A Stock Trend Prediction Model Based on Wavelet Transform and TCN Combined with Market Sentiment

Jian Zhang [a], Weidong Wang [b]

[a]e-mail:1057359832@qq.com, [b]e-mail:78653221@qq.com

Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, China

Abstract. Based on the existing stock trend prediction methods that combine market sentiment, combined with wavelet transform and TCN time convolution model, to improve the accuracy of stock trend prediction. [Method]Download stock numerical data through the interface, mine and analyze potential numerical data, and then use the GRU network model to perform sentiment analysis on the obtained stock news data, obtain sentiment indicators, and use them as training data. Train the stock price prediction model using wavelet transform and TCN time convolutional network model.[Results] After the introduction of wavelet transform and TCN network model, the accuracy of Gree Electric Appliance's stock trend prediction increased by 6.1%, ZTE's stock trend prediction increased by 5.92%, and AAPL's stock trend prediction accuracy increased by 7.45%. [Limitation] The stock market trades according to working days, and in the case of weekends or holidays in the middle, the previous data may not have a significant impact on the current data. [Conclusion] Stock fluctuations are affected by market sentiment. The use of wavelet transform to process stock data can minimize the impact of extreme situations on stock price trends, and the combination of TCN time convolution model can better predict the trend of stock data.

Keywords: TCN, Time, Convolutional, Wavelet Transform, Market Sentiment, Stock Trends

1. Introduction

Stock trading will produce huge economic gains, so the research of stock trend has always been a very hot direction. Since the change of stocks is not an irregular random walk [1], it is possible to find the change law through appropriate ways.

Early financial workers used linear prediction models to analyze stock trends, mainly including Generalized Autoregressive Conditional Heteroscedasticity (GARCH) [2] and Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) [3]. With the development of neural network modeling technology, we can use the algorithms of neural network models to find important factors that affect the trend of a certain stock among many influencing factors. Through deep learning of these factors, we can find the best model to fit the stock trend.

Hu Yuwen et al. [4] used LASSO and PCA to optimize the stock data and improve the LSTM long and short term series model, and proved that the classification and dimensionality reduction of stock data are conducive to better improving the accuracy of stock prediction. Liu

IEDM 2024, February 23-25, Kuala Lumpur, Malaysia

DOI 10.4108/eai.23-2-2024.2345895

Yuling et al. [5] analyzed the impact of market sentiment on stock market trends and proved that market sentiment plays an important role in stock trend prediction. Then, they combined sentiment data with stock numerical data and used GAN to generate adversarial network for adversarial learning to improve the accuracy of stock trend prediction. Wu Minghui et al. [6] proposed a time series prediction model based on multi-modal information fusion, and introduced a global attention mechanism into the model to better pay attention to the correlation between short -, medium - and long-term trends of stock data, thus improving the fit degree of stock prediction. Xu Yuemei et al. [7] proposed three indicators of financial events, market sentiment and numerical data, and adopted two-way LSTM model to prove that non-numerical values such as market sentiment and financial events are also of great significance to stock trend prediction.

2. Stock prediction model based on market sentiment

The prediction of stock trends is closely related to the basic financial data, potential trend indicators, and market sentiment of stocks. The prediction of stock trends can be divided into short-term, medium-term, and long-term forecasts. If you want to achieve positive returns in stock forecasting, short-term forecasting is difficult, long-term forecasting takes too long, and does not meet the target. Therefore, predicting medium-term trends is very important. We take two weeks as the interval for the mid-term forecast and predict the stock price on the 14th day through comprehensive analysis of the stock trends in the first 13 days.

The model architecture elaborated in this paper is shown in Figure 1, and the steps are described as follows:

Download basic financial data and related news data for predicting stocks through TuShare and Yahoo, respectively. Then, abnormal numerical processing is performed on the data to obtain financial datasets, news text datasets, relevant stock price data sets, and potential stock analysis numerical datasets for training.

By labeling the emotional polarity of news texts and analyzing certain types of stocks, it is found that their emotions are mainly reflected in relevant keywords, which can basically be summarized, and then programs are written to automatically label according to keywords. Finally, the labels of the training set are manually reviewed. After getting the correct news emotional labels, GRU neural network model is used for training.

Analyze basic financial data to obtain data such as 7-day moving average, 21-day moving average, MACD index, and related stock price data.

