2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)

Research on Stock Price Prediction Method Based on Convolutional Neural Network

Sayavong Lounnapha ,Wu Zhongdong ,Chalita Sookasame
School of electronics and Information Engineering, Lanzhou Jiaotong University, Gansu 730070, China.
noylucky@163.com

***Abstract*—In order to meet the needs of the financial industry and the financial market, effectively improve the rate of return on funds and avoid market risks, this paper proposes a stock price prediction model based on convolution neural network, which has obvious self-adaptability and self-learning ability. Combining the characteristics of CNN (Convolution Neural Network) and Thai stock market, the data set is trained and tested after pretreatment. On this basis, three stocks (BBL, CAPLL&PTT) listed on the Thai Stock Exchange are tested and compared with the actual stock price. The results show that the model based on CNN can effectively identify the changing trend of stock price and predict it which can provide valuable reference for stock price forecast. The prediction accuracy is high, and it is worth further promotion in the financial field.**

Keywords- Stock price; CNN; prediction; deep learning; algorithm

I. INTRODUCTION

Forecasting the direction of future stock prices is a widely studied topic in many fields including trading, finance, statistics and computer science. The motivation for which is naturally to predict the direction of future prices such that stocks can be bought and sold at profitable positions. Professional traders typically use fundamental and/or technical analysis to analyze stocks and make investment decisions [1-3].

Recently, a lot of interesting work has been done in the area of applying Machine Learning Algorithms for analyzing price patterns and predicting stock prices and index changes. Most stock traders nowadays depend on Intelligent Trading Systems which help them in predicting prices based on various situations and conditions, thereby helping them in making instantaneous investment decisions. Stock Prices are considered to be very dynamic and susceptible to quick changes because of the underlying nature of the financial domain and in part because of the mix of known parameters and unknown factors (like Election Results, Rumors etc.) An intelligent trader would predict the stock price and buy a stock before the price rises, or sell it before its value declines. Though it is very hard to replace the expertise that an experienced trader has gained, an accurate prediction algorithm can directly result into high profits for investment firms, indicating a direct relationship between the accuracy of the prediction algorithm and the profit made from using the algorithm.

Recently, Deep learning (DL) is a class of modern tools that is suitable for automatic features extraction and prediction. In many domains, such as machine vision and natural language processing, Deep learning methods have been shown to be able to gradually construct useful complex features from raw data or simpler features. Since the behavior of stock markets is complex, nonlinear and noisy, it seems that extracting features that are informative enough for making predictions is a core challenge, and Deep learning seems to be a promising approach to that [4].

II. RELATED WORKS

Different methods in stock prediction domain can be categorized into two groups. The first class includes algorithms try to improve the performance of prediction by enhancing the prediction models, while the second class of algorithms focuses on improving the features based on which the prediction is made. In the first class of the algorithms that focus on the prediction models, a variety of tools have been used, including Artificial Neural Networks (ANN), naive Bayes, SVM and random forests. The most popular tool for financial prediction seems to be ANN [5].

In 2011, a comparison between performance of ANN and SVM were done. Ten technical indicators were passed to these two classifiers in order to forecast directional movement of the Istanbul Stock Exchange (ISE) National 100 Index. Authors found that ANNs ability in prediction is significantly better than SVM. Feedforward ANNs are popular types of ANNs that are capable of predicting both price movement direction and price value [6].

In 2016, researchers used genetic algorithm and simulated annealing to find initial weights of an ANN, and then back-propagation algorithm is used to train the network [7]. This hybrid approach outperformed the standard ANN-based methods in prediction of Nikkei 225 index return. With slight modifications [8], genetic algorithm was successfully utilized to find optimized weights of an ANN in which technical indicators were utilized to predict the direction of Nikkei 225 index movement.

Deep ANNs, that are basically neural networks with more than one hidden layer, among the first deep methods used in the domain. In 2016, some researchers [9] predicted NASDAQ prices based on the historical price of four and nine days ago. ANNs with different structures, including both deep and shallow ones, were examined in order to find appropriate number of hidden layers and neurons inside them. The experiments proved the superiority of deep ANNs over shallow ones.

we summarize explained papers in terms of initial feature set, feature extraction algorithm and prediction method. As it can be seen there is a tendency toward deep learning models in recent publications, due to the capability of these algorithms in automatic feature extraction from raw data.

