IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 23 December 2023, accepted 10 January 2024, date of publication 25 January 2024, date of current version 6 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358406

RESEARCH ARTICLE

The Inclusion of the Volume-Price-Product Factor for the Trend Forecasting of Futures Time Series Data

YIPIAO CHEN AND XIAOGANG YUAN
School of Cyber Security, Gansu University of Political Science and Law, Lanzhou 730070, China

Corresponding author: Yipiao Chen (cyp7304@gsupl.edu.cn)

This work was supported in part by the Key Research Project at the University Level at Gansu University of Political Science and Law under Grant GZF2022XZD08.

ABSTRACT Predicting time series data involves extracting features and forecasting trends from observed phenomena. Although deep learning algorithms are widely used in this field, their emphasis on prediction accuracy may not be optimal for futures time series data. For a futures time series, achieving high prediction accuracy alone is not sufficient. This is because, in some cases, ten accurate predictions may not compensate for a single loss. Therefore, a high accuracy rate does not necessarily translate into good returns. Existing methods have yet to provide practical and reliable approaches for predicting futures time series data. The primary contributions of this study are as follows: First, we employ the Vapnik-Chervonenkis (VC) dimension and error function from the perspective of binary classification for futures time series data to elucidate the generalization ability of the simple moving average model. Furthermore, we offer theoretical guidance to enhance predictive performance by introducing effective factors (i.e., features) that positively impact prediction results. By incorporating influential features, the discrimination of the loss function can be enhanced, making it easier to adjust the parameters and minimize the overall loss function value. Consequently, this improves the overall return rate, which is achieved by introducing additional factors to minimize the error values in the loss function. This explains why the proposed moving average model, enhanced by the introduction of the volume-price-product factor, achieves good prediction performance.

INDEX TERMS Time series analysis, deep learning, average models, prediction methods, quantization signal trading.

I. INTRODUCTION

Research on time series data spans diverse fields, including finance [1], [2], energy [3], [4], [5], [6], transportation planning [7], [8], [9], and meteorology [10]. Within these domains, time series analysis employs various statistical and machine learning methods, including simple moving average (SMA), neural networks, and advanced deep learning models, such as recurrent neural networks (RNNs) [11], convolutional neural networks (CNNS) [12], and temporal convolutional networks (TCNs) [13]. From a survey perspective, [14], [15], [16] elaborated on the utilization of deep learning for time series data prediction and delineated the evolution of deep learning (DL) models in time series forecasting. DL models have been successfully applied in various domains, including speech recognition [17], audio recognition [18], and machine translation [19]. However, there are still gaps in the existing research within the futures market, which is an important component of the financial and economic fields. Challenges related to lag and uncertainty present significant obstacles to accurate prediction and effective decision-making. This lag and uncertainty are primarily attributed to unknown external factors and the non-rational behavior of market participants. Lag and uncertainty are impossible to eliminate, as the market often experiences events, such as black swans. These events are uncertain before they occur; however, once they occur,

The associate editor coordinating the review of this manuscript and approving it for publication was Sajid Ali.

they have a strong and transient impact on the trend changes. This inevitably leads to a lag and uncertainty in predictions, which is inherently unavoidable. Therefore, establishing a model should reduce the impact of the lag and uncertainty. This can be measured using practical indicators such as the most direct one, that is, the annualized return rate.

Our objective is to create universal and practical decision-making tools that can offer clear market signals, such as buying or selling signals, to support investors or decision-makers in making informed decisions. The research outcomes should be validated through empirical studies and testing in real economic environments to ensure practical feasibility and applicability. By achieving these goals, this study aims to furnish the finance and economics sectors with more potent and reliable predictive tools, ultimately enhancing the ability of investors or decision-makers to comprehend market trends accurately. The primary objectives of this study are as follows:

- *Explore effective models for predicting futures-type time series data.*
- *Provide theoretical explanations to elucidate the suitability and reliability for futures-type time series data forecasting.*

The major contributions of this paper are summarized as follows:

- *Proposing the application of VC dimension and error function and clarifying the generalization ability of SMA.*
- *Providing theoretical guidance for achieving better predictive performance.*
- *Proposing volume-price-product moving average (VPPMA), comparing it with SMA and DL, and explaining why the VPPMA achieves good prediction performance.*

The remainder of this paper is organized as follows. Section II provides an overview of the current state of research on time series data in financial and economic domains. Section III utilizes a deep learning model of Long Short-Term Memory (LSTM) to predict future trends. In Section IV, the VPPMA model is proposed and used for trend forecasting. Section V presents a comparison and analysis of the results obtained in Sections III and IV, respectively. Finally, Section VI provides a summary of the study and outlines future work.

II. OVERVIEW OF THE ACHIEVEMENTS IN THE FINANCIAL AND ECONOMIC DOMAINS OF TIME SERIES DATA

Time series data, through the analysis of its inherent characteristics, can predict future trends and hold significant value in financial and economic domains, such as financial market forecasting and stock market analysis. Jiang [1] integrated market information and trading features using a knowledge graph, fused multiple features with candlestick indicators and employed a bidirectional long short-term memory (LSTM) network for stock price trend prediction. However, specific profit outcomes were not provided. The daily trading volume and transaction count of the Internet financial market were used in a deep neural network-based prediction model for hierarchical time series learning but without actual trading results [2].

Reference [11] showed that LSTM and biLSTM models demonstrate higher accuracy in addressing time series prediction than regression-based modeling approaches. References [12] and [13] clarified that a combination of different deep learning architectures can enhance the predictive capabilities. However, whether it is an application in fields such as electricity and traffic [3], [4], [5], [6], [7], [8], [9] or applications in areas such as speech and machine translation [17], [18], [19], the former exhibits statistical regularities, and the latter adheres to grammatical rules. In contrast, futures time series data can be considered as the resonance of different individuals' psychology and behavior, and human nature, being the most complex, is challenging to define with rules.

Researchers have attempted to utilize deep learning algorithms in the futures and stock prediction markets, such as Catalin [20], who employed two traditional deep learning architectures, LSTM and TCN, to construct prediction models. Catalin proposed a threshold-based trading strategy and compared the profitability of the two models. However, it focuses solely on the gain and winning rates of the two models for various stocks, and the returns are below 20%. Another study by Kim and Kim [21] combined LSTM and CNN network models and presented stock prediction models based on time series and images. The combined model outperformed the individual models in terms of prediction performance. However, it affirmed the superior performance of the feature fusion LSTM-CNN model over individual models but did not specify gain figures.

Zhou et al. [22] introduced a two-stage end-to-end prediction model that combines empirical mode decomposition (EMD) and factorization machine (FM) neural network technology to handle nonstationary time series data. Zhou presented an annualized return level with achievable rates, factoring in transaction costs ranging from 11% to 14%. Jiang and Chen [1] utilized Bi-LSTM in conjunction with financial knowledge graphs and zigzag technology [23] to improve the accuracy of stock price prediction. However, it did not quantify potential profits, and methods such as zigzag, based on anticipating future market changes, could not ensure that actual future returns would be as favorable as currently adjusted. Articles [24], [25], [26], [27] aimed to reduce the prediction lag by forecasting inflection points to enhance the accuracy, reflecting a research direction. However, these studies primarily compared new and old methods, highlighting the enhanced accuracy of the new approach. Nonetheless, they lack indicators that reflect practical outcomes such as annualized returns.

It is worth noting that moving average theory is commonly employed by investors for trend forecasting using stocks or futures time series data. Neural networks and DL models utilize a parameter optimization approach, specifically by minimizing cumulative historical errors, to predict the output value at the next moment, whereas SMA models primarily concentrate on trend prediction. Therefore, the moving average model may be a good choice. Inspired by Li's moving average theory [28] and Wang's research [29], this study introduces a novel volume-price-product moving average (VPPMA) model. Furthermore, the feasibility and generalization ability of this model is discussed from the perspective of machine learning, utilizing statistical learning theory [30]. The experimental findings confirm that the VPPMA model better satisfies the objective of maximizing profitability for trend prediction using stocks or futures time series data. The results for stocks or futures time series data applications surpass those of their peers, achieving an annualized return rate higher than that of Warren Buffett [31].

Section III presents and validates the experimental results using DL models to forecast stocks or futures time series data. This performance is unreliable because of the inherent uncertainty of the stocks or futures time series data. For example, it is impossible to determine with certainty whether the eleventh day will rise based solely on the daily increases in the preceding ten days.

III. FUTURES TREND PREDICTION USING LSTM

In this study, an LSTM network was selected as the prediction model. LSTM is a classical recurrent neural network (RNN). However, when predicting futures time series data, there is the consistent issue of inaccurate predictions of the next moment because of the inherent volatility of such data, resulting in a lag behind the actual trend. The selection of the LSTM network did not impede the discussion on these lagging issues. To analyze the predictive performance of the DL model and understand the issue of prediction lag, we chose the fuel future as the input data for the LSTM networks. The data were sourced from a trading platform named "Trading Pioneer Futures Algorithmic Trading Software Platform."(TPFATSP) The data covered the period from August 25, 2004, to April 21, 2023, at a daily level, comprising six features: the highest price, lowest price, opening price, closing price, trading volume, and open interest. This dataset was used as the training set. The LSTM network uses these six indicators as inputs. The time step was set to 10, indicating that the model predicted the closing price on the 11th day using data from the previous ten days. The model employs the "relu" activation function. The model was constructed using TensorFlow and the Keras high-level API with a batch size of 30. Figure 1 shows the training results after 100 iterations.

