Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 4, April 2023

STOCK MARKET PREDICTION USING MACHINE LEARNING ALGORITHMS

Ms S.V.S. VALLI[1], Mr. S. AMRUTH RAJ[2], Mr. N. JEEVAN SAiI[3],Ms. K. LAVANYA[4], Mrs. Khadarunnisa[5]

1. BTECH, NADIMPALLI SATYANARAYANA RAJU INSTITUTE OF TECHNOLOGY, SONTYAM, VISAKHAPATNAM, ANDHRA PRADESH, INDIA - 531173
2. BTECH, NADIMPALLI SATYANARAYANA RAJU INSTITUTE OF TECHNOLOGY , SONTYAM, VISAKHAPATNAM, ANDHRA PRADESH, INDIA - 531173
3. BTECH, NADIMPALLI SATYANARAYANA RAJU INSTITUTE OF TECHNOLOGY , SONTYAM, VISAKHAPATNAM, ANDHRA PRADESH, INDIA - 531173
4. BTECH, NADIMPALLI SATYANARAYANA RAJU INSTITUTE OF TECHNOLOGY , SONTYAM, VISAKHAPATNAM, ANDHRA PRADESH, INDIA - 531173
5. Assistant Professor COMPUTER SCIENCE AND ENGINEERING, NADIMPALLI SATYANARAYANA RAJU INSTITUTE OF TECHNOLOGY, SONTYAM, VISAKHAPATNAM, ANDHRA PRADESH, INDIA-531173

ABSTRACT

Stock price prediction using machine learning makes it feasible to determine the future value of company stocks and other financial assets traded on an exchange. Making stock price forecasts is done with the intention of making substantial profits. Please be aware, however, that it is impossible to forecast the stock's exact price. The forecast takes into account additional factors, such as biological and psychological factors, as well as rational and irrational behaviour. These factors combine to produce a volatile and dynamic share market. As a result, making accurate stock price predictions is quite difficult. Here, we'll build a project that calculates stock values using a linear regression model and the Python programming language. A publicly traded company's stock prices can be used as historical information. We will employ a variety of machine learning techniques to forecast the future stock price of this company, starting with simple ones like linear regression. Because financial stock markets are unpredictable and nonlinear, it is difficult to predict stock market returns with any degree of accuracy. The development of artificial intelligence and improved processing power has made programmed prediction systems more effective at forecasting stock prices. In this study, the closing prices of five businesses from various industry sectors were predicted using Artificial Neural Network and Random Forest methodologies.

1 INTRODUCTION

In the investment world, analysing financial data in securities has been a significant and difficult problem. Due to the conflicting impacts of information rivalry among significant investors and the unfavourable selection costs imposed by their knowledge advantage, stock price efficiency for publicly traded companies is challenging to accomplish.

The two primary schools of thinking used to analyse the financial markets are as follows. The first strategy is referred to as fundamental analysis. Through qualitative and quantitative examination, the fundamental analysis approach determines a stock's intrinsic worth in order to appraise it. This method looks at the managerial, market, micro, and macroeconomic elements of an organisation.

Technical analysis is the name for the second strategy. The method utilised by technical analysis to predict price direction is the analysis of past market data. A number of charts are used in technical analysis to predict what is likely to happen. The several types of stock charts include candlestick, line, bar, point-and-figure, OHLC (open-high-low-close), and mountain charts. The charts can be viewed with price and volume in a variety of time intervals. The charts use a variety of indicators, including breakout, trending, momentum, resistance, support, and breakout.

There are a number of different ways to handle this kind of issue, ranging from conventional statistical models to approaches based on computational intelligence and machine learning. Vanstone and Tan conducted a review of the literature on using soft computing in financial trading and investment. They divided the papers under consideration into the following categories: classification, time series, hybrid approaches, optimization, and time series. The survey revealed that the majority of research in the discipline of financial trading was being done in the area of technical analysis. By concentrating on macroeconomic study, a model that integrates fundamental and technical analysis was looked at to assess the stock price patterns. Additionally, it examined how a company behaved in connection to its industry and the economy, which gives investors more knowledge when making investment decisions.

By combining the KNN methodology with technical analysis, a closest neighbour search (NNS) method obtained the desired result. This model utilised technical analysis to historical price and trade volume data from the stock market. Stop loss, stop gain, and RSI filters were used as technical indicators. The distance function was applied to the gathered data by the KNN algorithm portion.

Stop loss, stop gain, and RSI filters were used as technical indicators. The distance function was applied to the gathered data by the KNN algorithm portion. Using the fundamental analysis method, this model was contrasted with the buy-and-hold technique.

The Fast Library for Approximate Nearest Neighbors (FLANN) is used to carry out the searches to determine which algorithm among a number of algorithms in its library works best. The FLANN model was created by carrying out quick approximate nearest neighbour searches in high dimensional spaces, and Majhi tested it to predict the S&P 500 indices.

Comparing artificial neural networks (ANN) to traditional statistical methods, they show high generalisation capability. The traits of performing stocks can be identified by ANN by inferring from past data. Variables in the technical and financial realms reflect the information. As a result, ANN is employed as a statistical technique to investigate the complex connections between the performance of stocks and relevant financial and technical aspects.

