Received: 6 April 2021 | Revised: 27 October 2022 | Accepted: 4 January 2024

DOI: 10.1002/for.3071

RESEARCH ARTICLE

WILEY

Stock movement prediction: A multi-input LSTM approach

Pan Tang | **Cheng Tang** | **Keren Wang**

School of Economics and Management, Southeast University, Nanjing, 211189, China

Correspondence
Pan Tang, School of Economics and Management, Southeast University, Nanjing, Jiangsu 211189, China.
Email: pantang@seu.edu.cn

Funding information
National Social Science Fund of China

Abstract

Generally, the nonlinear and non-stationary financial time series becomes an obstacle in the process of stock movement prediction. But the recent theories of machine learning and deep learning have provided with some new solutions. Based on LSTM (long short-term memory), we propose a hybrid model of wavelet transform (WT) and multi-input LSTM to predict the trend of SSE composite index. It can mine valid data in time series and support different types of data as input. The whole model is divided into two stages. In the first stage, we adopt the level 1 decomposition with db4 mother wavelet to eliminate noise. In the second stage, combinative and qualitative analysis was made base on the data from Chinese stock market, US stock market, and technical indicators as input. According to the result, the proposed model, with the accuracy of 72.19%, performs better than single-input LSTM, decision tree, random forest, Support Vector Machine (SVM), and XGBoost.

KEYWORDS

deep learning, multi-input LSTM, stock movement prediction, wavelet transform

1 | INTRODUCTION

1.1 | Background

Stock market is an important driving force for the economic development of a country. As China continuously accelerates the pace of capital market opening and deepens the overall reform, its stock market provides opportunities for both domestic and foreign investors to raise their wealth through investments. Nevertheless, stock market is pretty volatile and it could bring losses to investors. It is necessary to construct stock price prediction models to tackle market risks. Therefore, many scholars devote themselves to the research of stock price forecasting, which is also the research direction of this paper.

Predicting stock prices is a classic problem. Previous studies use traditional statistical methods in time series prediction, which can only predict stock prices, rather than price movement, such as GARCH (Agnolucci, 2009; Awartani & Corradi, 2005; Bentes, 2015; Kriechbaumer et al., 2014), ARIMA (Torbat et al., 2018). The methods are mainly based on linear, stable, and normal distribution assumptions. However, because the financial time series is a nonlinear and non-stationary time series with low SNR (signal-to-noise ratio), the traditional statistical methods are sometimes ineffective in prediction problem, which have poor efficiency and accuracy of the forecast results. Compared with statistical methods, machine learning (Patel et al., 2015a; Vijh et al., 2020; Nevasalmi, 2020) can handle noisy financial time series. And with the rapid development of artificial intelligence, machine learning stands out in this area. Nayak et al. (2015) combine Support Vector Machine (SVM) with K-Nearest Neighbor (KNN) to Indian stock market indices prediction. Henrique et al. (2018) use Support Vector Regression (SVR) on daily and up to the minute prices to predict stock prices in three different markets. They demonstrate that the mean square error caused by linear and radial kernels is smaller than that of random walk model, especially when updating the model periodically. In particular, deep learning (Balaji et al., 2018; Chong et al., 2017; Liu & Long, 2020; Long et al., 2019) has robustness, self-learning ability, and a strong ability to fit nonlinear data, which can effectively improve the prediction accuracy. For example, Nikou et al. (2019) conclude that deep learning performs better than SVR and random forest. Saud and Shakya (2020) use three deep learning models with suitable look-back period, including Vanilla Recurrent Neural Networks (RNN), LSTM, and GRU, to predict stock prices of two representative banks listed on Nepal Stock Exchange (NEPSE). The results show that GRU has the best prediction performance.

Numerous studies over the years have focused on the direction of stock prices. Because prices are also influenced by political events, economic conditions, and investors' expectations, it is more difficult to predict the stock price movement than to predict prices alone (Huang et al., 2005). In order to solve the problem, machine learning and deep learning have been applied to improve prediction performance. Kara et al. (2011) compare performances of ANN and SVM in predicting movement directions of the daily Istanbul Stock Exchange (ISE) National 100 Index. The average performance of ANN model is 75.74%, better than that of SVM (71.52%). By using 10 technical parameters and inputting their trends, respectively, Patel et al. (2015b) compare four prediction models: ANN, random forest, naive-Bayes, and SVM. The results suggest that random forest outperforms the other three models with 10 technical parameters, and all of them improve with corresponding trends. When predicting stock price direction, another study indicates that random forest is the best algorithm via comparing ensemble methods (random forest, Adaboost, and Kernel Factory) with single classifier models (Neural Networks, Logistic Regression, SVM, and KNN) (Ballings et al., 2015). Zhang et al. (2018) utilize a combination of random forest, imbalance learning, and feature selection to forecast stock price movement and its interval of growth (or decline) rate. The trend of stock price is divided into four main classes (Up, Down, Flat, and Unknown). Basak et al. (2019) build a predictive model based on random forest and XGBoost classifiers, within a precision of 78% in forecasting the direction of stock movements. Compared to state-of-the-art baseline algorithms, Hoseinzade and Haratizadeh (2019) construct a CNN-based framework to predict stock market, with a significant improvement by about 3% to 11%, in terms of F-measure.

However, deep learning fails to obtain satisfactory results because it cannot extract multi-frequency features of stock prices. To solve the problem, wavelet transform (WT) can be utilized to process time series (Gençay et al., 2002; Percival & Walden, 2000). It can effectively separate the low-frequency information and high-frequency information of stock prices so as to reflect the useful information without noise. Consequently, many scholars combine WT with models. Fernandez (2007) compares four methods: multiplicative seasonal ARIMA, unobserved components (UC), SVM, and wavelet based. Based on the Granger–Newbold test, it appears that wavelet based outperforms than the others. Hsieh et al. (2011) combine WT and RNN based on artificial bee colony algorithm (ABC-RNN) to forecast stock prices. Ortega and Khashanah (2014) propose a wavelet neural network model which can achieve relatively accurate results for short-term stock returns. Lei (2018) constructs an integrated prediction method based on Rough Set (RS) and Wavelet Neural Network to predict stock price trends. The prediction results are better than other neural networks, SVM, WNN, and RS-WNN. Li and Tang (2020) propose WT-FCD-MLGRU model, which is better than ARIMA and SVR. Recently, some scholars combine WT with LSTM to predict financial time series and achieve significant improvements. Li and Tam (2017) find that the high degree of fluctuations are removed in the high volatility original indexes after wavelet denoising so that the combined model can get better results. Liang et al. (2019) propose MOCWT method to reduce the degree of distortion in signal reconstruction. The prediction results of the novel model have less oscillation. Hence, in this paper, we combine WT with a novel LSTM model in order to improve the accuracy of prediction.