978-1-7281-5050-5/19/$31.00 ©2019 IEEE
DOI 10.1109/ICVRIS.2019.00050

Based on this, this paper proposes a CNN-based in-depth learning method to predict stock prices, which can be used as a reference for relevant researchers.

III. CONVOLUTIONAL NEURAL NETWORK(CNN)

Convolutional neural network is a type of feedforward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex, whose individual neurons are arranged in such a way that they respond to overlapping regions tiling the visual field. There are several different theories about how to precisely define such a model, but all of the various implementations can be loosely described as involving the following process:

① convolve several small filters on the input image;

② subsample this space of filter activations;

③ repeat steps 1 and 2 until your left with sufficiently high-level features;

④ Use a standard a standard MLP to solve a particular task, using the results features as input.

The main advantage of convolutional neural networks is that we use convolutional and down sampling layers as learnable feature extractor, which allows us to feed neural network without sophisticated preprocessing, so that useful features will be learnt during the training. In our case, instead of considering the standard framework of 2D images, we apply convolutional network to 1D data, namely to (financial) time series. CNN has many layers which could be categorized into input layer, convolutional layers, pooling layers, fully connected layers and output layer.

A. *Convolutional layer*

The convolutional layer is supposed to do the convolution operation on the data. In fact, input could be considered as a function, filter applied to that is another function and convolution operation is an algorithm used to measure changes caused by applying filter on the input. Size of a filter shows the coverage of that filter. Each filter utilizes a shared set of weights to perform the convolutional operation. Weights are updated during the process of training. Let posit input of layer 1-1 is an N×N matrix and F×F convolutional filters are used. Then, input of layer 1 is calculated according to Equation (1).

$$V_{i,j}^{l} = \delta(\Sigma_{k=0}^{F-1}\Sigma_{m=0}^{F-1} w_{k,m} v_{i+k,j+m}^{l-1}) \tag{1}$$

In the Equation (1), $v_{i,j}^{l}$ is the value at row i column j of layer l, $w_{k,m}$ is the weight at row k, column m of filter and f(x) is the activation function.

$$f(x) = \max(0; x) \tag{2}$$

B. *Pooling layer*

Pooling layer is responsible for subsampling the data. This operation, not only reduces the computational cost of the learning process, but also it is a way for handling the overfitting problem in CNN. Overfitting is a situation that arises when a trained model makes too fit to the training data, such that it cannot generalize to the future unseen data. It has a connection to the number of parameters that are learned and the amount of data that the prediction model is learned from. Deep models, including CNNs, usually have many parameters so they are prone to overfitting more than shallow models. Some methods have been suggested to avoid overfitting. Using pooling layers in CNNs can help to reduce the risk of overfitting. All the values inside a pooling window are converted to only one value. This transformation reduces the size of the input of the following layers, and hence, reduces the number of the parameters that must be learned by the model, that in turn, lowers the risk of overfitting. Max pooling is the most common type of pooling in which the maximum value in a certain window is chosen.

C. *Fully connected layer*

At the final layer of a CNN, there is an MLP network which is called its fully connected layer. It is responsible for converting extracted features in the previous layers to the final output. The relation between two successive layers is defined by Equation (3).

$$v_i^j = \delta(\Sigma_k v_k^{j-1} w_{k,i}^{j-1}) \tag{3}$$

In Equation (3), v_i^j is the value of neuron i at the layer j, δ is activation function and weight of connection between neuron k from layer j-1and neuron i from layer j are shown by $w_{k,i}^{j-1}$.

D. *Dropout*

In addition to pooling, we have also used another technique called dropout that was first developed for training deep neural networks. The idea behind the dropout technique is to avoid the model from learning too much from the training data. So, in each learning cycle during the training phase, each neuron has a chance equal to some dropout rate, to not be trained in that cycle. This avoids the model from being too flexible, and so, helps the learning algorithm to converge to a model that is not too much fit to the training data, and instead, can be generalized well for prediction the unlabeled future data.

