International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:5 Volume No.9 September - October – 2025
DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data a Comparative Analysis

Agisam Bilwatej Kumar[1], M C Bhanu Prasad[2]
1M.Tech Student, Dept.of CSE, Tadipatri Engineering college,Tadipatri, Andhra Pradesh, India
2Associate professor & vice principal , Dept.of CSE, Tadipatri Engineering college,Tadipatri, Andhra Pradesh, India

Abstract
The nature of stock market movement has always been ambiguous for investors because of various influential factors. This study aims to significantly reduce the risk of trend prediction with machine learning and deep learning algorithms. Four stock market groups, namely diversified financials, petroleum, non-metallic minerals and basic metals from Tehran stock exchange, are chosen for experimental evaluations. This study compares nine machine learning models (Decision Tree, Random Forest, Adaptive Boosting (Adaboost), eXtreme Gradient Boosting (XGBoost), Support Vector Classifier (SVC), Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression and Artificial Neural Network (ANN)) and two powerful deep learning methods (Recurrent Neural Network (RNN) and Long short-term memory (LSTM). Ten technical indicators from ten years of historical data are our input values, and two ways are supposed for employing them. Firstly, calculating the indicators by stock trading values as continues data, and secondly converting indicators to binary data before using. Each prediction model is evaluated by three metrics based on the input ways. The evaluation results indicate that for the continues data, RNN and LSTM outperform other prediction models with a considerable difference. Also, results show that in the binary data evaluation, those deep learning methods are the best; however, the difference becomes less because of the noticeable improvement of models' performance in the second way.
Keywords: financials, SVC, ANN, LSTM

Introduction

The task of stock prediction has consistently presented a difficult challenge for finance professionals and statisticians. This prediction is primarily based on the assumption that stocks that are expected to appreciate in value will be purchased, while stocks that are expected to decline in value will be sold. In general, there are two methods for predicting the stock market. Fundamental analysis is one of them and is predicated on a company's technique and fundamental information, such as market position, expenses, and annual development rates. The second method is technical analysis, which emphasizes the analysis of historical stock prices and values. In order to forecast future prices, this analysis employs historical charts and patterns.
In the past, financial specialists typically foretold stock markets. Nevertheless, the advancement of learning techniques has prompted data scientists to address prediction problems. In an effort to optimize prediction models and enhance their precision, computer scientists have implemented machine learning methodologies. The subsequent phase in the enhancement of prediction models with superior performance was the implementation of deep learning [3&4]. Data scientists typically encounter challenges when attempting to develop a predictive model, as stock market prediction is rife with obstacles. Two primary obstacles are the instability of the stock market and the correlation between investment psychology and market behavior: complexity and nonlinearity.

It is evident that the trend of stock markets is consistently influenced by unpredictable factors, such as the public image of companies or the political situation of countries. Consequently, the trend of stock values and the index can be predicted by effectively preprocessing the data obtained from stock values and implementing appropriate algorithms. Machine learning and deep learning methodologies can assist investors and speculators in making informed judgments regarding stock market prediction systems. The objective of these methods is to autonomously identify and learn patterns among vast quantities of information. The algorithms are capable of effectively self-learning and can address the prediction of price fluctuations to enhance trading strategies. Six. Numerous methodologies have been enhanced in recent years to anticipate stock market trends. Hassan et al. [7] proposed the implementation of a model combination that combines Genetic Algorithms (GA), Artificial Neural Networks, and Hidden Markov Model (HMM). The objective of this model is to convert the daily stock prices into independent sets of values that serve as input to the HMM. Huang et al. evaluated the weekly trend of the NIKKEI 225 index to determine the predictability of financial trends using the SVM model. Eight. Their objective was to conduct a comparison of SVM, Linear Discriminant Analysis, Elman Backpropagation Neural Networks, and Quadratic Discriminant Analysis. The results suggested that SVM was the most effective classifier method. Sun et al. proposed a novel financial prediction algorithm that is based on the SVM ensemble. Nine. By considering both individual performance and diversity analysis, the method for selecting the base classifiers of an SVM ensemble from a list of candidates was put forth. The final results indicated that the SVM ensemble was significantly superior to the individual SVM in terms of classification. Ou et al. [10] implemented ten data mining techniques to forecast the value trends of the Hang index in the Hong Kong stock market. The methods include Bayesian classification, SVM, neural network, K-nearest neighbor, and tree-based classification. The results suggested that the SVM outperformed other predictive models. Liu et al. [11] predicted the price fluctuation of a developed Legendre neural network by analyzing the antecedent data on stock values and presuming the positions and decisions of investors. Additionally, they investigated a random function (time strength) within the forecasting model. To compare the results of the morphological rank linear forecasting approach with those of the time-delay added evolutionary forecasting approach and multilayer perceptron networks, Araújo et al. [12] proposed the approach.
It is evident from the aforementioned research background that each of the algorithms is capable of effectively resolving stock prediction issues. Nevertheless, it is crucial to acknowledge that each of them is subject to specific constraints. The representation of the input data and the prediction approach both influence the prediction results. Additionally, the prediction models' accuracy can be significantly enhanced by identifying only the most prominent features as input data, rather than all features.