Combine basic financial data, relevant news sentiment data, and potential stock analysis values based on transaction dates. After the combination is completed, perform wavelet denoising and then stack encoding. Then, the encoded data is trained using a TCN temporal convolutional model.

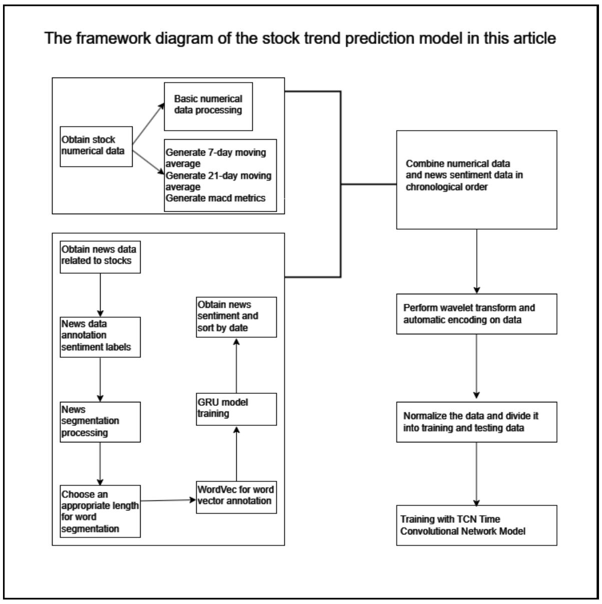

Figure 1. Prediction Model Structure

2.1 Analysis and processing of stock based financial data

Basic financial data include data on trading date, opening price, highest price, lowest price, closing price yesterday closing price, up or down, up or down, volume, turnover, etc. , since these data are not of the same order of magnitude, they need to be normalized. In this paper, we use maximum absolute normalization to scale the range of values to [-1,1] .The formula is as follows:

$$x' = \frac{x}{\|\max(x)\|} \quad (1)$$

2.2 Potential financial data analysis and generation

The MACD index is known in the industry as the difference and similarity moving average. It is based on the EMA moving average mechanism of prices. It is used to judge the long-short trend and the reversal signal. GET stock price data for relevant industries (including related industries, upstream and downstream companies) , and arrange the 7-day moving average, 21-day moving average and MACD index by date, it is then combined with the underlying financial data to form three columns of features.

2.3 Stock related news emotional polarity training and analysis

The rise and fall of a stock is often closely related to the positive and negative news. Judging the day's news shows that positive or negative market sentiment plays a big role in stock forecasts. If the market sentiment is high, the stock is likely to rise, whereas if the market sentiment is low, the stock is likely to fall. In addition, the degree of market sentiment high and low can be accumulated by the number of positive and negative news. The more positive news, the higher the market sentiment, and the more negative it is. If there is neither positive nor negative, the market sentiment for the day is 0. We obtain several news headline data of stock news through TuShare and Yahoo website interfaces, and then process the data and train the model. The detailed steps are as follows:

1) *Manually label the obtained title data.*

2) *Use the Jieba library to segment news headlines.*

3) *The length after word segmentation of all news headline samples is calculated and its length distribution is analyzed.*

4) *Use WordVec to annotate word vectors for word segmentation, generate training data, and divide the training and testing sets.*

The GRU[8] model is used to train the training set as follows:

Calculate the two GRU doors, the following writing ZT, the calculation method is as follows:

$$r_t = \sigma(W^r x_t + U^r h_{t-1}) \quad (2)$$

$$z_t = \sigma(W^z x_t + U^z h_{t-1}) \quad (3)$$

Calculate candidate hidden layer($\tilde{h}_t$) as follows:

$$\tilde{h}_t = \tanh(W\mathrm{x}_i + r_i U h_{t-1}) \quad (4)$$

Calculate ht as follows:

$$h_t = z_t \circ h_{t-1} + (1 - z_t) \circ \tilde{h}_t \quad (5)$$

The accuracy of the trained model is verified by the test set. After several rounds of parameter adjustment, a natural language emotion training model with an accuracy of 95.7% was obtained. By adding up the emotions of the same date by date, you get the emotional polarity of each day. Finally, the affective polarity is combined with the above stock affective data according to the date to form a new data dimension.