Comparing artificial neural networks (ANN) to traditional statistical methods, they show high generalisation capability. The traits of performing stocks can be identified by ANN by inferring from past data. Variables in the technical and financial realms reflect the information. As a result, ANN is employed as a statistical technique to investigate the complex connections between the performance of stocks and relevant financial and technical aspects.

There have been some studies done using both qualitative and quantitative analysis. Shynkevich investigated how the introduction of concurrent, correctly weighted news pieces with varying degrees of relevance to the target stock could enhance the effectiveness of a financial forecasting algorithm. The traders' and investors' decision-making was aided by the financial model. Techniques for text pre-processing were used to create the prediction system. To forecast the changes in stock price, a multiple kernel learning technique was used. While distinct kernels were used to analyse each news category, the technique combined data that was taken from various news categories. The news stories were divided up based on certain industry-specific variables and their applicability to the target stock. The tests were conducted on equities in the healthcare industry. The findings demonstrated that the financial forecasting model performed better when data sources included more categories of pertinent news. An improved model for this study included past prices as an additional data source and generated predictions using both textual and time series data. For various data sources, more kernels may be used. The performance of forecasting was enhanced by the addition of new category features. Financial forecasting and analysis frequently use linear regression. Numerous regression classifiers have proven beneficial for forecasting by analysing quantitative data and predicting model parameters.

In comparison to moving average, a regression driven fuzzy transform (RDFT) distributes a smoothing approximation of time series with a shorter delay. This capability is crucial for predicting tools where time is a crucial factor.

To examine stock market trend, a hybrid intelligent data mining methodology based on a Support Vector Machine-Genetic Algorithm model was examined. In this method, the genetic algorithm is utilised for variable selection in order to decrease the complexity of the model and increase the speed of the support vector machine. The historical data is then used to discover stock market trends. Due to the ANN-like black box technique's limitations, hybrid methodologies can be utilised to enhance the current forecasting models. To anticipate financial time series data, a combination of techniques including fuzzy rule-based systems, fuzzy neural networks, and Kalman filters with hybrid neuro-fuzzy architecture have been developed.

2. LITERATURE SURVEY AND RELATED WORK

A study on a KNN approach that made use of economic indicators and classification techniques to predict the stock price trends yielded considerable precision. Four indicators were identified and calculated based on the technical indicators and their formulas. The values of the indicators were normalized in the range between -1 and +1. Accuracy and F- measure were calculated to evaluate the performance of the model. Calculation of these evaluation measures required estimating Recall and Precision which were assessed from True Negative (TN), False Negative (FN), True Positive (TP) and False Positive (FP). The performance of the KNN model was improved by using the optimal value for the k parameter. The evaluation of the KNN model and its performance is illustrated in Table 1.

Measurement	K Parameter		
	25	45	70
Accuracy	0.8138	0.8059	0.8132
F-measure	0.8202	0.8135	0.8190

Table-1: KNN Model Performance.

A model known as Integrated Multiple Linear Regression-One Rule Classification was the subject of another investigation.Using the initial predicted outputs in regression form, the model (Regression-OneR) predicts the stock class outputs in classification form. The continuous outputs were predicted using regression classifiers. versus the categorization classifiers, which were employed to forecast the outputs in categorical values.As shown in Table 2, the results demonstrated that, when compared to a model that just included a typical classification algorithm, the hybrid regression-classification strategy generated a higher accuracy rate and reduced Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Classifier	Accuracy	MAE	RMSE
Regression-OneR	85.0746	0.14	0.3863
OneR	71.6418	0.2836	0.5325
ZeroR	64.1791	0.4619	0.4805
J48 Decision Tree	82.0896	0.2121	0.3796
REP Tree	76.1194	0.2764	0.4207

Table-2: Prediction Results of Classifiers.

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 4, April 2023
Sneha Kalra conducted research on stock market price fluctuations in 2019 for the authors of this report with regard to pertinent new company articles. They employed the Naive Bayes classifier to distinguish between utterances that were positive or negative for the goal of making predictions based on daily news variance, social media data, and blogs.
According to Aditya Menon et al., this research from 2019 is concentrated on an evaluation of neural models for forecasting stock performance. The long short term memory technique for forecasting the economic facts converging into the current fad would be given precedence.

Sharma, Ashish They conducted a survey of regression techniques for stock prediction using stock market data in 2017 and discovered that regression analysis is primarily employed for stock market trend prediction. More numerical variables could be used in the future to improve results.

2019's Andrea Picasso was the subject of this study's authors' efforts to combine economic and elemental analysis for the purpose of predicting market trends using a variety of applications and automated techniques. Neural networks are a machine learning tool for trend stock problems, and those are forecasting data charts. The sentiment of a news story is used as input data. Their analysis indicates that the use of news astral one-off information is the most troublesome task.
The appropriate feature fusion technique will be appropriate in the future to solve this issue.