The X-axis of Figure 1 represents the number of datasets or the length of the time series data, which is 3635 data points. The Y-axis represents the closing price, where the blue line represents the actual closing price, and the pink line represents the predicted closing price. Overall, when we display 3635 data points from August 25, 2004, to April 21, 2023, spanning nearly 19 years, the predicted values appear to closely align with the actual values. However, when we reduced the period to only one month, there was a significant disparity between the predicted and actual values. This can be observed through the subsequent predictions of the test dataset.

The test dataset was derived from Fuel Future, which covers approximately one month from April 22 to May 26, 2023, whereas the training dataset includes data up to April 21, 2023. The predictive performance of the trained model on the test set is illustrated in Figure 2.

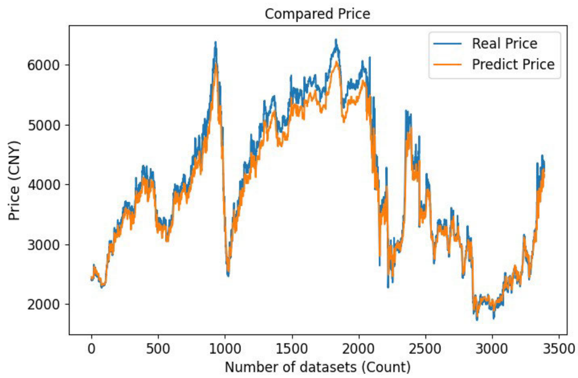

FIGURE 1. Prediction results of LSTM networks on the training set.

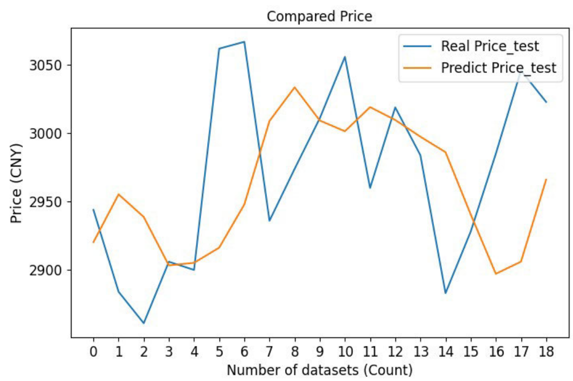

FIGURE 2. Prediction results of LSTM networks on the test set.

Figure 2 reveals a consistent pattern of predictive lag in which the predicted prices lag one step behind the actual prices. As illustrated in Figure 2, the blue line consistently leads to the pink line for one period. This pattern is particularly prominent under volatile market conditions, rendering it unreliable for guiding trading decisions. This issue arises from the DL prediction method, which involves training the parameters to minimize the value of the loss function and then using these optimized parameters to predict the value at the next time step. By contrast, conventional moving average models are commonly utilized for trend prediction in stocks and futures.

Time series data commonly exhibit characteristics such as tendency, periodicity, volatility, stationarity, and symmetry [18]. Stocks or futures time series data demonstrate heightened volatility and tendencies. Predicting short-term fluctuations is challenging and cannot be achieved solely by relying on historical data. When forecasting this type of data, it is crucial to prioritize trend analysis. Volatility characteristics can be analyzed using statistical theory, in which the price is perceived as fluctuating around a central point. The moving average line provides a visualization of the prices of an asset over a specific period by connecting the central points with lines. By analyzing the trend of the moving average, we can predict short- and long-term trends in future price movements. The moving average theory has been widely used in quantitative investment, using techniques called "golden cross" and "death cross" to predict fluctuations in the market. This is because the moving average model effectively captures the trend characteristics of the time series data.

Several scholars have conducted extensive research on enhancing their performance. Scholars have diverse focal points when analyzing time series data on stocks and futures. These focal points can be summarized in the following three research directions: First, major emphasis is placed on predicting price movements by abstracting the rise and fall of stocks as a binary classification problem. This involves utilizing machine learning models to minimize the probability of misclassification. Second, there is a particular focus on predicting inflection points. Jiang [24] proposed a method for forecasting inflection points in future data using the zigzag technology. Additionally, some scholars have integrated piecewise linear representation (PLR) with BP neural networks [25], Gaussian process classification [26], and weighted support vector machines [27] to predict stock inflection points. Third, there is an emphasis on researching quantitative investment models, which is distinct from the methods used to evaluate the model accuracy. This approach involves simulating historical data to derive optimal parameters and subsequently using these parameters to predict future market trends. The third strategy primarily emphasizes investment returns, particularly profitability, as the key factor.

In this study, we propose a volume-price integrated moving average model based on a third research direction. This model demonstrates superior annualized returns across multiple futures varieties and exhibits a certain level of universality and applicability.

IV. VPPMA MODEL

The SMA model is briefly reviewed before introducing the VPPMA model. The SMA has evolved as a fundamental tool in time series analysis, providing a basic yet effective means of understanding trends and patterns within data. The basic formula for calculating the SMA is as follows:

$$\text{SMA} = \frac{C_1 + C_2 + \ldots + C_N}{N} \tag{1}$$

where SMA represents the simple moving average value, C_1, $C_2, \ldots, and\ C_N$ represent the closing prices for the respective days, and N represents the number of periods or days considered for the moving average. This formula was improved by several researchers in subsequent studies. Several improvements have been made to the formula. For example, the composition of $C_1, C_2 \ldots C_N$ was adjusted, and the value of cycle N was modified. However, the most significant breakthrough was achieved by Li, who incorporated the trading volume factor into the equation. Li modified Equation (1) as follows:

$$\text{PVMA} = \frac{\sum_i^n V_i{*}\frac{O_i+C_i}{2}}{\sum_i^n V_i} \tag{2}$$

Equation (2) shows that the average is the mean value of the opening and closing prices over n days, with the weight determined by the daily trading volume. Days with higher trading volumes carry more weight in determining price, whereas days with lower trading volumes carry less weight. By incorporating trading volume as a weight, the model becomes more sensitive in identifying market trends and more accurate in forecasting market movement.

Wang [29] also emphasizes the significance of trading volume in moving averages and acknowledges that relying solely on trading volume is insufficient. Therefore, it is essential to conduct a comprehensive analysis that integrates trading volume and price. Building on the findings of previous studies, we propose a VPPMA model. The equation for this model is as follows:

$$\text{VPPMA} = \frac{\sum_i^n V_i{*}C_i}{N} \tag{3}$$

From the formula, it can be observed that, when compared to SMA, $V_i{*}C_i$ is directly substituted for C_i in SMA, and when compared to PVMA, both the numerator and denominator are simplified. In the numerator, C_i is directly used to replace the average C_i and O_i in PVMA, whereas in the denominator, it is simplified to N. The volume was not considered a weighting factor in the VPPMA. $V_i{*}C_i$ in the VPPMA expression is referred to as the VPP factor, denoted as S_i, where

$$S_i = V_i^* C_i \tag{4}$$

We then employ the VPPMA model to construct a trading model. Automated quantitative trading was successfully achieved by programming on the TPFATSP. It is a product of Trading Pioneer Technology Co., Ltd., designed specifically for the Chinese futures market. The TPFATSP is a robust formula support system that allows users to implement trading ideas and write trading strategies. It features a unique trading strategy testing engine for users to conveniently test and optimize trading strategies. With a leading strategy trading system, real-time data, and automated trading capabilities, automatic trading can be fully realized. The VPPMA algorithm was implemented on this platform.

Before introducing the VPPMA algorithm, let us consider a few simple "criteria" for trend prediction using the SMA.

- *When the SMA line is sloping upwards, it may indicate an upward market trend.*

- *When the SMA line is sloping downwards, it may indicate a downward market trend.*
- *When the short-term moving average crosses above the long-term moving average, typically with the short-term moving average trending upward through the long-term moving average. It is considered a buy signal, suggesting a potential uptrend. It's called a "golden cross," suggesting the beginning of an upward trend.*
- *When the short-term moving average crosses below the long-term moving average, typically, the short-term moving average trends downwards through the long-term moving average. It is considered a sell signal, indicating a potential downtrend. It's called a "death cross," suggesting the beginning of a downward trend.*

They can be expressed using mathematical formulas corresponding to Equations (5)–(6).:

$$\text{Trend} = \begin{cases} \text{uptrend} & \text{if SMA}[m] > \text{SMA}[n] \\ \text{downtrend} & \text{if SMA}[m] < \text{SMA}[n] \end{cases} \tag{5}$$

$$\text{Trend} = \begin{cases} \text{uptrend} & \text{if SMA}_{\text{short-term}} > \text{SMA}_{\text{long-term}} \\ \text{downtrend} & \text{if SMA}_{\text{short-term}} < \text{SMA}_{\text{long-term}} \end{cases} \tag{6}$$

In Formula 5, m and n represent the SMA values for the same period in the previous m and n days, respectively. Based on the definition in Equation (5), the default assumption is that m is less than n. For example, $m = 0$ denotes the current day, and $n = 1$ represents the previous day (yesterday). When $m = 0$, the current day is denoted as an SMA without SMA[0]. Using the concepts and mathematical models introduced above, we present the VPPMA algorithm. The VPPMA programming is described in Algorithm 1.

The variables VPPMA_1, VPPMA_2, VPPMA_3, VPPMA_4, MA_1, and MA_2 were calculated from the SMA function, which is defined in Equation (1). Parameter S in the function SMA represents the VPP factor, which is defined in formula (4), and the other parameter , C, refers to the closing price; both are arrays. Parameters N_1, N_2, N_3, and N_4 represent the number of periods. The parameter MarketPosition represents the position status. The position was opened only when it was empty. VPPMA and MA were also used as arrays. VPPMA[n] represents the VPPMA value for the previous n days. In the algorithm, $\text{VPPMA}_2[1]$ signifies yesterday's VPPMA_2 value, and $\text{VPPMA}_4[1]$ represents yesterday's VPPMA_4 value. The same applies to the SMA. The core of the algorithm consists of steps 8 to 17, where the trend is determined by the assessment of the VPPMA and MA. A bullish trend was observed when the golden cross was present, whereas a bearish trend was observed when the dead cross was present. The improvement of Algorithm 1 involves the incorporation of the VPPMA to forecast the trend of futures time series data.

