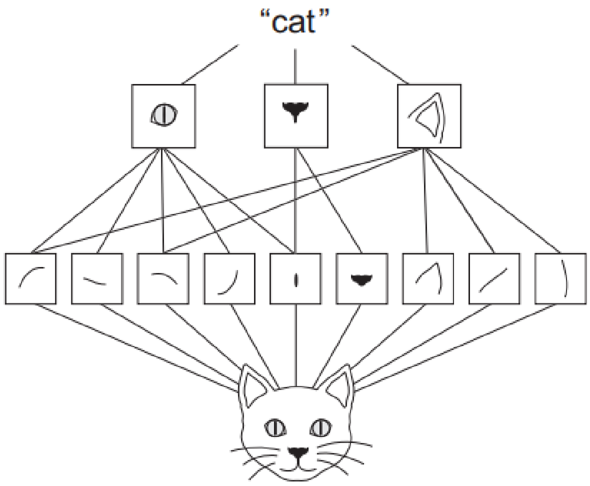

Figure 7: Larger local patterns in next layer

1D convolution [9] is a subset of convolution neural networks which can be used to learn time series data. It can recognize local patterns in a sequence. Because the same input transformation is performed on every patch, a pattern learned at a certain position in a sentence can later be recognized at a different position, making 1D convolution translation-invariant (for temporal translations). For instance, in Fig. 8, a 1D convolution processing sequence of characters using convolution windows of size five should be able to learn words or word fragments of length five or less, and it should be able to recognize these words in any context in an input sequence.

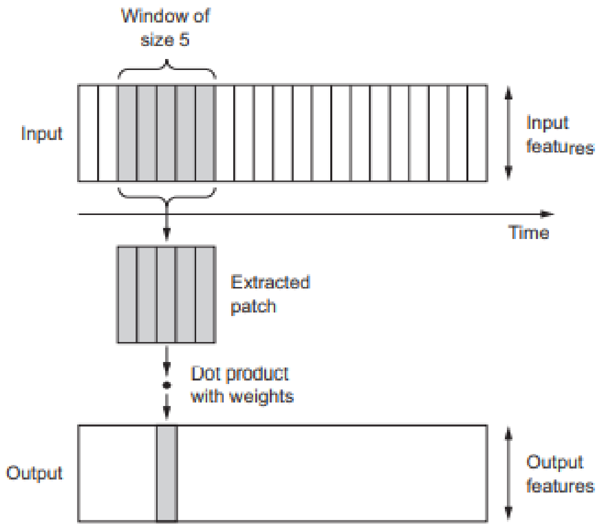

Figure 8: 1D Convolution

1D convolution with multiple input channels. Given we have 2 input channels with 3-time steps. The kernel will be initialized with 2 channels. So even though we are using a 1D Convolution, we have a 2D kernel. A 1D Convolution just means we slide the kernel along one dimension, it doesn't necessarily define the shape of the kernel, since that depends on the shape of the input channels.

2 Input channels, with 3-time steps

A1	A2	A3
B1	B2	B3

Kernel

C1	C2
C3	C4

The output (2x1) is calculated per below.

(C1xA1) + (C2xA2) + (C3xB1) + (C4xB2)	(C1xA2) + (C2xA3) + (C3xB2) + (C4xB3)

2.2.3 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) is different from other neural networks structure. The major characteristic of all neural networks we have seen so far, such as ANN and CNN, is that they have no memory. Each input shown to them is processed independently, with no state kept in between inputs. With such networks, in order to process a sequence or a temporal series of data points, you have to show the entire sequence to the network at once. In contrast, biological intelligence processes information incrementally while maintaining an internal model of what it's processing, built from past information and constantly updated as new information comes in. A recurrent neural network (RNN) adopts the same principle, albeit in an extremely simplified version.

It processes sequences by iterating through the sequence elements and maintaining a state containing information relative to what it has seen so far. In effect, an RNN is a type of neural network that has an internal loop (see Fig. 9). The state of the RNN is reset between processing two different, independent sequences, so you still consider one sequence a single data point: a single input to the network. What changes is that this data point is no longer processed in a single step; rather, the network internally loops over sequence elements (see Fig. 10).

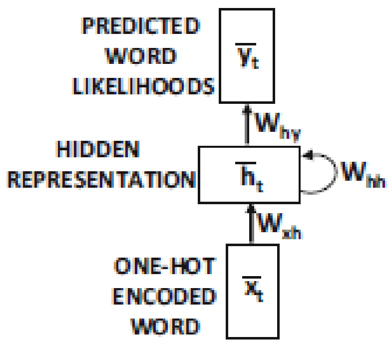

Figure 9: RNN with internal loops

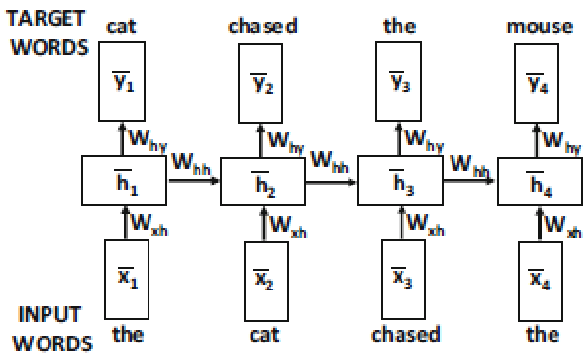

Figure 10: RNN loops over sequence elements

Hochreiter and Schmidhuber [10] proposed Long Short-Term Memory (LSTM), a Recurrent Neural Network (RNN) subset, to solve a problem in a recurrent neural network called the vanishing gradient. It could happen when the data sequence is too long, leading to a vanishingly small gradient to update the weight. LSTM learns long data sequences using four gates, as shown in Fig. 11. A forget gate controls the memory of the previous data sequence. Two input gates control new information. An output gate controls the output of hidden states.