A recent research endeavor involves the utilization of deep learning algorithms and tree-based ensemble methods to forecast stock and stock market trends. Tsai et al. [13] implemented heterogeneous and homogeneous ensemble classifiers due to their utilization of bagging and majority vote methods. Additionally, they evaluate the performance of models by analyzing macroeconomic features and financial ratios from the Taiwan stock market. The results indicated that ensemble classifiers were preferable to single classifiers in terms of prediction accuracy and investment returns. Ballings et al. [14] conducted a comparison of the performance of AdaBoost, Random Forest, and kernel factory with singular models that included SVM, KNN, Logistic Regression, and ANN. They anticipate the prices of European companies for the upcoming year. The final results indicated that Random Forest outperformed all other models. In order to predict the increase or decrease in stock based on previous values, Basak et al. [15] implemented the XGBoost and Random Forest algorithms for the classification problem. The results indicated that the prediction performances of numerous companies have improved in comparison to their current ones. Weng et al. [16] enhanced four ensemble models—the boosting regressor, bagging regressor, neural network ensemble regressor, and random forest regressor—to reliably predict the stock market for one month in advance by analyzing macroeconomic indicators. In fact, an additional objective was

DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

to demonstrate that macroeconomic characteristics are the most effective stock market predictors through the use of a hybrid LSTM approach.
Long et al. [17] evaluated the movement of stock prices by utilizing a deep neural network model with public market data and transaction records. This was achieved by utilizing deep learning algorithms. The experimental results demonstrated that bidirectional LSTM was capable of predicting the stock price for financial decisions, and it achieved the highest performance among other prediction models. Rekha et al. [18] utilized stock market data to compare the results of two algorithms with their actual performance using CNN and RNN. Pang et al. [19] endeavored to enhance an advanced neural network method in order to obtain more accurate stock market predictions. They suggested the use of an LSTM with an embedded layer and an LSTM with an automatic encoder to assess the movement of the stock market. The LSTM with imbedded layer demonstrated superior performance, with an accuracy of 57.2 and 56.9% for the Shanghai composite index, respectively. Kelotra and Pandey [20] employed the deep convolutional LSTM model as a predictor to effectively analyze stock market fluctuations. The model was trained using the Rider-based monarch butterfly optimization algorithm, resulting in a minimal MSE and RMSE of 7.2487 and 2.6923, respectively. An approach for stock market index forecasting was proposed by Baek and Kim [21]. This approach consisted of a prediction LSTM module and an overfitting prevention LSTM module. The proposed model's exceptional forecasting accuracy was verified by the results. In comparison to the model that lacks an overfitting prevention LSTM module. In order to enhance a novel stock market prediction model, Chung and Shin [22] implemented a hybrid approach that combined LSTM and GA. The benchmark model was outperformed by the hybrid model of LSTM network and GA, as indicated by the final results.
In summary, the aforementioned literature has been the subject of prior research, which has frequently focused on the detection of stock index or value movements by combining recent machine learning methods with macroeconomic or technical features, without taking into account the appropriate preprocessing methods.
The Tehran Price Index has experienced significant growth in recent decades, which has contributed to the recent surge in popularity of Iran's stock market. Additionally, the majority of state-owned enterprises are being privatized in accordance with the general policies of Article 44 of the Iranian constitution. Individuals are permitted to purchase shares of newly privatized firms under specific conditions. This market is distinguished from other countries' stock markets by its dealing price limitation of ±5% of the opening price of the day for all indexes. This restriction prevents abnormal market fluctuations and disperses market shocks, political issues, and other factors over a specific period, thereby enhancing market smoothness. Nevertheless, the impact of fundamental parameters on this market is relatively significant, and the task of predicting future movements is not straight forward.
This research focuses on the process of predicting future trends for stock market groups, which is essential for investors. In spite of the substantial progress that has been made in the Iranian stock market in recent years, there has been a dearth of research on the predictions and movements of stock prices using innovative machine learning methods.
This paper focuses on the comparison of the prediction performance of nine machine learning models (Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression, and ANN) and two deep learning methods (RNN and LSTM) in predicting stock market movement. Our models utilize ten technical indicators as input values. In order to examine the impact of preprocessing, our study incorporates two distinct input methods: binary data and continuous data. The former utilizes stock trading data (open, close, high, and low values), while the latter employs a preprocessing phase to convert continuous data to binary one. The inherent properties of the market determine the specific possibility of an upward or downward movement for each technical indicator. The performance of the aforementioned models is contrasted for both approaches using three classification metrics. The optimal tuning parameter for each model (excluding Naïve Bayes and Logistic Regression) is reported. The Tehran stock exchange provides

DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

ten years of historical data for four stock market groups (diversified financials, petroleum, non-metallic minerals, and fundamental metals) that are essential for investors. All experimental experiments are conducted using this data. We are of the opinion that this study is a novel research paper that integrates a variety of machine learning and deep learning techniques to enhance the prediction of the trend and movement of stock groups.

