Research on the Application of Convolutional Neural Network in Stock Market Forecasting

Zheyu Xu*
Central University of Finance and Economics, Beijing, China
*445407219@qq.com

Abstract. With the continuous development and complexity of financial markets, stock market forecasting has become a hot topic of concern for investors and decision-makers. This study aims to explore the application of Convolutional Neural Networks (CNN) in stock market prediction, and evaluate their performance in stock index prediction by comparing them with traditional methods such as BP Neural Networks (BPNN) and Support Vector Machines (SVM). In order to further improve the performance of the model, we introduce a model that combines CNN and SVM(CNN-SVM). Experiments show that CNN-SVM model has achieved remarkable advantages in accuracy and robustness, showing the potential value of combining deep learning with traditional machine learning methods. In the empirical analysis, it is found that using CNN-SVM to predict the stock market can achieve relatively superior performance. CNN-SVM can better capture nonlinear relations and complex patterns in time series data through its excellent feature extraction ability. Compared with traditional methods, CNN-SVM shows higher accuracy, precision and recall rate, which verifies its effectiveness in the financial field. The results of this study not only provide investors with more accurate stock market forecasting tools, but also provide a new perspective for the academic research on the combination of deep learning technology in the financial field.

Keywords: Convolutional Neural Network; Support Vector Machine; Stock Market Forecasting.

1. Introduction

The high complexity and uncertainty of the stock market make investors and decision makers face challenges. In this era of information explosion, it is increasingly difficult to process large-scale and high-dimensional financial data. With the continuous development of computer science and artificial intelligence, as a powerful deep learning tool, Convolutional Neural Network (CNN) has gradually attracted extensive attention in stock market forecasting [1-2].

In the era of rapid development of information technology, deep learning technology has gradually become a powerful tool to deal with large-scale data and complex models in the financial field. Through the deep learning model training of historical market data, researchers have successfully improved the accuracy of forecasting market volatility and systemic risks. The case in reference [3] shows that the deep learning model has advantages in identifying potential risk events in the financial market, and provides a more flexible and accurate tool for risk management. Portfolio optimization is another area where deep learning is widely used in the financial field. The research in reference [4] shows that deep learning technology can effectively analyze the complex relationship between a large number of assets, improve the return of portfolio and reduce risks. By applying the deep learning model to asset price forecasting and market trend analysis, investors can better adjust their investment portfolio to adapt to market changes. The application of deep learning technology in market forecasting has also attracted much attention. The research in reference [5] shows that deep learning models such as CNN and Long Short Term Memory (LSTM) perform well in processing time series data, helping to predict stock prices and market trends more accurately. The advantage of these models is that they can automatically capture complex market patterns and trends and provide investors with more reliable decision support. The application of deep learning technology in the financial field not only improves the understanding of markets and risks, but also provides financial practitioners with more powerful and innovative tools. However, there are still many challenges to be overcome, including data privacy, model interpretation and over-fitting.

Traditional financial models are usually based on fundamental and technical analysis, but these methods are often difficult to capture the complex nonlinear dynamics of the market [6-8]. CNN is famous for its successful application in image processing, however, its potential in time series data processing has been gradually recognized in recent years. The advantage of CNN is that it can automatically learn the features in the data and has the ability to hierarchically process the input data. This makes CNN perform well in capturing patterns and trends in time series data. By combining convolution layer and pooling layer, CNN can extract local and global features from the data, thus better understanding the complex dynamics in the stock market. The purpose of this study is to explore the application of CNN in stock market forecasting, and to improve the forecasting accuracy and market sensitivity by training a large number of historical stock data. Through this research, investors, financial analysts and decision makers can be provided with more powerful and reliable tools to better meet the challenges of the stock market and realize more intelligent and accurate investment decisions.