The author discussed three different types of machine learning techniques as well as various types of metrics like accuracy, confusion matrix, recall, RMSE, precision, and quintile of errors. Gangadhar Shobha published this paper in 2018 and it gave a thorough overview of machine learning techniques that will aid readers in using equations and concepts. Because most people are unsure whether to employ machine learning techniques for prediction or other purposes, the author believes that this review can be helpful for those who are new to the field.

Negahdari, Arash In 2018, as the stock prediction, numerous experiments and models have been created for prediction purposes on historical data, such as the HyS3 graph-based semi-supervised model presented in this study and through a network views Kruskal based graph method called ConKruG. They believe that in the future, social media and Twitter data could be utilised to predict stocks more accurately using these algorithms.

The least square support vector machine (LS SVM) and another algorithm called "PSO" (particle swarm optimisation) were both introduced by Akash in 2019. The "PSO" algorithm essentially chooses the best bounded parameter with the "LS SVM" to reduce overfitting and some technical indicators, which will essentially enhance the result accuracy. However, the proposed technique is also being contrasted with a concept of an artificial neural network.

Companies that provide financial services are introducing new tools to help with future forecasting. Stock prediction, also known as stock market mining, is one of the many sources of financial information in the world that can be useful research fields. Stock forecasting becomes more crucial, especially if several guidelines can be developed to aid in improving investing choices across various stock markets. In Sweden, Hellestrom and Homlstrom used a statistical analysis based on a modified KNN in 1998 to determine where correlated areas fall in the input space to improve the performance of prediction for the period 1987-1996. Shin had adopted the genetic algorithm in 2005; the number of trading rules was generated for the Korea Stock Price Index 200 (KOSPI 200). In 2004, the Zimbabwe Stock Exchange offered both of the aforementioned models—the Weightless Neural Network (WNN) model and the Single Exponential Smoothing (SES) model Mpofu—to forecast stock prices. Gavrilov offered a clustering stocks strategy in 2004 to group 500 equities from the Standard & Poor. The opening stock price was one of 252 numbers that made up the data. Cao proposed a fuzzy evolutionary algorithm in 1977 to find pair relationships in stock data based on user preferences. The study offered prospective guidelines to mine stock pairs, stock trading regulations, and markets. It also demonstrated the value of such a method for actual trading.

3 Implementation Study

Data Gathering

Data gathering is a highly fundamental module and the project's first stage. It typically has to do with gathering the appropriate dataset. The dataset that will be utilised to anticipate the market must be vetted based on a number of criteria. By incorporating more external data, data collecting also aids dataset enhancement. Stock prices from the prior year make up the majority of our data. In order to analyse the forecasts accurately, we will first analyse the live dataset and use the model with the data in accordance with the accuracy.

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 4, April 2023

Pre-processing Data

Pre-processing is the process of converting unstructured data into a more comprehensible form. Raw data typically contains numerous inaccuracies, is inconsistent, or is both. Checking for missing values, looking for categorical values, dividing the data set into a training and test set, and then doing feature scaling to narrow the range of variables so that they may be compared on similar environments are all part of the data preprocessing.

The Machine's Training

Feeding the data to the algorithm to improve the test data is analogous to training the machine. The models are adjusted and fitted using the training sets. The test sets are unaltered since it is improper to evaluate a model using hypothetical data. Cross-validation is a step in the model training process that allows us to obtain an approximation of the model's performance that is well-founded and is based on training data. The purpose of tuning models is to precisely adjust the hyperparameters, such as the number of nearest neighbours. On each set of hyperparameter values, the whole cross-validation cycle is run. We will next determine a cross-validated score for each set of hyperparameters.

A stock's price to fluctuate dependent on specific factors. It also demonstrates the weaknesses of a specific stock or organisation. To ensure that only authorised entities have access to the results, a user authentication system control is in place.scoring dataScoring the data is the process of using a predictive model on a set of data. The KNN Algorithm is the method used to process the dataset. We achieve intriguing outcomes using the learning models. Thus, the final module explains how the model's output might aid in probability prediction.

4 PROPOSED WORK

In addition, the forecast considers biological, psychological, and behavioural elements, as well as rational and irrational behaviour. These elements work together to create a volatile and active stock market. As a result, it can be challenging to predict stock prices accurately. Here, we'll create a project that uses the Python programming language and a linear regression model to determine stock values. Stock prices of publicly listed companies can be used as historical data. To predict the future stock price of this firm, we will use a range of machine learning approaches, starting with straightforward ones like linear regression. It is challenging to forecast stock market returns with any level of accuracy since financial stock markets are unpredictable and nonlinear. The advancement of artificial intelligence and increased computing power have increased the accuracy of programmed prediction systems' stock price forecasts. In this study, artificial neural networks and random forest techniques were used to forecast the closing prices of five enterprises from distinct industry sectors.

5 METHODOLOGIES

Both of the aforementioned models—the Weightless Neural Network (WNN) model and the Single Exponential Smoothing (SES) model Mpofu—were made available by the Zimbabwe Stock Exchange in 2004 for the purpose of predicting stock prices.

