JPX Tokyo Stock Exchange Prediction with Deep Neural Networks

Shurui Hu[1], Hua Hansen[2], Junchao Mao[3], Jiabing Zhang[*4]

[1]SteveHuShurui@163.com, [2]opticshhs@gmail.com,[3]ohmafohcs@outlook.com; [4]2125370465@qq.com

[1]Shandong University, Weihai, China
[2]Kyoto University, Kyoto, Japan
[3]Zhejiang Ocean University, Zhoushan, China
[4]Xihua University, Chengdu, China

Abstract. Quantitative investment is a quantitative analysis method, according to the historical related data, the use of computer technology combined with mathematical models to establish a quantitative analysis model of investment, to discover the price: the hidden trend behind the change, so as to guide the formulation of investment strategies, so it has been widely used in the field of investment. In this paper, we focus on the JPX stock exchange prediction using deep learning algorithms. We adopt the Deep Neural Network (DNN) as our model, and we do compared experiments with some classical models. Our DNN model owns the best performance in teams of the highest Sharpe Ratio score 0.158. On the contrary, The Xgboost and Lightgbm respectively owns 0.107, 0.112 Sharpe Ratio which are all lower than DNN's Sharpe Ratio.

Keywords: JPX Stock Exchange, Quantitative investment, Sharpe Ratio, Deep Neural Network

1 Introduction

In the financial field continues to pour in more and more investors, pay attention to the stock market and the price trend of various financial products, people always expect to be able to better predict the price changes of financial assets, so as to be able to obtain excess returns through trading strategies, and the quantitative investment method is an effective investment method. Quantitative investment is a quantitative analysis method, according to the historical related data, the use of computer technology combined with mathematical models to establish a quantitative analysis model of investment, to discover the price: the hidden trend behind the change, so as to guide the formulation of investment strategies, so it has been widely used in the field of investment. Compared with traditional investment strategies, quantitative investment has distinct characteristics, such as measurability, verifiability, objectivity and consistency. Quantitative investment in the process of building a strategy is a quantitative approach, so it can be accurately measured in the decision-making process. Quantitative investment is based on the results of model operations, rather than relying on the subjective feelings of investors, can get rid of the weakness of emotional fluctuations and fluke psychology, and can avoid the bias formed by subjective cognition, so as to be able to track the market more effectively and correct the model. From the perspective of data, quantitative investment can use massive multi-dimensional data combined with models to build investment

BDEDM 2024, January 12-14, Ningbo, People's Republic of China

DOI 10.4108/eai.12-1-2024.2347128

strategies, which can mine historical data to support investment. Quantitative investment can construct a model in a quantitative way to predict the trend of stock price fluctuations, which can guide investors to a certain extent in the selection of strategies.

In this paper, our attention is centered on the prediction of the JPX Tokyo Stock Exchange. Subsequent sections commence with a comprehensive review of the existing literature pertaining to Quantitative Investment. Detailed expositions of our methodologies and experimental procedures are presented in Sections III and IV, respectively. In the concluding section, a summary of our findings is provided, alongside proposed directions for future enhancements.

2 Related work

The Efficient Markets Hypothesis, which divides markets into three categories based on whether their prices can fully reflect all the information that affects prices: weak, semi-strong, and strong, was developed by economist Fama in the 1970s. [1] If you invest in a strong and efficient market, you will not be able to obtain excess returns.[2] In a weak and efficient market, the price of a security can fully reflect the historical price information, and the analysis based on the historical price prediction of the stock trend will not achieve excess profits; In a semi-strong efficient market, the price of a security not only fully reflects historical price information, but also includes all publicly available information, such as profitability and financial information; In a strong and effective market, the market price of a security contains the historical price and all information about the listed company, both public and unpublic, and no method can make investors profitable.

Jules Regnault first used quantitative methods to find universal laws from the market and profit from this investment, sowing the seeds for the development of quantitative theory later. Since then, more and more people have entered the field of quantitative related theories for research, and have also developed quantitative investment theories in various aspects such as quantitative stock selection and quantitative timing. Quantitative stock selection is generally the first step of quantitative investment, investors always hope to select high-quality stocks with potential, so as to better develop investment strategies.[3] The theoretical development of the multi-factor model in quantitative stock selection stems from the portfolio theory proposed by Markowitz (1952) in the twentieth century, which uses the mean variance model to measure returns and risks, laying the cornerstone for the development of quantitative financial theory[4]. Later, Sharp (1964) and others finally proposed the Capital Asset Pricing Model (CAPM) through research, which explained the intrinsic relationship between risk and return, and gradually improved and developed the theory of quantitative investment.[5] Merton (1973) proposed the intertemporal capital asset pricing model (ICAPM) theory on the basis of CAPM theory [6], Ross (1976) later created the arbitrage pricing theory (APT) in his research, which explained the return of risk assets according to various factors, and believed that when the market is in an unbalanced state, there are potential risk-free arbitrage opportunities [7], providing a theoretical basis for the proposal of multi-factor models.

With the rapid development of computer technology, more and more scholars began to try to combine machine learning with quantitative investment to predict the trend of the stock market or the stock price by establishing machine learning models.

