ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

Improved KNN-based Stock Price Prediction

Yiwen Wang[1,a,#], Yuchen Xie[2,b,#], Yuhao Wu[3,c,#], Yichen Yang[4,d,#]

[1]*Jilin Provincial Experimental School, Changchun, Jilin, China*
[2]*Maranatha Christian Academy, Windsor, Ontario, Canada*
[3]*Golden Sun Experimental High School, Nanchang, Jiangxi, China*
[4]*HD Beijing School, Beijing, China*
[a]*438199143@qq.com,* [b]*XYC7215@outlook.com,* [c]*2586713620@qq.com,* [d]*2637609536@qq.com*
[#]*Co-first author*

Abstract: *Accurate prediction of stock prices is of great significance as it provides critical information for investors, financial institutions, and government policymakers to manage risks, optimize investment portfolios, and formulate sound capital operation strategies. Addressing the challenges in stock price prediction, this paper proposes a novel stock price prediction model based on the K-Nearest Neighbors (KNN) algorithm. The model leverages the core idea of the KNN algorithm by identifying the nearest neighbor data samples to predict future trends in stock prices. Unlike traditional KNN approaches, the proposed model integrates an improved strategy incorporating the price change trends of the preceding N days in time-series data to forecast the price change of the subsequent day more accurately. Experimental results demonstrate that the proposed improved KNN model enhances the accuracy of stock price prediction. This model provides more reliable forecasting information for investors and financial institutions, assisting them in making wiser investment decisions and reducing investment risks. Additionally, this model offers valuable insights for government agencies in formulating monetary policies and risk management, thereby promoting economic stability and capital market development.*

Keywords: *Machine Learning, Stock Price Prediction, KNN*

1. Introduction

In today's global economic system, the stock market plays a pivotal role, with its fluctuations directly impacting the wealth of investors and reflecting the health of national and even global economies to a large extent. Therefore, accurate prediction of stock price trends holds significant importance for investors, governments, and businesses alike. However, due to the complexity and uncertainty of the stock market, stock price prediction has long been a challenging task in the financial domain. To address this issue, researchers have continuously explored various prediction models and algorithms, among which stock price prediction based on improved K-Nearest Neighbors (KNN) has become a hot research topic.

Firstly, understanding the impact of the stock market on the economy is crucial. Stock market movements not only affect corporate financing costs and investment returns directly but also influence consumer and business confidence, thereby impacting consumption and investment decisions. During periods of stock market prosperity, investor sentiment is high, leading to increased consumption and investment activities, often resulting in robust economic growth; conversely, when the stock market declines, investors may feel uneasy, potentially reducing investments and cutting back on consumer spending, thereby dragging down the overall economy. Therefore, accurate prediction of stock price trends is crucial for governments and businesses to formulate policies, manage risks, and enhance competitiveness effectively.

Secondly, stock price prediction holds significant importance for various stakeholders. For investors, accurate prediction of stock price trends enables better formulation of investment strategies, risk mitigation, and profit generation. For governments, effective stock price prediction facilitates timely adjustments of macroeconomic policies to maintain financial market stability. For businesses, accurate stock price prediction helps optimize financing structures, enhance enterprise value, and strengthen market competitiveness. Therefore, stock price prediction is not merely a technical issue in the financial domain but a critical issue involving the entire economic system.

However, stock price prediction faces numerous challenges and complexities. Firstly, the stock

ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

market is influenced by various factors, including but not limited to macroeconomic policies, industry cycles, and international political situations, rendering stock price fluctuations highly complex and uncertain. Secondly, market information is constantly changing, including company financial reports, industry news, and macroeconomic data, making it challenging to timely acquire and accurately analyze this information. Moreover, the stock market is characterized by significant noise and irrational factors, such as market sentiment and speculative behavior, further complicating stock price trend prediction. Therefore, overcoming these challenges and improving the accuracy and stability of stock price prediction has become a critical focus of current research.

Against this backdrop, the method of stock price prediction based on improved KNN has emerged. By leveraging the advantages of the KNN algorithm and introducing specific enhancement strategies, this approach effectively addresses the challenges in stock price prediction, enhancing prediction accuracy and stability. This paper will delve into the principles, advantages, and applications of stock price prediction based on improved KNN, providing more accurate and reliable decision support for investors, governments, and businesses.

2. Related work

Regarding the problem of stock price prediction, researchers have conducted a large number of related studies.

A study is described in the paper [1], which proposes an algorithm combining support vector machine (SVM) and K nearest neighbor (KNN) to solve the problem of stock price prediction. The study selected trading data and technical indicators that reflect stock changes, including trading volume, closing price, highest price, moving average (MA), etc., to predict the rise and fall trends and closing price index of the Shanghai Composite Index. This combination of algorithms may help improve the accuracy of stock prices and provide investors with more comprehensive information to support more informed decisions.

