A ML-based Stock Trading Model for Profit Predication

Jimmy Ming-Tai Wu[1], Lingyun Sun[1], Gautam Srivastava[2], and Jerry Chun-Wei Lin[3]⋆

[1] College of Computer Science and Engineering
Shandong University of Science and Technology, Qingdao, China
wmt@wmt35.idv.tw, 201983060020@sdust.edu.cn
[2] Department of Mathematics & Computer Science
Brandon University, Brandon, Canada
SRIVASTAVAG@brandonu.ca
[3] Department of Computer Science, Electrical Engineering and Mathematical Science
Western Norway University of Applied Sciences, Bergen, Norway
jerrylin@ieee.org

Abstract. This paper uses a new convolutional neural network framework to collect data on leading indicators including historical prices and their futures and options, and use arrays as the input map of the CNN framework for stock prices trend prediction. Experiments are then conducted by the stock markets of the United States and Taiwan using historical data, futures and options as data sets to predict the stock prices. After that, genetic algorithm is then utilized to find trading signals. Results showed that the designed model achieves good return of the investments.

Keywords: Convolutional neural network· Genetic algorithm· Trading signals· Leading indicators.

1 Introduction

The financial market [19] is a mechanism for determining the price of financial funds and trading financial assets. It is a market that enables the financing of securities and the trading of securities. The capital market is also called "long-term financial market", which mainly includes the stock market, fund market, and bond market. Its volatility can reflect the degree of risk of assets. The fluctuation of stock prices plays a considerable role in the appropriate timing of buying and selling stocks [13]. For investors, the true meaning of investing in the stock market is to obtain extraordinary returns by buying low and selling high, so the prediction of stock price fluctuations has become a special focus of private investors and investment companies [16].

Investors are not completely rational, for example, people may have positive or negative emotions at certain moments [2,6,10,11,14]. Therefore, after this

⋆ Corresponding author

hypothesis was raised, there were both support and opposition, thus becoming one of the most controversial investment theories [4,8]. Some researchers believe that if the trading signals [3] of the stocks can be found, the stocks can be bought and sold at the appropriate time to obtain relatively high profits.

Deep learning is based on traditional neural networks, including convolutional neural networks, recurrent neural networks, etc., and has good results in image recognition, text classification, machine translation and other fields. It forms the consistency of specific data and the consistency of goals on the basis of specific data, especially the rise of convolutional neural networks in deep learning research. At present, many researchers have applied convolutional neural networks to the prediction of the stock market [9,18].

The main contributions of this article are as follows. First, this study uses a two-dimensional tensor as the input of the SSACNN framework, and divides the output into three categories: rising, falling and unchanged. Also, the SSACNN framework and genetic algorithm are combined to propose a stock trading system. This system can find stock trading signals relatively well, so that investors can get a certain amount of income.

2 Related Work

The prediction of stock prices is mainly the analysis of historical behaviors, such as people's historical emotions, historical market information, etc., and then useful features are extracted from them to train better predictive models. The value of stock prices is a time series. Revealing the development and changes of stock prices is an objective record of stock historical behavior. In the early days of the stock market, a host of investors relied on their own experience to judge stock price movements, which seemed too subjective and lacked scientific basis. In addition, stock prices are also affected by many other factors. For instance, Zheng et al. [21] studied the relationship between exchange rates and stock prices on the Hong Kong stock market.

Intelligent optimization algorithms are used to find the optimal solution. In recent years, researchers have been more enthusiastic about using genetic algorithms to solve problems. In the financial market, many researchers also use genetic algorithms to find stock trading signals. Allen et al. [1] used the genetic algorithm to learn a technical trading rule based on the daily prices of the S&P 500 index from 1928 to 1995. Based on a series of technical indicators that generate buying and selling signals [17], a genetic algorithm is used to propose a trading strategy. Hirabayashi et al. [12] proposed a genetic algorithm system to find suitable trading signals and automatically generate trading rules based on technical indicators. The focus of this system is not to predict the price of the transaction, but to find the right trading opportunity. Lin et al. [15] used genetic algorithms to set optimal values for the parameters of the problem, and bought or sold stocks at the appropriate trading time.

3 Methodology

3.1 Designed Optimization Framework

The stock market has always been the focus of investors' attention. As the stock market is affected in many ways, finding trading signals for stocks is always a big problem. In the past, investors generally used trading strategies to obtain stock trading signals, which were generated by technical indicators or fundamentals [5,7]. CNN has been proven to have good image recognition capabilities, and many researchers have also used CNN for stock price prediction.