1.2 | Objectives

The objective of this paper is to reduce noise in financial time series and add different types of data as input to improve prediction accuracy. Based on the above analysis, we propose a hybrid model of WT and multi-input LSTM to predict the trend of SSE composite index. WT is an ideal algorithm to extract features contained in financial data, which can analyze data both in the time domain and the frequency domain. Furthermore, according to the existing studies, we are the first to use multi-input LSTM in forecasting stock price movements. The advantage of multiple inputs allows us to take the data from Chinese stock market, US stock market, and technical indicators as input. And with its well verified effectiveness, the accuracy of this data-rich method is up to 72.19%.

The remainder of this paper is structured as follows: Section 2 introduces the methods used in the research, including WT and LSTM. Section 3 briefly describes the research data. Section 4 provides details on the use of the above methods and presents the results. Finally, Section 5 concludes.

2 | METHODOLOGY

2.1 | Wavelet transform

Previously, WT was mainly used in image and signal processing. In recent years, because of the powerful ability of noise reduction and feature extraction, it is more widely used in stock price forecasting (Aussem, 1998; Alrumaih & Al-Fawzan, 2002; Caetano & Yoneyama, 2007; Huang, 2011). It first appeared in Mortlet's analysis of seismic data in 1984. Subsequently, Mallat proposed multi-resolution analysis (MRA) and simplified algorithm for wavelet coefficient calculation in 1989. Compared to other information extraction techniques, WT overcomes the problems of Fourier transform, which can only be used in stationary time series and may lose its efficiency in processing short-term series.

In wavelet transform, discrete wavelet transform (DWT) is the most suitable for financial time series analysis, which can capture the frequency and location information. The essence of WT is a filtering process. First, the given signal is passed through the low-pass filter to obtain approximation coefficients. And then detail coefficients are filtered out by the high-pass filter. By means of multi-resolution analysis and Mallat algorithm, effective information can be extracted from financial time series with low SNR. For example, signal x is passed through a half band high band pass filter h and a low-pass filter g. The filter output is subsampled by 2, which is given by

$$y_{\text{low}}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k], \tag{1}$$

$$y_{\text{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k], \tag{2}$$

Decomposition diagram of DWT is illustrated in Figure 1. Signal x is decomposed into CA_n and CD_n by Mallat algorithm. CA_n is the approximation coefficients, and CD_n is the detail coefficients (n is the number of decomposition levels).

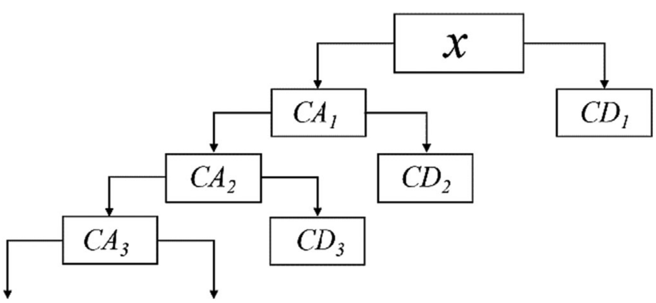


FIGURE 1 Decomposition diagram of DWT.

Various types of wavelets have different properties. Daubechies wavelet is widely used in time series prediction because of its advantages of compact support and orthogonality (Bai et al., 2015; Ramsey, 2002; Yousefi et al., 2005). Moreover, many previous studies have chosen db4 as the wavelet base and obtained better prediction results (Kao et al., 2013; Lahmiri, 2014; Li, 2015). According to the literatures, db4 mother wavelet is adopted in this paper.

2.2 | LSTM

Deep learning has strong self-learning ability to extract hidden information in massive data (Selvin et al., 2017). In particular, LSTM has memory ability because of its gating mechanism, so it is more suitable to forecast financial time series. Developed on the basis of Recurrent Neural Networks (RNN), LSTM solves the long-term dependence problem of RNN (vanishing and exploding gradients) by introducing the gating mechanism innovatively. As a result, LSTM can keep information over a long period of time. Furthermore, because financial time series is characterized by dynamic instability and long-term dependence. Therefore, LSTM can be used to predict stock price index because it can solve the long-term dependence (Lin et al., 2021).

The gating mechanism of LSTM consists of many memory blocks. As shown in Figure 2, each block has a cell state and three gates (forget gates f_t, input gates i_t, and output gates o_t). x_t denotes input and h_t denotes hidden state. C_t and $\tilde{C}_t$ denote candidate values.

The function of f_t is to determine the information discarded from the cell status. The mathematical expression of f_t is given by

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \tag{3}$$

where W denotes weight, b denotes bias, and σ denotes sigmoid function, which decides the output of the cell state.

The input gate consists of the following two formulas, which determine the information stored in the cell state:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{4}$$

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{5}$$

where $tanh$ is an activation function. And then the cell status is updated according to the below function:

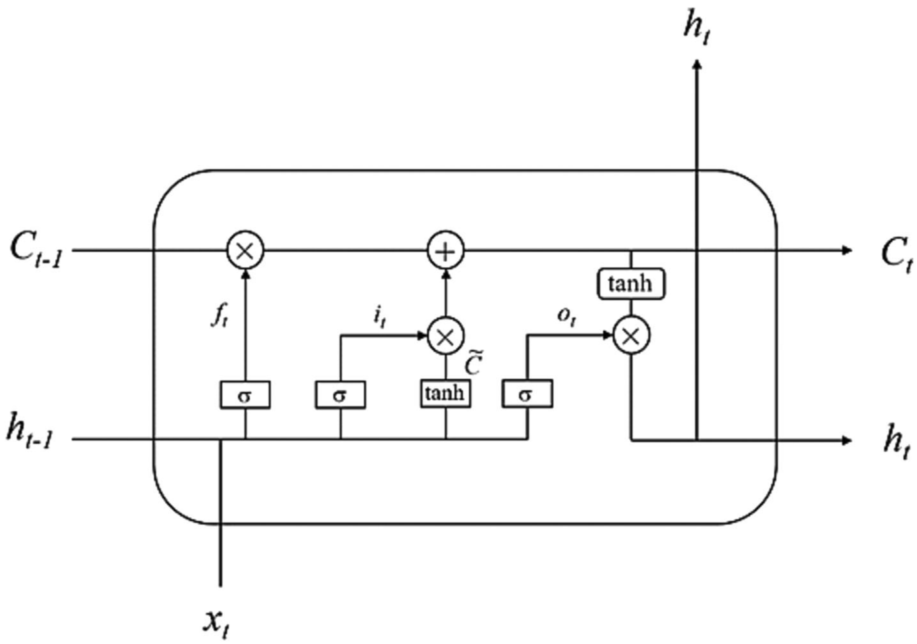


FIGURE 2 Structure diagram of LSTM. LSTM, long short-term memory.