In 2004, Gavrilov proposed a clustering stocks approach to group 500 Standard & Poor's stocks. The data consisted of 252 figures, including the starting stock price. For the purpose of discovering pair relationships in stock data based on customer preferences, Cao presented a fuzzy evolutionary method in 1977. The study provided future recommendations for mining stock pairs, stock trading rules, and markets. It also proved how useful a strategy like that would be for actual trading.

The probabilistic method's categorization procedures are illustrated as follows:

Steps:

Method of classification: probabilistic

From the data set, determine the prior probability for the Profit and Loss classes.

Based on the quantity of Profit nearest neighbours and the quantity of Loss nearest neighbors, determine the KNN's probabilities of the Profit class and the Loss class.

Compute the joint probabilities using the KNN's probabilities for the Profit and Loss classes along with the prior probabilities.

The joint probability of the Profit class and the Loss class are compared.

From the class values, choose the predictive value with the highest jointprobability.

.

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 4, April 2023

DJANGO

A high-level Python web framework called Django promotes quick iteration and logical, elegant design. It was created by seasoned programmers and handles a lot of the hassle associated with Web development, freeing you up to concentrate on building your app without having to invent the wheel. It is open source and free.

Django's main objective is to make it simpler to create intricate, database-driven websites. Rapid development, "pluggability," and the don't repeat yourself philosophy are all highlighted by Django. Everywhere, including in configuration files and data models, Python is used.