CNN includes convolutional layer, pooling layer and fully connected layer. The convolution layer mainly extracts local features of the input data. The researcher defines a convolution kernel inside the convolution layer. Its shape is a square matrix that is used to extract a certain feature. The convolution kernel is multiplied by the corresponding bits of the digital input matrix and then added to obtain the output value of a convolution layer. The calculation process is shown in Eq. 1.

$$V_{a,b}^{L} = \imath \left(\sum_{m=0}^{K-1} \sum_{n=0}^{K-1} w_{m,n} V_{a+m,b+n}^{L-1} + biase^{L-1} \right) \tag{1}$$

In the Eq. 1, $V_{a,b}^{L}$ is the value of layer L at row a, column b, $\imath$ is a activation function. $bias^{L-1}$ is represent the bias of $L-1$. $w_{m,n}$ is the weight of convolution filter at row m, column n. The formula for calculating the output image size of the convolutional layer is shown in Eq. 2.

$$w' = \frac{w + 2p - k}{s} + 1 \tag{2}$$

Among them, the size of the convolution kernel is k, the size of the input matrix is w, the number of zero-filling layers is p, and the step size is s. Give an example for this progress, the input layer $L-1$ is set as a 5×5 matrix and use the 3×3 convolutional filter. The layer of input L is calculated by Eq. 1, which is set as 3×3. Because one convolution kernel recognizes one feature, and the input data may have multiple features, there may be numerous convolution kernels in one convolution layer to extract multiple features. We then use the output of the obtained convolution layer as the input of the pooling layer.

However, before entering the value into the pooling layer, an activation function is usually added to solve the nonlinear problem. At present, the activation function Relu (Rectified Linear Unit) is commonly used, which is shown in Eq. 3.

$$f(x) = \begin{cases} 0, x \leq 0 \\ x, x > 0 \end{cases} \tag{3}$$

The pooling layer is mainly used to reduce the number of training parameters and reduce the dimension of the feature vector output by the convolution layer.

The most common pooling layers are maximum pooling and mean pooling. In this article, we choose maximum pooling; that is, the maximum value in a specified area is selected to represent the entire area. The output value of the pooling layer is then expanded as the input of the fully connected layer to generate the final output.

After several times of convolution, excitation, and pooling, the model will learn a high-quality feature map, and then input the feature map to the fully connected layer to get the final output. The calculation process is shown in Eq. 4.

$$V_a^b = \imath\left(\sum_K V_K^{b-1} w_{K,a}^{b-1} + biase^{b-1}\right) \tag{4}$$

In this formula, V_a^b is the value of layer b in neuron a, $\imath$ is an activation function, and $w_{K,a}^{b-1}$ is a weight which connects between neuron K from layer *b-1* and neuron a from layer b. $bias^{b-1}$ represents the bias of $b-1$. The pseudo-code of this process is shown in Algorithm 1.

Algorithm 1 Designed model

Require: d is the data of training; K is the data of testing; Z is the number of iteration; A is batch size; Algorithm SGD is named $Adam$.
Ensure: the train model m; evaluation result $accuracy$

```
Initialize algorithm
d ← Initializealgorithm
S ← (split d in equal parts of A)
for each round t = 1, 2, ..., z do
    {verify, train} ← {S_t, S − S_t}
    (tf, vf) ← (generate feature of train and verify)
    m_t ← modelFit(Adam, tf)
    r_t ← modelEvaluate(m_t, vf)
end for
m ← bestModel
K ← m
accuracy ← modelEvaluate(m, test)
```

In this study, the stock data for a period of time is converted into an image, and this image is used as the input of the CNN framework. The input here is 30-day stock data, and the generated "input image" is input to the convolutional layer, pooling layer, dropout layer, and norm layer. Then, loop this process three times. After a series of experiments, it is concluded that when the convolutional neural network is used for image recognition, the size of the convolution kernel is 3×3 and the size of the pooling layer is 2×2, the experimental effect obtained is the best it is good. Therefore, in order to achieve better results for this research, the size of the convolutional layer and pooling layer are set to 3×3 and 2×2, respectively. The specific structure of this framework is shown in Fig. 1.

