

STOCK MARKET ANALYSIS USING SUPERVISED MACHINE LEARNING

ABDUL MALIK AB, KABIL SANJAY S, INDIRAKUMAR S, VETRIVEL K
COMPUTER SCIENCE AND ENGINEERING
MAHENDRA INSTITUTE OF ENGINEERING AND TECHNOLOGY

Abstract- In Stock Market is the financial epitome of financial business and trading since it came into existence it has shown the impact of hits low and similarly when it is high. The stock market crash in 2008 showed the world that the business hit the low when the Dow Jones Industrial Average fell 777.68%. Several machine learning algorithms have shown that these stock prices can be predicted and these algorithms can be implemented using the approach of supervised learning. In Supervised Learning, we have test data using this we train the models. Although the results obtained after training the model may differ from the actual but it has been observed that in many cases accuracy is satisfactory. In this paper, the first task is to use web scrapping to collect datasets from stock data. Then we plot the data on the graph, from the graph we can analyse the stock prices going high or low. After this, we will predict stock prices using Linear Regression.

Index Terms-Stock Market Analysis, Stock prediction, Machine Learning, Deep Learning, Data Set, Predict the future moment of stock value.

I. INTRODUCTION

A correct prediction of stocks can lead to huge profits for the 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 the history of respective stock market. Machine learning is an efficient way to represent such processes. It predicts a market value close to the tangible value, thereby increasing the accuracy. Introduction of machine learning to the area of stock prediction has appealed to many researches because of its efficient and accurate measurements. The vital part of machine learning is the dataset used. The dataset should be as concrete as possible because a little change in the data can perpetuate massive changes in the outcome. In this project, supervised machine learning is employed on a dataset obtained from Yahoo Finance. This dataset comprises of following five variables: open, close, low, high and volume. Open, close,

low and high are different bid prices for the stock at separate times with nearly direct names. The volume is the number of shares that passed from one owner to another during the time period. The model is then tested on the test data. Regression and LSTM models are engaged for this conjecture separately. Regression involves minimizing error and LSTM contributes to remembering the data and results for the long run.

II. SYSTEMATIC REVIEW METHODOLOGY

Research method that has been adopted in this paper is Systematic Literature Review (SLR). A systematic review can be define as a research technique and process for identifying and critically appraising relevant research, as well as for collecting and analyzing data from said research[8]. Execution of a systematic review can be grouped into three main stages: planning, conducting the review and reporting the review[9].

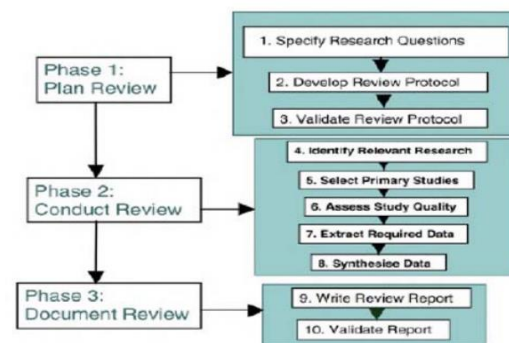


Fig. 1. Systematic Literature Review Process[9].

Literature review from all the journal publications and conference articles gathered and used to answer the research questions mentioned as follows: R1: What are the Supervised Learning Algorithms that are commonly used in Stock Market Prediction? R2: How Supervised Machine Learning techniques are applied in Stock Market Prediction? R3: Which technique is considered the best or frequently used for Stock Market Prediction among the Supervised Machine Learning techniques?

R4: What are the directions of future research on Stock

A. Supervised Machine Learning Techniques in Stock Market Prediction

Stock-market prediction using machine-learning technique which aims at developing effective and efficient models that can provide a better and higher rate of prediction accuracy [40]. Numerous classification and regression models have been used in stock market predictions for many years. Several supervised machine learning techniques applied in stock market prediction yielded a better outcome. For more than a decade, a number of supervised machine models and techniques has been proposed or implemented for stock market predictions. Plethora of researchers and technology practitioners tried their level best in order to come up with an algorithm that will give the best stock market prediction.

The majority of work published on supervised machine learning techniques used in stock market predictions are categorized as follows: are not included in the study:

- The comparison of classification algorithms
- The comparison of regression algorithms
- The comparison of classification and regression algorithms

B. Systemic Literature Review Approach

We collected journals and conference papers via electronic sources using credible and relevant databases such as IEEE Xplore, ScienceDirect, Google Scholar, Semantic Scholar, Open Access and Springer. The search keywords used to search publications are “ Stock Market Predictions using Machine Learning” , “ Stock Market Prediction using Supervised Machine Learning Techniques/Algorithms” and “ Stock Market Prediction using Classification/ Regression Algorithms” .

The study has address issues of Supervised Machine Learning and Stock Market Predictions therefore they were excluded. The articles has reduced the number from 84 to 38.

The following criteria has been used in eliminating papers that are not included in the study:

- Exclude if the paper do not emphasis on Machine Learning Predictions.
- Exclude if the paper do not discuss about Supervised Machine Learning in Stock Market Predictions or Stock Market Predictions.