6 RESULTS AND DISCUSSION SCREENSHOTS

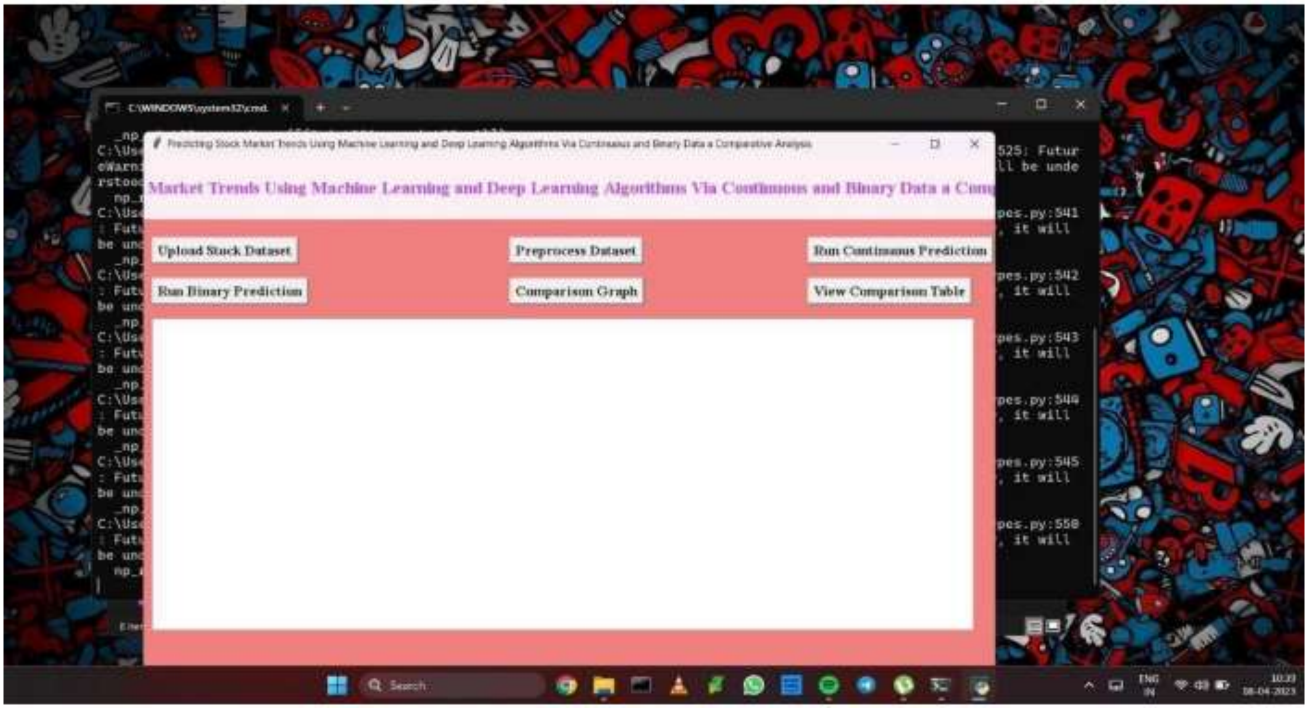


Fig-1: Upload screen

Fig-2: Preprocessing dataset

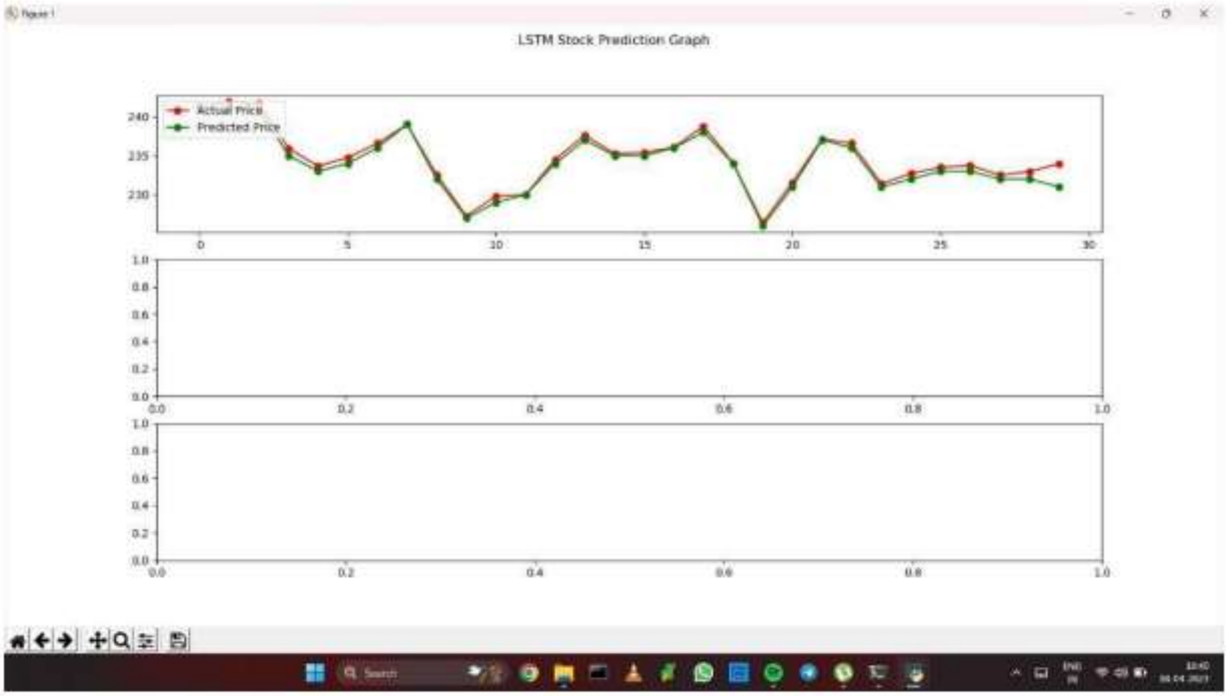