The paper [2] clearly describes a study that uses the improved K nearest neighbor (KNN) algorithm to predict stock price trends to determine the category of the data to be tested through majority voting, and uses the improved KNN algorithm to predict the stock price The analysis of historical data establishes a prediction model, and empirical results show that this method is the most effective in predicting stock price trends. The article mentions processing stock historical data into vector patterns and using the KNN algorithm to predict stock and price trends. The example analysis uses the stock data of Shanghai Stock Exchange A-share Huadong Technology (000727), and compares the improved KNN algorithm with the general KNN algorithm. The results show the average relative error of the improved KNN algorithm is smaller, indicating its effectiveness in stock prices. And the article concludes that the stock prediction model based on the KNN algorithm has advantages in terms of simplicity and effectiveness.

Paper [3] proposes a unique financial time series prediction model, which decomposes signals and extracts features through multiple steps and then outputs them. The output weighted K nearest neighbor method (KNN) is then reconstructed and the overall structure is used in the feature extraction process. All of them are innovative, and the model's prediction effect on China's stock index is certified. This model constructs a complex nonlinear feature extraction and regression prediction process by introducing methods such as FEMD, PCA and improved KNN, which has certain innovative and empirical results. This article reflects the rationality of weighted KNN and can provide better classification effects and prediction capabilities.

The paper [4] proposes to use the KNN algorithm to identify and predict the movement pattern of stock prices, emphasizing the effectiveness of the K-nearest neighbor (KNN) algorithm to predict stock prices, and finally constructs multiple KNN models to form a voting integrated model to improve prediction capabilities. And generalization ability. The article mentioned that it is more common to use hybrid models and improved models. Some studies combine KNN and technical analysis tools and believe that it is feasible to predict short-term stock price trends. The article introduces some examples of hybrid models and improved models, such as combining KNN and support vector machines (SVM), combining KNN and feature projection algorithms, etc., to improve prediction capabilities. The main ideas include the complexity of the stock market, technical analysis and machine learning, KNN algorithm, hybrid model and change model, sliding window method and inheritance model, as well as future research directions to improve the capture of nonlinear dynamic characteristics in financial markets. With computational efficiency, the model complexity is reduced and multiple models are introduced to make an innovative attempt to replace the traditional single model.

ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

The main ideas of the paper [5] are the analysis and determination of stock price indicators, the fusion and dimensionality reduction of stock price indicators, and the trend analysis and judgment based on K nearest neighbor classification. The KNN method is used to study the stock market price trend. This method uses historical data. Model training is more in line with actual market conditions and fully takes into account the impact of historical trends on the future. In this paper, by integrating stock price indicators and using statistics, the model is simplified, the interference of calculations is reduced, and the accuracy and generalization ability of the model are intentionally improved. The article mentioned that future research needs to incorporate more fundamental information and combine qualitative analysis with quantitative analysis. This shows that the author is aware of the impact of good news on stock price trends, and hopes to comprehensively consider factors for individual stocks in future research. .

The paper [6] mainly studies the stock data of listed companies in the stock market and changes in the company's stock price. It aims to establish an optimized version of the K-nearest neighbor (KNN) algorithm by analyzing these data and calculating characteristic values to predict listings. The type of stock price movement of a company. This article designs an algorithm suitable for the upward trend system model by optimizing the KNN algorithm, further optimizes the prediction model, and introduces big data technology and considers different characteristics to improve the accuracy of the prediction of stock price trends.

The paper [7] discusses the use of K nearest neighbor algorithm and nonlinear regression method to predict the stock prices of six major companies listed on Jordan Stock Education to help investors, management, decision-makers and users make correct and wise decisions. Investment decisions,,. According to the results, the KNN algorithm performed well with a small error ratio.

Paper [8] describes a stock market prediction method based on a machine learning model, which is the K-nearest neighbor (KNN) algorithm, in order to predict future stock values. In order to predict the next day's change in stock value, KNN is very powerful in numerical prediction problems because it can handle the relationship between data. By utilizing these algorithms in stock market prediction, the results proposed in this article and the experimental results show up to 70 % accuracy. The innovation points are feature extraction methods, weighted K nearest neighbor method, model integration, big data processing, and feature set selection. Improves the accuracy of stock price predictions and helps make correct and informed investment decisions.