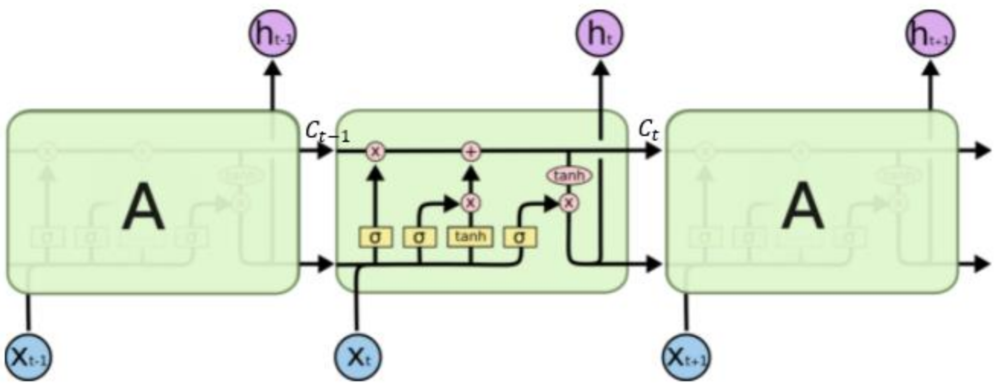

Figure 11: LSTM structure

Chapter 3

LITERATURE REVIEW

Considerable researchers used traditional statistic techniques to predict the price of stock, indexes, currency, and commodities. Zhao et al. [11] proposed Autoregressive Integrated Moving Average (ARIMA) model to predict world crude oil, and the results showed that the model performed well in short-term prediction. Several years later, machine learning techniques have been widely used, and Support Vector Machine (SVM) was one of the famous techniques used in stock price prediction. Kantavat and Kijsirikul [12] proposed the Support Vector Machine method by combining technical analysis and indicators to predict stock price. Yu et al. [2] also used the Support Vector Machine method in crude oil price forecasting.

In recent years after deep learning has been developed, many papers used this technique to develop a new trading strategy in various asset classes. Kietikul [13] conducted CNN architecture for candlestick chart pattern images to predict the direction of the short trend. The proposed architecture performs fairly well predicting the unseen pattern. Ju et al. [4] proposed an approach CNN with market profile (CNN-MPs) which is adapting Market Profile to covert financial time series data to a grey-scale image first, then generating two types of learning profile images. These stacked and sequential profiles that keep the interaction between continuous profiles, then build a CNN model that can classify future price trends. The accuracy of the proposed method was higher than the based LSTM model.

Some of the papers use more complexity in the deep learning model. Eapen et al. [14] proposed a multiple pipeline deep learning model with CNN and Bi-Directional LSTM to predict the S&P500 index. They used the daily S&P500 dataset from 2008 to 2014 as a training set and from 2015 to 2016 as a test set. The index data was segmented into sliding window sequences of 50 closing prices. They constructed a deep learning architecture consisting of three pipelines of separate layers (in Fig. 12).

Each pipeline consists of 1D CNN with a ReLU activation followed by Maxpooling before passing data to the Bi-directional LSTM neural layers. The output of each three pipelines is concatenated and given to the final layer consisting of a dense layer. The final output is a numeric value of the predicted price seven days into the future. The proposed model improved performance prediction by 9% compared to the based deep learning model.

Hao et al. [15] also proposed a multiple pipeline deep learning model to predict the price trend of the stock market index. The difference in this paper was all pipelines were based on multiple time frame feature learning. See Fig. 13, F1, F2, and F3 reflected relative short-term, medium-term, and long-term trend changes, respectively. The experimental results demonstrated that such a hybrid network could enhance the predictive performance compared with benchmark networks. Combining multiple time scale features can help promote accurate prediction of the proposed model and learned valuable information.

