Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

Comparative Analysis of Machine Learning and Deep Learning Algorithms for Predicting Stock Market Trends Using Continuous and Binary Data

Dr .M . KALPANA DEVI BAI, M .Tech,(Ph.D)[1], Ms. YARRAMSETTY PRAVALLIKA[2], Mr. AMKAM HARIBABU[3], Mr. CHATLA RAHUL[4], Mr. DUBBAKU NAGA RAMULU[5], Ms. SHAIK RESHMA SULTANA[6].

#1 Associate Professor, Department Of Computer Science And Engineering,PSCMR College Of Engineering And Technology.

#2 Student , DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,PSCMR College Of Engineering And Technology.

ABSTRACT: Investors have long struggled to pin down the precise nature of stock market movement for a variety of reasons. Finding techniques to make This research primarily aims to demonstrate how deep learning and machine learning algorithms may significantly improve the reliability of trend prediction. We test the varied financials, petroleum, non-metallic minerals, and basic metals markets on the Tehran Stock Exchange. A total of nine machine learning models, including Decision Tree, Random Forest, Naïve Bayes, K-Nearest Neighbors, Logistic Regression, and Artificial Neural Network, ANN, are compared in this research using two robust deep learning methods, Adaptive Boosting and eXtreme Gradient Boosting. Our input values, which consist of ten technical indicators, are a byproduct of ten years of data and have a dual purpose. Using stock trading values as continuous data is one stage in computing the indicators. Before using the indicators, the next step is to convert them to binary data. There are three input-way-based criteria that are used to assess each prediction model. When presented with continuous data, LSTM and RNN perform better than other prediction models, according to the assessment findings. Additionally, the findings demonstrate that both deep learning approaches excel at assessing binary input; yet, the effectiveness of the models is clearly enhanced by the second method, thus the disparity is diminished.

Keywords: LSTM, ML, STOCK MARKET.

1. INTRODUCTION

A lot of recent research in business has focused on the challenging and time-consuming task of predicting the future changes in stock values. Market movement forecasting for stock prices is a hot topic among academics, businesses, and curious people who think that the past and present dictate what happens next (Kim, 2003). However, due to its complexity, financial data is infamously hard to predict. Prices in the market may be hard to predict, says Fama's efficient market hypothesis (EMH) (1990). Financial data and the market are thought to be linked by the EMH. It goes on to say that market prices reflect all available information and that newly available

information is the only thing that can cause prices to change. The EMH states that inventors can never make money by betting on stock prices, since stock prices are always stable. Additionally, stock prices do not behave randomly, as stated by several authors [(Gallagher and Taylor, 2002; Walczack,
2001; Kavusssanos and Dockery, 2001; Lakonishok et.al, 1994; O'Connor et. al., 1997; Lo and MacKinlay, 1997; Kirt and Malaikah, 1992; Lo and MacKinlay, 1988])]. Several studies have attempted to forecast stock prices; for example, Subha and Nambi (2012), Qian and Rasheed (2007), Fama and French (1992), Cochrane (1988), Campbell (1987), Basu (1977), and Chen et al. (1986). Aside from being a place where stocks may be traded, the stock market also has other aspects like the closing price, which is the most essential factor in deciding a company's worth the following trading day. The interplay and unique behavior of the many elements that affect stock price variations throughout time are complex and multifaceted. Stock price predictions have taken into account a range of economic variables, including political stability and other unanticipated events, according to many studies (Ou, P. and Wang, H., 2009; Fama and French, 1993; Cochrane, 1988; Campel, 1987; Chen. et.al. 1986). Table 1 provides a concise overview of the key points discussed in this study and how they impact changes in stock prices.