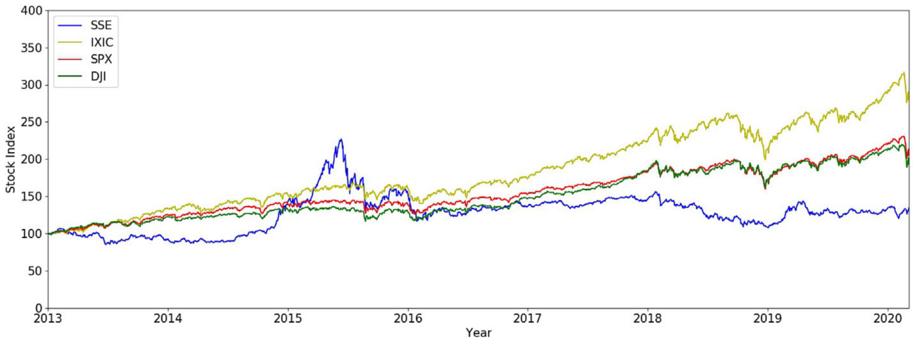


FIGURE 3 Trend of four stock indices.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t. \tag{6}$$

Finally, we get the output through the output gate. o_t and h_t are calculated by

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{7}$$

$$h_t = o_t \times tanh(C_t). \tag{8}$$

3 | DATA

3.1 | Description of data

The study employs original data from Chinese and US stock markets to predict the direction of SSE Composite Index movement. The reason for using US stock data is that stock market correlation among countries with closer economic cooperation is relatively high (Bekaert & Harvey, 1995; Gultekin et al., 1989). Before China joined the World Trade Organization (WTO), there seemed to be no causal relationship between Chinese and American stock markets (Huang et al., 2000). However, China has closer economic ties with the world after joining the WTO (Johansson, 2009; Rumbaugh & Blancher, 2004).

We compare the trends of daily close prices in China and the United States by drawing a time series line chart, which contains four indicators including daily closing prices from SSE, IXIC, S&P500, and DJI. For the convenience of comparison, we take the first trading day of 2013 as the base date, and each of the four stock indices is set at 1. Except for Chinese sudden change in the 2015 stock market, four indicators are about the same in trend (Figure 3).

In recent years, many studies have shown that there is a high correlation between Chinese and American stock markets. Goh et al. (2013) find that American economic variables have significant ability to forecast

TABLE 1 Descriptive statistics of data.

Index	Mean	Max	Min	SD	Skewness	Kurtosis
IXIC	5731.78	9817.18	3091.81	1610.64	0.38	−0.93
S&P 500	2289.64	3386.15	1457.15	454.27	0.31	−0.88
DJI	20295.32	29551.42	13328.85	4277.42	0.43	−1.22
SSE close	2899.70	5166.35	1950.01	562.93	0.54	1.18
SSE open	2895.80	5174.42	1935.52	562.16	0.55	1.22
SSE high	2919.71	5178.19	1959.16	571.13	0.58	1.25
SSE low	2872.97	5103.40	1849.65	550.78	0.48	1.01

TABLE 2 Variables for model building.

Data set	Name of indicators	Formulas
U.S. stock data	U_{IXIC}	$10*\left(\frac{IXIC_t}{IXIC_{t-1}}-1\right)$
	$U_{S\&P500}$	$10*\left(\frac{S\&P500_t}{S\&P500_{t-1}}-1\right)$
	U_{DJI}	$10*\left(\frac{DJI_t}{DJI_{t-1}}-1\right)$
SSE stock data	C_{close}	$\frac{close_t}{close_{t-1}}-1$
	C_{open}	$\frac{open_t}{open_{t-1}}-1$
	C_{high}	$\frac{high_t}{high_{t-1}}-1$
	C_{low}	$\frac{low_t}{low_{t-1}}-1$

Chinese stock market after China's admission into the WTO. Ye (2014) finds that the daily returns on the American stock market have strong predictive power for Chinese stock market openings since 2006. Dutta (2018) indicates high correlations among equity markets.

U.S. stock data here tracks the positioning of daily closing prices within the daily range of IXIC, S&P500, and DJI. Chinese counterpart takes typical daily prices of SSE into account, including open, high, low, and close prices. For comparison, the observation of data starts on January 4, 2013 and ends on March 5, 2020, 1742 days in all. The predictors, along with their descriptive statistical results, are shown in Table 1.

3.2 | Data preprocessing

We assess the direction of SSE movement via three variable groups, namely US stock data, SSE stock data, and 10 technical indicators. The first two are listed in Table 2. The data are presented in the form of yield, which reflects the daily movements of the index so as to predict whether the index will rise or fall on the following day. US stock data used in the model as input can measure the impact of the external environment, because the US stock market has influence on the Chinese stock market based on the above research. SSE stock data select the daily open, high, low, and close prices, which can reflect the internal influencing factors of the index. In particular, SSE stock data can be chosen with or without noise reduction before data preprocessing.

Technical indicators can indicate current market conditions and reflect price movements over a period of time. The following 10 technical indicators are often studied as determinations of stock price movements (Kim, 2003; Shynkevich et al., 2017). Aimed at predicting SSE

TABLE 3 Technical indicators (5 days).

Name of indicators	Formulas
Simple moving average (MA)	$\frac{close_{t-4}+close_{t-3}+\cdots+close_t}{5}$
Exponential moving average (EMA)	$EMA_{t-1}+\frac{2}{6}\cdot(close_t-EMA_{t-1})$
Average true range (ATR)	$\frac{TR_{t-4}+TR_{t-3}+\cdots+TR_t}{5}$, where $TR_t=\max\{H_t,\lvert close_{t-1}-low_t\rvert\}$, $H_t=\max\{high_t-low_t,\lvert close_{t-1}-high_t\rvert\}$
Average directional movement index (ADMI)	$\frac{DI_5^+-DI_5^-}{DI_5^++DI_5^-}\cdot 100$, where $DI_5^+=\frac{EMA_5(DM^+)}{ATR_5}\cdot 100$, $DI_5^-=\frac{EMA_5(DM^-)}{ATR_5}\cdot 100$
Commodity channel index (CCI)	$\frac{M_t-SM_t}{0.015D_t}$, where $M_t=\frac{high_t+low_t+close_t}{3}$, $SM_t=\frac{\sum_{i=1}^{5}M_{t-i+1}}{5}$, $D_t=\frac{\sum_{i=1}^{5}\lvert M_{t-i+1}-SM_t\rvert}{5}$
Rate of change (ROC)	$\frac{close_t-close_{t-5}}{close_{t-5}}\cdot 100$
Relative strength index (RSI)	$100-\frac{100}{1+\frac{EMA_5(DM^+)}{EMA_5(DM^-)}}$
The William's %R oscillator (%R)	$\frac{hp_t-close_t}{hp_t-lp_t}\cdot 100$
Stochastic K% (SK)	$\frac{p_t-lp_t}{hp_t-lp_t}\cdot 100$
Stochastic D% (SD)	$\frac{SK_{t-4}+SK_{t-3}+\cdots+SK_t}{5}$

Note: $lp_t=\min\{close_{t-4},close_{t-3},\cdots,close_t\}$,
$hp_t=\max\{close_{t-4},close_{t-3},\cdots,close_t\}$,
$DM^+=\max(close_t-close_{t-1},0)$,
$DM^-=\min(close_t-close_{t-1},0)$.

composite index, we cover these influences by using a 5-day time window to reflect short-term trends that consist of 10 technical indicators. The names of indicators, along with their formulas, are presented in Table 3.

