

A Technical Survey on Stock Price Prediction in Machine Learning

N.Anbarasu¹ and Dr.S.Karuppusamy²

¹PG Scholar and ²Associate Professor, Department of Computer Science and Engineering,

Nandha Engineering College (Autonomous), Erode, Tamilnadu

ABSTRACT — Machine learning is that the study of computer algorithms that improve automatically through experience and by the utilization of knowledge. It is seen as a neighborhood of Artificial Intelligence. Machine learning algorithms build a model supported sample data, referred to as training data, so as to form predictions without being explicitly programmed to try to so. A subset of machine learning is closely associated with computational statistics, which focuses on making predictions using computers. In this paper, we present a review of various stock price prediction algorithms in Machine Learning. Resources are allocated in better way to improve the efficiency of prediction.

Keywords -Machine learning, model stacking, sentiment analysis, stock movement direction prediction, textual features extraction, tweets mining.

1. INTRODUCTION

Stock price prediction has been at focus for years since it can yield significant profits. Fundamental and technical analyses were the primary two methods to forecast stock prices. The Raise and fall of economy of the country depends on the performance of the Stock market. There is a misunderstanding in trading shares is like gambling. For Economic growth of every Country Stock market plays a vital role. The estimation of the stock market index is of clear interest to various stakeholders in the market. A number of existing studies have shown that the events reported in news are important signals which will drive market fluctuations.

1.1 APPROACHES

There are three types of approaches present in machine learning which are given as follows:

Supervised learning: Supervised learning algorithms used to create a mathematical model of a set of data that consist of both the inputs and the required outputs. The result is known as training data, and which contains a set of training examples. Each training set has one or more inputs and the required output, also known as a supervisory signal.

Unsupervised learning: It picks a set of data that consist of only inputs, and find structure in the data, like combining or grouping of data points. Instead of reacting to feedback, unsupervised learning algorithms spot similarities in the data and respond to it based on the presence or absence of such similarities in each new set of data.

Reinforcement learning: It is a part of machine learning distributed with how software agents ought to take actions in an environment so as to increase some notion of total reward.



1.2 MACHINE LEARNING MODELS

Artificial Neural Network An ANN is design supported a gaggle of fastens nodes called artificial neurons, approximately organized like the neurons which are in abiological brain. Datapasses from the first layer to the last layer, possibly after crossing the layers many times.

Decision Tree It uses a decision tree as an anticipating model to go from study about an item to closure about the item's target value. It is one of the anticipating modeling approaches used in statistics, data mining, and machine learning.

Support-vector machines (SVMs) Also known as support-vector networks are a set of correlated supervised learning methods used for classify and lapse. Specified a group of sample data, each noted as belonging to at least one of two types, an SVM training algorithm builds a kind that predicts anyway a replacement example falls into one class or the other.

1.3 STOCK PREDICTION

TensorFlow is more popular in machine learning, but it's a learning curve. Scikit-learn and PyTorch also are popular tools for machine learning and both support Python programing language . Keras.io and TensorFlow are better for neural networks.

From the perspective of external information in economic markets, they use economic news to study the importance of stock market uncertainty. Since there are too many factors such as people opinions, general financial impacts, or impacts of political events, that have direct or indirect impacts on the future of financial time series. With the help of the machine learning tools the future prediction implemented from the time series data.

1.4 STOCK PREDICTION

This is a process that takes financial time series data and let the machine to predict the future of particular organization in finance market which helps to make decision in finance market. The predictor works along with the time series data then analysis the result approximately and gives it as the result.

2.LITERATURE SURVEY

Osman Hegazy, Omar S. Soliman and Mustafa Abdul Salam[1], analyzed that Stock market prediction is the way to determine the future price of an organization stock or other financial instrument traded on a financial market. The successful prediction of a stock's future price will increase investor's profits. The proposed LS-SVM-PSO model convergence to the global minimum. The proposed algorithm achieves the lowest error value followed by single LS-SVM.

Achyutasai Nikhil[2], presented the study which aims at Stock market price prediction for short time windows appears to be a random process. To get awareness about risk in share market Trading Company's used Machine Learning Techniques to predict the future stock market based on time series data. Support Vector Machine Algorithm is used to predict the share value based on the previous data because to estimate or to know the price of the stock will be lower or higher than the given date. In this paper, Stock price prediction analysis is studied using supervised machine learning algorithms for company data sets and calculated the loss percentage. The results shown that for large dataset LSTM performs better compared to other algorithms.