2. Overview of CNN principle

CNN is a kind of deep learning neural network, which is mainly used to process data with grid structure, such as images and videos. It has achieved great success in the field of computer vision, because it can effectively learn and extract features from images, and has the characteristics of parameter sharing and local receptive field. CNN structure is shown in figure 1:

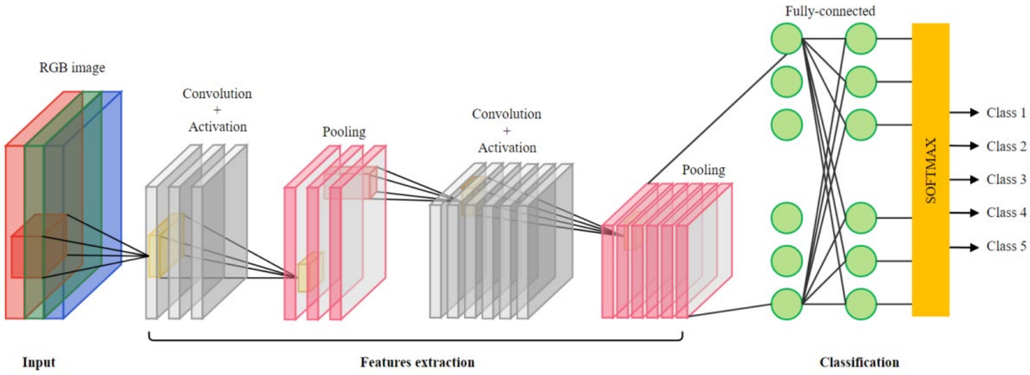

Figure 1 CNN structure

The core of CNN is convolution layer, which realizes feature extraction of input data through convolution operation. Convolution operation refers to sliding a small window (convolution kernel) on the input data, and calculating the product of the local area in the window and the convolution kernel, and then adding these products to get an output value. This process is carried out through the sliding window on the whole input, thus obtaining the characteristic diagram of the output.

The activation function is applied to the output of convolution layer, and ReLU is usually used to introduce nonlinearity. The introduction of activation function is helpful for network learning more complex patterns and representations. Pooling operation is used to reduce the spatial dimension of feature map, reduce the amount of calculation and retain the most important information. Commonly used pooling operations include maximum pooling and average pooling, which respectively select the maximum value or average value of local area as output. There are usually one or more fully connected layers between the convolution layer and the fully connected layer, which are used to integrate the features extracted from the convolution layer and make the final classification or regression. The neurons in the fully connected layer are connected with all the neurons in the previous layer.

The uniqueness of CNN lies in parameter sharing. In the convolution layer, each convolution kernel is used to extract different local features of the input, and the parameters of these convolution kernels are shared, thus reducing the number of parameters to be learned and improving the efficiency of the model. Through the design of local receptive field, CNN can better capture the local structure and mode of input data, rather than the overall characteristics. This is helpful to improve the networks ability to discriminate input data. CNN usually contains multiple convolution layers, and the hierarchical representation of input data is realized by extracting and combining features layer by layer. This deep structure enables the network to learn more abstract and complex features.

Through the above mechanism, CNN can effectively learn and understand the features in images, thus achieving excellent performance in image classification, object detection, semantic segmentation and other tasks. Its success in the field of computer vision also reveals the potential of applying CNN in other fields.

3. Establishment of stock market forecast model

Forecasting the stock market is a challenging task in the financial field. The stock market is affected by many factors, including macroeconomic indicators, corporate financial situation, international political situation, etc. The complex interaction of these factors leads to the nonlinear and unstable market behavior. It is difficult for traditional linear models to capture these complex relationships, which makes prediction more difficult. The price of the stock market is affected by a lot of noise and random fluctuations, which makes it extremely difficult to accurately predict the price changes in the short term. Even the prediction based on a large number of data and complex models is vulnerable to random fluctuations in the short term and leads to misleading results. The stock market involves a lot of information, and not all participants can get it at the same time. Sometimes, some market participants may have some sensitive information, which leads to the asymmetry of information in the market, which makes the prediction based on public information more difficult.