III. FINDING AND DISCUSSION

The results of the first research question are covered in part 4.1, while the other research questions are addressed in Sections III.A, III.B, and III.C, respectively. In this part, we have offered detailed replies to the questions stated in

Section II. The selected case studies provide in-depth responses to the study topics. The Application of Supervised Machine Learning (SML) to the Forecasting of Stock Prices The most popular supervised machine learning methods for stock market forecasting are shown in Table I. Depending on the study, researchers may have employed one or more than one.

TABLE I. SUPERVISED MACHINE LEARNING TECHNIQUES

Journal/Conference Paper	Stock Prediction Models	Algorithm Type
Robert P. Shumaker et al [10]	SVM	Regression
Han Lock Slew et al [11]	Linear Regression (LR)	Regression
Khalid Alkhatib et al [12]	KNN and None Linear Regression	Classification and Regression
Carol Hargreaves et al [13]	Logistic Regression	Regression
Usman Hegazy et al [2]	LS-SVM	Classification
Abubakar S. Magaji et al [14]	SVM-SMO	Classification
Saahil Madge et al [1]	SVM	Classification
Shashaank D.S et al [15]	Hybrid future selection	Classification
Hakob Grigorian et al [4]	SVM	Classification
Khan, W. Ghazanfar et al [16]	SVM, KNN and Naïve Bayes classifiers	Classification
Makram Zaidi et al [17]	ANN	Regression
Aparna Nayak et al [18]	Decision Tree, Logistic Regression and SVM	Classification and Regression
Mustansar Ali Ghazanfar et al [5]	SVM	Classification
Mr. Pramod Mali et al [19]	LR and Multiple Regression	Regression
Hakan Gunduz et al [20]	SVM and ANN	Classification
D. A Puspitasari et al [21]	SVM and KNN	Classification

Table II. AI Algorithms with Supervision

Supervised Machine Learning Algorithms	Case Studies
Classification	50%
Regression	24%
Comparison between Classification and Regression	24%

Classification was shown to be the most used approach for stock market forecasting in Table II. Using the aforementioned research questions to provide a clear response to the first research question. Publications organized by method(s) used over a decade are shown in the following chart. While the chart remains static as a result of the articles that are available for this research. Articles published in 2013 make use of classification, regression, and comparisons between the two, as seen in Figure 2, while journals and conference papers published in 2006 and 2012 used regression techniques. Both 2014 and 2015 saw the writers publishing works that made use of categorization techniques. The number of articles comparing the two algorithms increased from one per year in 2016 to four in 2019 and three in 2020.

B. Using SML for Stock Market Forecasting

It all comes down to the performance of the supervised machine learning algorithms used for stock market prediction. Figure 2 and Table III provided a summary of the responses to the second research question of this paper. The second question sought to clarify the most common supervised machine learning algorithm used for stock market prediction from the 38 case studies, specifically in the classification and regression categories.

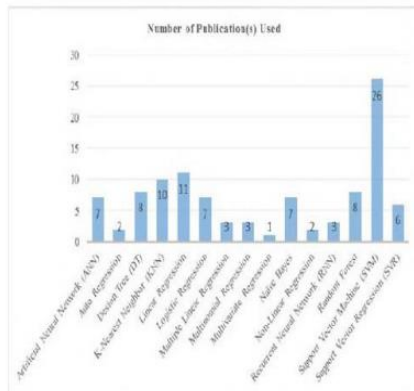


Fig. 3. SML Techniques used in 38 case studies



C. Where to take stock market prediction studies from here

Several computational models grounded in soft-computing and machine learning paradigms have been used in stock market analysis, prediction, and trading in order to circumvent the obstacles inherent in stock market analysis [40]. The use of methods such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Support Vector Regression (SVR), Linear Regression, and Artificial Neural Network (ANN) has led to more accurate stock market forecasts. An effective method or methods are required to circumvent the obstacles encountered while using the current methods. Regardless of the resilience, high accuracy, and capacity to deal with noisy and missing data. One of the challenges of utilizing ANN for stock market prediction is over fitting. Outliers and parameter selection may also affect SVM. Even though training RNN

models is challenging, the results of its predictions have been excellent. The free parameters that the user defines also have an impact on SVR. Research into supervised machine learning for stock market prediction should focus on fixing the aforementioned issues with current methods.

IV. CONCLUSION

This paper was an attempt to determine the future prices of the stocks of a company with greater accuracy and reliability using machine learning techniques. The primary contribution of the researchers being the application of the novel LSTM Model as a means of determining the stock prices. Both the techniques have shown an improvement in the accuracy of predictions, thereby yielding positive results with the LSTM model proving to be more efficient. The results are quite promising and has led to the conclusion that it is possible to predict stock market with more accuracy and efficiency using machine learning techniques.

REFERENCES

- M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.
- K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.
- H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
- M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.
- H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012

International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1- 5.

- K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.
- S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modelling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2787-2790.
- T. Gao, Y. Chai and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.
- K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156.