Fig-3: Run Continuous Prediction

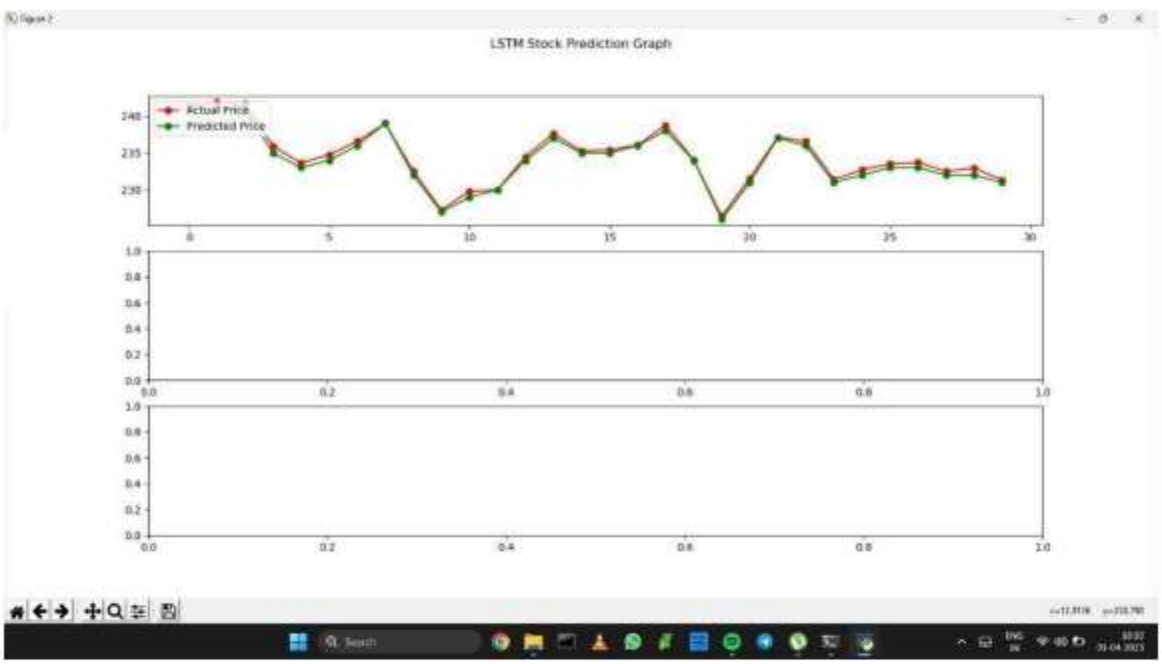


Fig-4: Run Binary Prediction

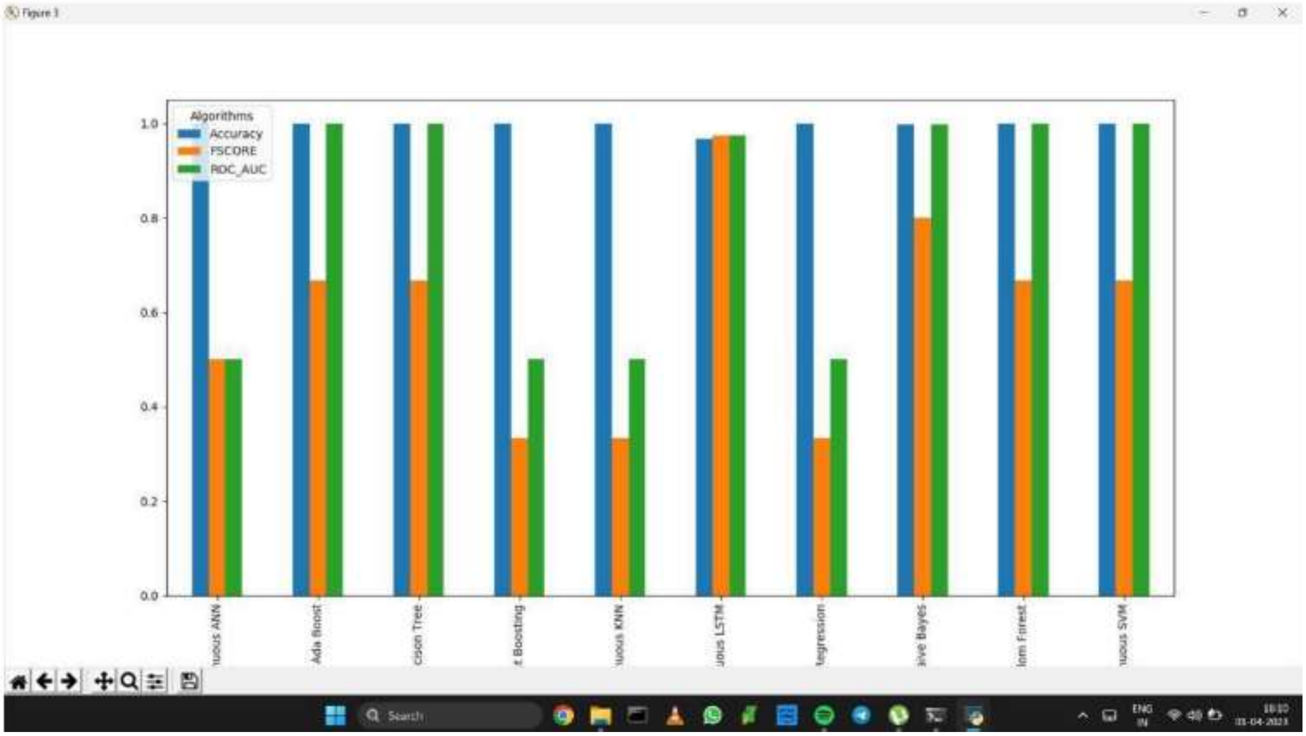
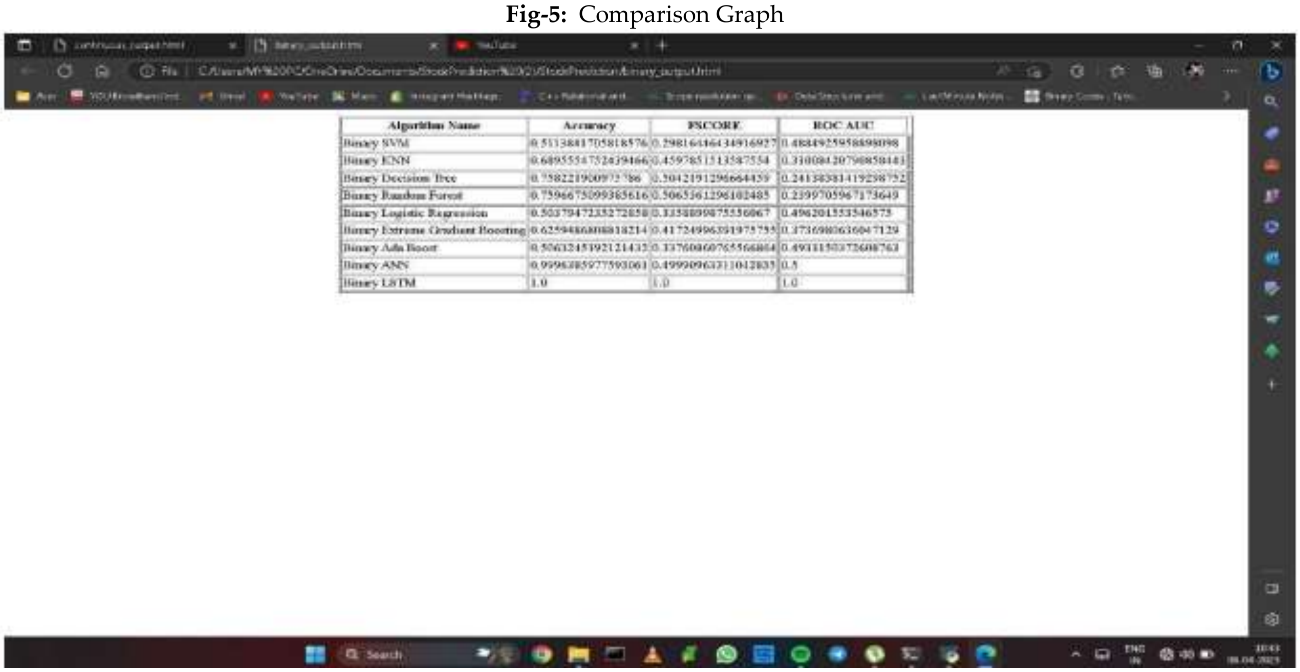