In order to uncover changes in stock prices, data mining technology sorts through massive amounts of financial and economic data. When you need to examine how different pieces of information interact with one another over time, temporal stock market mining is a must-have tool. To foretell how much stocks will be worth in the future, analysts rely on derived data, fundamental data, and purely technical data. Firm actions and market conditions provide the fundamental data, while historical stock data provides the technical data. By combining data mining categorization methods with stock prediction, it is possible to use past data to forecast the future values of unknown entities of firms' stocks. This forecast makes use of a number of categorization techniques, including k-Nearest Neighbors (kNN), prediction trees, evolutionary algorithms, neural networks, and regression. Separating a dataset into a training set and a testing set is a common approach in classification algorithms. To assess items, the k-NN algorithm compares them to the training dataset using similarity measures. An n- attribute record is represented by a data entity. When given an unlabeled record, kNN will look for the k records in the training set that are the most similar to it in order to make a class label prediction.

2. LITERATURE SURVEY

[1] Methods for identifying stock market trends over time that combine neural networks with genetic algorithms Featuring Hyun-jung Kim and Kyung-shik Shin

The goal of this study is to find out how well a mixed method that uses ANNs for time series characteristics and genetic algorithms (GAs) for stock market prediction jobs performs. We take a closer look at TDNNs and ATNNs, or adaptive time delay neural networks. In addition to the few control variables included in an ANN design, ATNN and TDNN provide an additional estimate of the amount of time delays by including time-delayed network connections into a multi-layer feed-forward network. One more layer is added to this design with every iteration. This research introduces a generic method for estimating ATNN and TDNN designs using a combination of statistical methods, trial and error, and various heuristics. We optimize the network architecture parameters and the amount of time delays for both the ATNN and TDNN models simultaneously using GAs. This is because these variables do not always improve the modeling results when used alone. The findings demonstrate that the suggested integrated method outperforms the conventional ATNN, TDNN, and RNN methods in terms of accuracy. Yes, Elsevier In 2006, the magazine was under the management of B.V. The terms genetic algorithms, time series prediction, stock market prediction, adaptive time delay neural networks, and time delay neural networks may all be grouped together under a single umbrella. 1. Background Businesses, investors, and speculators face a formidable obstacle when trying to predict stock prices. They base their efforts on the assumption that current and historical facts and occurrences provide the groundwork for future forecasts. But the most challenging and "noisiest" signals to predict are financial time series. The majority of the early research on stock market forecasting relied on statistical methods. However, when non-linearities are present in the dataset, investigations that rely only on classic statistical methods for prediction encounter difficulties. When it comes to dealing with non-linear situations, a hidden-layer ANN is much superior than statistical approaches. An ANN technique known as back-propagation neural networks (BPNs) has several uses, including pattern recognition, forecasting, and classification. When it comes to learning input-output mappings of time-independent spatial or static patterns, BPN has several serious limitations. One time-based approach exploits time-delayed connections, while the other makes advantage of recurrent links, to circumvent this restriction. Combining genetic algorithms (GAs), time delay neural networks (TDNNs), and adaptive time delay neural networks (ATNNs) is a novel strategy that has emerged as a promising method for tackling problems related to stock market prediction. In order to provide ANNs with memory, TDNN and ATNN offer an additional estimate of the amount of delays via time-delayed links for each unit. In addition to the many variables that may be controlled in an ANN design, such as the number of hidden nodes, network topologies, activation function choices, etc. This research introduces a generic method for estimating

ATNN and TDNN designs using a combination of statistical methods, trial and error, and various heuristics. To properly characterize the temporal patterns, it is necessary to use a method that picks the amount of time delays and TDNN and ATNN jointly, rather than executing each part of the network design individually. Consequently, the performance of the ATNN and TDNN models will be improved. With GAs, the ATNN and TDNN models may optimize the network architectural factors and delay amounts simultaneously. We show that our approach works by applying it to the realm of stock market predictions. Here is how the remainder of the paper is structured: Prior studies on stock market prediction are summarized in Section 2. Section 3 presents an ANN-proposed time series feature. In Section 4, GAs are described in simple terms. Section 5 details the studies and data needed to develop the GA-ATNN and GA-TDNN hybrid methods. Section 6 presents a summary and analysis of the empirical data.

The study's limitations and findings are detailed in the third part.