IV. EXPERIMENTAL DESIGN

The experimental design includes data processing, program design, results and analysis. Experimental platform: CPU intel i7-2630QM, python3.7.3

A. *Data Preparation and data processing*

1) The data Set

The dataset available to learn and evaluate the performance of the implemented system includes daily historical prices available for the stocks currently listed on the Stock Exchange of Thailand (SET). The Stock Exchange of Thailand Index is a weighted index of a representative selection of the stocks listed on the Stock Exchange of Thailand. All stocks listed on the index are easily transferable which makes our model easier to validate as we can assume that selected stocks can be bought and sold at any time. Stock Exchange of Thailand (SET) is Thailand's only stock market. Trading is carried out through computer networks, and trading starts at 08:00 and ends at 17.00 Monday through Friday. The Exchange serves as a center for the trading of listed securities, a system provider for facilitating securities trading, business undertaking or clearing house, securities depository, securities registrar, and other similar activities, as well as any other business approved by the Securities and Exchange Commission (SEC). The stocks used can be used to download the data from https://www.set.or.th. The data used is all data available from 2003/01/01 to 2019/03/01. The data is always separated in a training set and test set where the separation point is configurable and will be noted in the report when needed. Although our focus in this thesis is on stocks listed on the Stock Exchange of Thailand, the developed model can just as easily be used for any other stock exchange where a sufficient amount of daily historical prices is available. In future work, we will examine different strategic selections of stocks. Possible strategies include only looking into one industry at once or including the largest representative from each industry. However, in this work, we wanted to include a random selection to then examine the information flow and see whether this matches the intuition one might have about interdependence of certain stocks. We choose three companies listed on SET50 which are CPALL, PTT and BBL to do this research.

2) Data Preparation

The stock prices is a time series of length N, defined as $P_0, P_1, \cdots, P_{N-1}$ in which P_i is the close price on day $i, 0 \le i < N$ .A sliding window approach is used in this paper. A sliding window of a fixed size w that we refer to this as input size and every time we move the window to the right by size w shown in Figure 1, so that there is no overlap between data in all the sliding windows. We use content in one sliding windows to make prediction for the next, while there is no overlap between two consecutive windows.

Figure 1. The SET50 prices in time

3) Data processing

In what follows we focus our attention on the SET50 index, exploiting related data from 2003 to 2019 (16706 data points). Usually stock market data looks like on Figure 2, which reports the close prices for every day in the aforementioned time interval.

Our goal is to predict the movements of the three listed in SET50 index, exploiting some information from previous data. Suppose we are going to predict if the close price of the next day is larger or smaller than the previous one based on the last days of observations. Appropriate time window and prediction horizon should be chosen during hyper parameter optimization stage. First, we have to scale or normalize our input data. The input vectors of the training data are normalized in such a way that all the features have zero-mean, and unitary variance. Usually data are scaled to be in the range [-1;1], or [0;1]. In the neural network approaches such renormalization strongly depends on the activation function of every neuron. A different approach, is about considering not raw open or close ticks, but calculating return during the day, resp. during the minute, and then using this data as the training set.

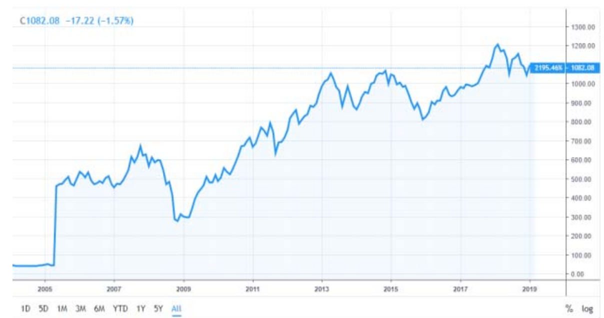


Figure 2. SET50 index data from 2005 to 2019

In our research we use return prices as more representative data with normalization for stock price movement forecasting problem. In particular, we normalized our time series to have zero mean and unit variance.

B. Experiment on Convolutional neural network architecture

We use the Convolutional neural network architecture (CNN), as a sequential combination of 1D convolutional and max-pooling layers, choosing hyper parameters as follows:

① Number of filters = 64;

②Pool length = 2;

③ Subsample length = 1;

④ Activation function ReLU.

We provide experiment results with 1 and 2 hidden convolutional layers. The architecture of this model is shown in Figure 3.

C. Results and Analysis

The experiment was done for CNN models by using the three companies listed on SET50 index. The predicted results and actual stock prices of the three companies are shown in Figure 4.