Dixon, Klabjan and Jin (2016) [8] applied the deep neural network to the prediction of futures price changes, and obtained a high accuracy. Basak et al. [9] CananXiao et al. (2019) [10] established a multi factor stock selection model using an enhanced particle swarm support vector machine algorithm. First, they used the nearest neighbor algorithm to remove the samples with insignificant effect and optimize feature selection. Then, they used the improved discrete particle swarm algorithm to select the parameters of SVM, which can greatly reduce the operation time.

3 Methods

The Deep Neural Network (DNN) constitutes a foundational architecture within the realm of deep learning, characterized by the incorporation of at least one hidden layer. Mirroring the capabilities of shallow neural networks, DNNs similarly facilitate the modeling of intricate nonlinear systems. However, the inclusion of additional layers affords a more sophisticated level of abstraction, thereby augmenting the model's effectiveness.

In the context of our research, we employ a DNN framework to advance the domain of quantitative investment. The comprehensive architecture of our proposed model is delineated in Figure 1.

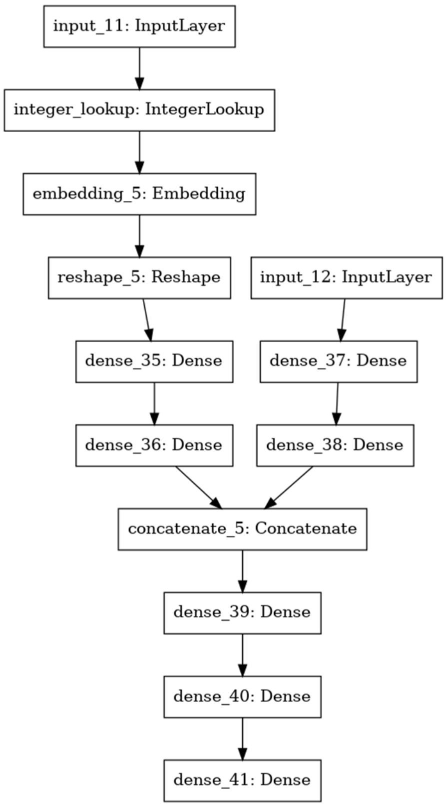

Figure 1: the structure of DNN model.

4 Experiments

- **Experimental Data**

The dataset employed in our study is sourced from Japan Exchange Group, Inc. (JPX), a preeminent holding entity that administers several of the world's leading financial exchanges, including the Tokyo Stock Exchange (TSE), along with derivatives exchanges such as the Osaka Exchange (OSE) and the Tokyo Commodity Exchange (TOCOM). This dataset encompasses historical records pertaining to an extensive array of Japanese equities and derivatives.

The table 1 shows the dataset for this task about the Tokyo stock..

Table 1: dataset files.

stock_prices.csv	The primary file under examination, which comprises the daily closing prices for each stock along with the target column.
options.csv	Information pertaining to the condition of numerous options within the wider market context. A significant proportion of these options embed implicit forecasts concerning future stock market prices, rendering them potentially valuable despite not being directly evaluated.
secondary_stock_prices.csv	The central dataset focuses on the top 2,000 equities by trading volume; however, a multitude of less frequently transacted securities also exists within the Tokyo market. This file encompasses information pertaining to these additional securities, which, while not evaluated for scoring purposes, may provide valuable insights into the broader market dynamics.
trades.csv	Consolidated recapitulation of trade volumes from the preceding week of business activity.
financials.csv	Outcomes derived from quarterly financial statements.
stock_list.csv	The quantity of shares at the highest/second-highest buy level in terms of competitiveness.

The equity pricing data encompasses a dozen columns, serving as the fundamental input for our model. The columns are shown in table 2. And we draw three different windows simple moving average curve showing the input intuitively in figure 2. Figure 3 shows the distribution of opening prices of stocks with different stock codes.

Table 2: important inputs.

RowId	Date
Code	Low
High Securities	Open
Close	Volume
AdjustmentFactor	ExpectedDividend
SupervisionFlag	Target

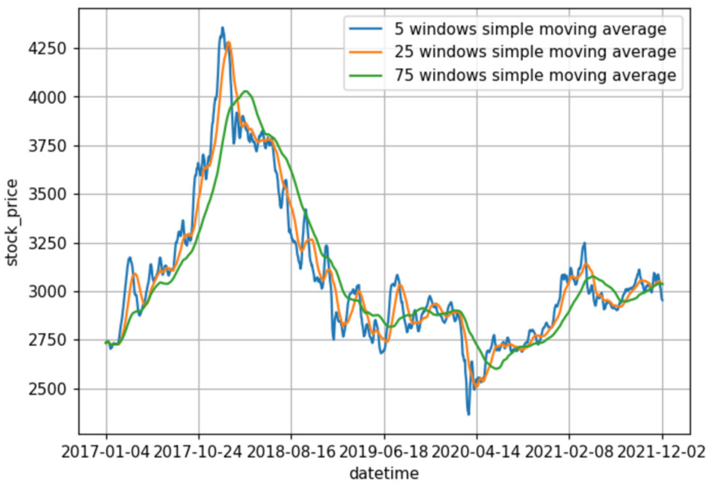

Figure 2: Stock prices in different time windows.