[2] An Artificial Neural Network-Based System for Pattern Recognition in Stocks Visit icst.pku.edu.cn to learn more about the XinyuGuo Institute of Computer Science and Technology at Peking University in Beijing, China 100871. Xun Liang Institute of Computer Science and Technology at Peking University, Beijing, China 100871
Liang Xun@icst.pku.edu.cn Capital of China 100871 China National University's Xiang Li Institute of Computer Science and Technology For correspondence, please use lixiang@icst.pku.edu.cn.

According to recent studies, stock trends could provide valuable insights for estimating future stock values. Computer methods are becoming more popular in the area of stock price pattern recognition as a consequence of time restrictions that prevent human researchers from discovering all underlying patterns in price time series. There are now two main categories of algorithms used to detect patterns in stock prices: rulematching algorithms and template-matching algorithms. However, both algorithms have significant limitations, such as a lack of learning capability and an over-reliance on domain knowledge. To tackle these challenges, we provide a technique in this paper that uses artificial neural networks (ANNs) to identify stock price patterns. High precision pattern identification and learning of pattern characteristics are both shown by the findings. 1. Context Technical
pattern
analysis is one of the many approaches to technical analysis that researchers have examined in depth as a stock investment technique. Analysts who have spent time monitoring the stock market have noticed two primary kinds of technical patterns: continuation patterns and reversal patterns. Better investment judgments may be made by investors by observing these tendencies. In the case of a continuation pattern, the stock price is expected to continue in its current trend, whereas a reversal pattern indicates a change in direction. There are sixty-three noteworthy technical patterns, as stated in [1]. A total of eighteen prevalent technical patterns—ten continuation patterns and eight reversal patterns—are the focus of this paper's research.

[3] The prediction of stock market trends using supervised machine learning algorithms: a comparative study Kumar Indu Dogra Kiran ChetnaUtreja is on the verge ofThe following email addresses are associated with Premlata Yadav: kumarindu22@gmail.com, kirandogra1@gmail.com, and chetna4395@gmail.com; she is a computer engineering department faculty member at NIT Kurukshetra.

yadav.premlata15@gmail.com

Many variables impact business pricing, making it a complicated and challenging undertaking to predict. Our research gets around these problems by using machine learning to forecast stock prices. In this article, we compare five different models and assess their accuracy in predicting the future moves of the stock market. These models are built on top of five supervised learning methods: Naive Bayes, Support

Vector Machine (SVM), Random Forest, K-Nearest Neighbor (KNN), and Softmax. If your dataset is big, use Random Forest; if it's little, use Naive Bayesian Classifier, which is the most difficult method. In addition, the findings show that when the quantity of technical signals decreases, the accuracy of each algorithm also diminishes. Machine learning is associated with terms such as classifier, random forest, support vector machine, k-nearest neighbor, naíve bayes, and softmax. Initial Part: Synopsis For emerging nations like India to have rapid economic development, the stock market is very essential. Our nation's and other emerging nations' futures are precariously balanced, much like the stock market. Strong economic growth is possible if the stock market rises. When stock prices drop, the GDP of a nation decreases. To restate, the nation's economic growth is proportional to the stock market's success. Due to its inherent unpredictability, the stock market attracts only 10% of investors. This is true on a global scale. Many people see stock trading as being similar to gambling. Therefore, increasing awareness of this may assist in clearing this misconception. Methods used to forecast the stock market could entice new investors while strengthening ties with existing ones. Machine learning has become popular due to its capacity to identify stock trends in massive datasets that mirror the ever- changing dynamics of stock prices. Here, we used supervised learning methods to foretell how the stock market would behave in the future. The structure of the article is described in full below.

3. PROPOSED SYSTEM The author of this work uses four stock datasets, each of which contains both continuous (normal) values and binary (converted from stock values to binary by indicators that check) data, to assess the efficacy of several machine learning algorithms in predicting stock prices. We shall add 1 to the dataset if the stock's price has decreased since its last update, and -1 otherwise.