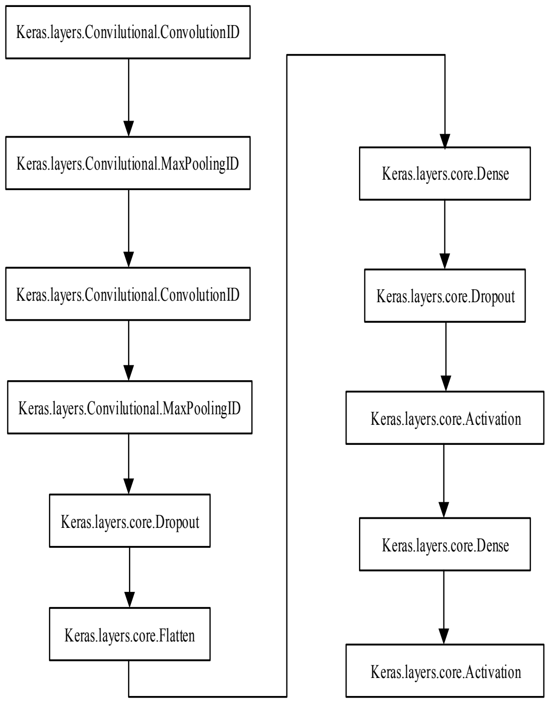
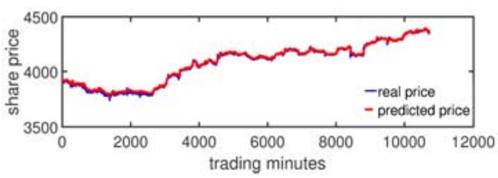
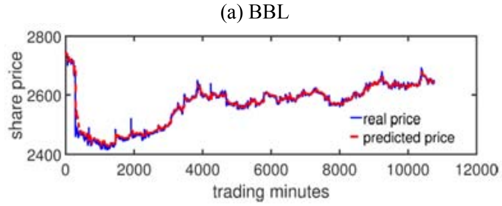


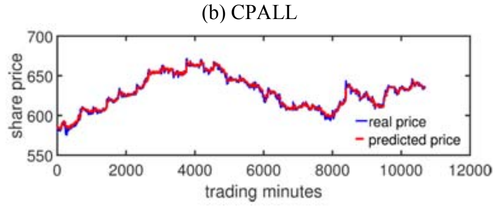
Figure 3. Architecture of ANN model



(a) BBL



(b) CPALL



(c) PTT

Figure 4. Predicted results vs actual stock price

As can be seen from Figure 4, the in-depth learning model based on CNN is very accurate in forecasting the stock of three different companies. This is due to the reason that CNN does not depend on any previous information for prediction. It uses only the current window for prediction Which enables the model to understand the dynamical changes and patterns occurring in the current window.

V. CONCLUSIONS

The CNN stock price forecasting method in this paper has high accuracy and high application value. In similar analysis and prediction, we should use CNN as the basic algorithm to obtain higher accurate results. Although our focus in this thesis is on stocks listed on the Stock Exchange of Thailand, the developed model can just as easily be used for any other stock exchange where a sufficient amount of daily historical prices is available. In the future work, we will try to apply this model to other different stocks, and try to combine different algorithms to build a more accurate prediction model.

REFERENCES

[1] Ma W, Wang Y, Dong N. Study on stock price prediction based on BP Neural Network[C]// IEEE International Conference on Emergency Management & Management Sciences. 2010.

[2] Zhao Q Y, Zhao X, Duan F. Prediction Model of Stock Prices Based on Correlative Analysis and Neural Networks[C]// International Conference on Information & Computing Science. 2009.

[3] Fei Y, Xiao J, Ying C, et al. The Combined Stock Price Prediction Model Based on BP Neural Network and Grey Theory[M]// Computer, Informatics, Cybernetics and Applications. 2012.

[4] Lv X Y, Sun S L, Liu H. Stock Price Prediction Model Based on BA Neural Network and its Applications[J]. Advanced Materials Research, 2014, 989-994(9):1646-1651.

[5] Krollner, B., Vanstone, B., & Finnie, G. (2010). Financial time series forecasting with machine learning techniques: A survey.

[6] Kara Y, Boyacioglu M A,etc. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange[J]. Expert Systems with Applications, 2011, 38(5):5311-5319.

[7] Qiu M, Song Y, Akagi F. Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market[J]. Chaos, Solitons & Fractals, 2016, 85:1-7.

[8] Qiu M, Song Y. Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model[J]. PLoS ONE, 2016, 11(5).

[9] Moghaddam A H , Moghaddam M H , Esfandyari M . Stock market index prediction using artificial neural network:[J]. Journal of Economics Finance & Administrative Science, 2016.