The relationships between N_1, N_2, N_3, and N_4 do not need to be predetermined; they are automatically optimized based on the relationships between VPPMA_1, VPPMA_2,

Algorithm 1 The VPPMA Model for the Trend Forecasting of Futures Time Series Data

Input: VPPMA periods of N_1, N_2, N_3, N_4;
Output: trend forecasting by trading signals;

1: $\text{VPPMA}_1 \leftarrow \text{SMA}(S, N_1)$;
2: $\text{VPPMA}_2 \leftarrow \text{SMA}(S, N_2)$;
3: $\text{VPPMA}_3 \leftarrow \text{SMA}(S, N_3)$;
4: $\text{VPPMA}_4 \leftarrow \text{SMA}(S, N_4)$;
5: $\text{MA}_1 \leftarrow \text{SMA}(C, 5)$;
6: $\text{MA}_2 \leftarrow \text{SMA}(C, 20)$;
7: **if** MarketPosition = *0* **then**
8:     **if** $\text{VPPMA}_1 < \text{VPPMA}_2$ and $\text{VPPMA}_3 < \text{VPPMA}_4$ and $\text{MA}_1 < \text{MA}_2$ **then**
9:         Short selling; #downtrend
10:     **else if** $\text{VPPMA}_1 > \text{VPPMA}_2$ and $\text{VPPMA}_3 > \text{VPPMA}_4$ and $\text{MA}_1 > \text{MA}_2$ **then**
11:         Long buying; #uptrend
12:     **end if**
13: **else if** MarketPosition $\neq$ *0* **then**
14:     **if** $\text{VPPMA}_1 > \text{VPPMA}_1[1]$ and $\text{VPPMA}_3 > \text{VPPMA}_3[1]$ and $\text{MA}_1 > \text{MA}_2$ **then**
15:         Closing short position; #trend changes upward
16:     **else if** $\text{VPPMA}_2 < \text{VPPMA}_2[1]$ and $\text{VPPMA}_4 < \text{VPPMA}_4[1]$ and $\text{MA}_1 < \text{MA}_2$ **then**
17:         Closing long position; #trend changes downward
18:     **end if**
19: **end if**

VPPMA_3, and VPPMA_4. Similarly, the relationships among VPPMA_1, VPPMA_2, VPPMA_3, and VPPMA_4 need not be predetermined; they can be relationships other than those specified in the algorithm. When the relationships among VPPMA_1, VPPMA_2, VPPMA_3, and VPPMA_4 changed, the values of N_1, N_2, N_3, and N_4 also changed. However, different combinations of relationships among VPPMA_1, VPPMA_2, VPPMA_3, and VPPMA_4 result in different profit levels for different varieties. The other relationship combination, compared to the relationships provided by the algorithm, may demonstrate better profit performance in certain varieties, whereas, in other varieties, the profit performance may be worse. In summary, there is no single combination of VPPMA_1, VPPMA_2, VPPMA_3, and VPPMA_4 that exhibits optimal performance across all varieties.

To enhance clarity in illustrating Algorithm1, we have depicted it in a flowchart, as depicted in Figure 3. The TPFATSP triggers this program at regular intervals propelled by time. Consequently, during trading hours, the provided code undergoes a continuous execution on the platform. The code is executed not only once, although there is an absence of a loop.

In future research, we will choose several popular futures contracts to prioritize maximizing annual returns as the

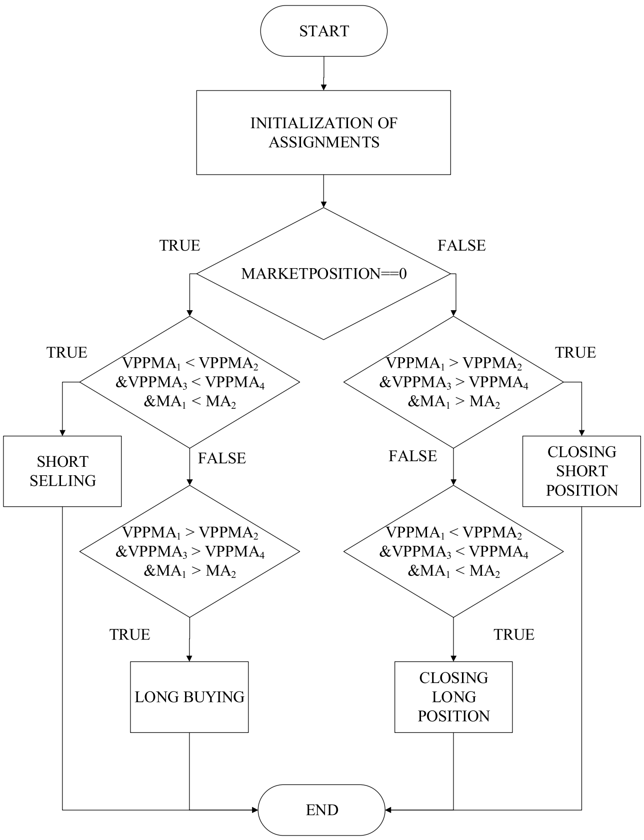

FIGURE 3. The flowchart of Algorithm 1.

optimization objective instead of total returns. There were two reasons for this finding. First, the variation in the listing durations of different futures makes it difficult to compare only total returns. The higher ranking of the total returns may be solely due to the longer listing time. Second, variations in contract prices among different futures markets also impact total returns. Similarly, the higher ranking of the total returns could be due to the higher per-contract price. The annualized return remains unaffected by these two factors.

To assess the efficacy of Algorithm 1, we deployed Algorithm 1 on the TPFATSP. Six futures contracts are selected from this platform for testing. These contracts comprise three chemical products (fuel, methanol, and PTA), two agricultural commodities (cotton and palm oil), and a black-metal commodity (rebar). The datasets for these six commodities cover the period from the listing date of each futures contract to April 21, 2023. Therefore, the dataset for fuel was identical to the training set used in the LSTM deep learning method. The datasets for the other five commodities were new and directly obtained from the TPFATSP. The test period begins on the listing date of each contract and ends on April 21, 2023. Each contract is bought in one lot, with transaction fees calculated at 10 CNY per lot on a single side. The chosen contracts represent the main continuous contracts for each commodity, and their trading volumes and closing prices are derived from the daily K-line. The input range for variables N_1-N_4 was set between 10 and 50 with increments of 10 units for each variable, resulting in 625 combinations of data. The number 625 is obtained as follows: Because N_1–N_4 are set from 10 to 50 in steps of 10, each variable (N_1 to N_4) can only take five values: 10, 20, 30, 40, and 50. The total number of combinations is 5^4 = 625. Only the top and bottom five rankings were provided for each commodity to conserve space without compromising the clarity of the problem explanation. The test results are presented in detail in Tables 1–6 respectively.

TABLE 1. The annualized return of fuel future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	10	10	50	40	94.29%
2	10	10	40	40	89.00%
3	10	10	40	50	84.14%
4	10	10	40	20	82.29%
5	10	10	30	50	81.20%
...	…	…	…	…	...
621	40	30	30	50	-23.52%
622	40	30	20	50	-28.25%
623	30	40	50	30	-28.97%
624	50	30	20	50	-35.99%
625	50	30	10	50	-36.05%

TABLE 2. The annualized return of methanol future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	30	40	20	20	146.91%
2	30	20	20	10	126.24%
3	30	50	20	20	125.80%
4	30	30	20	20	123.81%
5	20	30	10	10	122.35%
...	…	…	…	…	...
621	50	20	20	40	-4.06%
622	40	10	50	30	-6.79%
623	40	10	50	40	-11.64%
624	40	10	10	30	-16.68%
625	10	20	40	30	-23.32%

TABLE 3. The annualized return of PTA future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	10	10	20	30	98.76%
2	10	10	20	40	93.48%
3	10	10	20	20	89.44%
4	20	20	30	30	86.98%
5	20	40	50	10	85.29%
...	…	…	…	…	...
621	50	50	40	40	-8.41%
622	50	30	40	10	-9.22%
623	50	40	30	10	-9.22%
624	50	30	30	10	-9.64%
625	50	40	30	50	-9.72%

Notably, upon analyzing the data in Tables 1 to 6, there is a significant disparity in the annualized returns between the highest- and lowest-ranked investments. For example, the first ranking of methanol exhibits an annualized return of

TABLE 4. The annualized return of cotton future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	40	30	50	50	80.73%
2	30	30	50	50	78.37%
3	30	20	20	10	77.68%
4	30	30	20	10	75.26%
5	40	30	20	10	73.38%
...	…	…	…	…	...
621	50	50	50	30	-0.00%
622	50	50	50	40	-0.00%
623	50	50	50	50	-0.00%
624	20	50	40	20	-1.07%
625	50	20	20	40	-2.61%

TABLE 5. The annualized return of palm future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	10	10	50	50	102.98%
2	10	10	20	50	95.64%
3	10	10	40	50	90.59%
4	10	20	50	30	89.11%
5	10	50	50	40	87.95%
...	…	…	…	…	...
621	40	50	30	40	-1.83%
622	40	50	20	40	-1.97%
623	50	20	40	10	-2.37%
624	50	20	30	10	-3.48%
625	50	30	30	10	-8.98%

TABLE 6. The annualized return of rebar future.