3.1 IMPLEMENTATION

Data Collection and Preprocessing Module:

Collect the selected industries' stock market data from the Tehran Stock Exchange. During the preparation phase, address issues with data quality such as outliers and missing values.
In order to train the model, you must convert the data into a format that takes into consideration both the continuous and binary versions of technical indicators.

Feature Engineering Module:

Using tools like trend-predicting technical indicators, extract actionable insights from the raw data.
It may be required to transform continuous attributes into binary representations in order to do binary data analysis.

Model Selection and Training Module:

Configure many machine learning algorithms, including Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression, and ANN.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

Make a comparison between two deep learning models—RNN and LSTM—by creating your own.
Put each model through its paces on the supplied dataset before putting it to the test on both continuous and binary data.

Evaluation Metrics Module:

You may create assessment criteria like recall, accuracy, precision, F1-score, and others to see how each model does. Find the metrics for the models by using their predictions on the test data.

Results Analysis Module:

Determine and compare the performance of deep learning and ML models using both continuous and binary data types. Applying the results, identify the top models for forecasting future stock market movements.

Visualization Module:

4. EXPERIMENTAL RESULTS

Present the results in an effective manner by visualizing them using plots, charts, and tables.
Show how different models and data formats function differently.

Optimization and Fine-Tuning Module:

To get even better results, tweak the hyperparameters of the top models. Use the results of the preliminary study to fine-tune the models.

Deployment Module:

To predict stock market movements in batch or real-time, implement the selected model or models.
Install monitoring and updating mechanisms to monitor the deployed models.