Above all, the meaning of each indicator is shown as follows.

- Simple Moving Average (SMA) is the most widely used one. It is the average of prices over a period of time, which can reflect the trend of stock prices.
- Exponential Moving Average (EMA) is also used to observe price trends. It is calculated as an exponential and degressive weighted moving average over a particular period. Compared to SMA, EMA places a greater weight on the most recent data points. As a result, it is more sensitive to short-term price changes.
- Average True Range (ATR) is a moving average of stock price fluctuations within a certain period of time, which is mainly used to study the timing of buying and selling. The higher the value of the indicator, the more likely the trend will change, and vice versa.
- Average Directional Movement Index (ADMI) is a kind of technical indices to indicate the strength of a trend by analyzing the change of equilibrium point between buyers and sellers in the process of price change.

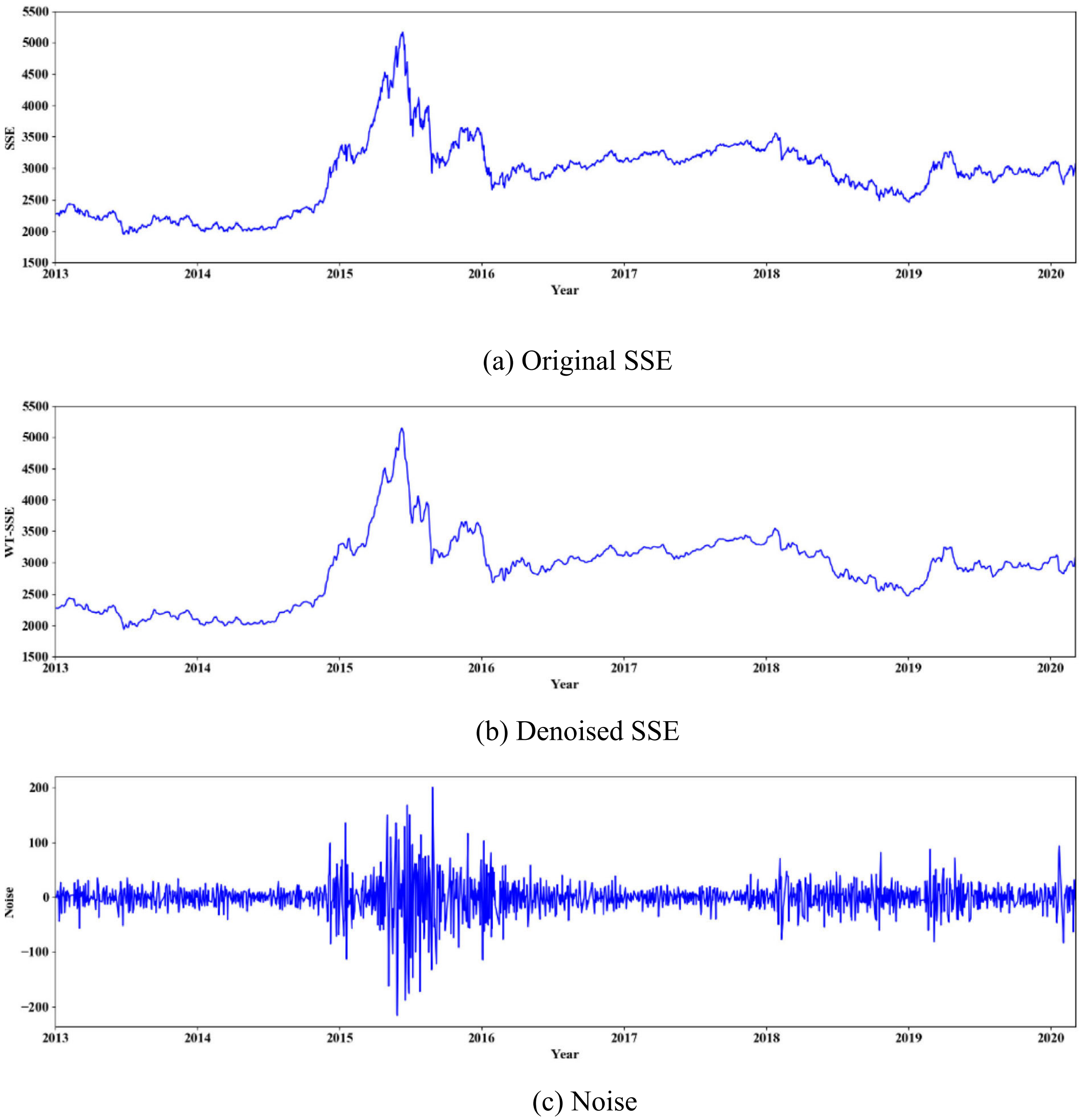


FIGURE 4 Difference between original data and denoised data.

- Commodity Channel Index (CCI) is an index that measures whether the stock price has exceeded the normal distribution range. It belongs to a special kind of overbought and oversold index, which has an infinite range of fluctuations.
- Rate of change (ROC) is to observe the speed changes in the stock market by calculating rate of price change over a certain period.
- Relative strength index (RSI) is a momentum indicator taking the speed and changes of price into account, which tries to find out whether the stock is overbought or oversold. RSI oscillates between 0 and 100. In general, when RSI is below 30, it indicates a buy signal, and when RSI is above 70, it indicates a sell signal.
- The William's %R oscillator (%R) is also a momentum indicator. It indicates the relationship between a market's closing price and the highest price over the past period. It ranges from −100 to 0. When its value is below −80, it may show that the stock is oversold, and when its value is above −20, it may indicate the stock is overbought.
- Stochastic K% (SK) is used to measure the idea of momentum and compares a close price and its price interval during a period of particular past days.
- Stochastic D% (SD) is the moving average of SK. They both belong to Stochastic Oscillator (SO) and measure the level of closing price relative to low–high range over a period of time.

Later, in order to improve model accuracy and accelerate model convergence, we adopt z-score to scale SSE stock data and technical indicators. The standardized method can scale data of different sizes and dimensions to the same interval and range, so that it can reduce the impact of size, characteristics, and distribution differences on the model. The specific formula is expressed as follows:

$$z = \frac{x-\mu}{\sigma}, \tag{9}$$

where z is the result of standardization, and μ,σ are the mean and standard deviation of x, respectively.