The psychological factors of investors have an important influence on the market trend. Investors emotional fluctuation, group behavior and psychological deviation will cause irrational fluctuations in the market, which are difficult to be captured by traditional mathematical models. The stock market is easily affected by unexpected events and external factors, such as natural disasters, political events, financial crisis and so on. These unpredictable factors have a huge impact on the market, making it difficult for the forecasting model to adapt to this complex and dynamic environment. The data of the stock market is usually high-dimensional, including various technical indicators and financial data. Processing and analyzing such multi-dimensional data requires powerful computing resources and appropriate algorithms, and is easily affected by dimensional disasters.

There are cycles and trends in the stock market, but the changes of these trends and cycles are not fixed and are influenced by many factors. It is a very complicated problem to capture trends and periodicity on different time scales and to accurately predict them. Generally speaking, stock market forecasting is a challenging task, and its difficulty stems from the complexity, uncertainty and changeable factors of the market. An effective stock market forecasting model needs to fully consider these difficulties, and combine advanced technology and profound market understanding to improve the forecasting accuracy.

The prediction of stock market has always been a research direction of great concern in the financial field. Traditional time series analysis and technical indicators are often difficult to capture complex market patterns and trends. In recent years, the rise of deep learning technology provides a new idea for stock market forecasting.This paper will discuss how to use CNN and support vector machine (SVM) to establish a stock market forecasting model to improve the forecasting accuracy and market sensitivity.

Through the combination of convolution layer and pooling layer, CNN can automatically learn local and global features in data and effectively capture patterns in time series data [9-10]. Stock market data often contain complex nonlinear relationships, and CNN can better fit these complex data structures through multi-level nonlinear processing. Stock market data is time-series and spatial. CNN can process multi-channel time-series data at the same time and learn the correlation between different time points and different stocks [11].

By introducing kernel function, SVM can map data to high-dimensional space, so as to better deal with nonlinear relations, which is suitable for complex market scenarios. Stock market data usually contains a large number of features, and SVM performs well on high-dimensional data and can effectively handle complex market information [12]. When dealing with relatively small-scale training data, SVM still has strong generalization ability, and it can also achieve good results in the case of insufficient data in the stock market.

This paper collects and cleans up historical stock market data, including stock price, trading volume and other information, and standardizes and normalizes them. The hierarchical features of time series data are extracted by CNN and transformed into a higher level abstract representation. Input the features extracted by CNN into SVM model, classify them, and predict the trend of stock rise and fall. The structure of CNN-SVM hybrid prediction model is shown in Figure 2.

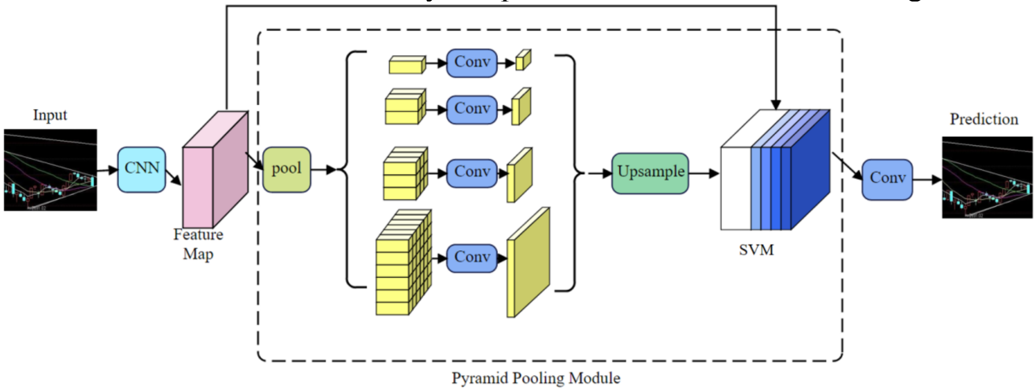

Figure 2 Stock market forecasting model based on CNN-SVM

The training process of CNN-SVM hybrid model is divided into two processes. First, CNN training, after several times of training, until the model converges. Then the features of the input samples are automatically extracted by the obtained CNN, and the extracted features are used as the input of SVM. Once the SVM training is completed, it can be used in the prediction process.