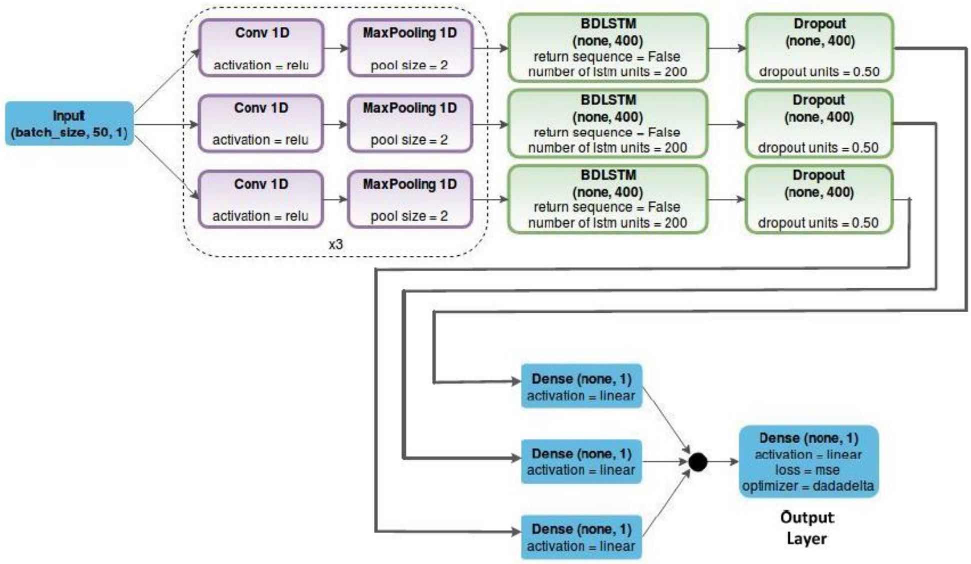

Figure 12: A multiple pipeline deep learning model

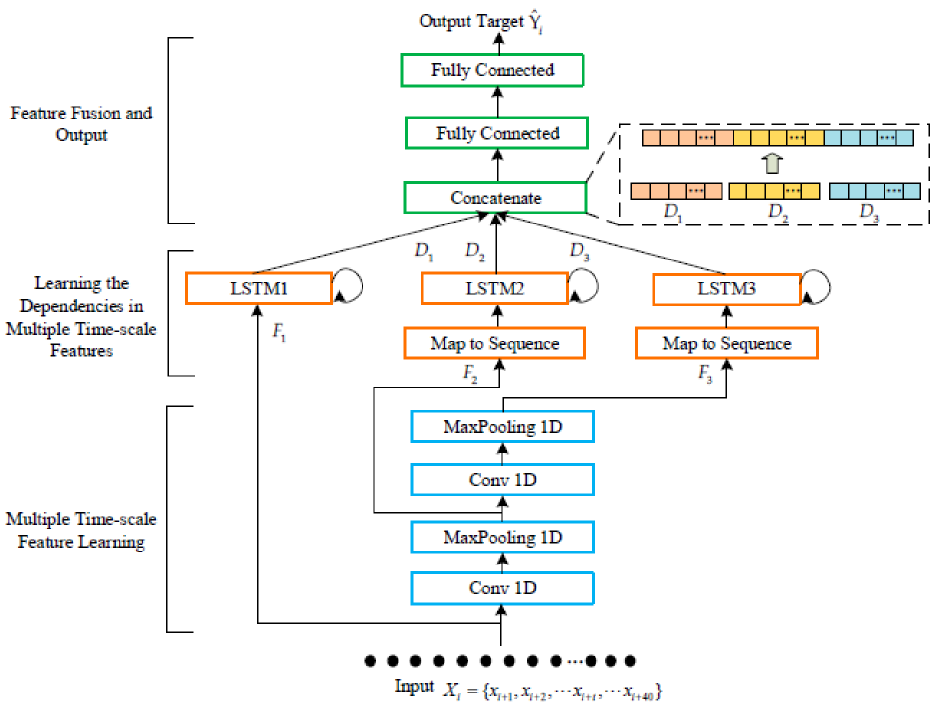

Figure 13: Time frame based multiple pipeline deep learning model

Chapter 4

PROPOSED METHOD

We proposed CNN – LSTM model to learn movements of Price Action features. Important Price Action features that were selected as input are 1. Price features 2. The slope of SMA (angle of the dashed line in Fig. 14) 3. Distance of close price to SMA (length of arrow line in Fig. 14).

Two examples in Fig. 14 (a.) and (b.), with a 32-period sliding window, explain how the model could learn these features' movements and predict the daily gold price. At (a), the price is moving sideways with low volatility, the slope of SMA is continuing upward, and the distance line length becomes shorter, predicting the next day's price to go up. At (b), the price is moving sideways with medium volatility, the slope of SMA is changing from upward to sideways, and the length of the Distance line is changing from positive to negative, which predicts the price of the next day to go down.

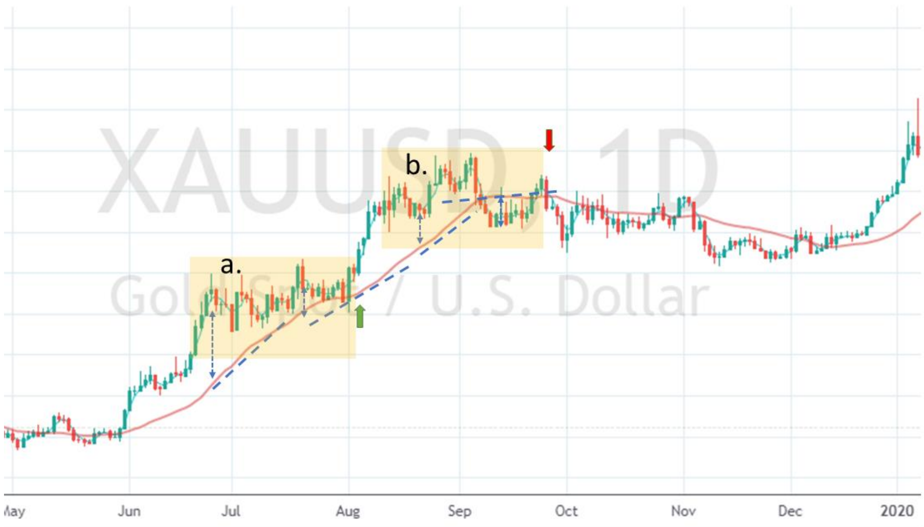

Figure 14: Example of Price Actions Trading ideas

The proposed CNN + LSTM model in Fig. 15 will receive inputs as a set of timesteps of Price Action features listed in Table 1. All the features listed in the table represent three important Price Action features we have explained in the first paragraph The prices are a combination of $ret1_t$, $ret5_t$, $std5_t$, $volume\_by\_ave12_t$, $volume\_by\_ave26_t$, $(H-L)_t$ and $(O-C)_t$. Slope of SMA is a combination of $slope12_t$ and $slope26_t$. The distance of Close price to SMA is a combination of $distance12_t$ and $distance26_t$.