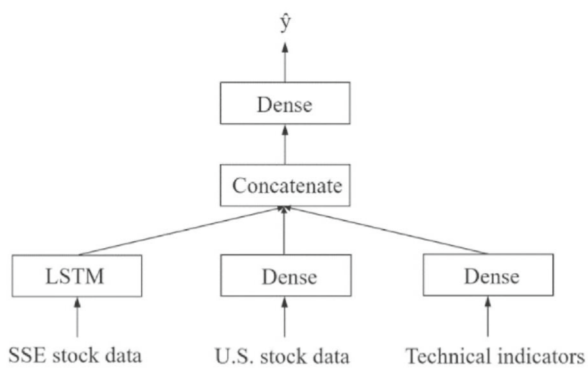
4 | EMPIRICAL ANALYSIS

4.1 | Wavelet denoising

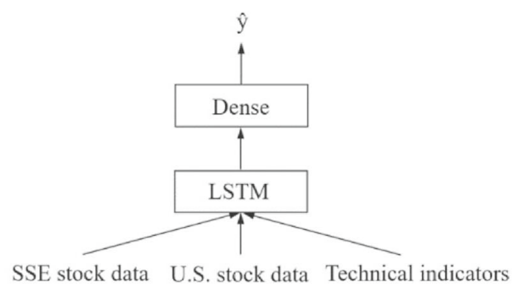
We reduce the noise of financial time series through wavelet transform. In this work, the level 1 decomposition with db4 mother wavelet is used to process SSE close. Moreover, the high-frequency coefficient is set to 0 to reconstruct the data, so as to obtain the SSE after denoising. Figure 4 represents the original and denoised SSE. In comparison with the first graph, the local jitter phenomenon in the second graph is significantly reduced after denoising. And as shown below, the modified figure looks smoother, which means the white noise is noticeably reduced. Later, we feed the reconstructed denoised lower-frequency data into multi-input LSTM so that prediction performance can be improved.

4.2 | Multi-input LSTM

After wavelet denoising, a novel multi-input LSTM method that we propose is applied to predict stock index. Multi-input LSTM is a variant of LSTM model, and it is an indispensable part for the analysis of multidimensional data. Compared with single-input LSTM, multi-input LSTM performs more accurately in prediction because it takes different types of data into account. Therefore, it has been widely used in many fields where a large amount of data needs to be processed, such as



(a) Multi-input LSTM



(b) Single-input LSTM

FIGURE 5 Difference between multi-input and single-input LSTM. LSTM, long short-term memory.

TABLE 4 Confusion matrix.

Actual	Predicted Positive	Negative
True	TP	FN
False	FP	TN

TABLE 5 Evaluation metrics.

Evaluation metrics	Formulas
Accuracy	$\frac{TP+TN}{TP+FP+FN+TN}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{FP+TN}$
Prec rise	$\frac{TP}{TP+FP}$
Prec fall	$\frac{TN}{FN+TN}$

computer vision and natural language processing. Because it involves three types of data, including information from SSE stock market, US stock market ,and technical indicators of SSE composite index, in this research, it is logical to adopt to multi-input LSTM method.

The model takes SSE stock data of five trading days, US stock data of one trading day, and technical indicators of one trading day as input. Then the three types of data go through LSTM and the full connection layer. Finally, the output of the previous layer is concatenated and then fed into the full connection layer to get the result. In contrast, single-input LSTM directly feeds all the data from the previous 5 days into the LSTM layer to predict SSE movement. Figure 5 shows the difference in the input structure.

In order to predict the trend of stock index, the rise and fall of the stock index are divided into class 1 and class 0, respectively. The assignment of labels to each data is performed according to the formulas below. Label 1 indicates an increase in the stock index. Label 0 indicates that the index will go down.

$$Lable_{t+1} = \begin{cases} 1, if\left(C_{t+1} - C_t\right)/C_t > 0; \\ 0, if\left(C_{t+1} - C_t\right)/C_t \leq 0. \end{cases} \tag{10}$$

where C_t means price of stock index on day t.

Because of rules of thumb, we adopt mini-batch gradient descent to train LSTM network, of which the size is set to 32. Mini-batch method can reduce the amount of calculation and reduce the randomness of gradient descent. We use Binary Crossentropy as loss function aiming to forecast the next-day movement of the stock index. To train LSTM model, it may have an excellent performance by introducing RMSProp optimizer, which can realize the adaptive adjustment of the learning rate of each parameter. In addition, we consider the full sample that ranges from January 4, 2013 to March 5, 2020, of which the first 80% composes the training set and the rest makes up the test set. The training epoch is set to 200. The whole experiment is conducted under Keras framework.

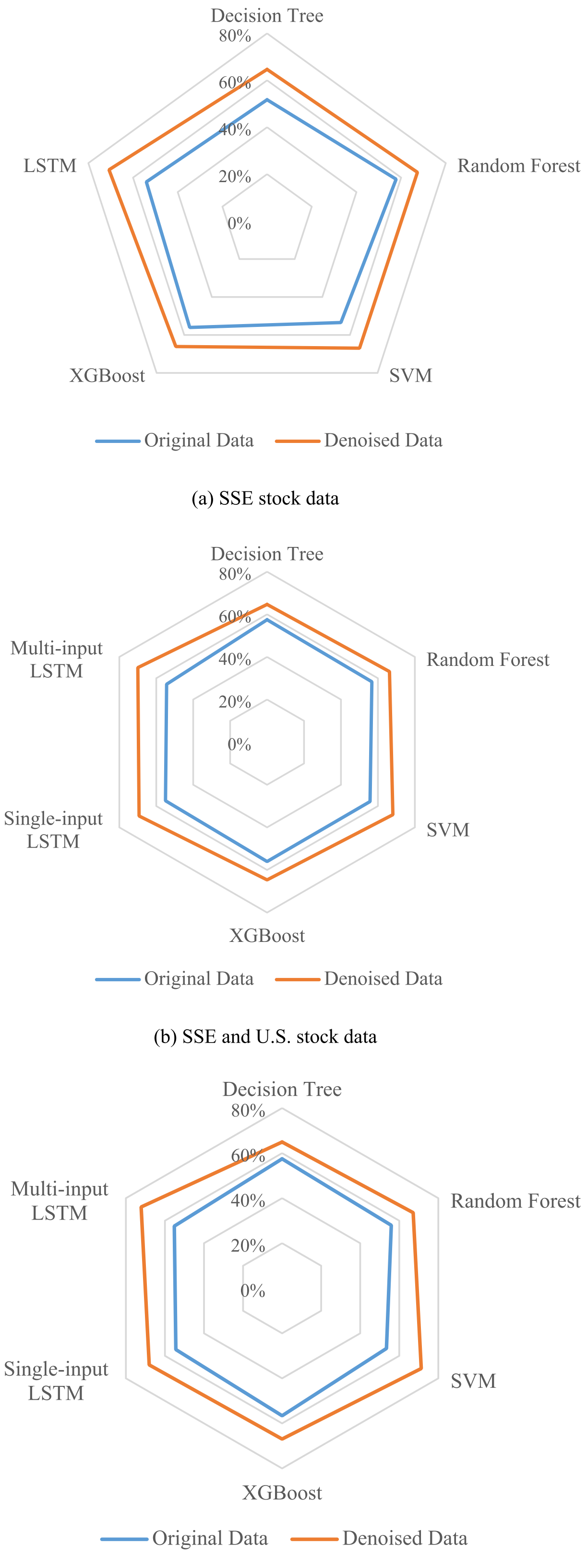
4.3 | Evaluation metrics

Four typical evaluation metrics for predictive performance are applied to measure forecasting accuracy of each model, including accuracy, sensitivity, specificity, and precision for class 1 and class 0. Symbols are defined in Table 4, and metrics are calculated in Table 5.