Xi Zhang, (Member, Ieee), SiyuQu, Jieyun Huang, Binxing Fang, And Philip Yu[3],proposes a method to improve the prediction for stock market composite index movements, we exploit the consistencies among different data, and develop a multi-source multiple instance model which willproductively combine events, sentiments, also the quantitative data into a complete framework. In this work aim to find out a predictive model for describing the fluctuations with in the stock exchange index by utilizing various sources of knowledge, involving the historical quantitative data, the social media and Web news. In addition to events, investors emotions even have great impacts on the stock exchange. First use the sentiment analyzer to obtain the collective sentiments from social media, and extract effective event representations from the online news. Then the extracted sentiments, movementsalso the stock quantitative data are fed into the M-MI model. In this paper, a Multi-source Multiple Instance model is proposed which will predict the stock exchange movement and identify the principal of the information simultaneously.

Min Wen, Ping Li, Lingfei Zhang, And Yan Chen[4], Introduced a replacement to simplify noisy-filled financial temporal series via sequence reconstruction by leveraging motifs then utilize a convolutional neural network to capture spatial structure of time series.In this work, we proposed an algorithm combining motif-based sequence reconstruction with CNN for stock statistic trend prediction. The underlying patterns within the reconstructed sequence are learned employing a convolutional neural network, which gives useful data for ups and downs prediction.This method sheds light on macroscopic pattern find in financial statistic and provides a complete unique solution for price prediction.

MojtabaNabipour, PooyanNayyeri, HamedJabani, Shahab S.(Senior Member, Ieee), And Amir Mosavi[5], Aims to significantly minimize the risk of price prediction with machine learning and deep learning algorithms. Research works clear that there was a significant improvement in the performance of models when they use binary data instead of continuous one.

Jennifer S. Raj[6], Provides a reliable prediction model for many applications. Basically, back-propagation algorithms were used for training RNN. This paper predicts system robustness by applying SVM learning algorithm to RNN. Comparison of the given model is done with the existing systems for analysis of forecast of performance. These results show that the performance of proposed system exceeds that of the existing ones.

Aline de Oliveira Machado, CaioJordao de Lima Carvalho , Antonio Carlos dos Santos Souza, MarcioCerqueira de FariasMacedo[7],proposes a neural networks and genetic algorithms to buy or sell financial assets forcefully based on the stock value variation. The forecast of stock index is themain objective aimed by users. It allows making the best decision while buying or selling of stocks. Experiment result shows the techniques and fields of artificial intelligence, besides ANNs and Genetic Algorithms, such as sentiment analysis and Deep Learning that can improve the decision-making process.

HarshitAgarwal, GauravJariwala and Akshit Shah[8], proposes a progressive result on the application of recurrent neural networks in stock price forecasting. Machine learning and deep learning strategies are being used by enormous quantitative hedge funds to maximize their returns. Finance data belongs to time series data.The proposed system is accurately following the pattern of the stock prices where the predicted values are in close vicinity to the actual market values.

JaiprakashVerma, Sanjay Garg, Ishit Gandhi and Nikita H. Bachani[9],the research work proposed in this paper is based on a project which consists of creating a system for stock market trend prediction based on the machine learning algorithms. It is also managing stocks for particular user profiles with the functionality of add update and delete share as per current market prices.

Yang Liu, JelenaTrajkovic, Hen-Geul (Henry) Yeh, Wenlu Zhang [10], presents a review that there are many factors that affect performance of stock market, such as world and local finance, political events, supply and demand, and out of the normal events, as COVID-19 pandemic. Analyze the stock market activity performance using daily world news headlines from Reddit. In similar, propose novel CNN and RNN architectures to find out the result of global big events to the stock price movements.

Leonardo dos Santos Pinheiro and Mark Dras [11], proposed a work to explore RNN with character-level language model pre-training for both intraday and inter day stock market forecasting. This paper shows the use of a simple LSTM neural network with character level sink for stock forecasting using only economic news as parameter.

Uma Gurav and Nandini Sidnal [12], attempt to do the analysis of various problems pertaining to dynamic stock market price prediction, based on the fact that minimization of stock Market investment risk is strongly correlated to minimization of forecasting errors. The analysis made in the paper can be used as an EML which is a correct union of all predictive models in order to make a good choice better. This forms the basis for subsequent research.

Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan [13], focuses on the use of Regression and LSTM based Machine learning to predict stock values. Parameter taken are open, close, low, high and volume. LSTM and Regression have been utilized in finance data set. This will improve the performance by using bigger data sets. It helps to improve the accuracy in stock prediction.

Mehak Usmani, Syed Hasan Adil, Kamran Raza and Syed Saad Azhar Ali [14], proposed a new algorithm approach to predict a particular stock exchange by its closing time. The results of this study confirm that machine learning techniques are capable of predicting the stock market performance.