Table 1: The input features calculation

Features	Details
$\mathrm{ret1_t}$	1 Day return or Percent Change of Close price $\mathrm{ret1_t} = \frac{\mathrm{C_t - C_{t-1}}}{\mathrm{C_{t-1}}}$ (1)
$\mathrm{ret5_t}$	5 Day return or Percent Change of Close price $\mathrm{ret5_t} = \frac{\mathrm{C_t - C_{t-5}}}{\mathrm{C_{t-5}}}$ (2)
$\mathrm{std5_t}$	Standard deviation of 5 Day return $\mathrm{ave\_ret5_t} = \frac{1}{5}\sum_{\mathrm{x=0}}^{4} \mathrm{ret1_{t-x}}$ (3) $std5_t = \sqrt{\frac{\sum_{x=0}^{4}(ret1_{t-x} - ave\_ret5_t)}{5}}$ (4)
$\mathrm{volume\_by\_ave12_t}$	Volume devided by 12 days average of Volume $\mathrm{volume12_t} = \frac{1}{12}\sum_{\mathrm{x=0}}^{11} \mathrm{V_{t-x}}$ (5) $\mathrm{volume\_by\_ave12_t} = \frac{\mathrm{V_t}}{\mathrm{volume12_t}}$ (6)
$\mathrm{volume\_by\_ave26_t}$	Volume devided by 26 days average of Volume $volume26_t = \frac{1}{26}\sum_{x=0}^{25} V_{t-x}$ (7) $\mathrm{volume\_by\_ave26_t} = \frac{\mathrm{V_t}}{\mathrm{volume26_t}}$ (8)
$\mathrm{(H-L)_t}$	$\mathrm{(H-L)_t = high_t - low_t}$ (9)
$\mathrm{(O-C)_t}$	$\mathrm{(O-C)_t = open_t - close_t}$ (10)
$\mathrm{distance12_t}$	$SMA12_t = \frac{1}{12}\sum_{x=0}^{11} C_{t-x}$ (11)

	$\text{distance12}_t = \text{close}_t - \text{SMA12}_t$ (12)
distance26_t	$SMA26_t = \frac{1}{26}\sum_{x=0}^{25} C_{t-x}$ (13) $\text{distance}_t = \text{close}_t - \text{SMA26}_t$ (14)
slope12_t	Simple Moving Average period = 12 $SMA12_t = \frac{1}{12}\sum_{x=0}^{11} C_{t-x}$ (15) $\text{slope}_t = \text{SMA12}_t - \text{SMA12}_{t-1}$ (16)
slope26_t	Simple Moving Average period = 26 $\text{SMA26}_t = \frac{1}{26}\sum_{x=0}^{25} \text{C}_{t-x}$ (17) $\text{slope}_t = \text{SMA26}_t - \text{SMA26}_{t-1}$ (18)

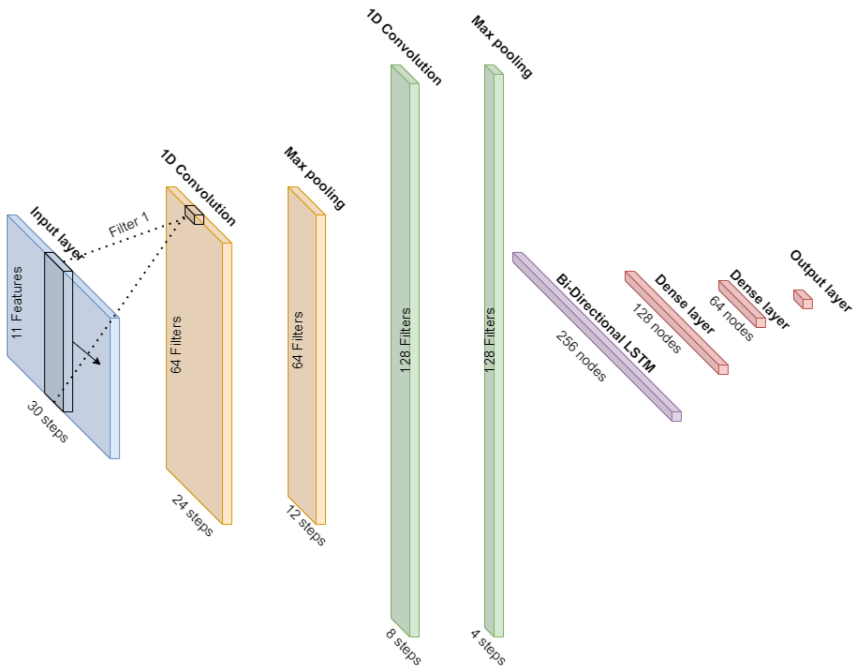

Figure 15: Proposed CNN – LSTM architecture

Chapter 5

EXPERIMENTAL DATASET

We use the daily gold price during the period 2006 to 2021 from Yahoo! Finance. The daily historical data includes open, high, low, close, and volume. Then, we calculate those pricing data into price action features, as shown in Table 1. The data is then separated into six groups by sliding windows. We set the name of each group corresponding to the trading back-testing year from 2016 to 2021. Each group is split into training data (9 years), validation data (1 year after training period), and test data (1 year after validation period), as demonstrated in Table 2.

Table 2: Training, validation, and test data for the experiment.

Dataset	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
2016																
2017																
2018																
2019																
2020																
2021																

Training Validation Test

Chapter 6

DATA PREPROCESSING AND LABELING

We transform the data using two techniques to make the time series stationary. Firstly, we apply relative change compared to the previous time-step to get 1-day and 5-day price returns for ret1 (1) and ret5 (2), respectively. In addition, we calculate std5 as the standard deviation of the past five days' return (4) to give information on how the return is changing due to the market's changes. We then calculate volume_by_ave12 (6) and volume_by_ave26 (8) to extract the trend of the volume in the short term and long term. H-L (9) is the range of daily price, and O-C (10) is the price change within a day.

We then calculate important price action features, which are the distance of the close price to the short-term moving average as distance12 (12) and the distance of the close price to the long-term moving average as distance26 (14). The slope of the short-term trend and the slope of the long-term trend is calculated as slope12 (16) and slope26 (18), respectively.

After that, all data are normalized in the range [0, 1] before training neural network models.

$$Scaled\ X_t = \left(\frac{X_t - X_{min}}{X_{max} - X_{min}}\right) \tag{19}$$

The dataset is then processed into sliding windows as a set of 30-day look-back time steps before being split into the training and validation set. See Table 3 for calculation of output as a next-day return. This time series represents the next-day return of gold price.