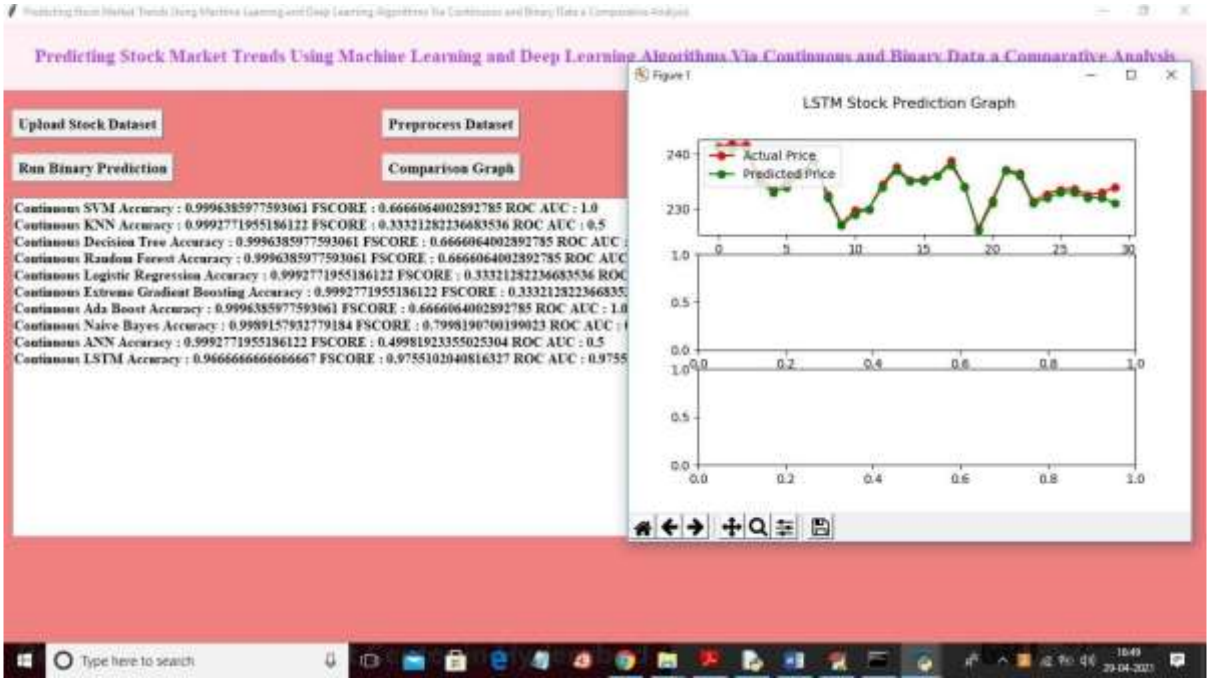
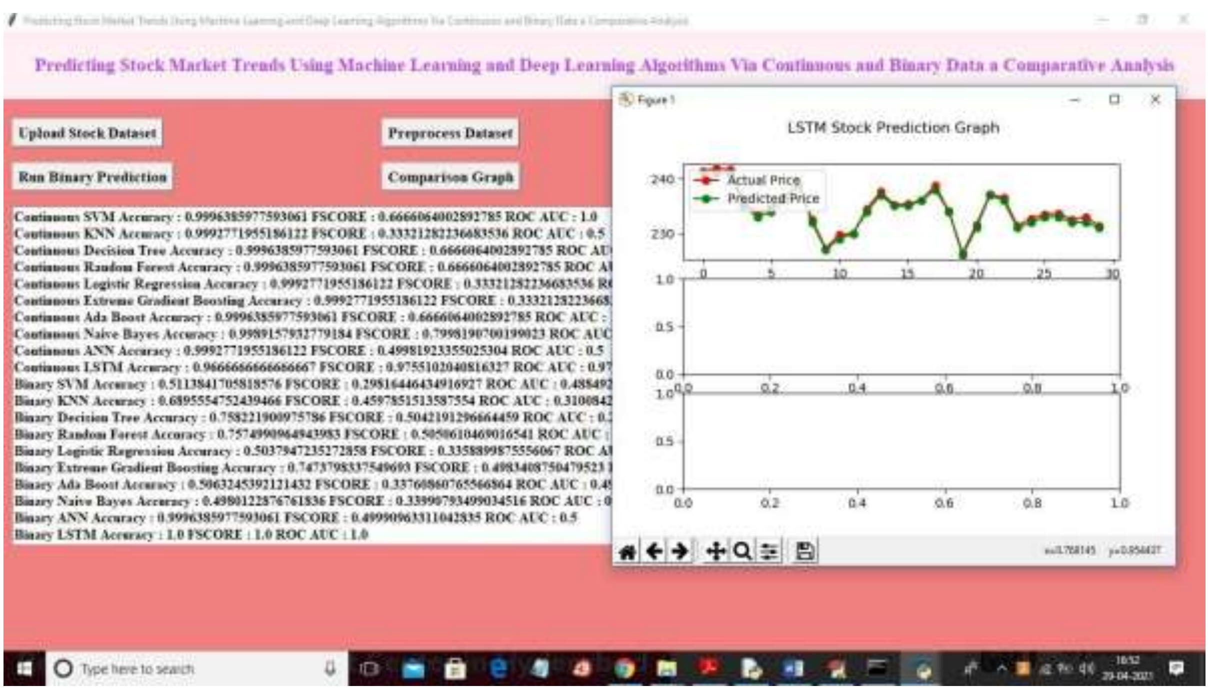


Fig 1: In the text section of the above screen, you can see the accuracy, FSCORE, and ROC AUC values for all the methods that use continuous data. Both the stock price and the number of days are shown on the graph up top. The green line shows the expected price and the red line shows the actual pricing. Good LSTM performance is shown by the small gap between the two



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

lines. Press the "Run Binary Prediction" button to execute the prediction after converting the dataset to binary values.

Fig 2: Binary prediction likewise yields the best results on the screen shown before, and the LSTM accuracy is 1.0, which signifies 100% accuracy, as can be seen in the text section. A graph comparing each method is available when you click the "Comparison Graph" button.