For the binary classification problem, the decision function of SVM is:

$$f(x) = sign\left(\sum_{i=1}^{m} \alpha_i y_i K(x,x_i) + b\right) \tag{1}$$

Where α_i is the Lagrange multiplier of the support vector, y_i is the label and $K(x,x_i)$ is the application of the kernel function.

When integrating CNN and SVM, the output of CNN is used as the input of SVM. If the output of CNN is A_{final}, the input of SVM is:

$$X_{SVM} = Flatten(A_{final}) \tag{2}$$

Among them, $Flatten$ means flattening the high-dimensional feature Tula into a one-dimensional vector.

Global loss function of CNN-SVM model

$$J = J_{CNN} + \lambda \cdot J_{SVM} \tag{3}$$

Where J_{CNN} is the loss function of CNN, J_{SVM} is the loss function of SVM, and λ is the trade-off parameter between them.

4. Empirical analysis

This section compares the stock market forecasting model based on CNN-SVM proposed in this paper with the traditional stock market forecasting methods. The validity of the prediction model proposed in this paper is verified. In this experiment, NASDAQ market index is selected as experimental data for model training and testing. The reason why the market index is chosen as the experimental data is because the market is the concentrated expression of the fluctuation of the overall financial market, which can rule out the contingency of the violent fluctuation of individual stocks

[13-14]. The data is downloaded from the financial website (Yahoo Finance: http://finance.yahoo.com/). The index selects the highest price as the forecast index.

Obtain historical trading data of NASDAQ, including opening price, closing price, highest price, lowest price and trading volume. Data preprocessing, including data cleaning, standardization and division into training set and test set.

CNN-SVM, BP neural network (BPNN) and SVM are used for model training respectively. The test set is used to evaluate each model and compare their prediction performance. Figure 3 shows the forecast results of NASDAQ stock index using three models.

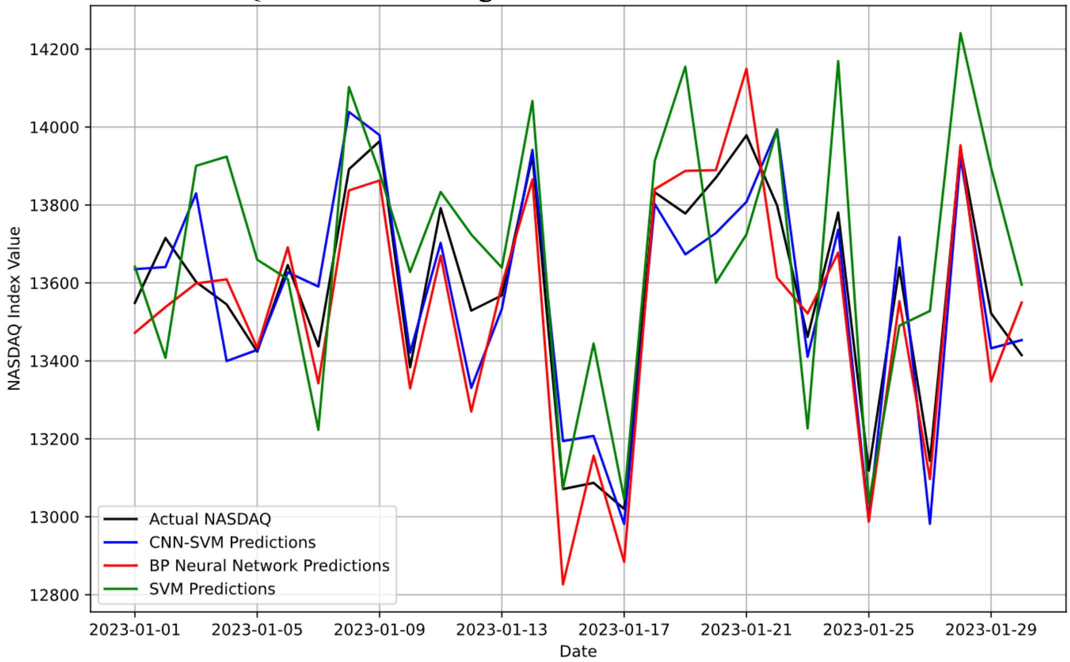

Figure 3 Forecast results of NASDAQ stock index

As shown in the above figure, we use data to compare the performance of CNN-SVM, BPNN and SVM models in predicting NASDAQ. Table 1 below shows the data for the first five days:

Table 1 The data of the first five days

date	Actual NASDAQ	CNN-SVM	BPNN	SVM
2023-01-01	13548.81	13635.26	13472.19	13641.95
2023-01-02	13715.19	13640.97	13538.09	13407.94
2023-01-03	13602.76	13829.74	13598.54	13900.41
2023-01-04	13544.88	13399.45	13609.13	13924.06
2023-01-05	13423.65	13428.23	13433.63	13659.41

As can be seen from the chart, the predicted values of each model closely follow the changing trend of the actual NASDAQ index, but there are differences in volatility and accuracy. Specifically, the prediction value of CNN-SVM model is the closest to the actual value, which shows that it has certain advantages in stock index prediction. In contrast, the predicted values of BPNN and SVM models fluctuate greatly, which may be due to the limitations of instability and complexity in dealing with financial market data.

Compare the accuracy, precision and recall of different models, and analyze the performance advantages of CNN-SVM compared with traditional methods. Figure 4 shows the performance indicators of accuracy, precision and recall of different models.

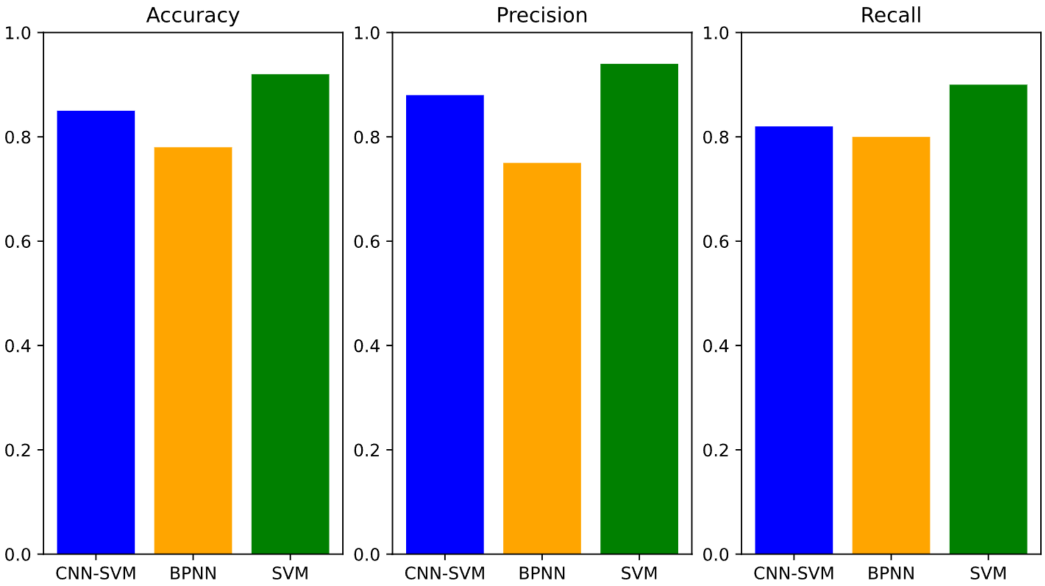

Figure 4 Performance indicators of different models

By observing Figure 4 and model performance indicators, we can find that CNN-SVM has achieved the best performance in accuracy, precision and recall. Compared with BPNN and SVM, CNN-SVM model is more robust and captures the ups and downs of the stock market more comprehensively. BPNN is average in accuracy and recall, and may be sensitive to noise in data. SVM performs well in accuracy and precision, but it is slightly insufficient compared with CNN-SVM. SVM may need more adjustment or feature engineering to adapt to the complexity of stock market data.