Fig-5: Comparison Graph



Algorithm Name	Accuracy	FSCORE	ROC AUC
Binary SVM	0.5113881705818576	0.2981646434916927	0.4884925958898098
Binary KNN	0.6895554752439466	0.4597851513587554	0.3100842079858143
Binary Decision Tree	0.7582219009772786	0.3042191296664439	0.24138381419298752
Binary Random Forest	0.7596675099385616	0.5065361296102485	0.2399705967173649
Binary Logistic Regression	0.5037947235272858	0.3358899875556067	0.496201553546575
Binary Extreme Gradient Boosting	0.6259486808818214	0.41724996391975755	0.3736980636047129
Binary Ada Boost	0.5063245192121432	0.33760860765566864	0.4931150372608763
Binary ANN	0.9996385977593061	0.49990963311042835	0.5
Binary LSTM	1.0	1.0	1.0

Fig-6: Comparison Table

7. CONCLUSION AND FUTURE WORK

The purpose of this study is to enhance the statistical fitness of the suggested model to address a computation issue with the Machine Learning Algorithm. The empirical distribution of the class values for Profit and Loss can be computed by the machine learning classifier using the k nearest neighbours.
Due to the lack of data, the result is, nonetheless, insufficient. Due to its inability to generalise sparse data outside of its immediate neighborhood, machine learning algorithms classifiers have an underfitting problem.

On the issue of forecasting stock price movements, we contrasted a hybrid machine learning algorithm-probabilistic model with four common algorithms. Our findings demonstrated that the suggested KNN-Probabilistic model outperforms the traditional KNN algorithm and other classification algorithms by a wide margin.

The suggested model has a drawback because it uses a binary classification method. This binary classification model's real output is a prediction score with two classes. The score represents how certain the model is that the provided observation belongs to the Profit class or the Loss class. The knowledge component for upcoming work entails converting binary classification to multiclass classification. The multiclass classification entails observation and analysis of more statistical class values than the two already present. To provide more detailed information about each class value, additional study will apply the probabilistic model to multiclass data. There will be five class labels in the newly created multiclass categorization, with the names "Sell," "Underperform," "Hold," "Outperform," and "Buy." In numerical mapping values In order to achieve our goal, we will translate "Sell" to -2, which denotes extremely unfavourable; "Underperform" to -1, denoting moderately unfavourable; "Hold" to 0 to denote neutral; "Outperform" to 1 to denote moderately favourable; and "Buy" to 2 to denote very favourable.

8. REFERENCES

[1] Benjamin Graham, Jason Zweig, and Warren E. Buffett, The Intelligent Investor, Publisher: Harper Collins Publishers Inc, 2003.

[2] Charles D. Kirkpatrick II and Julie R. Dahlquist, Technical Analysis: The Complete Resource for Financial Market Technicians (3rd Edition), Pearson Education, Inc., 2015.

[3] Bruce Vanstone and Clarence Tan, A Survey of the Application of Soft Computing to Investment and Financial Trading, Proceedings of the Australian and New Zealand Intelligent Information Systems Conference.

[4] Vol.1, Issue 1, http://epublications.bond.edu.au/infotech_pubs/ 13/, 2003, pp. 211–216.

[5] Monica Tirea and ViorelNegru, Intelligent Stock Market Analysis System - A Fundamentaland Macro-economic Analysis Approach, IEEE, 2014.

[6] Kian-Ping Lim, Chee-Wooi Hooy, Kwok-Boon Chang, and Robert Brooks, Foreign investors and stock price efficiency: Thresholds, underlying channels and investor heterogeneity, The North American Journal of Economics and Finance. Vol. 36, http://linkinghub.elsevier.com/retrieve/pii/S106 2940815001230, 2016, pp. 1–28.

[7] Lamartine Almeida Teixeira and Adriano Lorena Inácio de Oliveira, A method for automatic stock trading combining technical analysis and nearest neighbor classification, Expert Systems with Applications, http://linkinghub.elsevier.com/retrieve/pii/S095 7417410002149, 2010, pp. 6885–6890.

[8] Banshidhar Majhi, Hasan Shalabi, and Mowafak Fathi, FLANN Based Forecasting of S&P 500 Index. Information Technology Journal, Vol. 4, Issue 3, http://www.scialert.net/abstract/?doi=itj.2005.289.292, 2005, pp. 289–292.