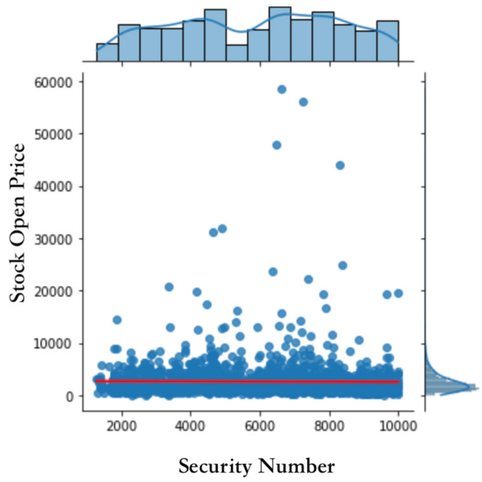

Figure 3: open price distribution.

- **Designed Experimental hyperparamters and Experimental results**

The configurations for our experiments are detailed in Table 3, wherein we utilize Pytorch for model training.

Table 3: experimental settings.

Epoch	3
batch size	4096
Adam	0.05

In order to benchmark our model against others, we assess our outcomes employing the Sharpe Ratio. In the domain of finance, the Sharpe Ratio quantifies an investment's yield (be it securities or portfolios) in relation to risk-free assets, factoring in the associated risk. It quantifies the mean of the discrepancy between the returns on investments and those risk-free, normalized by the standard deviation of the investment (representing its volatility). Essentially, it denotes the excess return acquired for every unit of risk undertaken by the investor.

$$S_a = \frac{E[R_a - R_b]}{\sigma_a} = \frac{E[R_a - R_b]}{\sqrt{var[R_a - R_b]}}$$

The experimental results of competing models for this task and our proposed model are shown in table 4.

Table 4: performance of different models.

Models	Sharpe Ratio
Xgboost	0.107
Lightgbm	0.112
DNN	0.158

Table 4 reveals that our DNN model outperforms all others, achieving the highest Sharpe Ratio at 0.158. In comparison, the Xgboost and Lightgbm models yield Sharpe Ratios of 0.107 and 0.112, respectively, both of which fall short of the DNN's performance.

5 Conclusion

In the financial field continues to pour in more and more investors, pay attention to the stock market and the price trend of various financial products, people always expect to be able to better predict the price changes of financial assets, so as to be able to obtain excess returns through trading strategies, and the quantitative investment method is an effective investment method. Quantitative investment is a quantitative analysis method. This paper is centered on predicting the JPX Tokyo Stock Exchange. Initially, we delve into existing research related to Quantitative Investment. Sections III and IV outline our methodologies and experimental procedures. We conclude by summarizing our findings and suggesting areas for future enhancement. Our results demonstrate that the DNN model outshines competitors, boasting the top Sharpe Ratio, scored at 0.158, significantly higher than the scores achieved by the Xgboost and Lightgbm models, which are 0.107 and 0.112, respectively.

Acknowledgments: Our gratitude extends to the Kaggle platform, which has been instrumental in facilitating our extended period of research and enabling us to compile and present our findings in this document.

References

[1] Cao L, F. E. Tay. Modified support vector machines in financial time series forecasting[J]. Neurocomputing,2002,48(1-4):847-861.

[2] Kim K J. Financial time series forecasting using support vector machines[J]. Neurocomputing,2003,55(3):307-319.
[3] Min S.H, J.Lee, I.Han. Hybrid genetic algorithms and support vector machines for bankruptcy prediction[J]. Expert systems with applications,2006,3 1(3):652-660.
[4] Sorensen E, Miller K L, Ooi C K, et al. The Decision Tree Approach to Stock Seletion[J]. The Journal of Portfolio Management,2009,27(1): 42-52.
[5] Nair B B, Mohandas V P, Sakthivel N R, et al. A Decision Tree- Rough Set Hybrid System for Stock Market Trend Predicion[J]. International Journal of
Computer Applications,2010,6(9):34-41.
[6] Ticknor J L. A bayesian regularized artificial neural network for stock market forecasting[J]. Expert Systems with Applications,2013,40(14):5501 -5506.
[7] Fan A., Palaniswami M.. Stock selection using support vector machines[A]. Neural Networks, 2001, Proceedings. IJCNN '01. International Joint Conference on.
[8] Cao L, F. E. Tay. Financial forecasting using support vector machines[J]. . Neural Computing& Applications,2001,10(2):184-19.
[9] Cao L, F. E. Tay. Application of support vector machines in financial time series forecasting[J]. Omega,2001 ,29(4):309-3 17.
[10] Canran Xiao , Liwei Hou , Jun Huang. Research on Multi- factor Stock Selection Strategy based on Improved Particle Swarm Support Vector Machine[A]. Proceedings of the 1st International Symposium on Economic Development and Management Innovation (EDMI 2019).