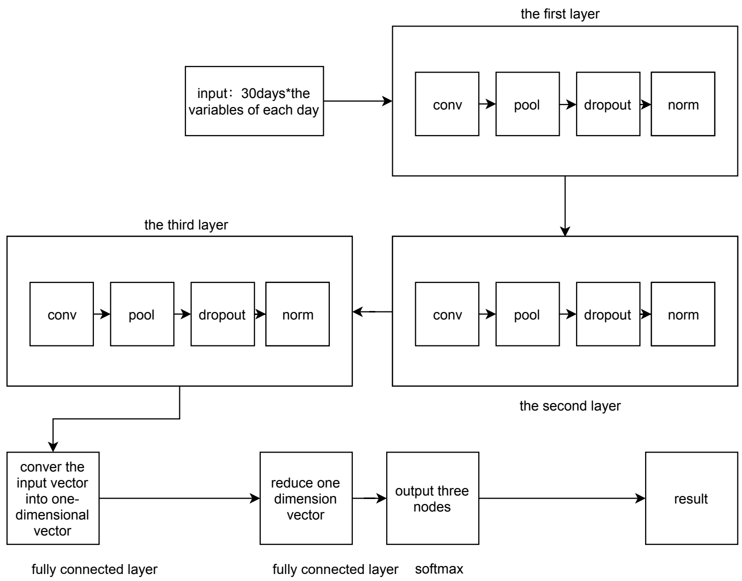

Fig. 1. The designed framework for stock trading prediction.

3.2 Genetic Algorithm

In this section, the components of the GA-based method are described as follows.

Chromosome Representation: The main goal of this method is to find trading signals for stocks, namely buy signals and sell signals. Therefore, in this method, the chromosome is composed of two parts, as shown in Fig. 2.

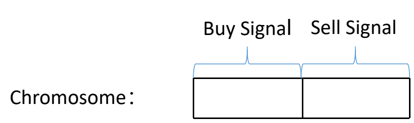

Fig. 2. Representation of chromosome.

The first part represents the threshold when the stock has a buy signal, and the second part represents the threshold when the stock has a sell signal.

Fitness Evaluation: The fitness function is used to evaluate the quality of chromosomes, that is, to evaluate whether the threshold value of the generated buying and selling signals is the optimal solution. In this method, the fitness function is defined as the cumulative return over a period of time, which is shown in Eq. 5.

$$fitness = \sum_{i=0}^{n} (SClose_i - BClose_i) \tag{5}$$

Among them, n indicates that there have been n buying and selling signals in total during this period. $SClose_i$ represents the closing price of the stock when the i-th sell signal appears, and $BClose_i$ represents the closing price of the stock when the i-th buy signal appears.

Crossover: The genetic algorithm performs crossover and mutation operations to generate new solutions. The crossover operation needs to find two chromosomes from the chromosomes of the previous generation, one as the father and one as the mother. Then the two chromosomes are cut and spliced together at a certain position to generate a new chromosome. Part of the new chromosome is the father's genes, and the remaining part is the mother's genes.

Mutation: The crossover operation is only to operate on the original chromosomes, only to exchange their gene order. This can only guarantee a local optimal solution after multiple evolutions. In order to achieve the global optimal solution, a mutation operation is added. Introduce new genes into existing chromosomes by randomly modifying genes.

4 Proposed a Stock Trading System Combining SSACNN and GA

4.1 Dataset

Futures trading, as a special trading method, has undergone two complex evolutionary processes from the beginning of spot trading to forward trading, and then from forwarding trading to futures trading. To put it simply, futures are not a spot, but a standardized contract. The purpose of futures trading is generally not to obtain physical objects at maturity, but to buy and sell futures contracts. Time, quantity, and quality are the three elements of this standardization, that is, a contract that delivers a fixed quantity and a certain quality of a certain quality at a specific time. Futures contracts are uniformly formulated by the futures exchange. The delivery period of futures is placed in the future.

Options are similar to futures and are also a contract. The option is generated on the basis of futures. When the option is traded, the party who buys the option is called the buyer, the assignee of the right, and the party who sells the option is called the seller, the obliger who must perform the buyer's exercise of the right. The difference between futures and options is that options give the buyer of the contract the right to buy or sell a predetermined number of commodities at the agreed price within the agreed period of the parties. It is a right to choose whether to execute or not in the future.

Five important indicators of historical price are often used: *opening price*, *lowest price*, *highest price*, *closing price* and *volume* for stock price analysis. The futures indicators uses in our experiments include *opening price*, *highest value*, *lowest value*, *closing price* and *volume*. The indicators of the options the test uses include *volume*, *open interest*, *closing price* and *settlement price*.