Table 3: The output features calculation.

output variable	*Details*
$retNextDay1_t$	$Next\ Day\ return\ or\ Percent\ Change\ of\ next\ day\ Close\ price$ $$retNextDay1_t = \frac{C_{t+1} - C_t}{C_t} \quad (20)$$

Chapter 7

EXPERIMENTAL RESULTS

This section shows the experimental results and evaluation method. These performances are assessed and compared to other deep learning techniques and traditional strategies, including Buy and Hold, RSI, Stochastics, and MACD.

There are 2 Evaluation Methods 1. Prediction Error and 2. Trading Performance

7.1 Prediction Error

The models are trained using training data and evaluated using Mean Square Error (MSE) (21). We set the number of 1st layer convolution filter equal to 64 with a length of 7, and 2nd layer convolution filters equal to 128 with a length of 5. The LSTM is a Bi-Directional type with numbers of nodes equal to 128. The epoch is set to 100 with a batch size equal to 80. After that, we compare trading performance evaluation using test data. The structures of the single CNN and single LSTM are shown in Fig. 16 and Fig. 17, respectively.

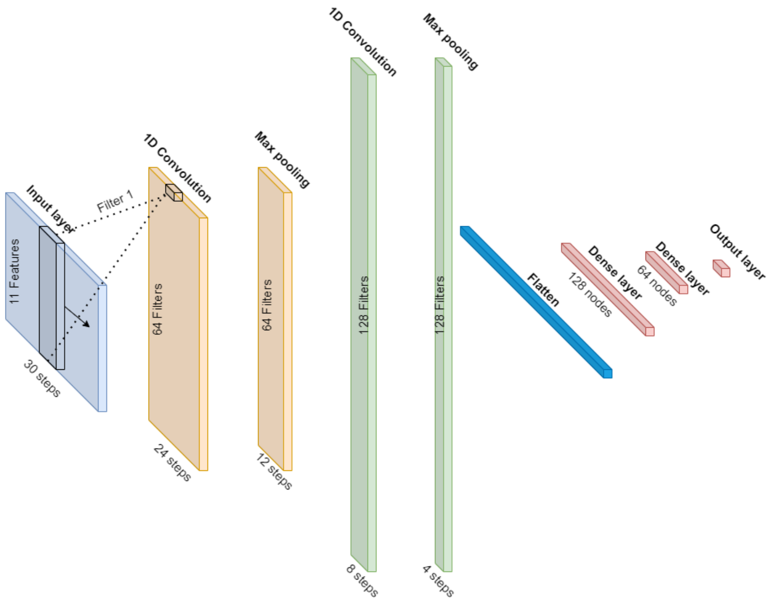

Figure 16: Structure of single CNN model.

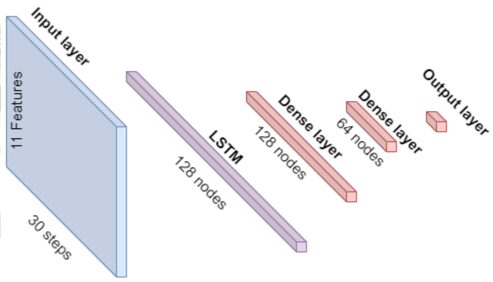

Figure 17: Structure of single LSTM model.

$$MSE = \frac{\sum_{t=1}^{n} (\hat{y}_t - y_t)^2}{n} \qquad (21)$$

The MSEs of the test data for each group are shown in Table 4. The bold texts indicate the lowest MSE of each year. The results show that single LSTM outperforms single CNN and CNN-LSTM structures. However, the graph plot between the predicted value and actual value showed that LSTM predicts as an average medium line value (in Fig. 18) compared to the plot of single CNN (in Fig. 19) and CNN-LSTM (in Fig. 20), which are more realistic and able to capture the price movement better.

Table 4: Average MSE of the test set for each data group

	CNN-LSTM	CNN	LSTM
2016	3.89	4.56	**3.11**
2017	2.12	2.05	**1.26**
2018	1.66	1.90	**1.36**
2019	1.86	2.32	**1.59**
2020	7.78	7.89	**5.45**
2021	3.82	3.76	**2.55**
Avg.	3.52	3.75	**2.55**

The units of MSE on the table are 10^{-3}

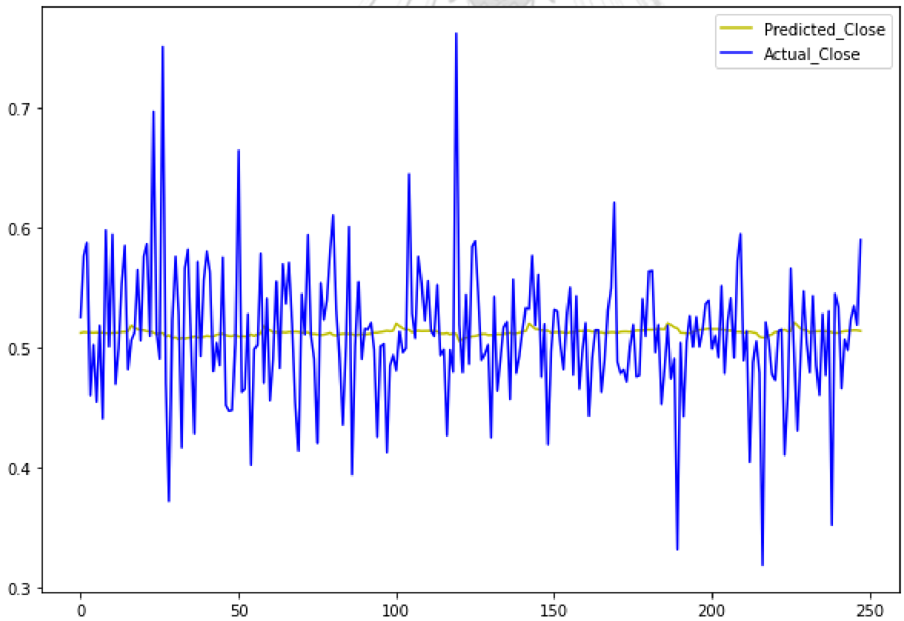

Figure 18: Comparing LSTM predicted and actual change of close price.