4.4 | Result

Another five models, including single-input LSTM, decision tree, random forest, SVM, and XGBoost, have been used to compare the prediction results with that of multi-input LSTM. Tables 8 and 9 present the results. Train acc stands for average accuracy on the training set, while Test acc stands for average accuracy on the test set. Sens, Spec, Prec rise, and Prec fall are averages for sensitivity, specificity, and precision of rise and fall on the test set, respectively.

Comparison of the two tables shows that by using the denoised SSE to predict, the performance of all models is improved, indicating that WT can effectively eliminate the noise of time series. The test accuracy without denoising among models ranges from 51.72% to 57.76%, whereas the test accuracy with denoising oscillates between 64.66% and 72.19%. By utilizing wavelet transform, there is an increase of more than 10%. Figure 6 can intuitively show the influence of WT on the models. As seen in all three pictures, orange lines representing results of models using denoised data as input are outside blue lines representing original data. Blue lines are all below 60%, whereas orange lines are above 60%, which shows that all models achieve better results after wavelet denoising. Moreover, the positive influence of WT on LSTM can be verified in Figure 6. Comparing picture (b) with picture (c), we can find that when using original data, results of multi-input LSTM have no significant change with the increase of input values. While using denoised data as input, multi-input LSTM achieves significant improvements from 69.98% to 72.19% and performs



(a) SSE stock data

(b) SSE and U.S. stock data

(c) SSE, U.S. stock data and technical indicators

FIGURE 6 Comparison between original and denoised data.

best among all models. In picture (a) and (b), we can also find that except for multi-input LSTM, the accuracy of single-input LSTM increases more than that of other models after wavelet denoising. The better results achieved by combining WT with LSTM are mainly because of the following reasons. Owing to slow convergence, easy inclination to local minimum, and lack of ability to capture information in the frequency domain, LSTM cannot extract time frequency information of time series. However, WT can make up for the defects of LSTM by scaling and shifting in many respects to achieve the goal of frequency segmentation. Thus, the combination of WT and LSTM can fully extract time–frequency information of data and refine a model for time series frequency.

There are more details about the effect of wavelet denoising on labels. Original data have 938 Label 1 and 804 Label 0 in original data. In denoised data, there are 927 Label 1 and 815 Label 0. The ratio of Label 1 before denoising is 53.85%, and that of Label 1 after denoising is 53.21%. The proportion is similar, both around 53%. It shows that WT has no significant influence on the labels.

Multi-input LSTM is another important reason for the improvement of performance. When input values are the same, the performance of multi-input LSTM is better than that of single-input LSTM. Tables 6 and 7 show the forecasting performance. In the tables, loss stands for average binary cross-entropy. Train loss of multi-input LSTM is slightly higher than that of the single-input LSTM, but test loss is significantly reduced. The results show that multi-input LSTM improves the generalization ability of the model. Therefore, the model can be used effectively to predict time series. As shown in the tables, when input values are the same, the prediction accuracy of multi-input LSTM is higher in most cases. Moreover, with the increase of types of data sets, the prediction accuracy of multi-input LSTM is improved.

Results in Tables 8 and 9 show that multi-input LSTM with all data as input after wavelet denoising has the highest test accuracy, reaching 72.19%, which is significantly higher than that of other models. Sensitivity, specificity, precision for class 1, and precision for class 0 of the model are all above 70%. The overall performance of the proposed classifiers is better than the others. Furthermore, taking three types of data sets as input, respectively, the test accuracy reaches 70.76%, 69.98%, and 72.19%. All the results rank in the top four, indicating that multi-input LSTM is effective in predicting the trend of index and has great application value. In addition, we have carried out repeated experiments and found that the results are robust.

TABLE 6 The forecasting performance without denoising.

	SSE and US stock data		SSE, US stock data, and technical indicators	
	Single-input LSTM	**Multi-input LSTM**	**Single-input LSTM**	**Multi-input LSTM**
Train loss	0.57	0.58	0.45	0.59
Test loss	0.92	0.87	1.09	0.78
Train acc	67.43%	66.60%	76.80%	66.75%
Test acc	55.00%	54.39%	54.42%	55.20%

Note: LSTM, long short-term memory.

TABLE 7 The forecasting performance with denoising.

	SSE and US stock data		SSE, US stock data, and technical indicators	
	Single-input LSTM	**Multi-input LSTM**	**Single-input LSTM**	**Multi-input LSTM**
Train loss	0.43	0.44	0.28	0.42
Test loss	0.73	0.69	0.93	0.64
Train acc	79.48%	79.08%	88.15%	81.07%
Test acc	69.30%	69.98%	68.05%	72.19%

Note: LSTM, long short-term memory.

TABLE 8 The forecasting performance without denoising.

Data set	Model	Train acc	Test acc	Sens	Spec	Prec rise	Prec fall
SSE stock data	Decision tree	54.09%	51.72%	33.87%	72.22%	58.33%	48.75%
	Random Forest	64.12%	57.76%	79.14%	33.21%	57.64%	58.14%
	SVM	56.03%	53.45%	98.92%	1.23%	53.49%	50.00%
	XGBoost	56.82%	56.03%	92.47%	14.20%	55.31%	62.16%
	LSTM	64.14%	54.04%	69.73%	36.00%	55.59%	51.21%
SSE and U.S. stock data	Decision tree	58.98%	57.47%	72.58%	40.12%	58.19%	56.03%
	Random Forest	65.98%	56.72%	71.40%	39.88%	57.69%	54.85%
	SVM	56.75%	55.75%	93.55%	12.35%	55.06%	62.50%
	XGBoost	60.42%	56.03%	76.88%	32.10%	56.52%	54.74%
	Single-input LSTM	67.43%	55.00%	64.95%	43.56%	57.00%	51.90%
	Multi-input LSTM	66.60%	54.39%	64.89%	42.31%	56.47%	51.25%
SSE, US stock data, and technical indicators	Decision tree	58.98%	57.47%	72.58%	40.12%	58.19%	56.03%
	Random Forest	67.59%	55.92%	67.69%	42.41%	57.43%	53.37%
	SVM	59.77%	53.45%	81.18%	21.60%	54.32%	50.00%
	XGBoost	60.42%	56.61%	73.12%	37.65%	57.38%	54.95%
	Single-input LSTM	76.80%	54.42%	58.53%	49.69%	57.24%	51.05%
	Multi-input LSTM	66.75%	55.20%	56.97%	52.34%	66.52%	42.19%

Note: LSTM, long short-term memory; SVM, Support Vector Machine.