A. Akash, R. Shanthi, R. Aravindh, V. Kurunjivendhan, D. Veerapandi [15], In the proposed system Least Square Support Vector Machine (LS-SVM) and Particle Swarm Optimization (PSO) algorithm is used. PSO algorithm selects best free data combination for LS-SVM to avoid over-fitting and local minimization issue and improve prediction correctness. The proposed system is trust worthy than other stock brokers.

Akira Yoshihara, Kazuki Fujikawa, Kazuhiro Seki, and Kuniaki Uehara [16], proposes a study to predict the trend of stock prices, which is influenced by miscellaneous events happening around the world. This paper proposed an approach to predicting the trend of stock prices by focusing on news events with long-term effects. It is crucial to model temporal properties of news events expressed in natural language which have short- or long-term influences on stock prices. The results indicate the effectiveness of the deep learning models in general in the financial domain and also suggest the potential of the recurrent model to capture the properties of past significant events with long-term effects on the stock exchange.

Ghaith Abdulsattar A. Jabbar Alkubaisi, Siti Sakira Kamaruddin & Husniza Husni [17], proposed a sentiment analysis based stock price prediction method which is focus on consumer reaction. The first phase of this model is data collection, and therefore the second involves the filtration and transformation, which are conducted to urge only relevant data. This method has mainly been founded on sentiment analysis approach, by employing expert labeling technique and features, namely, spatial and temporal. This study proposes HNBCs as a machine learning method for stock market classification. A hybrid algorithm has been adapted from NB; these hybrid algorithms incorporate two different NB algorithms supported their specific functionalities.

YanhuiGuo, Siming Han, ChuanheShen, Ying Li, Xijie Yin, And Yu Bai[18], Aiming at the discreteness, non-normality, high-noise in high-frequency data, a support vector machine regression (SVR) algorithm is introduced in this paper. To effectively mitigate the danger and to realize high investment return, an outsized number of prediction models are proposed. An adaptive SVR based on PSO is proposed to enhance the versatility of the model and to avoid suffering from adjusting parameters of SVR. The results showed that the adaptive SVR has better adaptability and better prediction results than the traditional SVR and BPNN.

ParaskeviNousi, AvraamTsantekidis, NikolaosPassalis, AdamantiosNtakaris, JuhoKanniainen, AnastasiosTefas, MoncefGabbouj And AlexandrosIosifidis,[19],presents a machine learning approach for the prediction of future price movements using limit order book data. Furthermore, the distribution of limit order events changes rapidly, not just from one day to the next but also within the same day. Machine learning has been used very greatly to analyze the economic market from many different aspects. Introduce the classifications methods that are used to predict the misprices movements.

Salah Bouktif, Ali Fiaz, And MamounAwad[20], contributes constructively in this debate by empirically investigating the predictability of stock market movement direction using an enhanced method of sentiments analysis. Sentiment analysis (SA), opinion mining, natural language processing (NLP), information retrieval, and structured/unstructured data mining have been utilized to analyze and discover sentiment from texts and other communication mediums. The analysis rich with sentiment analysis papers discussing the usage of tweets, financial news along with other relevant information to predict the stock movement. Sentiment classification is traditionally performed using both supervised and unsupervised methods, namely, machine learning and lexicon-based approach. In general, a machine learning algorithm attempts to minimize a cost function.

S. Karuppusamy, G. Singaravel [21], Analysis of faults in the code phase is detected through integral methods that identify the error in software. The repositories of the data set are collected during the software product development life cycle model, which is then integrated with a machine learning algorithm namely Bayesian decision theory to detect the error probabilities and to predict unbound error during the prediction of the software faults.

3. COMPARITIVE ANALYSIS

Title	Techniques & Mechanisms	Parameter Analysis	Future Work
A Machine Learning Model for Stock Market Prediction	LS-SVM, Artificial neural network,Levenberg-Marquardt algorithm	Predict stock price	Improve result from fluctuations in the time series function
A Predictive Analytic study on stock market trend by supervised machine learning algorithms	Supervised machine learning algorithms, artificial neural networks, SVM	Price trend forecasting	predict long-term stock price trend with high accuracy
Stock Market Prediction via Multi-Source Multiple Instance Learning	Multiple Instance Learning algorithm, sentiment analysis, SVM	sentiments and historical data analysis	Further improving the effectiveness of the model
Stock Market Trend Prediction Using High-Order Information of Time Series	Trend prediction, convolutional neural network, ANN, Deep Learning	Trend prediction scheme on latest trends	Improve effectiveness on capturing trend information of stock shares
Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data	Decision Tree, Random Forest, Adaptive Boosting, ANN, RNN, Naïve Bayes, KNN XGBoost, SVC	Market Trend prediction	Improving efficiency with binary data.