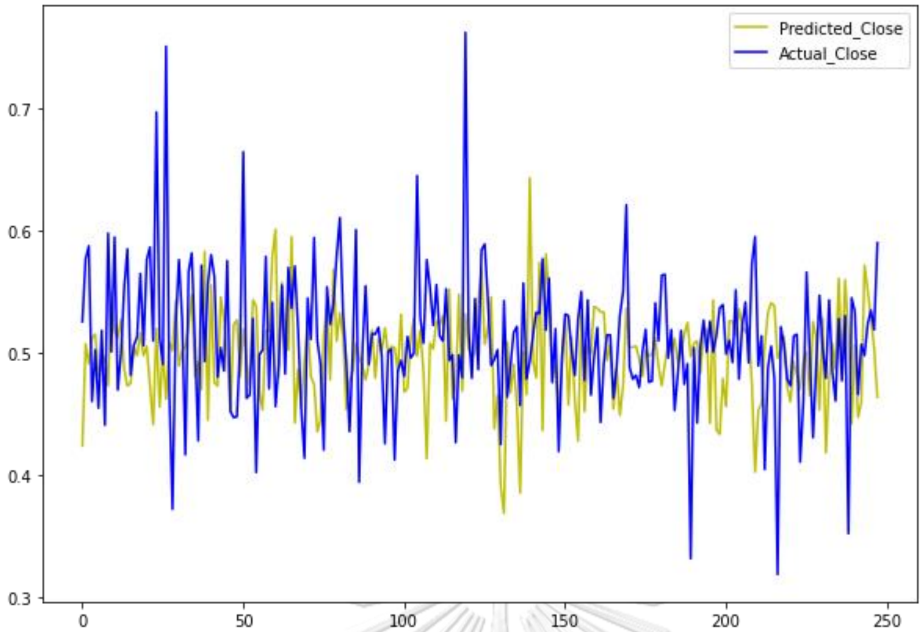

Figure 19: Comparing CNN predicted and actual change of close price.

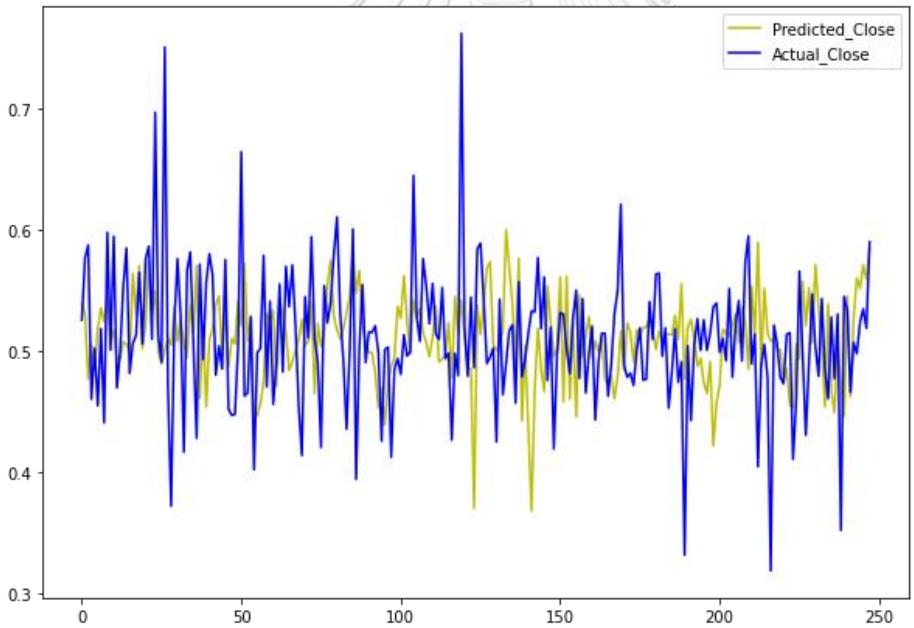

Figure 20: Comparing CNN-LSTM predicted and actual change of close price.

7.2 Trading Performance

We deploy the trained models, including CNN, LSTM, and CNN-LSTM, to perform trading backtest. A backtest is a trading simulation based on buying signals and selling signals. The trading signal is triggered by the predicted daily return from the model. If the predicted daily return is positive, it will trigger the buy signal. In this experiment, we set the signal threshold as zero.

The performance was evaluated by three trading evaluation metrics, including 1. Return on investment (ROI), 2. Sharpe ratio and 3. Treynor and Mazuy method. The ROI measure profitability of an investment calculated by (22). The Sharpe ratio [16] or the reward-to-variability ratio is a widely used metric for risk-adjusted return of portfolios, calculated by the ratio of return and risk as shown in (23). The Treynor and Mazuy method [17] is a quadratic version of the CAPM (24). R_f is the risk-free rate which is the simple average return of the 10-year Treasury rate during 2016 – 2021, equals to 2%. Since this research focuses only at the timing ability of gold, R_M represents the return from buy-and-hold strategy which equals to the gold spot price return in the same period. β_2 is the coefficient of our interest, as it would be statistically significant if our proposed strategy exhibits the market timing skills.

$$ROI = \frac{Portfolio\,Value_t - Portfolio\,Value_{t-1}}{Portfolio\,Value_{t-1}} \tag{22}$$

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p} \tag{23}$$

$$R_p - R_f = \alpha_p + \beta_1(R_M - R_f) + \beta_2(R_M - R_f)^2 \tag{24}$$

ROI for each group is shown in Table 5. The bold texts indicate the highest ROI of each year. The green cells indicate positive ROI, and the red cell indicates negative ROI. The average ROI of all groups is shown in the lowest row of the table. The average annualized ROI of CNN-LSTM at 27% is the highest among all strategies, while single CNN provided a lower annualized ROI at 16%. CNN-LSTM and single CNN are the only two strategies that provide 2-digit ROI for all years. The average annualized ROI from the single LSTM underperforms CNN-LSTM, single CNN, and some traditional strategies with an annualized ROI of 4%.