2.4 Wavelet denoising

The basic idea of wavelet threshold de-noising is that the wavelet coefficient of signal contains the important information of signal after it is transformed by wavelet (Mallat algorithm) , the wavelet coefficient of noise is smaller than that of signal, and the wavelet coefficient of noise is smaller than that of signal. By choosing a suitable threshold, the wavelet coefficient larger than the threshold is considered to be generated by signal and should be preserved, less than the threshold is considered to be generated by noise, set to zero to achieve the goal of de-noising.

From this, it can be seen that wavelet denoising[9] is actually a combination of feature extraction and low-pass filtering. A noisy model can be represented as follows:

$$S(\mathrm{k}) = f(k) + \varepsilon * e(k) \quad (6)$$

Among them, f (k) is a useful signal, s (k) is a noisy signal, and e (k) is noise, ε is the standard deviation of noise coefficient.

2.5 TCN sequential convolutional network model

Temporal Convolutional Network (TCN) is a method based on the structure of convolutional neural networks, which is specially designed to process time series data. TCN[10] has advantages such as strong parallel computing ability, strong modeling ability of long-term dependence, and small number of parameters.

The core idea of TCN is to capture time dependence in time series by Convolution operation, and two techniques of Causal Convolution and Dilated Convolution are built in. The network diagram is shown in Figure 2.

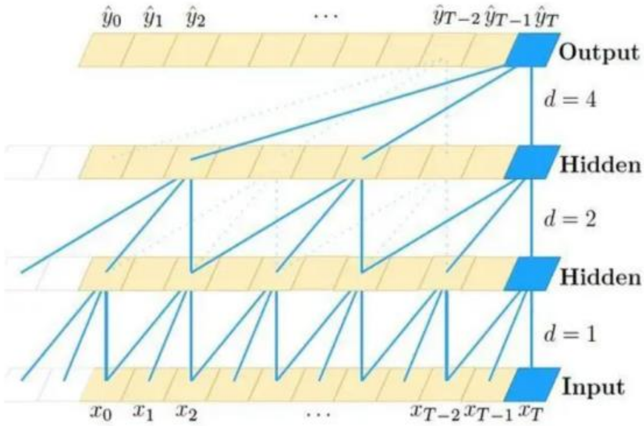

Figure 2. TCN model structure

Use the self encoded dataset above to predict the closing price at time t+1. Define a window partitioning function that divides the input data and its labels into delay time intervals. In this article, a single step prediction is performed. Therefore, assuming the input is from day 1 to day 13, the output predicted price is the price from day 14. Then, the training and testing sets are divided according to time intervals, with the first 70% of the data used for model training and the last 30% for model testing. TCN uses causal convolution to process time series data, which ensures that future information is not leaked during prediction. This means that when predicting the output of a certain time step, only the input information before that time step can be used to run the model, and dilated convolution expands the receptive field, thereby capturing long-term time dependencies. Dilated convolution allows the convolutional kernel to perform interval sampling on the input sequence, allowing each convolutional layer to cover a wider time range. MSE was used as the loss, Adam as the optimizer, and a scheduler for setting the learning rate. Finally, 200 epochs were run, and the model was evaluated on the test set every 10 epochs. Finally, based on the first t data, predict the rise and fall of the stock on day t+1, and achieve the prediction of the future trend of the stock.

3. Results & Discussion

3.1 Experimental data set

Select two A-share representative of Gree Electronics and ZTE shares and AAPL (Apple) stock, a-share selection range of January 1,2009 to January 1,2022 a total of 2,448 financial data, the TuShare crawler pulled headline data for all industries and a total of 38,796 headlines for Gree Electronics and ZTE from January 1,2009, to January 1,2009, the 2023 date. AAPL selected a total of 4,716 data points between January 3,2000 and September 28,2018, and obtained 26,125 related news headlines from Yahoo! Finance.

The financial and financial data selected 12 characteristics, including opening price, highest price, lowest price, closing price, yesterday's closing price, fluctuation amount, fluctuation range, trading volume, transaction volume, 7-day moving average, 21 day moving average, and MACD index.

3.2 Experimental analysis of news headlines corresponding to market sentiment

Since the orientation of news emotion analysis will be used as a dimension of the later model training, the positive and negative emotions need to be highly accurate. If the accuracy of emotion is low, it will inevitably lead to the deviation of the model training results. Therefore, for the analysis of the emotion of news headlines, the training model adopts the accuracy rate to judge its effect. In multiple tuning and comparison experiments, it is found that the optimal experimental effect can be obtained by using the following model parameter Settings. Details of parameters are shown in Table 1.

