Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.199

IJCSMC, Vol. 9, Issue. 3, March 2020, pg.106 – 111

Pattern Recognition in Stock Market

D. Sorna Shanthi[*1]; T. Aarthi[*2]; A.K. Bhuvanesh[*3]; RA. Chooriya Prabha[*4]

[1]Associate Professor, Rajalakshmi Engineering College, Computer Science and Engineering, Chennai, India
[2]Student, Rajalakshmi Engineering College, Computer Science and Engineering, Chennai, India
[3]Student, Rajalakshmi Engineering College, Computer Science and Engineering, Chennai, India
[4]Student, Rajalakshmi Engineering College, Computer Science and Engineering, Chennai, India
Email: [1*] sornashanthi.d@rajalakshmi.edu.in, [2*] aarthi.thina@gmail.com, [3*] akbhuvanesh123@gmail.com, [4*] chooriyaprabhara@gmail.com

Abstract— Financial markets have a significant role in the development of contemporary society. They allow the deployment of economic resources. The increase/decrease of stock prices affects the stock market movement. In this study, we focus on predicting stock prices by machine learning model. This paper presents feasible method to predict stock movement with more accuracy. Identifying patterns in high dimensional data are difficult because data cannot be visualized easily. High dimensional data can into various machine learning methods to classify and predict data from the specified datasets. The frequent fluctuation in Stock Market Price makes the prediction process difficult. Correct prediction of Stocks can yield more profit. This paper presents the tactics to spot the pattern hidden within historical data and provides the means to predict the long term values.

Keywords- Stock prediction, Machine Learning, Pattern Recognition

I. INTRODUCTION

The Stock Market is the platform where buyers and sellers come together to trade. A share in general represents claim of ownership on a business for a particular individual or group of people. The prices of stocks are governed by the principles of demand and supply, and the ultimate goal of buying shares is to make money by buying stocks in companies whose anticipated value (i.e., share price) is expected to rise. With the rises in demand, prices of the stock tend to grow and with the decrease in demand, they tend to fall. The process of determining the future value of the stock market is known as stock market prediction. The prediction expected to be robust and accurate. A correct prediction of stock can yield huge profit for seller and the broker. Frequently it is brought out that prediction is chaotic rather than random, which means it can be predicted by carefully analyzing history of the respected stock market.

The significant part of machine learning is the dataset used. The dataset should be as concrete as possible because a little change in the data can bring massive changes in the outcome.

Patterns are recurring series found in Open Close High Low candlestick charts which traders have regularly used as buy and sell alert. Although to a limited extent, these studies have found some correlation between patterns and future trends. The correlations values were found to be between 50 and 60%, 50% being no better than random. Technical Analysis is the field where many traders are using stock chart patterns, sometimes combined with other trading techniques, to take their trading decisions. Our

goal is to automate the detection of these patterns and to evaluate how a Machine Learning based recognizer behaves compared to manual predictions. Automation would simplify the process of finding patterns which vary in scale and length.

From pragmatic to Machine Learning, several types of detection algorithms exist. The solutions differ in efficiency, re-usability and speed.

The solution we propose to study is based on Machine Learning. There exists several ways to detect patterns. Here we detect a pattern based on the percent difference between the prices in the stock and minimum percent similarity considered. These patterns provide us a way to predict the stock price (outcome of a pattern).

II. RELATED WORKS

The application of machine learning techniques to stock market prediction is being undertaken thoroughly throughout the world. These techniques are proving to be much more accurate and faster as compared to contemporary prediction techniques. Here, we present a detailed summary of the related papers done by significant researchers.

A. A Stock Pattern Recognition Algorithm Based on Neural Networks

The pattern recognition for stock price is based on the artificial neural networks. It effectively learns the characteristics of the patterns and recognizes the patterns accurately. It uses the recurrent neural network algorithm to recognize the triangular patterns. The stock investment decision making depends on two processes, continuation and reversal patterns. The continuation patterns depicts the stock prices changes with respective to current movement trend, whereas reversal patterns moves in opposite trend. Searching for patterns from the original stock prices becomes a time consuming processes. The stock price time series is reduced to multi line segments, to extract the features and to recognize the pattern in the stock prices feedforward neural networks. Therefore this methodology is not applicable for real time pattern recognition.

B. Stock Chart Pattern recognition with Deep Learning

Hard coded algorithm is used for recognition of common charts patterns in a historical stock data. It automates the detection of these patterns and to evaluate how a deep learning based recognizer to compared hard corded algorithm. Then deep learning models are used to train the data sets. Those patterns are not enough to predict the trends. Human examining candlestick charts are required for extending the number of pattern recognition. The pattern recognition varies with size and length. In Long Short Term Memory recurrent network, the past states are required to make future predictions. A single pattern of time series is obtained as an output for stock chart pattern recognition. It also compares hard coded algorithm with several models.