Ranking	N_1	N_2	N_3	N_4	Annualized return
1	50	40	10	20	66.94%
2	20	30	10	20	66.10%
3	20	40	10	20	64.79%
4	30	30	20	40	64.55%
5	50	30	10	20	64.17%
...	…	…	…	…	...
621	50	50	50	50	-0.00%
622	10	40	50	30	-0.25%
623	40	10	50	40	-2.19%
624	20	10	50	40	-2.70%
625	30	10	50	40	-3.54%

146.91%, whereas its last rank results in an annualized return of −23.32%. A horizontal comparison of the six types of futures commodities shows that rebar, at a rate of 66.94%, exhibits the weakest performance among the six top-ranked annualized returns. Correspondingly, Fuel Futures demonstrate the lowest performance at −36.05% when considering the bottom-ranked annualized returns for each type. Among the top-ranked annualized returns, the lowest value is 66.94%. This annualized return rate of 66.94% is remarkably higher than the approximately 20% annualized return achieved by the renowned investor Warren Buffett, as reported online [31]. This exceeded Buffett's return by approximately 200%, which is an impressive achievement.

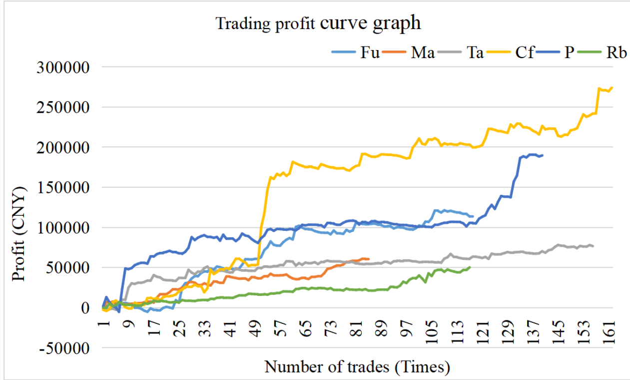

FIGURE 4. Trading profit curve graph of six futures commodities.

Figure 4 presents the profit curves of the six top-ranked annualized returns for fuel, methanol, PTA, cotton, palm oil, and rebar.

The horizontal axis in Figure 4 represents the number of trades, and the vertical axis represents profit. As the listing dates of the different varieties are different, the number of trades on the horizontal axis is also different. At the same time, the unit price of each variety is different; thus, cumulative profit varies. However, regardless of variety, the profit curves oscillate upward, indicating that the model can capture trends. There is a question: Although this moving average model performs well in training with historical data, can it also be applicable in the future? Does it possess a generalization capability?

The performance of model f learned on the training set, evaluated on the overall distribution, is referred to as the generalization performance of model f [30]. Is model f learned from sample dataset D applicable to all the samples in the overall distribution? If the performance of the learned model in the overall distribution is consistent with its performance in the training set, it is referred to as being consistent with the learning process. In the field of machine learning, statistical learning theory focuses on evaluating the generalization performance of learning models and exploring the consistency of the learning process. For predictive models, the goal is not to achieve zero error but to ensure that the predicted results are consistent with the empirical error, and the smaller the upper limit of the error, the better. According to statistical theory, the upper limit of the generalization error is proportional to the empirical error and VC dimension of the hypothesis space and inversely proportional to the sample capacity of the training set, whose formula is as follows:

$$R(\theta) \leq R_{\text{emp}}(\theta) + \Delta(\frac{d_{vc}}{m}) \tag{7}$$

where $R(\theta)$ stands for the generalization error, $R_{emp}(\theta)$ represents the empirical error, $\Delta(\frac{d_{vc}}{m})$ and denotes the confidence range. The confidence range, or confidence interval, provides a measure of this uncertainty by providing a range of values within which the true generalization error is likely to fall. The sample capacity of the training set, denoted by m, is dependent on the length of the selected historical data period. The larger the sample capacity m used for training, the more it can avoid underfitting caused by insufficient training data. The formula provides insights into how the generalization performance of learning models is affected by factors, including the complexity of the hypothesis space, the amount of training data available, and the empirical error. By minimizing empirical error and managing the complexity of the hypothesis space, we can endeavor to enhance the consistency and generalization capability of our learning models.

The intuitive definition of the VC dimension is that if there exist m samples that can be separated into all possible 2^m dichotomies by the hypothesis space H formed by a set of functions, then the hypothesis space H is said to be able to shatter m samples, and the VC dimension of the hypothesis space is the maximum number m of samples that it can shatter. If, for any m, there always exists a set of samples that can be shattered by the hypothesis space, then the VC dimension of the function set is infinite. In other words, the VC dimension denotes the highest number of points that the hypothesis space can classify in each possible manner. It plays a crucial role in determining the complexity of a model and its ability to fit different training data patterns. A larger VC dimension indicates a more versatile model that is capable of accommodating diverse patterns. However, they may also be more susceptible to overfitting. Conversely, a smaller VC dimension implies a more restricted model with reduced flexibility; however, it often exhibits better generalization capability.

In the moving average model, the closing price is either above or below the moving average. It is unlikely that the closing price is exactly equal to the average and can, therefore, be categorized as above-average. Thus, the moving average model can be considered a binary classification model. Consider the function set of the hypothesis space H, which is composed of moving averages with different periods. Then, the hypothesis space H has an infinite VC dimension because, irrespective of the data points, a moving average can be calculated based on their close price in any given period, classifying the data points as either above or below the average. This function is expressed as an indicator.

$$I(\mathrm{SMA}(C,N) \leq C_0) = \begin{cases} 1 & \mathrm{SMA}(C,N) \leq C_0 \\ 0 & \mathrm{SMA}(C,N) > C_0 \end{cases} \tag{8}$$

where C is an array comprising the elements C_1, C_2, C_3, $\ldots$, and C_N. C_0 is the closing price of the current day. SMA and N were consistent with those previously defined. A more general expression for Equation (8) can be written as:

$$I(f(x;\theta) = y) = \begin{cases} 1 & f(x;\theta) = y \\ 0 & f(x;\theta) \neq y \end{cases} \tag{9}$$

where f represents the learning model, and θ is the set of parameters.

The principle of empirical risk minimization aims to reduce the empirical error, whereas that of structural risk minimization aims to reduce the generalization error. The VC dimension of the hypothesis space defined by the classification function based on Equations (8) and (9) was infinite. An infinite VC dimension indicates that the hypothesis space can fit any dataset flexibly and accurately. The moving average line, which is a smooth curve formed by connecting average prices, consists of straight line segments. Line segments can be viewed as univariate functions that are defined at specified intervals. The univariate function is the simplest. Therefore, the function space formed by the moving average line can be decomposed into a hypothesis space composed of a series of univariate functions. The complexity of the hypothesis space formed by univariate functions was the lowest. Therefore, the structural risk of the functional space defined on this basis was minimal. Therefore, to reduce the upper bound of the error, it is necessary to reduce the first term of Equation (7), which is the empirical risk. If the empirical error is minimized, then the upper bound of the error in Equation (7) decreases.

The formula for calculating the empirical risk, $R_{emp}(\theta)$, is presented below:

$$R_{\mathrm{emp}}(\theta) = \frac{1}{m}\sum_{i=1}^{m} L(f(x_i;\theta) - y_i) \tag{10}$$

where $L(f(x_i;\theta) - y_i)$ is the loss function. The loss function plays a crucial role in the training of machine-learning models. The main purpose is to evaluate the performance of the model by measuring the degree of mismatch between the predicted and actual outputs of the training data. Different types of loss functions can be used depending on the nature of the problem being solved. The loss function was redefined as the loss function for binary classification problems, and its value was determined based on the correctness of the classification. When the classification is correct, the loss function value is assigned a value of zero; otherwise, it is set to one. Based on this definition, the loss function in the summation term of Equation (10) can be expressed as follows:

$$L(f(x_i;\theta) - y_i) = \begin{cases} 0 & f(x_i;\theta) = y_i \\ 1 & f(x_i;\theta) \neq y_i \end{cases} \tag{11}$$

where $f(x;\theta) = y$ represents the correct classification and $f(x;\theta) \neq y$ represents incorrect classification. Accordingly, the loss function acts as an indicator, and Equation (11) resembles Equation (9): As previously mentioned, any moving average line can effectively classify data points as above or below the average. Therefore, according to Equation (11), leading to a loss function value of zero for each data point. Consequently, the empirical error $R_{\mathrm{emp}}(\theta)$ was equal to zero. In this case, $R(\theta) \leq \Delta(\frac{d_{\mathrm{vc}}}{m})$ indicates a small generalization error and strong generalization ability. Algorithm 1 was developed based on the moving average model. The algorithm compares pairs of moving averages (including VPPMA and SMA) for different periods to determine the opening and closing conditions. The compared two moving averages in algorithm 1, One of the moving averages is considered as a classification average, while the other moving average is seen as individual discrete data points. The Golden Cross and Death Cross of the two moving averages represent the relative positions of points on one moving average compared to the other (above or below the moving average). This constitutes a binary classification task. Hence, Algorithm 1 adheres to the application conditions outlined in Equations (8)–(11), which results in a small generalization error, thereby ensuring that the model can generalize its predictions.