Table 1. The optimal training parameters for GRU model

Parameter name	Parameter value
Percentage of training set	70%
Share of verifier	10%
Test set percentage	20%
Length of training words	20
Learning rate	0.001
Dropout	0.5
Number of iterations	200
BatchSize	64
Embedding_dim	400

In order to verify the advantages of GRU network model in Chinese semantic sentiment analysis, two neural network models are selected for comparative experiments, and the experimental results are compared in Table 2:

Table 2. Comparison of training accuracy of different models

	BP	RNN	GRU
Training set	70.2%	85.5%	99.5%
Validation set	58.6%	79.0%	95.3%
Test set	59.3%	78.6%	95.7%

3.3 Experimental Analysis of Stock Trends

The 12 financial numerical indicators were combined with the news emotion indicators corresponding to the date to form 13 dimensional features. The 13 features were transformed by wavelet and then automatically encoded to form the data set, which was trained by TCN network. In order to verify that the above models can improve the performance of stock prediction, the following experiments will be conducted for comparison. Then RNN neural network, LSTM neural network and TCN model network are used for comparison experiments.

5) *Using 12 financial numerical data as Data Set and LSTM as training model.*

6) *Using 12 financial data and 1 Emotion Index, LSTM was used as training model.*

7) *Using 12 financial data and 1 emotion index, TCN was used as training model.*

8) *Using 12 financial data and 1 emotion index, combined with wavelet transform and automatic coding, and then using TCN as the training model for training. Compared with the third training model, the training data set is filtered by wavelet transform to filter out the values and emotional anomalies that cause sharp changes in stock value data due to statistical errors or special factors, so as to rationalize the stock trend and improve the accuracy of prediction.*

Through the analysis of stock trading, it can be found that the stock closing price of the next day will be affected by the stock trading data of the previous days, and the closer the time is, the greater the influence relationship will be. For the data with a particularly distant time, the impact can be almost ignored. So to sum up, we take 13 days as a window and use the stock data of the previous 13 days to predict the closing price of the stock on the 14th day. In order to reflect the accuracy of the model prediction, we will forecast the stock price, output the predicted price according to the time, and then analyze the rise and fall of the predicted price. If the closing price is higher than or equal to the closing price of the previous day, it is an upward trend, and the value is 1; if it is lower than the opening data of the previous day, it is a downward trend, and the value is -1. Form the forecast rise and fall data set. Then do the same data analysis on the original numerical data, get the rising and falling trend, and then form the rise and fall data set of the original data. Finally, the prediction accuracy is obtained by comparing the fluctuation data set according to the date. The accuracy of the above four training models is shown in Table 3 below.

Table 3. The prediction accuracy of different models on three stocks

Stock data	LSTM with financial values only	LSTM adopts financial value and news sentiment	TCN using financial values and news sentiment	TCN after wavelet transform using financial value and news emotion
Gree Electric	54.2%	56.33%	57.65%	60.3%
ZTE	53.7%	55.9%	56.4%	59.62%
AAPL	54.7%	57.52%	59.79%	62.15%

From Table 3 above, we can see that compared with the LSTM model which only uses the financial numerical training, the LSTM model which uses the stock numerical value and the emotion, the accuracy of the stock rising and falling forecast in the GREE electric appliance, ZTE and AAPL, they increased by 2.13% , 2.2% and 2.82% respectively, indicating that stock prices are also affected by market sentiment. When market sentiment is positive, stock prices tend to rise, and when market sentiment is negative, stock prices tend to fall. Using financial data plus emotional value, combined with wavelet transform and TCN training model, the accuracy of stock price prediction in GREE, ZTE and AAPL increased by 6.1% , 5.92% and 7.45% respectively.

In order to more intuitively demonstrate the effect and accuracy of stock price prediction, the following figures show the stock price trend prediction of Gree Electric Appliances, ZTE and Apple [AAPL], respectively, as Figure 3, Figure 4 and Figure 5.