TABLE 9 The forecasting performance with denoising.

Data set	Model	Train acc	Test acc	Sens	Spec	Prec rise	Prec fall
SSE stock data	Decision tree	67.39%	64.66%	76.88%	50.62%	64.13%	65.60%
	Random Forest	70.57%	67.30%	76.56%	56.67%	66.98%	67.83%
	SVM	70.11%	66.95%	76.34%	56.17%	66.67%	67.41%
	XGBoost	67.60%	66.09%	77.96%	52.47%	65.32%	67.46%
	LSTM	78.00%	70.76%	72.39%	68.88%	72.94%	68.57%
SSE and US stock data	Decision tree	67.39%	64.66%	76.88%	50.62%	64.13%	65.60%
	Random forest	70.62%	66.26%	74.73%	56.54%	66.38%	66.10%
	SVM	69.97%	68.10%	79.03%	55.56%	67.12%	69.77%
	XGBoost	67.89%	64.66%	72.04%	56.17%	65.37%	63.64%
	Single-input LSTM	79.48%	69.30%	69.78%	68.75%	72.10%	66.54%
	Multi-input LSTM	79.08%	69.98%	70.95%	68.86%	72.43%	67.48%
SSE, US stock data, and technical indicators	Decision tree	66.24%	64.94%	72.04%	56.79%	65.69%	63.89%
	Random forest	73.64%	67.13%	75.11%	57.96%	67.23%	66.99%
	SVM	73.13%	71.26%	80.65%	60.49%	70.09%	73.13%
	XGBoost	69.40%	66.95%	78.49%	53.70%	66.06%	68.50%
	Single-input LSTM	88.15%	68.05%	70.54%	65.19%	70.04%	66.00%
	Multi-input LSTM	81.07%	**72.19%**	74.66%	69.67%	72.75%	71.54%

Note: LSTM, long short-term memory; SVM, Support Vector Machine.

5 | CONCLUSION

We hold the opinion that the financial time series is a nonlinear and non-stationary time series, which can weaken the forecasting ability of LSTM. To solve this problem, we propose a method based on WT and multi-input LSTM to predict the trend of SSE composite index. Before applying the model, we first process the data with WT to extract the implicit information. Moreover, multi-input LSTM can take data with multiple dimensions as input to increase amounts of data and enrich data selection. The combination of the two new methods can effectively improve the accuracy of prediction. Compared with single-input LSTM, decision tree, random forest, SVM, and XGBoost, our proposed model with SSE stock data, US stock data, and technical indicators after denoising as input has the highest accuracy. The hybrid model can find nonlinear and non-stationary information in financial time series, which proves that the final model has a bright future in stock price forecasting.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the National Social Science Fund of China (No. 19CJL028).

DATA AVAILABILITY STATEMENT

Data is available upon request.

REFERENCES

Agnolucci, P. (2009). Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics*, *31*(2), 316–321. https://doi.org/10.1016/j.eneco.2008.11.001

Alrumaih, R. M., & Al-Fawzan, M. A. (2002). Time series forecasting using wavelet denoising an application to Saudi stock index. *Journal of King Saud University-Engineering Sciences*, *14*(2), 221–233. https://doi.org/10.1016/S1018-3639(18)30755-4

Aussem, A. (1998). Wavelet based feature extraction and decomposition strategies for financial forecasting. *International Journal of Computational Intelligence in Finance*, *6*, 5–12.

Awartani, B. M. A., & Corradi, V. (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: The role of asymmetries. *International Journal of Forecasting*, *21*(1), 167–183. https://doi.org/10.1016/j.ijforecast.2004.08.003

Bai, L., Yan, S., Zheng, X., & Chen, B. M. (2015). Market turning points forecasting using wavelet analysis. *Physica A Statal Mechanics & its Applications*, *437*, 184–197. https://doi.org/10.1016/j.physa.2015.05.027

Balaji, A. J., Ram, D. S. H., & Nair, B. B. (2018). Applicability of deep learning models for stock price forecasting an empirical study on BANKEX data. *Procedia Computer Science*, *143*, 947–953. https://doi.org/10.1016/j.procs.2018.10.340

Ballings, M., Dirk, V. D. P., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, *42*(20), 7046–7056. https://doi.org/10.1016/j.eswa.2015.05.013

Basak, S., Kar, S., Saha, S., Khaidem, L., & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, *47*, 552–567. https://doi.org/10.1016/j.najef.2018.06.013

Bekaert, G., & Harvey, C. R. (1995). Time-varying world market integration. *Journal of Finance*, *50*, 403–444.

Bentes, R. S. (2015). A comparative analysis of the predictive power of implied volatility indices and GARCH forecasted volatility. *Physica a: Statistical Mechanics and its Applications*, *424*, 105–112. https://doi.org/10.1016/j.physa.2015.01.020

Caetano, M. A. L., & Yoneyama, T. (2007). Characterizing abrupt changes in the stock prices using a wavelet decomposition method. *Physica a Statal Mechanics & its Applications*, *383*(2), 519–526. https://doi.org/10.1016/j.physa.2007.03.027

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, *83*, 187-205.

Dutta, A. (2018). Implied volatility linkages between the U.S. and emerging equity markets: A note. *Global Finance Journal*, *35*, 138–146. https://doi.org/10.1016/j.gfj.2017.09.002

Fernandez, V. (2007). Wavelet- and SVM-based forecasts: An analysis of the U.S. metal and materials manufacturing industry. *Resources Policy*, *32*(1–2), 80–89. https://doi.org/10.1016/j.resourpol.2007.06.002

Gençay, R., Selçuk, F., & Whitcher, B. (2002). *An introduction to wavelets and other filtering methods in finance and economics*. Academic Press.