In practical application, choosing the most suitable model needs to consider many factors, including the accuracy, stability, computational complexity and adaptability to new data. CNN-SVM model shows good accuracy and stability in this example, and may be more suitable for the prediction of NASDAQ index. However, the actual situation may be different because of the specific characteristics of the data and changes in the market environment.

5. Conclusion

The purpose of this study is to explore the application of CNN in stock market forecasting, and compare CNN-SVM with traditional methods (such as BPNN and SVM) to evaluate its performance in stock index forecasting. By comparing the experimental results, it is found that CNN-SVM can achieve relatively superior performance in stock market forecasting. CNN-SVM can better capture complex time series patterns and nonlinear relationships, and it shows a more powerful performance in the field with highly nonlinear characteristics such as the stock market. Compared with BPNN and SVM, CNN-SVMN has achieved better performance in accuracy, precision and recall. This shows that CNN-SVM has stronger expressive ability and generalization ability in dealing with financial time series data, and has more advantages than traditional methods. Although this research has achieved some success in stock market forecasting, there are still some potential improvement spaces. Future research can consider larger data sets, more complex model structures and more detailed feature engineering to further improve the performance of the model.

References

[1] Zhang, Y. J. , & Wang, J. L. (2019). Do high-frequency stock market data help forecast crude oil prices? evidence from the midas models. Energy Economics, 78(11), 192-201.

[2] Vlah Jeric, S. , & Andelinovic, M. (2019). Evaluating croatian stock index forecasts. Empirical

Economics, 56(4), 1325-1339.

[3] Meng, & Kyle, C. (2017). Using a free permit rule to forecast the marginal abatement cost of proposed climate policy. American Economic Review, 107(3), 748-784.

[4] Lee, J. J. , Lim, Y. , & Oh, H. I. (2018). Does customer satisfaction matter to managers earnings forecasts and stock returns. European Journal of Marketing, 52(10), 2026-2051.

[5] Salisu, A. A. , & Isah, K. O. (2017). Revisiting the oil price and stock market nexus: a nonlinear panel ardl approach. Economic Modelling, 66(10), 258-271.

[6] Xiao, D. (2021). Research on the application of generative adversarial networks in the generation of stock market forecast trend images. Scientific programming, 2021(13), 2021.

[7] Bustos, O. , & Pomares-Quimbaya, A. (2020). Stock market movement forecast: a systematic review. Expert Systems with Applications, 156(32), 113464.

[8] Zhou, X. , Pan, Z. , Hu, G. , Tang, S. , & Zhao, C. (2018). Stock market prediction on high-frequency data using generative adversarial nets. Mathematical Problems in Engineering, 2018(4), 1-11.

[9] Wang, H. , Lu, S. , & Zhao, J. (2019). Aggregating multiple types of complex data in stock market prediction: a model-independent framework. Knowledge-Based Systems, 164(15), 193-204.

[10] Wang, J. S. (2017). Research on the market demand risk and sharing mechanism of ppp projects. Journal of Railway Engineering Society, 34(11), 95-98.

[11] Rubio, A. , Bermudez, J. D. , & Vercher, E. (2017). Improving stock index forecasts by using a new weighted fuzzy-trend time series method. Expert Systems with Applications, 76(7), 12-20.

[12] Yan, Y. , & Yang, D. (2021). A stock trend forecast algorithm based on deep neural networks. Scientific Programming, 2021(2), 1-7.

[13] Zhang, X. , Zhang, Y. , Wang, S. , Yao, Y. , Fang, B. , & Yu, P. S. (2018). Improving stock market prediction via heterogeneous information fusion. Knowledge-based systems, 2018(1), 143.

[14] Wu, L. , & Li, Y. (2017). A novel stock market risk prediction method based on support vector machine optimized by multi-objective optimization. Revista de la Facultad de Ingenieria, 32(10), 593-598.