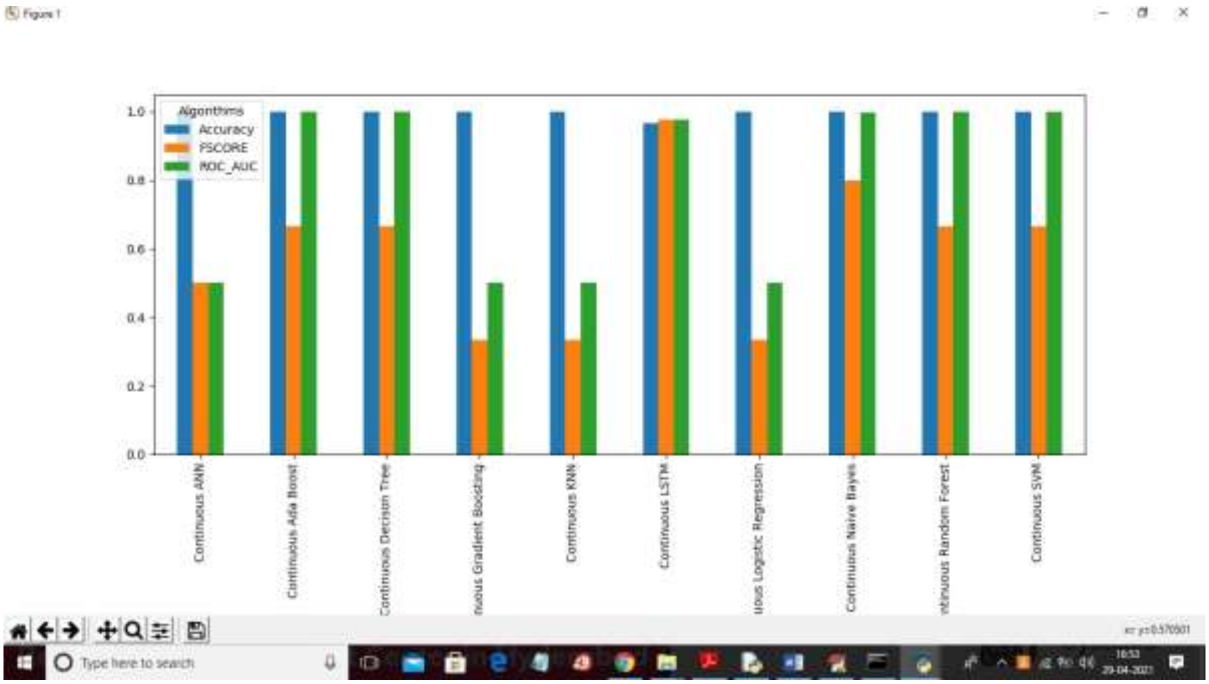
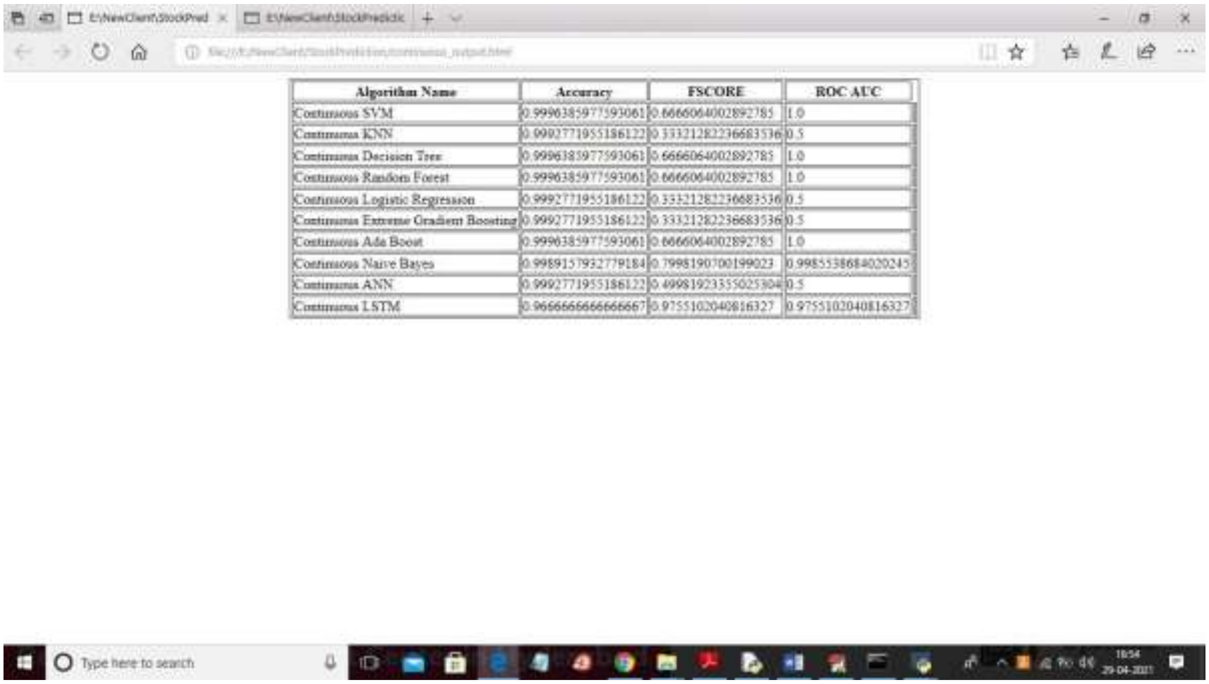