Recurrent Neural Networks And Nonlinear Prediction In Support Vector Machines	Support vector machines, RNN, pattern recognition, back-propagation algorithm	Predict system reliability	Comparing more optimization algorithms and extracting their best features
Expert Advisor using Artificial Neural Networks and Genetic Algorithms to Predict Stock Market Trends	Genetic Algorithm, ANN	Predict stock market trends	Use DL for higher amount of data.
Analysis and Prediction of Stock Market Trends Using Deep Learning	Recurrent neural network, Random forest classifier	Analysis of stock prediction	Adding more dimensionality to the data set
Evaluation of Pattern Based Customized Approach for Stock Market Trend Prediction With Big Data and Machine Learning Techniques	ANN,SVM, Big Data Analytics	New pattern-finding algorithm	Unstructured and semi-structured input as for a future enhancement
Machine Learning for Predicting Stock Market Movement with News Headlines	CNN, SVM, RNN,K-Nearest Neighbor	Analyze the stock market activity	Plan to employ word2vec word embedding and deep transformers to improve the performance
Stock Market Prediction with Deep Learning: A Character-based Neural Language Model for Event-based Trading	Bag-of-words and support machines, CNN, LSTM	Forecasting using only financial news as predictors	To create richer feature sets
Predict Stock Market Behavior:Role of Machine learning algorithms	ML algorithms, six sigma	Minimization of stock Market investment risk	Improve Performance
Stock Market Prediction Using Machine Learning	ML algorithms, SVM,RBF	Gain intelligence for an accurate prediction	Gain efficiency of the stock index movements
Stock Market Prediction using Machine Learning Techniques	Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF) and Support Vector Machine (SVM)	To predict the market performance of KSE	Modify the algorithm to suitable for large number of resource
Stock Market Trend Prediction using Machine Learning	LS-SVM,Particle Swarm Optimization,Levenberg - Marquardt algorithm, ANN	Forecast stock trends	Improve the accuracy of the prediction
Predicting the Trend of the Stock Market by Recurrent Deep Neural Networks	RNN, RBM	To model temporal effects of past events	Events with long-term resource
Stock Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naive Bayes Classifiers	Hybrid Naïve Bayes Classifiers	Sentiment analysis	Employing expert labeling technique
An Adaptive SVR for High-Frequency Stock Price Forecasting	Support vector machineregression ANN, particle swarm optimization	Dynamic optimization of learning parameters	A weighted adaptive SVR will be introduced in our future work
Machine Learning for Forecasting Mid-Price Movements Using Limit Order Book Data	Machine Learning Algorithms, SVM	Future price movements using limit order book data	Systematically improve the time-wise performance
Augmented Textual Features-Based Stock Market Prediction	Machine learning, model stacking, NLP,sentiment analysis	Predict stock market movement by sentiment analysis	Future research involves mainly the investigation of variations of techniques for both sentiment features engineering and prediction modeling of the stock movement

Investigation Analysis for Software Fault Prediction using Error Probabilities and Integral Methods	Software fault prediction, repository mining, error probability, integral method	improve the accuracy of predicting software defects	To produce a faultless product
A Survey on Sentiment Analysis Technique in Web Opinion Mining	Opinion mining, sentiment analysis, sentiment lexicon, feature extraction, sentiment classification	opinion-oriented information-seeking	Improved discussion of available resources and benchmark datasets

4. CONCLUSION

Machine Learning plays an important role in future prediction of stock market. This paper reviewed algorithms, techniques used for stock prediction and parameter analysis used in stock prediction from an existing time series mechanisms. Major papers reviewed about the stock prediction algorithms based on future prediction, cost, priority, and money flow and resource utilization. Considering this survey my work will be related to stock price prediction by implementing a hybrid algorithm.