C. Survey of Stock Market Prediction Using Machine Learning Approach

Stock market plays a important role in business which is complicated to predict the future prices. For prediction regression analysis is used. In the field of machine learning these regression analysis is widely used for prediction and forecasting. It also understands the relationship between independent and dependent variables. The performance of regression analysis depends on the form of data that are generated, and it relates to the variables. Many type of regression models are being used to predict the future prices.

D. Stock Trend Prediction Using Regression Analysis – A Data Mining Approach

Multiple regression technique is used in this paper which is the extension of linear regression. In the prediction process it almost uses values of two or more variables. The value which is predicted is called dependent variables. Those training data set is used to train a model and calculate the unknown coefficients of the auto regression values. Regression analysis model is used as data mining technique and develops a time series data to forecast periodical changes. The demand for data analysis will be increasing. Back propagation will reduce the errors and improve accuracy.

E. Pattern Recognition in Stock Data

The goal of this paper is being able to apply specific expert knowledge in a field and have a computer program learn to identify the patterns within the stock data. The only consideration here is the stock data is high-dimensional which makes it a tedious task to find patterns among them. Although, many different machine learning methods are able to fit this high dimensional data in order to predict and classify future data but there is typically a large expense on having the machine learn the fit for a certain part of the data set. This papers uses a form of supervised learning and is different than other methods because the pattern definitions will be defined by geometry that is based on how the patterns look to the expert which extends the idea behind the paper in a geometric way of defining different patterns in data that is invariant under size and rotation so it is not so dependent on the input datasets. The predictions are made from the pattern is found within stock market data using Gaussian Process. Therefore, the algorithm will be learning geometrical shapes rather than numbers.

III. METHODOLOGY

In this proposed system, we focus on predicting the closing price for Apple Inc. stock with the help of Machine Learning Algorithms. To make a future prediction we train the machine from the various data points from the past to make a future prediction. We took data from the previous year stocks to identify the pattern and making use of available patterns to predict future patterns

A. Data Pre-processing

In Machine Learning, data preprocessing is an important step. The ability of our model is to learn the quality of data and the useful information that are derived from it. Therefore before feeding the data into our model we preprocess them. Data is preprocessed by normalization. The main idea of normalization is to change the numeric values in in the columns of the dataset to a common scale, without distorting differences in the ranges of values.

B. Pattern Recognition

This module is the most important module which is a composition of many computations. It contains many decision making parameters such as deciding whether a pattern is similar to the one already in the existing in the pattern list or not and or not and computing the percent changes between two patterns, etc. It also prints out all the patterns in the list even in at least one patterns is similar to the predicted patterns. The predicted outcome is calculated for each one of it and then the average predicted outcome is computed based on which rise or drop in the prediction is decided.