Fig 3: You can see that ANN and LSTM perform better on continuous data in the following graph; to see the results side by side, click the "View Comparison Table" button.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024



Algorithm Name	Accuracy	FSCORE	ROC AUC
Continuous SVM	0.9996385977593061	0.6666064002892785	1.0
Continuous KNN	0.9992771955186122	0.3332128223668353 6	0.5
Continuous Decision Tree	0.9996385977593061	0.6666064002892785	1.0
Continuous Random Forest	0.9996385977593061	0.6666064002892785	1.0
Continuous Logistic Regression	0.9992771955186122	0.3332128223668353 6	0.5
Continuous Extreme Gradient Boosting	0.9992771955186122	0.3332128223668353 6	0.5
Continuous Ada Boost	0.9996385977593061	0.6666064002892785	1.0
Continuous Naive Bayes	0.9989157932779184	0.7998190700199023	0.9985538684020245
Continuous ANN	0.9992771955186122	0.4998192335502530 4	0.5
Continuous LSTM	0.9666666666666667	0.9755102040816327	0.9755102040816327

Fig 4: Viewing continuous data on the screen above The binary data result is below, and the LSTM FSCORE is high.

4. CONCLUSION

The primary goal of this study was to develop an algorithm for predicting stock market movements in the future using ML

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

and DL techniques. This dataset was compiled from four different market segments on the Tehran Stock Exchange: diversified financials, petroleum, non-metallic minerals, and basic metals. It is based on ten years of historical records and has ten technical features. To add to that, ANN, Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, and Logistic Regression were among the nine machine learning models used as predictors, along with two deep learning approaches, LSTM and RNN. Assuming both continuous and binary data types for model input values, we used three different categorization measures for our assessments. In contrast to continuous data, which significantly degraded model performance, binary data significantly enhanced it. In every case, we discovered that the deep learning algorithms using RNN and LSTM yielded the best models.

REFERENCES

[1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, GiridharMaji, Narayan C. Debnath, Soumya Sen

[2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman - Stock price prediction using LSTM, RNN and CNN-sliding window model - 2017.

[3] MurtazaRoondiwala, Harshal Patel, Shraddha Varma, “Predicting Stock Prices Using LSTM” in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, “An innovative neural network approach for stock market prediction”, 2018

[5] IshitaParmar, Navanshu Agarwal, SheirshSaxena, Ridam Arora, Shikhin Gupta, HimanshuDhiman, LokeshChouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, 37Pune.SavitribaiPhule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.

[8] V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India - Stock Market Prediction Using Machine Learning.

[9] Asset Durmagambetov currently works at the mathematics, CNTFI- 'The Riemann HypothesisMillennium Prize Problems' - stock market predictions.

[10] Mariam MoukalledWassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut - Automated Stock Price Prediction Using Machine Learning.