REFERENCES

- [1] Osman Hegazy, Omar S. Soliman and Mustafa Abdul Salam "A Machine Learning Model for Stock Market Prediction" International Journal of Computer Science and Telecommunications [Volume 4, Issue 12, December 2013] ISSN 2047-3338.
- [2] Achyutasai Nikhil "A Predictive Analytic study on stock market trend by supervised machine learning algorithms" Alohana Chakra Journal, Volume IX, Issue V, May/2020, ISSN NO:2231-3990.
- [3] XI Zhang (member, IEEE), Siyu Qu, Jieyun Huang, Binxing Fang, And Philip Yu "Stock Market Prediction via Multi-Source Multiple Instance Learning" The University of Illinois at Chicago, Chicago, October 8, 2018.
- [4] Min Wen, Ping Li, Lingfei Zhang, And Yan Chen "Stock Market Trend Prediction Using High-Order Information of Time Series" Special Section On Big Data Learning And Discovery, March 18, 2019.
- [5] Mojtaba Nabipour, Pooyan Nayyeri, Hamed Jabani, Shahab S "Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis" Journal of Communications Volume 8, 2020.
- [6] Dr. Jennifer S. Raj, J. Vijitha Ananthi "Recurrent Neural Networks And Nonlinear Prediction In Support Vector Machines" Journal of Soft Computing Paradigm (JSCP) (2019) Vol.01/ No. 01
- [7] Aline de Oliveira Machado, Caio Jordao de Lima Carvalho, Antonio Carlos dos Santos Souza, Marcio Cerqueira de Farias Macedo "AlphaB3 – Expert Advisor using Artificial Neural Networks and Genetic Algorithms to Predict Stock Market Trends" 17 December 2020.
- [8] Harshit Agarwal, Gaurav Jariwala and Akshit Shah "Analysis and Prediction of Stock Market Trends Using Deep Learning" Future Computing and Informatics Journal 10 May 2020.
- [9] Sudeep Tanwar, Sanjay Garg, Sanjay Garg, Nikita H. Bachani "Evaluation of Pattern Based Customized Approach for Stock Market Trend Prediction With Big Data and Machine Learning Techniques" International Journal of Business Analytics Volume 6 • Issue 3 • July-September 2019
- [10] Yang Liu, Jelena Trajkovic, Hen-Geul (Henry) Yeh, Wenlu Zhang "Machine Learning for Predicting Stock Market Movement with News Headlines" Conference Paper • November 2020.
- [11] Leonardo dos Santos Pinheiro, Mark Dras "Stock Market Prediction with Deep Learning: A Character-based Neural Language Model for Event-based Trading" In Proceedings of Australasian Language Technology Association Workshop 2017.
- [12] Uma Gurav and Prof. Dr. Nandini Sidnal "Predict Stock Market Behavior: Role of Machine learning algorithms" K.I.T's College of engineering, Department of Information Technology, 20 July 2020.
- [13] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan "Stock Market Prediction Using Machine Learning" Conference Paper • December 2018.

-
- [14] MehakUsmani, Syed HasanAdil, Kamran Raza, Syed SaadAzhar Ali "Stock Market Prediction using Machine Learning Techniques" Iqra University Karachi, 23 May 2018.
- [15] A.Akash, R.Shanthi, R.Aravinth, V.Kurunjivendhan , D.Veerapandi "Stock Market Trend Prediction using Machine Learning" Stock Market Trend Prediction using Machine Learning, Vol. 7, Issue 2, February 2019.
- [16] Akira Yoshihara, KazukiFujikawa, Kazuhiro Seki, and KuniakiUeharaPredicting the Trend of the Stock Market by Recurrent Deep Neural Networks"Faculty of Intelligence and Informatics, Issue 7, July 2014.
- [17] GhaithAbdulsattar A. JabbarAlkubaisi, SitiSakiraKamaruddin&HusnizaHusniStock Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naive Bayes Classifiers" Computer and Information Science; Vol. 11, No. 1; 2018 ISSN 1913-8989 E-ISSN 1913-8997.
- [18] YanhuiGuo, Siming Han, ChuanheShen, Ying Li, Xijie Yin, And Yu Bai"An Adaptive SVR for High-Frequency Stock Price Forecasting" International Journal of Intelligent Engineering and Systems, Volume 6, 2018.
- [19] ParaskeviNousi, AvraamTsantekidis, NikolaosPassalis, AdamantiosNtakaris, JuhoKanniainen, AnastasiosTefas, (Member, Ieee), MoncefGabbouj, (Fellow, Ieee), And AlexandrosIosifidis"Machine Learning for Forecasting Mid-Price Movements Using Limit Order Book Data" Volume 7, 2019.
- [20] Salah Bouktif , (Member, Ieee), Ali Fiaz, And MamounAwadAugmented Textual Features-Based Stock Market Prediction" Volume 8, 2020.
- [21] S. Karuppusamy and G. Singaravel, "Investigation Analysis for Software Fault Prediction using Error Probabilities and Integral Methods", Applied Mathematics & Information Sciences An International Journal, Volume13, 2019.
- [22] S. Veeramani, S. Karuppusamy, "A Survey on Sentiment Analysis Technique in Web Opinion Mining", International Journal of Science and Research (IJSR), Volume 3, 2012.