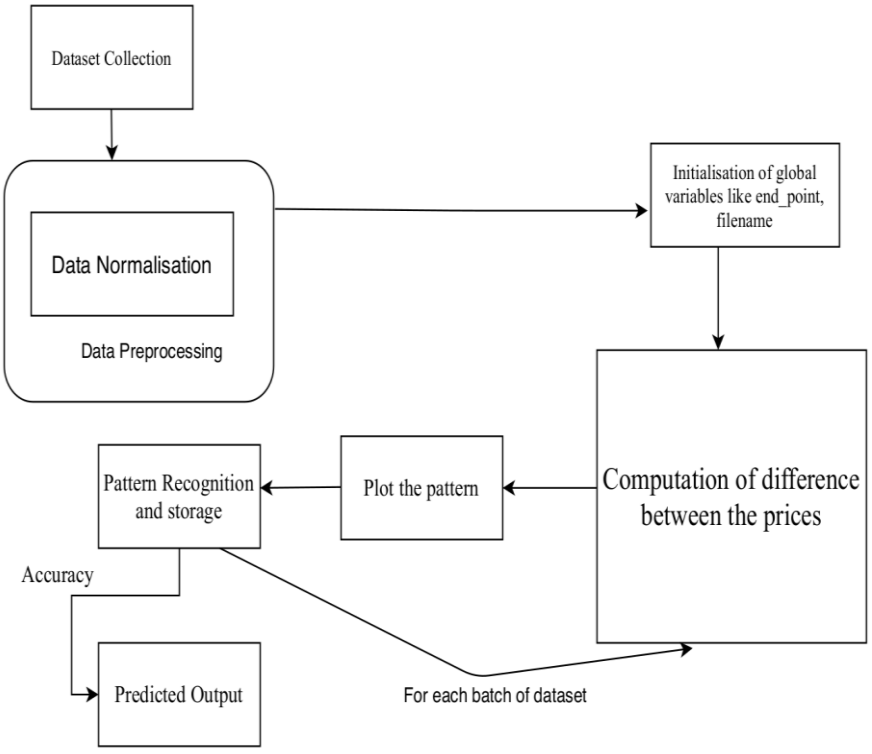
C. Back Testing

Backtesting is widely used in quantitative finance. It is natively sequential and as close to real-life situations as possible. Any performance measure, from standard prediction errors to complex cost functions a can be backtests accommodate. This is very helpful for directing your actual industrial goal directly - for example, maximizing/minimizing the gains/losses resulting from good/bad inventory predictions, instead of minimizing an inventory prediction error. Since every data point is already (and consecutively) part of the testing and training sets, backtests don't use validation sets.

D. System Design

The system architecture consists of following modules namely,

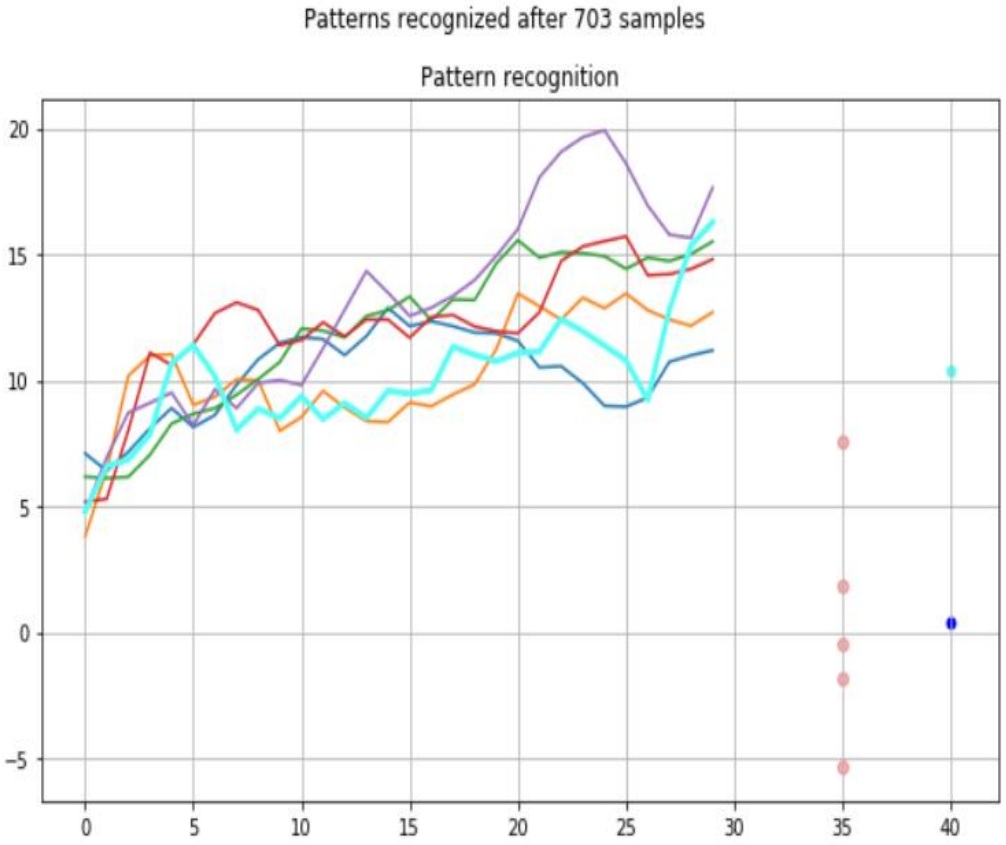
1. Data Collection
2. Pattern Storage
3. Current Pattern Prediction
4. Pattern Recognition
5. Backtest Accuracy Computation



IV. IMPLEMENTATION AND RESULTS

The model first collects the data sets i.e., Apple stock data from Yahoo finance for a required time period. Once the data is collected it is then normalized and passed to the all upcoming modules with a fixed batch size. First the batch is passed to pattern storage module. The pattern storage module does the computations necessary to find the pattern array values to plot in the graph and get the average outcome and future outcome from it. The future outcome is then appended to the performance array. Then current pattern is predicted for every last dots_for_patterns entries of data. After this the final and major step is performed which is the pattern recognition. Patterns are recognized by the help of algorithms used in Machine Learning. Recognizing patterns is the process of classifying the data based on the model that is created by training data, which then detects patterns and characteristics from the patterns. In this paper the new patterns are recognized by comparing the patterns with the already existing patterns list and using a threshold for checking on the similarities between them. Based on the similarities of values between them it is either as either new pattern or an existing pattern.