V. COMPARATIVE ANALYSIS OF THE PERFORMANCES OF DIFFERENT MODELS

The two metrics commonly used in regression analysis to evaluate the performance and accuracy of predictive models are the mean squared error (MSE) and R-squared value. The MSE measures the average squared difference between the predicted and observed values in a model, providing a quantitative assessment of the model's fit to the data. A smaller MSE suggests a reduced discrepancy between the actual and predicted values, indicating a stronger predictive performance of the model. The R-squared value, also referred to as the coefficient of determination, is a statistical measure used to evaluate the goodness-of-fit of a regression model. This indicates the extent to which the values predicted by the model were aligned with the observed data. The R-squared value ranges from 0 to 1, with a value of 1 indicating a perfect fit, indicating that the model explains all variability in the data. Conversely, a value of zero suggests that the model fails to capture any variability and serves as the baseline for the model prediction performance. Negative R-squared values indicate poor model performance, which is generally deemed unacceptable. Subsequently, the mean squared error and R-squared values were calculated using the following formulae:

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2 \tag{12}$$

$$R^2 = 1 - \frac{\mathrm{SS_{res}}}{\mathrm{SS_{tot}}} \tag{13}$$

where in Equation (13), $\mathrm{SS_{res}}$ denotes the sum of the squared residuals, and $\mathrm{SS_{tot}}$ represents the total sum of squares. The calculations for $\mathrm{SS_{res}}$ and $\mathrm{SS_{tot}}$ are as follows:

$$\mathrm{SS_{res}} = \sum_{i=1}^{m}(y_i - f(x_i))^2 \tag{14}$$

$$\mathrm{SS_{tot}} = \sum_{i=1}^{m}(y_i - \bar{y})^2 \tag{15}$$

where $\bar{y}$ is defined as follows:

$$\bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i \tag{16}$$

In Equations (12), (14), and (15), y_i represents the actual value, and $f(x_i)$ represents the predicted value. Based on the experimental data in Section III, we calculate the mean squared error (MSE) and coefficient of determination R-squared value for fuel futures under the same experimental conditions. The calculation methods for the MSE and R-squared values were as follows:

$$\mathrm{MSE} = 0.00013;$$
$$R^2 = 0.20.$$

MSE and coefficient of determination (R^2) were obtained from a series of experiments. The experimental results were derived by averaging the outcomes of the three trials. The experimental results suggest that although the mean squared error is relatively small, indicating a reasonably accurate level of prediction, the performance of the model is unsatisfactory and unacceptable because the coefficient of determination is approximately 0.2, indicating weak or no fitting. This indicates that the deep learning model may not be suitable for forecasting future series of data. Importantly, high accuracy alone or a small MSE alone does not guarantee a good predictive performance. Even with a smaller MSE, the model may be unreliable if the coefficient of determination is approximately 0.2. Hence, it is essential to consider both the coefficient of determination and mean squared error (MSE) simultaneously rather than solely focusing on achieving a small MSE or high accuracy.

What is the performance of the proposed VPPMA model? We do not calculate the MSE of the VPPMA model here because we treat the moving average model as a binary classification problem. We also did not use evaluation metrics for binary classification problems, such as precision and recall rate, because, in the preceding discussion, the moving averages correctly classified the data based on our definition. Instead, we focus on the accuracy of trend prediction, which is demonstrated by the accuracy of the trades executed based on this trend. Therefore, we selected trade accuracy as the evaluation metric for the moving average model. The profitability ratio is defined as the number of profitable trades divided by the total number of trades, as shown in Equation (17).

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{17}$$

Equation (17) is similar to the calculation formula for the precision index; however, its meaning is distinct. In Equation (17), TP denotes the number of profitable trades, FP denotes the number of loss trades, and TP+FP denotes the total number of trades. The R-squared value was calculated using data obtained from the trade data in the moving average model. It should be noted that these data differ from those used in the DL model to calculate the R-squared value. Utilizing the aforementioned definitions, we can calculate the trade accuracy and R-squared value of the VPPMA. Likewise, the six commodity futures markets were individually tested under the same experimental conditions as in Section IV. The experimental results for the profitability ratio and R-squared value for the six commodity futures markets are presented in Tables 7 and 12, respectively. Similarly, the tables only

TABLE 7. Fuel performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	94.29%	43.22%	0.9457
2	89.00%	41.46%	0.9299
3	84.14%	42.50%	0.9241
4	82.29%	41.59%	0.9291
5	81.20%	40.68%	0.9304
...	...	...	...
621	-23.52%	31.34%	0.7153
622	-28.25%	38.16%	0.4496
623	-28.97%	34.85%	0.7806
624	-35.99%	36.36%	0.5865
625	-36.05%	32.88%	0.7010

TABLE 8. Methanol performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	146.91%	58.82%	0.9257
2	126.24%	60.00%	0.8886
3	125.80%	53.09%	0.8981
4	123.81%	53.61%	0.8648
5	122.35%	54.02%	0.9358
...	...	...	...
621	-4.06%	28.13%	0.0001
622	-6.79%	40.58%	0.0608
623	-11.64%	41.89%	0.0040
624	-16.68%	32.14%	0.5872
625	-23.32%	33.73%	0.0092

TABLE 9. PTA performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	98.76%	46.15%	0.8470
2	93.48%	44.59%	0.7946
3	89.44%	46.75%	0.7277
4	86.98%	41.25%	0.7774
5	85.29%	58.02%	0.8490
...	...	...	...
621	-8.41%	44.30%	0.3129
622	-9.22%	43.44%	0.2612
623	-9.22%	41.18%	0.2919
624	-9.64%	42.86%	0.2482
625	-9.72%	38.64%	0.4240

TABLE 10. Cotton performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	80.73%	49.38%	0.9158
2	78.37%	47.40%	0.9070
3	77.68%	49.40%	0.9307
4	75.26%	51.63%	0.8927
5	73.38%	52.76%	0.8779
...	...	...	...
621	-0.00%	0.00%	0.0000
622	-0.00%	0.00%	0.0000
623	-0.00%	0.00%	0.0000
624	-1.07%	37.65%	0.0807
625	-2.61%	39.08%	0.0768

TABLE 11. Palm performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	102.98%	49.29%	0.7935
2	95.64%	51.52%	0.7607
3	90.59%	48.59%	0.7843
4	89.11%	46.09%	0.6407
5	87.95%	49.02%	0.7806
...	...	...	...
621	-1.83%	41.04%	0.1451
622	-1.97%	44.26%	0.2860
623	-2.37%	39.34%	0.3205
624	-3.48%	37.82%	0.0056
625	-8.98%	34.17%	0.0120

TABLE 12. Rebar performance using VPPMA model.

Ranking	Annualized return	Profitability ratio	R-squared Value
1	66.94%	46.15%	0.9612
2	66.10%	46.83%	0.9415
3	64.79%	48.15%	0.9652
4	64.55%	50.00%	0.9525
5	64.17%	48.39%	0.9422
...	...	...	...
621	-0.00%	0.00%	0.0000
622	-0.25%	38.78%	0.5357
623	-2.19%	37.07%	0.5515
624	-2.70%	40.50%	0.6683
625	-3.54%	37.93%	0.1925

display the rankings for the first five positions and last five positions, which are consistent with the rankings observed in Tables 1–6. Because the table cannot display many columns of information, we omit the values of the parameters N_1, N_2, N_3, and N_4 in Tables 7–12, which are identical to those in Tables 1–6.

Upon analyzing the data from Tables 7 to 12, it is evident that the majority of the accuracy rates range from 40% to 50%, with methanol being one of the few varieties showing a profitability ratio between 50% and 60%. When observing the R-squared value column, it is evident that all R-squared values surpass those of the deep learning model. This phenomenon is observed not only for the top five rankings but also for the bottom five rankings. Among the top five rankings of the six futures varieties, the lowest R-squared value achieved was 0.64, and the palm variety and remaining R-squared values were all above 0.7. It can be observed that there is a consistent pattern in which the R-squared values of the bottom five ranked varieties are lower than those of the top five. A similar pattern was observed for the accuracy rate, except for PTA. Combining the two metrics of the profitability ratio and R-squared value, we can conclude that higher accuracy rates and larger R-squared values are correlated with higher overall profit capabilities. This also indicates the significance of the accuracy rate and R-squared values as crucial factors that influence the overall profit level of the algorithm. It is important to note that higher R-squared values do not necessarily guarantee higher profitability because they are related to the accuracy rate. Similarly, higher accuracy rates do not necessarily ensure higher profitability as they may be accompanied by smaller R-squared values. In specific cases, if the profitability ratio and R-squared value are both zero, it signifies the absence of transactions.

In general, the VPPMA model proposed in this study showed an accuracy rate ranging from 40% to 60% for the top five ranked varieties. Moreover, the VPPMA model exhibited notable superiority over the deep learning model in terms of R-squared values. This observation indicates that the VPPMA model is better suited for analyzing time series data of futures and stocks than the deep learning model.

How can the performance of the VPPMA be compared to that of the SMA? Is the VPPMA superior in performance? A performance comparison between the VPPMA and SMA models must be conducted under identical experimental conditions, including the same platform, data, time, commission setting, and transaction volume for each lot. The only difference is that the SMA model does not incorporate the volume price factor. The SMA model can be obtained by replacing parameter S with C in the SMA function from steps 1 to 4 in Algorithm 1 while maintaining all other codes unchanged. The experimental results obtained by the SMA model are presented in Tables 13 to 18, which show the metrics of annualized return, profitability ratio, and R-squared value.