The Sharpe ratio of each group is shown in Table 6. The average Sharpe ratio of each group is shown in the lowest row of the table. CNN-LSTM yields the highest Sharpe ratio at 1.57. Subsequently, single CNN and Buy and Hold provide Sharpe ratios at 0.96 and 0.38, respectively.

The results of Treynor and Mazuy are shown in Table 7. We plotted a graph between $R_p - R_f$ and $R_M - R_f$, then fit the polynomial regression curve, then computed and compared the value of β_2 of each group. The average β_2 of each group is shown in the lowest row of the table. CNN-LSTM gives the highest β_2 among all strategies at 22.5, while single CNN provided a lower β_2 at 10.9; this showed that CNN-LSTM gave the best market-timing capacity according to Treynor and Mazuy method.

According to Table 5, Table 6, and Table 7, a combination of CNN-LSTM can perform the highest ROI, Sharpe ratio, and β_2 among all strategies. The ROI and Sharped ratio of CNN-LSTM outperformed all strategies almost yearly except in 2017 and 2021. Although the MSE metrics performance of CNN-LSTM is not distinguishable compared to LSTM, which provides a lower MSE, from visualization prediction vs. actual plot, LSTM provided a lower accuracy and more medium line result compared to CNN-LSTM and single CNN.

The performance of the combinations of CNN and LSTM could enhance performance from a single CNN model or single LSTM model. As the CNN - LSTM has more capability to learn and understand the complex relationship between movements of Price Action features and output.

Table 5: ROI of all trading strategies

	BH	CNNLSTM	CNN	LSTM	MACD	STO	RSI
2016	6%	**40%**	16%	5%	-10%	35%	-10%
2017	9%	19%	11%	-7%	**22%**	-12%	-5%
2018	-4%	**16%**	14%	3%	-18%	4%	7%
2019	18%	**22%**	19%	-2%	-1%	-3%	-10%
2020	23%	**47%**	33%	20%	-6%	35%	-12%
2021	-7%	16%	2%	5%	5%	-3%	**34%**
Avg.	8%	**27%**	16%	4%	-1%	9%	1%

Table 6: Sharpe Ratio of all trading strategies

	BH	CNNLSTM	CNN	LSTM	MACD	STO	RSI
2016	0.32	**2.10**	0.89	0.27	-0.69	1.87	-0.70
2017	0.72	1.56	0.86	-0.81	**1.82**	-1.43	-0.61
2018	-0.51	**1.32**	1.14	0.14	-2.09	0.28	0.50
2019	1.32	**1.65**	1.46	-0.28	-0.19	-0.39	-1.06
2020	0.99	**1.83**	1.35	0.86	-0.27	1.43	-0.58
2021	-0.55	0.95	0.08	0.28	0.28	-0.34	**1.94**
Avg.	0.38	**1.57**	0.96	0.08	-0.19	0.24	-0.09

Table 7: β_2 of all trading strategies

	CNNLSTM	CNN	LSTM	MACD	STO	RSI
2016	**15.3**	11.5	7.4	7.4	12.6	-3.8
2017	**38.7**	2.9	1.5	11.0	-17.4	5.4
2018	**25.5**	9.0	0.0	-30.7	-7.8	5.6
2019	**29.4**	17.1	-10.2	13.2	-9.4	-0.3
2020	6.7	**11.3**	3.3	-9.6	6.9	5.4
2021	**19.7**	13.5	0.0	14.2	11.0	13.2
Avg.	**22.5**	10.9	0.3	0.9	-0.7	4.2
Std.	*11.2*	*4.8*	*5.9*	*17.8*	*12.4*	*5.8*