[9] Ritanjali Majhi, G. Panda, and G. Sahoo, Development and performance evaluationof FLANN based model for forecasting of stock markets, Expert Systems with Applications, Vol. 36, Issue 3, http://linkinghub.elsevier.com/retrieve/pii/S095 7417408005526, 2009, pp. 6800–6808.

[10] Tong-Seng Quah and Bobby Srinivasan, Improving returns on stock investment through neural network selection, Expert Systems with Applications, Vol. 17, Issue 4, http://linkinghub.elsevier.com/retrieve/pii/S095741749900041X, 1999, pp. 295–30

[11] Luigi Troiano, Pravesh Kriplani, and Irene Díaz, Regression Driven F-Transform and Application to Smoothing of Financial Time Series, IEEE, 2017. 74

[12] Tong-Seng Quah and Bobby Srinivasan, Improving returns on stock investment through neural network selection, Expert Systems with Applications, Vol. 17, Issue 4, http://linkinghub.elsevier.com/retrieve/pii/S095 741749900041X, 1999, pp. 295–301.

[13] Halbert White, Economic prediction using neural networks: the case of IBM daily stock returns, IEEE International Conference on Neural Networks, http://ieeexplore.ieee.org/document/23959/, IEEE, 1988, pp.451–458.

[14] Yauheniya Shynkevicha, T.M. McGinnity, Sonya A. Coleman, and Ammar Belatreche, Forecasting movements of health-care stock prices based on different categories of news articles using multiple kernel learning, Decision Support Systems, Vol. 85, http://linkinghub.elsevier.com/retrieve/pii/S016 7923616300252, 2016, pp. 74–83.

[15] Han Lock Siew and Md Jan Nordin, Regression Techniques for the Prediction of Stock Price Trend,2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), IEEE, 2012.

[16] Chi Ma, Junnan Liu, Hongyan Sun, and Haibin Jin, A hybrid financial time series model based onneural networks, IEEE, 2017.

[17] Lean Yu, Shouyang Wang, and Kin Keung Lai, Mining Stock Market Tendency Using GABased Support Vector Machines, Internet and Network Economics,http://link.springer.com/10.1007/11600930_33, 2005, pp. 336–345.

[18] Fu-Yuan Huang, Integration of an Improved Particle Swarm Algorithm and Fuzzy Neural Network forShanghai Stock Market Prediction, 2008 Workshop on Power Electronics and Intelligent Transportation System.[Online].August 2008, IEEE, http://ieeexplore.ieee.org/document/4634852/, 2008, pp. 242–247.

[19] Ching-hsue Cheng, Tai-liang Chen, and Hiajong Teoh, Multiple-Period Modified Fuzzy TimeSeries for Forecasting TAIEX, Fourth International Conference on Fuzzy Systems and Knowledge Discovery(FSKD 2007), http://ieeexplore.ieee.org/document/4406190/, IEEE, 2007, pp. 2–6. 75

[20] Pei-Chann Chang, Chin-Yuan Fan, and ShihHsin Chen, Financial Time Series Data Forecasting by Wavelet and TSK Fuzzy Rule Based System, Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), http://ieeexplore.ieee.org/document/4406255/, IEEE, 2007, pp. 331– 335.

[21] Lixin Yu and Yan-Qing Zhang, Evolutionary Fuzzy Neural Networks for Hybrid Financial Prediction,IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews), Vol. 35, Issue2, http://ieeexplore.ieee.org/document/1424198/, IEEE, 2005, pp. 244–249.

[22] Chokri Slim, Neuro-fuzzy network based on extended Kalman filtering for financial time series, World Academy of Science, Engineering and Technology, Vol. 15, http://citeseerx.ist.psu.edu/viewdoc/download?d oi=10.1.1.193.4464&rep=rep1&type=pdf, 2006, pp.

134–139.

[23] Gérard Biau and Luc Devroye, Lectures on the Nearest Neighbor Method, Springer, 2015.

[24] Saed Sayad, K Nearest Neighbors - Classification, [Online]. 2017,http://www.saedsayad.com/k_nearest_neighbors .htm, 2017, [Accessed: 10 January 2017].

[25] Dan Morris, Bayes' Theorem: A Visual Introduction for Beginners, Blue Windmill Media, 2017.
[26]James V Stone, Bayes' Rule With R: A Tutorial Introduction to Bayesian Analysis, Sebtel Press, 2016.

[27] Marco Scutari and Jean-Baptiste Denis, Bayesian Networks: With Examples in R, CRC Press, 2014.

[28] Peter Bruce and Andrew Bruce, Practical Statistics for Data Scientists: 50 Essential Concepts, O'ReillyMedia, 2017.

[29]Ciaran Walsh, Key Management Ratios, 4th Edition (Financial Times Series), Prentice Hall, 2009.

[30] Seyed Enayatolah Alavi, Hasanali Sinaei, and Elham Afsharirad, Predict the Trend of StockPrices Using Machine Learning Techniques, IAIEST, 2015