If a pattern satisfies the similarity value, remember that a patterns was found and append that pattern to the list of patterns to plot. It at least one similar pattern was found then print all the pattern that are in the array of patterns to be plotted. The color of the prediction is determined based on the comparison of the prediction pattern with the value of the pattern. Also the predicted average decides whether there is a rise or fall in the predictions.




```
Backtested accuracy is 87.5 % after  774 samples
[ ] Backtested accuracy is 87.5 % after  775 samples
Backtested accuracy is 87.5 % after  776 samples
Backtested accuracy is 87.5 % after  777 samples
Backtested accuracy is 87.5 % after  778 samples
Backtested accuracy is 87.5 % after  779 samples
Backtested accuracy is 87.5 % after  780 samples
Backtested accuracy is 87.5 % after  781 samples
Backtested accuracy is 87.5 % after  782 samples
Backtested accuracy is 87.5 % after  783 samples
Backtested accuracy is 87.5 % after  784 samples
Backtested accuracy is 87.5 % after  785 samples
Backtested accuracy is 87.5 % after  786 samples
Backtested accuracy is 87.5 % after  787 samples
Backtested accuracy is 87.5 % after  788 samples
Backtested accuracy is 87.5 % after  789 samples
[-1.0]
-1.0
Drop predicted
16.174471699768393
2.2249116004910614
Backtested accuracy is 88.23529411764706 % after  790 samples
[-1.0, -1.0, -1.0, -1.0]
-1.0
Drop predicted
15.163113798934747
2.04572003945398
Backtested accuracy is 88.88888888888889 % after  791 samples
Backtested accuracy is 88.88888888888889 % after  792 samples
Backtested accuracy is 88.88888888888889 % after  793 samples
Backtested accuracy is 88.88888888888889 % after  794 samples
Backtested accuracy is 88.88888888888889 % after  795 samples
Backtested accuracy is 88.88888888888889 % after  796 samples
Backtested accuracy is 88.88888888888889 % after  797 samples
Backtested accuracy is 88.88888888888889 % after  798 samples
[1.0]
```

V. CONCLUSION & FUTURE ENHANCEMENT

The proposed model identifies the pattern hidden in the historical Stock Market data and predicts the future scope. Due to the uncertainties in the input stock data, the accuracy in obtained in this model is nearly 80%. The accuracy of the stock price prediction is affected by various factors. Generally the stock prices are affected by news feeds, supply demand, historical data etc. Here we have considered only one of the main criteria which affects the stock prices i.e., historical stock prices to predict the future outcome. Hence a slight decrease in accuracy is encountered.

This model can be improved by considering several other factors such as news which can be categorized using sentiment analysis to enhance the accuracy of the prediction. It can also use Generative Adversarial Networks (GAN) along with Long Short Term Memory (LSTM), a Recurrent Neural Network (RNN) to build the prediction model and use Convolutional Neural Network (CNN) as a discriminator. We can also use Bayesian Optimization and Reinforcement Learning to produce better results as they decide when and how to change GAN's hyperparameters.

REFERENCES

[1]. Xinyu Guo, Xun Liang, Xiang Li "A Stock Pattern Recognition Algorithm Based on Neural Networks" Third International Conference on Natural Computation (ICNC)2007.
[2]. Marc Velay and Fabrice Daniel, "Stock Chart Pattern Recognition with Deep Learning".
[3]. Ashish sharma, Dinesh Bhuriya, Upendra Singh, "Survey of Stock Market Prediction Using Machine Learning Approach" International Conference on Electronics, Communication and Aerospace Technology 2017.
[4]. Bhagyashree Nigade, Aishwarya Pawar, Varsha Bavane, Siddhant Navaratna, "Stock Trend Prediction Using Regression Analysis – A Data Mining Approach" International Journal of Innovative Research in Computer and Communication Engineering 2017.
[5]. Min Wen, Ping Li, Yan Chen "Stock Market Trend Prediction Using High-Order Information of Time Series".
[6]. Piotr Czekalski, Michał Niezabitowski and Rafał Stybliński "ANN for FOREX Forecasting and Trading" International Conference on Control Systems and Science 2015.
[7]. Kathryn Dover "Pattern Recognition in Stock Data", Theses in Harvey Mudd College 2017.
[8]. Tao Xing, Yuan Sun, Qian Wang, Guo Yu "The Analysis and Prediction of Stock Price" IEEE International Conference on Granular Computing 2013.
[9]. Batyrkhan Omarov, Young Im Cho "Machine Learning based Pattern Recognition and Classification Framework Development" 17th International Conference on Control, Automation and Systems (ICCAS 2017).