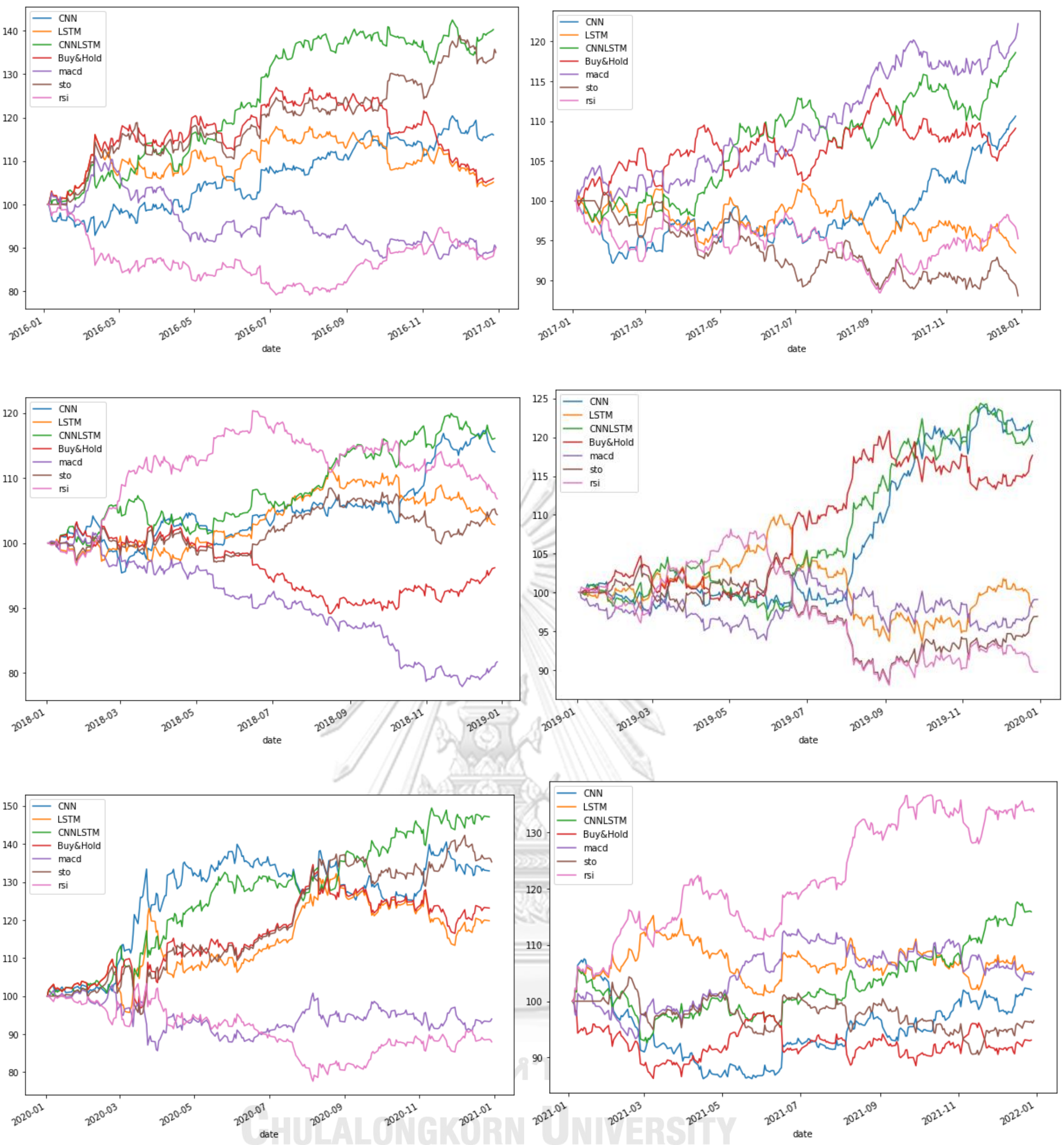

Figure 21: Portfolio of all trading strategies during 2016 to 2021

Chapter 8

CONCLUSION

Various deep learning combinations based on CNN and LSTM are proposed in this paper to perform gold price prediction. We evaluated the models' performance in prediction error and trading performance. We also conduct a portfolio back-testing during 2016 – 2021.

From the experimental results, CNN-LSTM yields the highest performance among all the models because it contains both CNN and LSTM components and has more capability to learn and understand the complex relationship between movements of Price Action features and output better than other techniques.

REFERENCES

1. *What Drives the Price of Gold?* . Available from: https://www.investopedia.com/financial-edge/0311/what-drives-the-price-of-gold.aspx.
2. Yu, L., X. Zhang, and S. Wang, *Assessing potentiality of support vector machine method in crude oil price forecasting.* EURASIA Journal of Mathematics, Science and Technology Education, 2017. **13**(12): p. 7893-7904.
3. Kaboudan, M. *Compumetric forecasting of crude oil prices*. in *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No. 01TH8546)*. 2001. IEEE.
4. Ju, C.-B., M.-C. Hung, and A.-P. Chen. *Market Profile with Convolutional Neural Networks: Learning the Structure of Price Activities*. in *2020 International Symposium on Computer, Consumer and Control (IS3C)*. 2020. IEEE.
5. JuHyok, U., et al., *A new LSTM based reversal point prediction method using upward/downward reversal point feature sets.* Chaos, Solitons & Fractals, 2020. **132**: p. 109559.
6. *An Introduction to Price Action Trading Strategies* Available from: https://www.investopedia.com/articles/active-trading/110714/introduction-price-action-trading-strategies.asp.
7. *Moving Averages as Dynamic Support and Resistance* Available from: https://dailypriceaction.com/blog/moving-averages-as-dynamic-support-and-resistance/.
8. Aggarwal, C.C., *Neural networks and deep learning.* Springer, 2018. **10**: p. 978-3.
9. Chollet, F., *Deep learning with Python (Shelter Island, New York*. 2018, Manning Publications Co), oCLC: ocn982650571.
10. Olah, C. *Understanding LSTM Networks.*; Available from: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

11. Zhao, C.-l. and B. Wang, *Forecasting crude oil price with an autoregressive integrated moving average (ARIMA) model*, in *Fuzzy information & engineering and operations research & management*. 2014, Springer. p. 275-286.

12. Kantavat, P. and B. Kijsirikul. *Combining technical analysis and support vector machine for stock trading*. in *2008 Eighth International Conference on Hybrid Intelligent Systems*. 2008. IEEE.

13. Jearanaitanakij, K. and B. Passaya. *Predicting short trend of stocks by using convolutional neural network and candlestick patterns*. in *2019 4th International Conference on Information Technology (InCIT)*. 2019. IEEE.

14. Eapen, J., D. Bein, and A. Verma. *Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction*. in *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*. 2019. IEEE.

15. Hao, Y. and Q. Gao, *Predicting the trend of stock market index using the hybrid neural network based on multiple time scale feature learning.* Applied Sciences, 2020. **10**(11): p. 3961.

16. Treynor, J., *i K. Mazuy 1966. Can Mutual Funds Outguess the Market.* Harvard Business Review, 1966. **44**.

17. Sharpe, W.F., *Capital asset prices: A theory of market equilibrium under conditions of risk.* The journal of finance, 1964. **19**(3): p. 425-442.

VITA

NAME	Kwan Boon-long
DATE OF BIRTH	30 Dec 1985
PLACE OF BIRTH	Bangkok
INSTITUTIONS ATTENDED	B.Eng. in Automotive Engineering, MBA and M.Sc. in Computer Science from Chulalongkorn University in 2009, 2019 and 2022, respectively.