Goh, J. C., Jiang, F., Tu, J., & Wang, Y. (2013). Can US economic variables predict the Chinese stock market? *Pacific-Basin Finance Journal*, *22*, 69–87. https://doi.org/10.1016/j.pacfin.2012.10.002

Gultekin, M. N., Gultekin, N. B., & Penati, A. (1989). Capital controls and international capital market segmentation: The evidence from the Japanese and American stock markets. *Journal of Finance*, *44*, 849–869. https://doi.org/10.1111/j.1540-6261.1989.tb02627.x

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of Finance and Data Science*, *4*(3), 183–201. https://doi.org/10.1016/j.jfds.2018.04.003

Hoseinzade, E., & Haratizadeh, S. (2019). Cnnpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, *129*, 273–285. https://doi.org/10.1016/j.eswa.2019.03.029

Hsieh, T. J., Hsiao, H. F., & Yeh, W. C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied Soft Computing*, *11*(2), 2510–2525. https://doi.org/10.1016/j.asoc.2010.09.007

Huang, B. N., Yang, C. W., & Hu, J. W. S. (2000). Causality and cointegration of stock markets among the United States, Japan and the South China growth triangle. *International Review of Financial Analysis*, *9*, 281–297. https://doi.org/10.1016/S1057-5219(00)00031-4

Huang, S. C. (2011). Forecasting stock indices with wavelet domain kernel partial least square regressions. *Applied Soft Computing*, *11*(8), 5433–5443. https://doi.org/10.1016/j.asoc.2011.05.015

Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, *32*(10), 2513–2522. https://doi.org/10.1016/j.cor.2004.03.016

Johansson, A. C. (2009). China's financial market integration with the world. *Journal of Chinese Economics and Business Studies*, *8*(3), 293–314.

Kao, L. J., Chiu, C. C., Lu, C. J., & Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems*, *54*(3), 1228–1244. https://doi.org/10.1016/j.dss.2012.11.012

Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange. *Expert Systems with Applications*, *38*(5), 5311–5319. https://doi.org/10.1016/j.eswa.2010.10.027

Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1–2), 307–319. https://doi.org/10.1016/S0925-2312(03)00372-2

Kriechbaumer, T., Angus, A., Parsons, D., & Casado, R. M. (2014). An improved wavelet–ARIMA approach for forecasting metal prices. *Resources Policy*, *39*(C), 32–41. https://doi.org/10.1016/j.resourpol.2013.10.005

Lahmiri, S. (2014). Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks. *Journal of King Saud University Computer & Information Sciences*, *26*(2), 218–227. https://doi.org/10.1016/j.jksuci.2013.12.001

Lei, L. (2018). Wavelet neural network prediction method of stock price trend based on rough set attribute reduction. *Applied Soft Computing*, *62*, 923–932. https://doi.org/10.1016/j.asoc.2017.09.029

Li, S. (2015). Volatility spillovers in the CSI300 futures and spot markets in China: Empirical study based on discrete wavelet transform and VAR-BEKK-bivariate GARCH model. *Procedia Computer Science*, *55*, 380–387. https://doi.org/10.1016/j.procs.2015.07.085

Li, X., & Tang, P. (2020). Stock index prediction based on wavelet transform and FCD-MLGRU. *Journal of Forecasting*, *39*, 1–9. https://doi.org/10.1002/for.2682

Li, Z., & Tam, V. (2017, November). Combining the real-time wavelet denoising and long-short-term-memory neural network for predicting stock indexes. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1–8). IEEE.

Liang, X., Ge, Z., Sun, L., He, M., & Chen, H. (2019). LSTM with wavelet transform based data preprocessing for stock price prediction. *Mathematical Problems in Engineering*, *2019*, 1–8. https://doi.org/10.1155/2019/1340174

Lin, Y., Yan, Y., Xu, J. L., Liao, Y., & Ma, F. (2021). Forecasting stock index price using the CEEMDAN-LSTM model. *The North American Journal of Economics and Finance*, *57*, 101421. https://doi.org/10.1016/j.najef.2021.101421

Liu, H., & Long, Z. (2020). An improved deep learning model for predicting stock market price time series. *Digital Signal Processing*, *102*, 102741. https://doi.org/10.1016/j.dsp.2020.102741

Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, *164*, 163–173. https://doi.org/10.1016/j.knosys.2018.10.034

Nayak, R. K., Mishra, D., & Rath, A. K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, *35*, 670–680. https://doi.org/10.1016/j.asoc.2015.06.040

Nevasalmi, L. (2020). Forecasting multinomial stock returns using machine learning methods. *The Journal of Finance and Data Science*, *6*, 86–106. https://doi.org/10.1016/j.jfds.2020.09.001

Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management*, *26*(4), 164–174. https://doi.org/10.1002/isaf.1459

Ortega, L., & Khashanah, K. (2014). A neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns. *Journal of Forecasting*, *33*(2), 134–146. https://doi.org/10.1002/for.2270

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015a). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications an International Journal*, *42*(4), 2162–2172. https://doi.org/10.1016/j.eswa.2014.10.031

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015b). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, *42*(1), 259–268. https://doi.org/10.1016/j.eswa.2014.07.040

Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis*. Cambridge University Press. https://doi.org/10.1017/CBO9780511841040

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Social Science Electronic Publishing*, *6*(3), 1090–1090. https://doi.org/10.2202/1558-3708.1090

Rumbaugh, T., & Blancher, N. (2004). *China: International trade and WTO accession*. IMF Working Paper 04/36. International Monetary Fund.

Saud, A. S., & Shakya, S. (2020). Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE. *Procedia Computer Science*, *167*, 788–798. https://doi.org/10.1016/j.procs.2020.03.419

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1643–1647). IEEE.

Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. H. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, *264*, 71–88. https://doi.org/10.1016/j.neucom.2016.11.095

Torbat, S., Khashei, M., & Bijari, M. (2018). A hybrid probabilistic fuzzy ARIMA model for consumption forecasting in commodity markets. *Economic Analysis and Policy*, *58*, 22–31. https://doi.org/10.1016/j.eap.2017.12.003

Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia Computer Science*, *167*, 599–606. https://doi.org/10.1016/j.procs.2020.03.326

Ye, G. L. (2014). The interactions between China and US stock markets: New perspectives. Journal of international financial markets. *Institutions and Money*, *31*, 331–342. https://doi.org/10.1016/j.intfin.2014.04.008

Yousefi, S., Weinreich, I., & Reinarz, D. (2005). Wavelet-based prediction of oil prices. *Chaos, Solitons & Fractals*, *25*(2), 265–275. https://doi.org/10.1016/j.chaos.2004.11.015

Zhang, J., Cui, S. C., Xu, Y., Li, Q. M., & Li, T. (2018). A novel data-driven stock price trend pre-diction system. *Expert Systems with Application*, *97*, 60–69. https://doi.org/10.1016/j.eswa.2017.12.026

AUTHOR BIOGRAPHIES

Pan Tang is currently an Associate Professor at the School of Economics and Management, Southeast University. He received his PhD degree in quantum finance from the National University of Singapore. His research interests are financial artificial intelligence, deep learning in finance, quantitative investment and financial big data analysis.

Cheng Tang is a master's student studying at the School of Economics and Management, Southeast University. He obtained his Bachelor's degrees from the School of Finance, Nanjing University of Finance and Economics. His research field is financial big data analysis and financial prediction.

Keren Wang is an undergraduate student studying at the School of Economics and Management, Southeast University. Her research field is financial engineering.

How to cite this article: Tang, P., Tang, C., & Wang, K. (2024). Stock movement prediction: A multi-input LSTM approach. *Journal of Forecasting*, *43*(5), 1199–1211. https://doi.org/10.1002/for.3071