TABLE 13. Fuel performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	81.29%	65.00%	0.9378
2	71.65%	55.06%	0.7311
3	69.21%	56.82%	0.8047
4	66.86%	50.00%	0.6455
5	63.74%	50.53%	0.7584
...	...	...	...
621	-79.69%	31.25%	0.9559
622	-80.16%	46.30%	0.9158
623	-82.37%	44.44%	0.9522
624	-83.02%	46.55%	0.8952
625	-85.36%	45.45%	0.8401

TABLE 14. Methanol performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	112.15%	50.00%	0.9290
2	109.53%	52.63%	0.9254
3	103.34%	56.52%	0.9013
4	101.22%	46.67%	0.8504
5	99.74%	46.91%	0.7700
...	...	...	...
621	-56.54%	46.00%	0.6110
622	-58.77%	43.40%	0.3446
623	-58.85%	46.94%	0.7613
624	-67.03%	47.73%	0.8320
625	-68.49%	44.00%	0.7757

We compared the VPPMA and SMA models by evaluating the metrics of annualized return, profitability ratio, and R-squared value, the values of which can be retrieved from Tables 1–18. Line charts were used to provide a thorough and lucid comparison of the forecasting performances of the VPPMA and SMA models. These charts effectively illustrate

TABLE 15. PTA performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	102.05%	65.96%	0.9215
2	84.83%	60.61%	0.8578
3	80.59%	64.04%	0.8963
4	80.05%	69.31%	0.9320
5	79.54%	61.54%	0.9187
...	...	...	...
621	-28.87%	52.04%	0.0520
622	-29.15%	54.95%	0.0596
623	-29.69%	53.27%	0.0451
624	-34.06%	53.26%	0.1128
625	-49.98%	53.19%	0.4619

TABLE 16. Cotton performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	76.81%	57.76%	0.8960
2	75.11%	60.56%	0.9429
3	74.66%	57.53%	0.9329
4	73.65%	56.46%	0.9460
5	72.88%	57.24%	0.9503
...	...	...	...
621	-32.23%	48.62%	0.7267
622	-32.66%	45.00%	0.6780
623	-33.05%	51.33%	0.6599
624	-36.66%	40.66%	0.9547
625	-38.72%	50.47%	0.7083

TABLE 17. Palm performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	73.63%	50.56%	0.2646
2	65.06%	52.75%	0.1913
3	64.28%	50.49%	0.4267
4	62.76%	46.51%	0.0458
5	61.81%	50.00%	0.1710
...	...	...	...
621	-31.69%	33.85%	0.4982
622	-34.91%	41.51%	0.2386
623	-35.81%	22.86%	0.8331
624	-35.91%	37.37%	0.6497
625	-38.32%	49.40%	0.6613

TABLE 18. Rebar performance using SMA model.

Ranking	Annualized Return	Profitability ratio	R-squared Value
1	52.39%	56.63%	0.9095
2	49.02%	53.78%	0.9476
3	48.64%	69.64%	0.9277
4	47.26%	54.55%	0.9308
5	47.23%	73.47%	0.9341
...	...	...	...
621	-17.53%	51.09%	0.6069
622	-18.33%	53.25%	0.6211
623	-20.33%	50.00%	0.6114
624	-20.36%	52.94%	0.4927
625	-21.63%	52.81%	0.6212

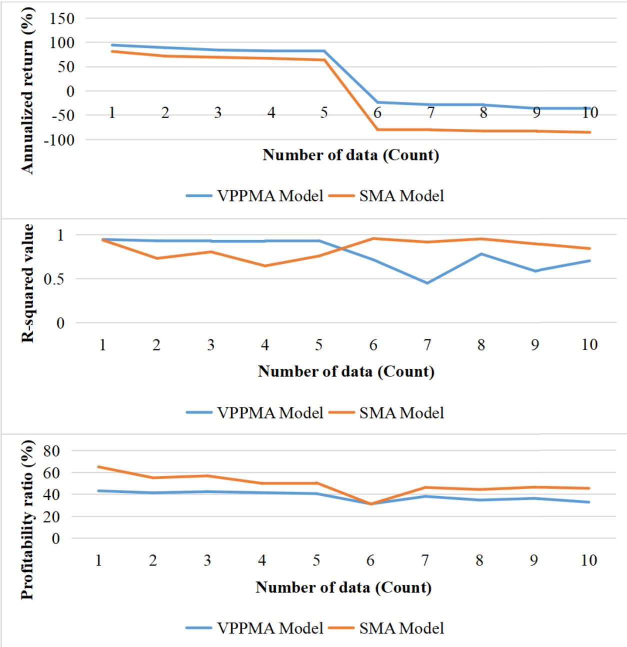

FIGURE 5. Performance of fuel futures in VPPMA and SMA models.

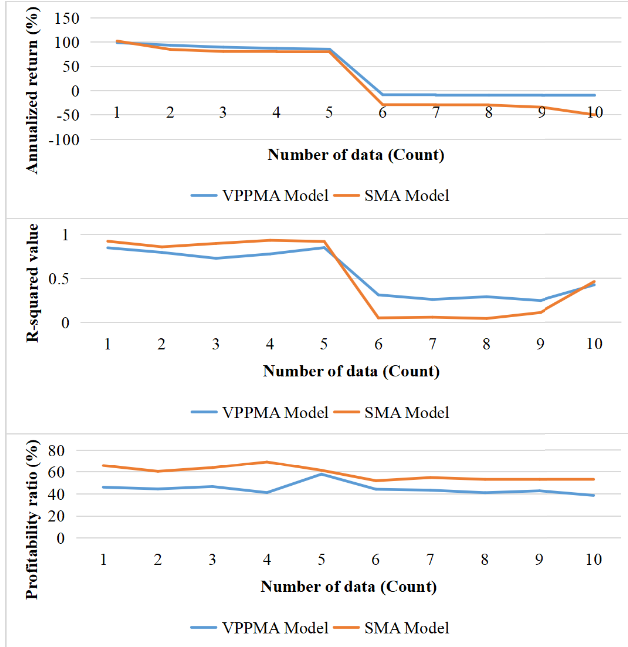

FIGURE 7. Performance of PTA futures in VPPMA and SMA models.

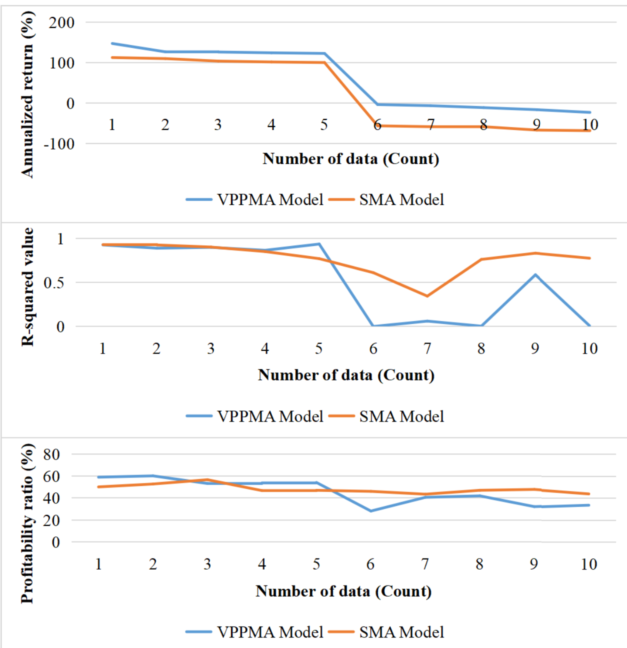

FIGURE 6. Performance of methanol futures in VPPMA and SMA models.

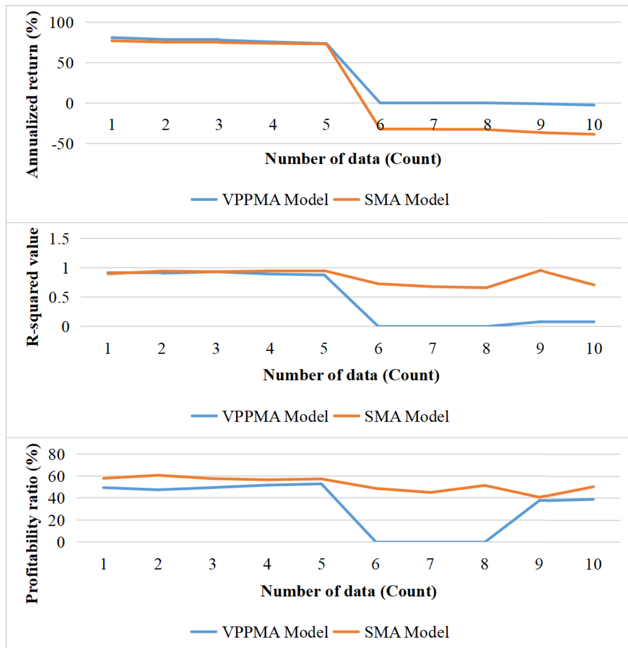

FIGURE 8. Performance of cotton futures in VPPMA and SMA models.

the comparison of the predictive performances of VPPMA and SMA, as shown in Figures 5-10.