First, let's look at a dataset of two stocks in the Taiwan market. The two stocks are: DJO and DVO Table 1, Table 2, and Table 3 are the historical data, futures data, and option data of the two Taiwan stocks, respectively.

Table 1. The historical prices of the two stocks

	d_{i1}	d_{i2}	d_{i3}	d_{i4}	$d_{i.}$
DJO	260	264	259	264	...
DVO	244.5	246.5	243	246.5	...

Table 2. Futures data of the two stocks

	t_{i1}	t_{i2}	t_{i3}	t_{i4}	$t_{i.}$
DJO	163.5	262	263.5	258	...
DVO	246	244	246.5	243.5	...

Table 3. Option data of the two stocks

	z_{i1}	z_{i2}	z_{i3}	z_{i4}	$z_{i.}$
DJO	14.85	1	0.27	1	...
DVO	3.4	2	6.85	2	...

The d_i, t_i, z_i are the various factors that affect the stock price. d_i represents the open, high, low, close or volatility attributes of the stock. t_i represents the current price, the opening price, the highest price, and the closing price of the futures. z_i represents the attributes of open interest and settlement price of options. Among them, options include buy options and sell options.

4.2 Data Initialization

Before the experiment, input data needs to be further processed, that is, to standardize the data. Because the experiment may be affected by the size of the data and the results are not ideal. First, standardize the original data to make the range of all data consistent. The processed data conforms to the standard normal distribution and helps improve the accuracy of the prediction, which is shown in Eq. 6.

$$x^{*} = (x - \mu)/\sigma, \tag{6}$$

where μ is the mean of all sample data and σ is the standard deviation of all sample data.

4.3 Proposed Stock Trading System

In this section, the proposed stock trading system is introduced in detail. This system consists of SSACNN framework and GA. In order to make the predicted stock price closer to the actual value, we uses three data sets, including historical prices, futures and options. First, input three data sets into the SSACNN framework to get the values of three nodes. These three nodes represent three meanings: the first node represents a buy signal, the second node represents a hold signal, and the third node represents a sell signal. Here, only the first node and the third node are used. In the designed model, you only need to know the buy signal and sell signal of the stock. Next, use two nodes to find the optimal threshold of stock trading signals. The nodes are input into the genetic algorithm, and the thresholds of the two buying and selling signals can be obtained through operations such as crossover, mutation and iteration. Finally, the value of the node obtained through the test data set is compared with the threshold to obtain a more correct buying and selling point, and this rule is used to calculate the cumulative income over a period of time.

5 Experimental Results

In this study, a new SSACNN framework with leading indicators is used. Combining this framework with genetic algorithms can find a pair of trading signals. Using this trading signal for stock trading can get a certain amount of income. In the experimental part, a total of two stocks were used, including International Business Machines (IBM) and Facebook, Inc. (FB) in the US market. Two stocks including MediaTek (DVO) and Asustek (DJO) in the Taiwan market. In the experiment, three data sets are used, which are historical stock prices, futures and options. The data set is divided into two parts, one part is used for training and the other part is used for testing. This research compares the proposed stock trading system with another trading strategy (Deep Q-Learning combined with Technical Index). The specific experimental process is shown below.

First, conduct a study on two stocks in the US stock market shown in Fig. 3 and Fig. 4. Among them, the abscissa represents the data set, “HP” represents historical prices, “O” represents options, and “HP+O” represents historical prices and options. Because there is no leading indicator of futures in the US stock market, only two data sets, historical prices and options, are used. In

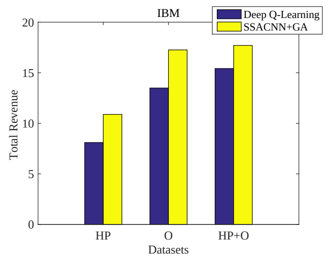

Fig. 3. The total income of IBM stock.

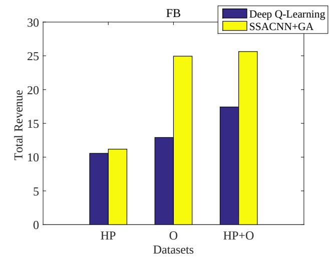

Fig. 4. The total income of FB stock.

addition, the two data sets of historical price and option rights are combined to form a third data set, namely "HP+O". The ordinate represents the cumulative income during the test period. The experimental results show that the stock trading system we proposed is better than the trading strategy composed of Deep Q-Learning and technical index. The benefits of training with the "HP+O" and "O" data sets are better than the benefits of training with the "HP" data sets. The benefits of training with the "HP+O" data set are better than the benefits of training with the "O" data set.