The paper [9] improves the traditional KNN algorithm and introduces a new concept. The stock price information is collectively called a sample and put into the KNN model for learning. The improved KNN algorithm can make more comprehensive use of historical data, making the model better Capturing potential trends and patterns improves prediction performance, which means that in practical applications, the improved KNN algorithm can more accurately capture the changing trend of stock prices and provide more reference information for investment decisions.

The paper [10] investigates a hybrid model that combines K-Nearest Neighbors (KNN) with a probabilistic approach to improve the accuracy of stock price trend predictions. The study addresses the limitations of traditional KNN methods, which primarily focus on central data points, by incorporating Bayesian probabilities to consider both central and peripheral data points. Although the proposed model demonstrates notable improvements in binary classification, future research aims to expand it to a multiclass classification system. This expansion would categorize predictions into five distinct classes: “Sell,” “Underperform,” “Hold,” “Outperform,” and “Buy” offering more comprehensive investment insights. The multiclass approach is expected to further enhance the model’s predictive capabilities, providing more detailed guidance for stock market investments.

3. Methods of this Article

The usual method of predicting stock prices is to learn and model the historical data of stocks on the basis of deep neural network (represent by LSTM), so that the change trend of stock prices can be predicted. Firstly, this method needs to model a large amount of historical data, and secondly, the calculation amount is relatively large, which can not meet the demand of real-time stock trading decision.

Moreover, the stock price is complex and dynamic, mainly because the change of stock price is not only affected by the performance of the stock issuing company, but also affected by the game between buyers and sellers, bookmakers, retail investors and even economic and political factors. Therefore, it is difficult to predict the short-term trend of stock prices based solely on historical data.

This paper intends to realize real-time prediction of stock price base on KNN algorithm, which can

ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

solve the time-consuming problem of conventional methods as much as possible. KNN algorithm is easy to implement, and the calculation amount is controllable within a certain range, which can meet the real-time demand of stock trading decision through certain optimization. The tradition KNN algorithm takes the characteristics of a single data object as input, and then makes prediction by comparing the change trend of the next day with the data of one day. If the selected data is very similar to the sample data, there will be different trends on the next day, such as upward, downward, unchanged, etc. Therefore, after the K most similar samples are voted on. The result may lose its own meaning.

To solve this problem, an improved method based on KNN is purposed. This method is based on the assumption that the share price of the next day depends not only on the previous day's share price, but on the share price of the previous N days, which can be better reflect the trend of the next day based on the trend of the previous N days. Based on this assumption, the method in this paper combines the stock price information of the previous N days into a sample and injects it into the KNN model for learning.

1	Attribute characteristics 1
2	Attribute characteristics 2
3	Attribute characteristics 3
4	Attribute characteristics 4
5	Attribute characteristics 5
6	Attribute characteristics 6
7	Attribute characteristics 7
8	Attribute characteristics 8
...	.....

Figure 1: Input of traditional KNN

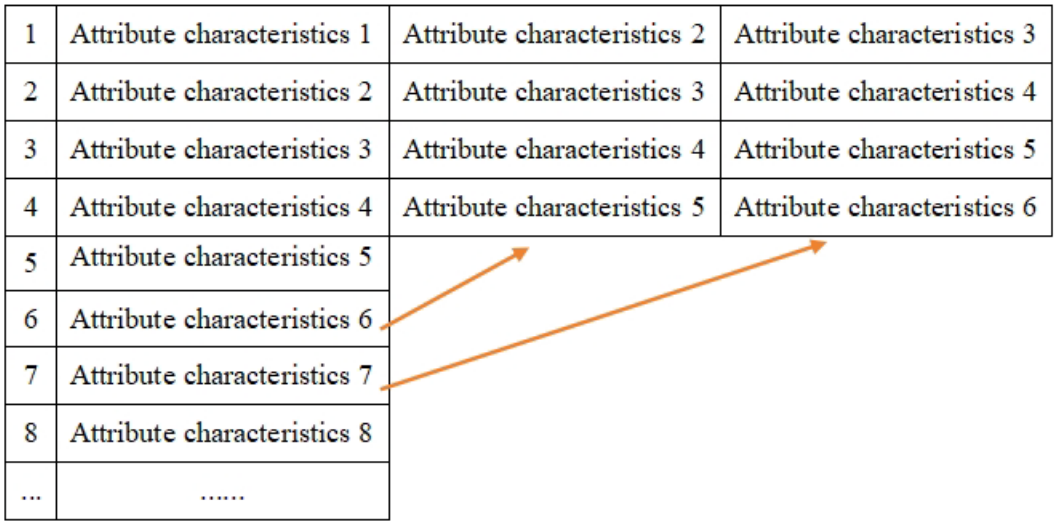

Figure 2: This paper improves the input of KNN

As shown in the Figure 1-2, in the improved stock price prediction based on KNN, for the stock price information of each day, the stock price information of the previous N days (the first two days in the figure) is added as the stock price information of the extended day and extended information is used for the stock price prediction for the following day.