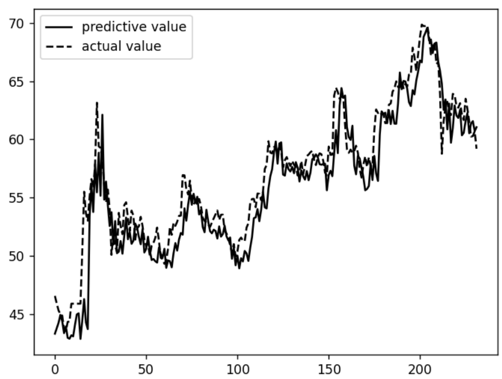

Figure 3. Gree Stock Trend Prediction Chart

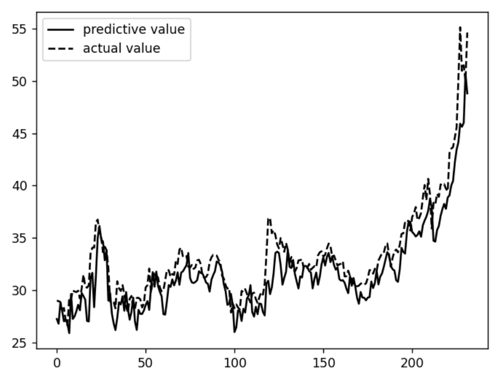

Figure 4. ZTE Stock Trend Forecast Chart

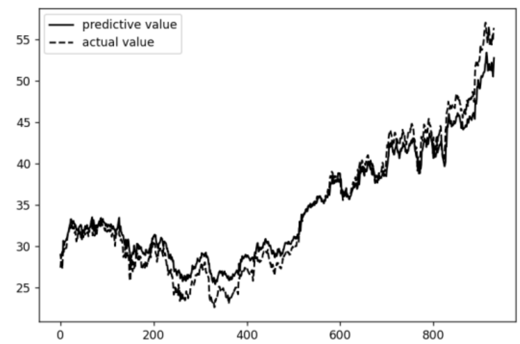

Figure 5. AAPL Stock Trend Forecast Chart

4. Conclusions

It can be proved by the previous experiments that the combination of wavelet transform and TCN has better performance than LSTM. Wavelet transform can remove noise from data and prevent data overfitting. The combination of causal convolution and expansive convolution unique to TCN can better analyze the trend of time series data.

References

[1] Sun x Q,Shen H w,Cheng X Q. Trading Network PredictsStock Price[J]. Scientific Reports, 2014,4:Article No.3711.

[2] YANG Q,CAO X B.Analysis and prediction of stock price based on arma-garch model [J].Practice and Understanding ofMathematics,2016,46(6):80-86.

[3] HUANG LJ,JIN T X.Stability of Islamic stock market basedon egarch-m model[J].Gansu Academic Journal of theory,2019(6):107-115.

[4] HU Yu-wen.Stock Forecast Based on Optimized LSTM Model[J].Computer Science,2021,Vol. 48(201): 151-157

[5] LIU Yuling,ZHAO Guolong,ZOU Ziran,WU Shengting.Stock Price Prediction Method Based on Sentiment Analysis and Generative Adversarial Network[J].Journal of Hunan University (Natural Science Edition),2022,Vol.49(10): 111-118

[6] WUMinghui,ZHANGGuangjie,JINCanghong.Time series prediction model based on multimodal information fusion[J].ComputerApplication,2022,Vol.42(8): 2326-2332

[7] Xu Yuemei;Wang Zihou;WuZixin.A stock trend prediction model based on CNN-BiLstm multi-feature fusion[J].Data Analysis and Knowledge Discovery,2021, Vol.5(7): 126-137

[8] Agbessi Akuété Pierre;Salami Adekunlé Akim;Agbosse Kodjovi Semenyo;Birregah Babiga.Peak Electrical Energy Consumption Prediction by ARIMA, Lstm, GRU, ARIMA-Lstm and ARIMA-GRU Approaches[J].Energies,2023,Vol.16: 4739

[9] HUANG Chao;GAO Xiang;DING Yongmei;LIU Lijun,Research on Meteorological Time Series Prediction Based on Wavelet Transform and LSTM Hybrid Model.Advances in Applied Mathematics,2022,Vol.11(8): 6107-6117

[10] Pandey, A.;Wang, D..TCNN: Temporal Convolutional Neural Network for Real-time Speech Enhancement in the Time Domain(Conference Paper)[J].ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings,2019,: 6875-6879