Figures 5–10 show that VPPMA consistently outperforms SMA in terms of annualized return indicators from the top five rankings to the bottom five rankings across all varieties. The profitability ratio and R-squared value do not provide a clear indication of which model–VPPMA or SMA–is superior. Regarding the performance of the SMA model, although its profitability ratio does not significantly decrease among the bottom five rankings, the annualized return has already become negative. Similarly, when the annualized return became negative, the R-squared value of the SMA model among the bottom five rankings did not decrease significantly. These findings suggest a weak correlation between the profitability ratio and the R-squared value metrics and annualized return metric in the SMA model. By contrast, VPPMA captures the impact of the profitability ratio and R-squared value on annualized returns. The introduction

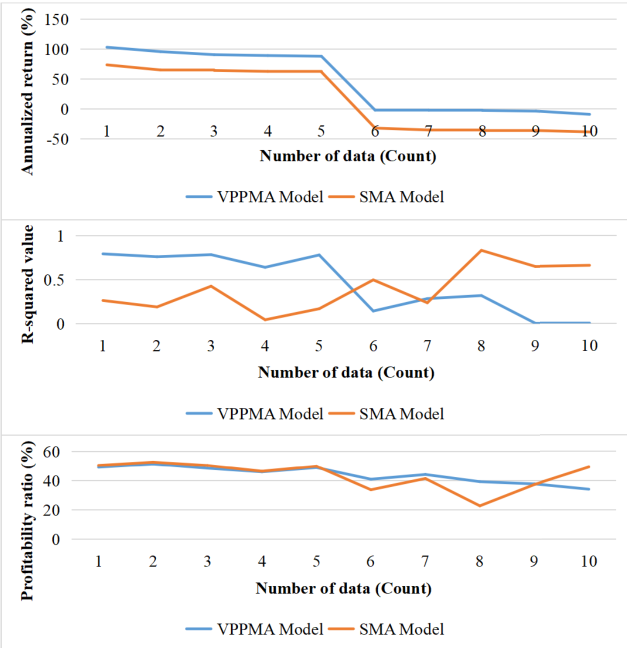

FIGURE 9. Performance of palm futures in VPPMA and SMA models.

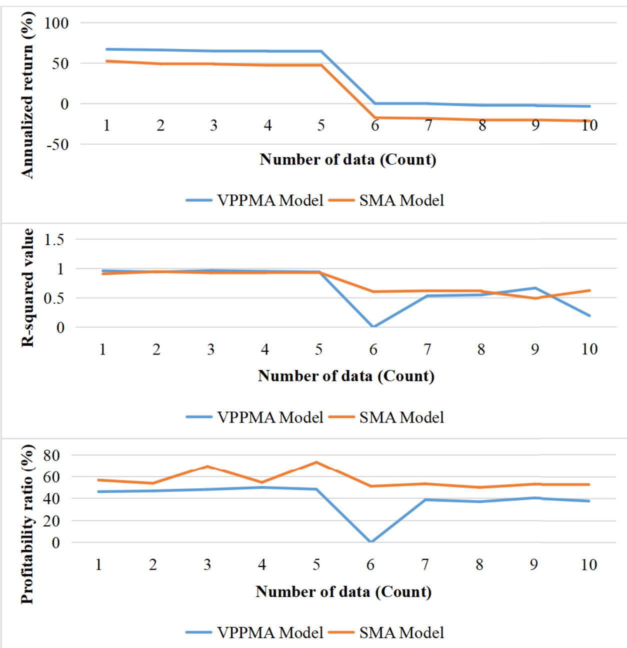

FIGURE 10. Performance of rebar futures in VPPMA and SMA models.

of the VPPMA model with the volume-price product factor enhances the accuracy and regularity of evaluating time series data of futures and has a better annualized return performance.

The VPPMA model demonstrates superior performance, and we attempt to elucidate its effectiveness in Section IV. In the absence of the VPP factor, according to the definition of the loss function in Equation (11), the value is always zero. Consequently, each period of the moving average yields a loss function value of zero. However, in cases where multiple moving averages are used for trend prediction, the same data point can be classified as above certain moving averages or below certain other moving averages. Consequently, the overall loss function value for the combined moving average is not always zero. Algorithm 1 then transforms into selecting a combination of four SMA periods to minimize the overall loss function value, that is, to maximize the annualized return. Upon introducing the VPP factor, there is a higher likelihood that the respective loss functions of the VPPMA and SMA will not be equal to zero simultaneously. This discrepancy can be attributed to the asynchrony between the VPPMA and SMA, leading to amplified divergence. Consequently, there is a higher probability that the overall loss function is non-zero. Ultimately, the inclusion of the VPP factor enhances the discriminability of the moving average model. The process of determining the optimal parameters in Algorithm 1, whether using VPPMA or SMA, can be understood as the process of minimizing the overall loss function. Minimizing the overall loss function indicates an improved classification ability that can promptly generate bullish and bearish crossovers, thereby enabling a timely trend prediction.

According to the proposed theory, a single moving average (SMA) can consistently and accurately classify points at any price. Thus, irrespective of the SMA period, the corresponding classification loss function remains at zero. This implies that, for any SMA period, the returns are identical. Can our model adapt to this extreme scenario? The answer is yes because there is no opening position when there is only one line. When not trading, profit remains constant at zero, and the loss function is zero. Having a loss function of zero is easily explained because doing nothing ensures that there are no mistakes. However, once the conditions for opening a position, such as $MA_1 > MA_2$ or $VPPMA_1 > VPPMA_2$, are satisfied, indicating the presence of at least two lines, the loss function no longer consistently equals zero. At this point, the returns vary owing to the different SMA periods, illustrating the impact of distinct moving averages on returns.

Let us now discuss another scenario: Are more moving averages better? We consider an extreme case assuming an infinite number of moving averages. We hypothesized that an infinite number of SMA lines would exist. In this case, if the condition is set as "$MA_1 > MA_2$, $MA_1 > MA_3, \ldots$" continuing indefinitely, it is equivalent to the condition "$MA_1 > \mathrm{Max}(MA_2, MA_3, \ldots)$" and essentially boils down to comparing only the two lines. Therefore, increasing the number of SMA lines does not necessarily improve the performance. Second, even if we pair an infinite number of SMA lines for comparison, such as "$MA_1 > MA_2$, $MA_3 > MA_4$, $MA5 > MA6, \ldots$," it is equivalent to only the pair of SMAs with the minimum loss function value being effective. Furthermore, considering the condition "$MA_1 > MA_2 > MA_3 > \ldots$," it can be transformed into $MA_1 > MA_2$, $MA_2 > MA_3, \ldots$, so it becomes the second condition.

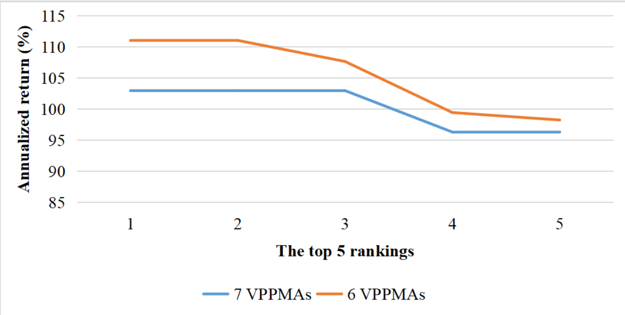

FIGURE 11. The comparison of the top 5 rankings between 7 SMAs and 6 SMAs.

In summary, higher moving averages did not necessarily yield better results. We conducted experiments using six and seven moving averages, with input parameters ranging from N_1 to N_6 and N_1 to N_7, respectively. Figure 11 shows the annualized returns for the top five performers. Surprisingly, the first-ranked performer was a combination of six VPPMAs, surpassing seven VPPMAs. This demonstrates that increasing the number of moving averages does not necessarily lead to an increase in the returns. Compared to the four moving averages provided by Algorithm 1, using six or seven moving averages resulted in higher annualized returns. This is because the opening conditions for the four moving averages in Algorithm 1 (as the conditions for SMA remain unchanged, only the conditions for VPPMA are provided) are subsets of those with six or seven VPPMAs. The opening and closing conditions are presented in Table 19.

TABLE 19. Open and close conditions for 4, 6, and 7 VPPMAs.

Trading	4 VPPMAs	6 VPPMAs	7 VPPMAs
Short selling	$VPPMA_1$ < $VPPMA_2$ and $VPPMA_3$ < $VPPMA_4$	$VPPMA_1$ < $VPPMA_2$ and $VPPMA_3$ < $VPPMA_4$ and **$VPPMA_5$ < $VPPMA_6$**	$VPPMA_1$ < $VPPMA_2$ and $VPPMA_3$ < $VPPMA_4$ and **$VPPMA_5$ < $VPPMA_7$**
Long buying	$VPPMA_1$ > $VPPMA_2$ and $VPPMA_3$ > $VPPMA_4$	$VPPMA_1$ > $VPPMA_2$ and $VPPMA_3$ > $VPPMA_4$ and **$VPPMA_5$ > $VPPMA_6$**	$VPPMA_1$ > $VPPMA_2$ and $VPPMA_3$ > $VPPMA_4$ and **$VPPMA_5$ > $VPPMA_7$**
Closing short position	$VPPMA_1$ > $VPPMA_1$[1] and $VPPMA_3$ > $VPPMA_3$[1]	$VPPMA_1$ > $VPPMA_1$[1] and $VPPMA_3$ > $VPPMA_3$[1] and $VPPMA_5$ > $VPPMA_5$[1]	$VPPMA_1$ > $VPPMA_1$[1] and $VPPMA_3$ > $VPPMA_3$[1] and $VPPMA_5$ > $VPPMA_5$[1]
Closing long position	$VPPMA_2$ < $VPPMA_2$[1] and $VPPMA_4$ < $VPPMA_4$[1]	$VPPMA_2$ < $VPPMA_2$[1] and $VPPMA_4$ < $VPPMA_4$[1] and $VPPMA_6$ < $VPPMA_6$[1]	$VPPMA_2$ < $VPPMA_2$[1] and $VPPMA_4$ < $VPPMA_4$[1] and $VPPMA_6$ < $VPPMA_6$[1]

According to our proposed perspective, combinations of six or seven moving averages are more numerous than combinations of four moving averages, thereby increasing the likelihood of achieving a smaller loss function. Consequently, they outperform the four moving averages. Although the closing conditions for the six and seven moving averages are identical, the opening conditions differ, which is expressed in black font in Table 19. The opening conditions are not mutually inclusive. Incidentally, the combination of six moving averages has a smaller loss function than the combination of seven moving averages, resulting in superior performance of the six moving averages over the seven moving averages. The experimental results agree with the theoretical description, and there is no contradiction between them. (In this experiment, apart from the varying number of moving averages, leading to different combinations for opening and closing positions, all other conditions remain the same).