This method takes in account the stock price trend of the previous N days, rather than the encouraging point of a single day, and should theoretically have better predictive performance.

ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

4. Experimentation

In order to verify the effectiveness of the proposed method in this paper, the stock data of Moutai is obtained using the akshare library, and the proposed method as well as the traditional KNN method are tested in comparison. The stock data of Moutai contains attributes such as date, opening price, closing price, high price, low price and volume. In this paper, RMSE is used as a measure of stock price prediction accuracy.

Table 1: Test results

	Traditional KNN	The improved KNN
RMSE	4.176	4.04

As shown in Table 1, the RMSE 4.176 of the traditional KNN, which only uses the current closing price to predict the next day's closing price, may be able to reveal the trend and pattern of the closing price if it is predicted using the previous days' closing prices, which will help us to predict the "trend" of the closing price. Therefore, a possible improvement would be to add more historical closing price information to the feature (The RMSE of the improved KNN is 4.040, which is an improvement over the traditional KNN).

5. Conclusion

This paper has extensively investigated the stock price prediction problem and proposed a method based on an enhanced K-Nearest Neighbors (KNN) algorithm. Considering the volatility and complexity of the stock market, we enhanced the predictive capability of the model by improving feature processing and optimizing parameters within the KNN algorithm. By integrating historical stock price data with other technical indicators, the improved KNN model aims to more accurately capture stock price trends.

The article initially described in detail the importance of the stock market to the economy and the necessity of accurately predicting stock prices. Subsequently, it analyzed the dynamic nature and uncertainty of the stock market, outlining the primary challenges faced in stock price prediction.

In the methodology section, an enhanced KNN model was employed that not only considered historical stock price data but also introduced additional key technical indicators such as trading volume and moving averages. Improvement strategies included optimizing distance calculations within KNN, introducing feature selection, and implementing automatic parameter adjustment mechanisms.

In specific experiments, we utilized data from the Shanghai A-share market for model training and testing. By comparing with traditional KNN models and other benchmark prediction models, the improved KNN model demonstrated lower average relative errors and better prediction stability.

Ultimately, the research results demonstrate that through the introduction of advanced feature processing and algorithm adjustment strategies, the improved KNN model exhibits high accuracy and practicality in stock price prediction. This achievement not only provides a powerful decision-making tool for stock market investors but also opens up new research directions and perspectives for future stock price prediction studies.

References

[1] Peiqi Zhang, Navy Liu, Dongxiao Pei, Xiyue Wang. Stock price prediction based on SVM-KNN. Statistics and Applications, 08.6(2019):859-871..
[2] Bo Wang, and Fuyun Cheng. Application of KNN algorithm in stock prediction. Technology Entrepreneurship Monthly, 28.16(2015):2.
[3] Li Tang, Heping Pan, Yiyong Yao. A prediction model based on nonlinear feature extraction and weighted K nearest neighbour meta-regression. Journal of Information Systems, 2(2019):10.
[4] Wei-Nan Zhang, Uni-Yu Lu, & Jian-Ming Sun. K-nearest neighbour and its integrated model for stock price prediction. Electronic Technology Applications, 45.5(2019):6.
[5] Xiaoying Liu. Research on stock price trend prediction based on KNN method. Heilongjiang: Northeast Agricultural University, 2015.
[6] Tong Yang. A stock ranking method based on machine learning. Tianjin University of Technology, 2017.
[7] Alkhatib, Khalid, et al. "Stock price prediction using k-nearest neighbor (kNN) algorithm."

ISSN 2616-5775 Vol. 7, Issue 6: 38-43, DOI: 10.25236/AJCIS.2024.070606

International Journal of Business, Humanities and Technology 3.3 (2013): 32-44.
[8] Latha, R. S., Sreekanth, G. R., Suganthe, R. C., Geetha, M., Selvaraj, R. E., Balaji, S., ... & Ponnusamy, P. P. Stock Movement Prediction using KNN Machine Learning Algorithm, 2022 International Conference on Computer Communication and Informatics (ICCCI), IEEE, (2022), pp. 1-5.
[9] Yunneng, Qian, A new stock price prediction model based on improved KNN, 2020 7th International Conference on Information Science and Control Engineering (ICISCE), IEEE, (2020), pp. 77-80.
[10] Sarala, V., and GNV PHANI BHUSHAN. "Stock Market Trend Prediction Using K-Nearest Neighbor (KNN) Algorithm." J. Eng. Sci. 13.8 (2022): 249-256.