6 Conclusion

This paper mainly proposes a stock trading system by CNN and GA algorithms for obtaining stock price prediction. Genetic algorithms are used to find the optimal value of stock trading signals. This paper integrates the data into a matrix, and then uses the matrix as input instead of inputting it into the model one by one. In addition, the output of the model adopts a classification method to divide the stock price into three categories of the predicted stock price. The output of the designed model is three nodes. In this study, only the first node and the third node are concerned, that is, only the buy signal and the sell signal are concerned. The experimental results show that the proposed stock trading system can help investors obtain certain returns within a period of time.

References

1. Allen, F., Karjalainen, R.: Using genetic algorithms to find technical trading rules. Journal of financial Economics **51**(2), 245–271 (1999)
2. Baker, M., Wurgler, J.: Investor sentiment and the cross-section of stock returns. The journal of Finance **61**(4), 1645–1680 (2006)
3. Bao, W., Yue, J., Rao, Y.: A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLOS ONE **12**(7), 0180944 (2017)
4. Borovkova, S., Tsiamas, I.: An ensemble of lstm neural networks for high-frequency stock market classification. Journal of Forecasting **38**(6), 600–619 (2019)

5. Chen, C.H., Lu, C.Y., Lin, C.B.: An intelligence approach for group stock portfolio optimization with a trading mechanism. Knowledge and Information Systems pp. 1–30 (2019)
6. Chong, T.T.L., Cao, B., Wong, W.K.: A new principal-component approach to measure the investor sentiment (2014)
7. Chou, Y.H., Kuo, S.Y., Chen, C.Y., Chao, H.C.: A rule-based dynamic decision-making stock trading system based on quantum-inspired tabu search algorithm. IEEE Access **2**, 883–896 (2014)
8. Cowles 3rd, A.: Can stock market forecasters forecast? Econometrica: Journal of the Econometric Society pp. 309–324 (1933)
9. Gunduz, H., Yaslan, Y., Cataltepe, Z.: Intraday prediction of borsa istanbul using convolutional neural networks and feature correlations. Knowledge-Based Systems **137**, 138–148 (2017)
10. Han, B.: Investor sentiment and option prices. The Review of Financial Studies **21**(1), 387–414 (2008)
11. Hiew, J.Z.G., Huang, X., Mou, H., Li, D., Wu, Q., Xu, Y.: Bert-based financial sentiment index and lstm-based stock return predictability. arXiv preprint arXiv:1906.09024 (2019)
12. Hirabayashi, A., Aranha, C., Iba, H.: Optimization of the trading rule in foreign exchange using genetic algorithm. In: Proceedings of the 11th Annual conference on Genetic and evolutionary computation. pp. 1529–1536 (2009)
13. Kim, T., Kim, H.Y.: Forecasting stock prices with a feature fusion lstm-cnn model using different representations of the same data. PloS one **14**(2) (2019)
14. Lee, C.M., Shleifer, A., Thaler, R.H.: Investor sentiment and the closed-end fund puzzle. The journal of finance **46**(1), 75–109 (1991)
15. Lin, L., Cao, L., Wang, J., Zhang, C.: The applications of genetic algorithms in stock market data mining optimisation. Management Information Systems (2004)
16. Samuelson, P.A.: Lifetime portfolio selection by dynamic stochastic programming. In: Stochastic Optimization Models in Finance, pp. 517–524. Elsevier (1975)
17. Schoreels, C., Logan, B., Garibaldi, J.M.: Agent based genetic algorithm employing financial technical analysis for making trading decisions using historical equity market data. In: Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004. (IAT 2004). pp. 421–424 (2004)
18. Siripurapu, A.: Convolutional networks for stock trading. Stanford Univ Dep Comput Sci (2014)
19. Tsai, H. H., Wu, M. E., Wu, W. H.: The information content of implied volatility skew: evidence on Taiwan stock index options. Data Science and Pattern Recognition, **1**(1), 48–53 (2017)
20. Wu, J.M.T., Li, Z., Srivastava, G., Tasi, M.H., Lin, J.C.W.: A graph-based convolutional neural network stock price prediction with leading indicators. Software: Practice and Experience (2020)
21. Zheng, L., Jiang, Y., Long, H.: Exchange rates change, asset-denominated currency difference and stock price fluctuation. Applied Economics **51**(60), 6517–6534 (2019)