VI. CONCLUSION AND OUTLOOK

This study explores the prediction of time series data for stocks or futures using two approaches: a deep learning model and a moving average model. The article emphasizes that relying solely on the accuracy, the commonly used comparative indicator of the deep learning model does not guarantee profitability. Additionally, an analysis of the R-squared value shows that the moving average model is more suitable than the deep learning model for predicting stocks or futures data. The commonly used SMA model considers only the price factors. However, trading volume is also an important feature of the time series data. It was found that considering volume alone is insufficient and that the average price line is a fundamental component for predicting such time series data. This study advocates the VPPMA model, which introduces the volume-price product factor into the moving average model. The generalization capability of the moving average model was demonstrated. In conclusion, this study compared the performance of the VPPMA and deep learning models as well as that of the VPPMA and SMA models in predicting futures time series data.

However, further investigation and exploration are required in some areas. In the next step, we focused on two aspects. First, we redefine the loss function of the DL model to enhance its suitability for handling time series data related to stocks or futures. By redefining the loss function, the DL model can optimize the training process and improve its ability to capture unique characteristics and patterns of financial time series data. Second, we continue to explore the factors or features that influence stocks or futures time series data to enhance the model's return rate. For the moving average model, the value of the overall loss function was directly computed for multiple moving average models. Typically, when working with financial data, the goal is to maximize annualized returns. Using the loss function value as an optimization metric enables us to examine its influence on annualized returns. This offers a fresh outlook for enhancing the trend prediction in futures time series data using the moving average model.

REFERENCES

[1] H. M. Jiang and F. Chen, "Modeling and forecasting of financial time series based on deep learning and knowledge graphs," *J. Miniaturized Microcomputer Syst.*, vol. 43, pp. 1793–1800, Jun. 2022.

[2] Z. Hongke, W. Likang, L. Zhi, Z. Xi, L. Qi, and C. Enhong, "Predicting the dynamics in internet finance based on deep neural network structure," *J. Comput. Res. Develop.*, vol. 56, pp. 1621–1631, 2019, doi: 10.7544/issn1000-1239.2019.20190330.

[3] L. Miao, Q. Li, Y. Jiang, J. R. Cui, and Y. X. Wang, "A survey of power system prediction based on deep learning," *Chin. J. Eng.*, vol. 45, no. 4, pp. 663–672, 2023, doi: 10.13374/j.issn2095-9389.2021.12.21.006.

[4] F. Zhiguo, J. Ri, L. Chong, and L. Ang, "Power transformer state monitoring based on transformer deep neural network," *Foreign Electron. Meas. Technol.*, vol. 2, pp. 145–150, 2023.

[5] L. Li, K. Ota, and M. Dong, "Everything is image: CNN-based short-term electrical load forecasting for smart grid," in *Proc. 14th Int. Symp. Pervasive Syst., Algorithms Netw. 11th Int. Conf. Frontier Comput. Sci. Technol. 3rd Int. Symp. Creative Comput. (ISPAN-FCST-ISCC)*, Jun. 2017, pp. 344–351, doi: 10.1109/ISPAN-FCST-ISCC.2017.78.

[6] A. Xu, M.-W. Tian, B. Firouzi, K. A. Alattas, A. Mohammadzadeh, and E. Ghaderpour, "A new deep learning restricted Boltzmann machine for energy consumption forecasting," *Sustainability*, vol. 14, no. 16, p. 10081, Aug. 2022.

[7] Y. Wang and Y. Tang, "Traffic forecasting network combining autocorrelation attention and dynamic convolutions," *J. Miniaturized Microcomput. Syst., Eng. Sci. Technol. Compilation II*, vol. 4, pp. 1–9, 2023.

[8] H. Shao and B.-H. Soong, "Traffic flow prediction with long short-term memory networks (LSTMs)," in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2016, pp. 2986–2989, doi: 10.1109/TENCON.2016.7848593.

[9] K. Jin, J. Wi, E. Lee, S. Kang, S. Kim, and Y. Kim, "TrafficBERT: Pre-trained model with large-scale data for long-range traffic flow forecasting," *Exp. Syst. Appl.*, vol. 186, Dec. 2021, Art. no. 115738.

[10] O. Surakhi, M. A. Zaidan, P. L. Fung, N. Hossein Motlagh, S. Serhan, M. AlKhanafseh, R. M. Ghoniem, and T. Hussein, "Time-lag selection for time-series forecasting using neural network and heuristic algorithm," *Electronics*, vol. 10, no. 20, p. 2518, Oct. 2021.

[11] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of LSTM and BiLSTM in forecasting time series," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2019, pp. 3285–3292.

[12] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, Dec. 2020.

[13] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.

[14] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 379, no. 2194, Apr. 2021, Art. no. 20200209, doi: 10.1098/rsta.2020.0209.

[15] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso, "Deep learning for time series forecasting: A survey," *Big Data*, vol. 9, no. 1, pp. 3–21, 2021, doi: 10.1089/big.2020.0159.

[16] H. T. Liang, S. Liu, J. W. Du, Q. Hu, and X. Yu, "A comprehensive review of deep learning applied in time series forecasting research," *Comput. Sci. Explor.*, vol. 6, pp. 1285–1300, 2023.

[17] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[18] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "WaveNet: A generative model for raw audio," 2016, *arXiv:1609.03499*.

[19] J. Gehring, M. Auli, D. Grangier, and Y. N. Dauphin, "A convolutional encoder model for neural machine translation," 2016, *arXiv:1611.02344*.

[20] C. Stoean, W. Paja, R. Stoean, and A. Sandita, "Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations," *PLoS ONE*, vol. 14, no. 10, Oct. 2019, Art. no. e0223593, doi: 10.1371/journal.pone.0223593.

[21] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0212320, doi: 10.1371/journal.pone.0212320.

[22] F. Zhou, H.-M. Zhou, Z. Yang, and L. Yang, "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction," *Exp. Syst. Appl.*, vol. 115, pp. 136–151, Jan. 2019.

[23] C. Phetking, M. N. M. Sap, and A. Selamat, "Identifying zigzag based perceptually important points for indexing financial time series," in *Proc. Int. Conf. Cogn. Inform.*, 2009, pp. 295–301, doi: 10.1109/COGINF.2009.5250725.

[24] H. M. Jiang, "Design and implementation of Zigzag-based K-line trend turning point recognition algorithm," *Comput. Technol. Autom.*, vol. 36, no. 4, pp. 68–71, 2017, doi: 10.16339/j.cnki.jsjsyzdh.2017.04.016.

[25] P.-C. Chang, C.-Y. Fan, and C.-H. Liu, "Integrating a piecewise linear representation method and a neural network model for stock trading points prediction," *IEEE Trans. Syst., Man, Cybern., C Appl. Rev.*, vol. 39, no. 1, pp. 80–92, Jan. 2009.

[26] F. Li, F. Gao, and P. Kou, "Integrating piecewise linear representation and Gaussian process classification for stock turning points prediction," *Comput. Appl.*, vol. 8, pp. 2397–2403, 2015, doi: 10.11772/j.issn.1001-9081.2015.08.2397.

[27] H. Tang, P. Dong, and Y. Shi, "A new approach of integrating piecewise linear representation and weighted support vector machine for forecasting stock turning points," *Appl. Soft Comput.*, vol. 78, pp. 685–696, May 2019, doi: 10.1016/j.asoc.2019.02.039.

[28] W. H. Li, "A new moving average line—PVMA," *J. Hunan Univ. Commerce*, vol. 6, pp. 93–94, 2002, doi: 10.3969/j.issn.1008-2107.2002.06.031.

[29] Y. B. Wang, "Research on stock investment strategies based on volume-price moving averages," M.S. thesis, College Sci., Minzu Univ. China, 2021, doi: 10.27667/d.cnki.gzymu.2021.000384.

[30] D. H. Kan, *Machine Learning and Its Python Practice* (Micro-Course Video Edition). Beijing, China: Tsinghua Univ. Press, 2022, pp. 152–184.

[31] Interface News. (2023). *The Annualized Return of Berkshire Hathaway is 19. 8 % for 57 Consecutive Years. Is It As High As You Expected?* [Online]. Available: https://baijiahao.baidu.com/s?id=1758877996860231317&wfr=spider&for=pc

YIPIAO CHEN was born in April 1987. He received the master's degree in computer software and theory from Lanzhou University. He started working in July 2013 and joined the School of Cyberspace Security, Gansu University of Political Science and Law, in June 2021. He is currently a Senior Engineer (not hired by the school) and a Consulting Engineer (investment). He is also engaged in teaching courses, such as computer network and Python data analysis and application. He is participating in one provincial project and one departmental project.

XIAOGANG YUAN was born in November 1980. He received the degree from the Communications and Information System Program, Air Force Engineering University, in June 2009, and the Ph.D. degree in engineering. He is currently an Associate Professor and a Master's Supervisor. He is also engaged in research work in network security and data security and is responsible for the teaching of courses, such as information content security, digital forensics technology, C language programming, and content understanding and public opinion analysis. He is leading one provincial-level project and two department-level projects and has published more than 30 academic papers. He has guided students to participate in various academic competitions and won two national and three provincial awards. He was awarded a third-class merit citation in 2010 and transferred to the College of Cyberspace Security, Gansu University of Political Science and Law, in October 2021.