Literature Survey

1) Kim (2003) — *SVM for financial time series / direction*

Citation / source: Kim, 2003. MQL5
Problem: Binary prediction of index direction (up / down) and price forecasting experiments.
Methods: Support Vector Machine (classification & regression) vs ANN baseline.
Data: KOSPI (Korean index) daily series (historical windowing).
Key findings: SVM performed competitively and often better than backprop ANN for directional forecasting. MQL5
Merits: Early rigorous application of SVM, clear experiments comparing classification vs regression framing.
Demerits / gaps: Limited feature engineering (mostly price lags), small set of markets; limited exploration of deep architectures or alternative data (news/sentiment).

2) Patel et al. (2015) — *Comparative ML models on Indian indices*

Citation / source: Patel et al., 2015 — predicting stock/index using ANN, SVM, Random Forest, Naïve Bayes. ResearchGateScienceDirect
Problem: Both continuous (index level) and directional forecasting experiments.
Methods: Several ML classifiers/regressors, technical indicators as inputs.
Data: Indian market indices (e.g., CNX Nifty) — multi-year daily data.
Key findings: Ensemble/tree-based (RF) and SVM often beat basic ANN and naive baselines in standard metrics. ScienceDirectResearchGate
Merits: Practical, comparative evaluation across classic ML methods.
Demerits / gaps: Not deep-learning focused; evaluation sometimes limited to accuracy/MSE without trading/financial metrics.

3) Fischer & Krauss (2018) — *LSTM / deep nets for directional trading signals*

Citation / source: Fischer & Krauss, 2018 (LSTM on S&P 500 constituents). IDEAS/RePEceconpapers.repec.org
Problem: Binary one-day-ahead direction (outperform market or not) across many S&P500 stocks.
Methods: LSTM deep networks, gradient-boosted trees, random forests, ensembles; training on lagged returns of all stocks, careful survivor-bias elimination.
Data: S&P 500 constituents (1992–2015), daily returns.
Key findings: LSTMs / deep models can generate useful directional signals and can be combined into statistical-arbitrage strategies; careful cross-sectional training helps. IDEAS/RePEc
Merits: Large-scale, long-horizon dataset; robust experimental design; focuses on **binary directional** signals used in trading.
Demerits / gaps: Directional framing misses continuous price error analysis; practical transaction costs / market impact not always fully modelled.

4) Bao, Yue & Rao (2017) — *Stacked Autoencoders + LSTM for price forecasting*

Citation / source: Bao et al., 2017 (PLOS ONE). PMC
Problem: Continuous price forecasting (next-day closing price).
Methods: Wavelet Transform (denoising) → Stacked Autoencoders for feature extraction → LSTM for forecasting.
Data: Multiple market indices and index futures; several years.
Key findings: Hierarchical deep features (SAE) combined with LSTM improved price RMSE vs baselines. PMC
Merits: Good pipeline: denoising → representation learning → temporal model.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:5 Volume No.9 September - October – 2025
DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

Demerits / gaps: Focuses on prediction error (RMSE) but less on profitability / directional accuracy; may be sensitive to hyperparameters.

5) Di Persio & Honchar (2016) — *ANN architectures comparison*

Citation / source: Di Persio & Honchar, 2016 (comparative study of ANN architectures). COREAcademia

Problem: Both directional and price forecasting across different ANN / RNN variants.
Methods: Multi-layer perceptrons, RNN, LSTM, GRU comparisons across horizons.
Data: Typical benchmark stocks (e.g., Google) and indices.
Key findings: RNN variants (LSTM/GRU) outperform simple MLPs for medium-term horizons; architecture choice depends on forecast horizon. CORE
Merits: Useful comparison of architectures and horizons.
Demerits / gaps: Limited exploration of external data (news, sentiment); experimental reproducibility sometimes limited.

6) CEEMD-CNN-LSTM hybrids (2020ish) — *frequency-domain + hybrid DL*

Citation / source: CEEMD-CNN-LSTM studies and variants (hybrid frequency decomposition + CNN + LSTM). ScienceDirect
Problem: Continuous price prediction enhanced by decomposing signal into frequency components.
Methods: CEEMD/EMD/FFT → CNN (feature extraction) → LSTM (temporal modeling).
Data: Multiple stock indices; rolling-window evaluation.
Key findings: Frequency decomposition + hybrid DL often reduces noise and improves RMSE vs single models. ScienceDirect
Merits: Tackles non-stationarity and noise explicitly.
Demerits / gaps: Added complexity; risk of overfitting; less focus on trading metrics.

Existing System

In this study, ten years of historical data of four stock market groups (diversified financials, petroleum, non-metallic minerals and basic metals) from November 2009 to November 2019 is employed, and all data is gained from www.tsetmc.com website. Figures 1-4 show the number of increase or decrease cases for each group during ten years.

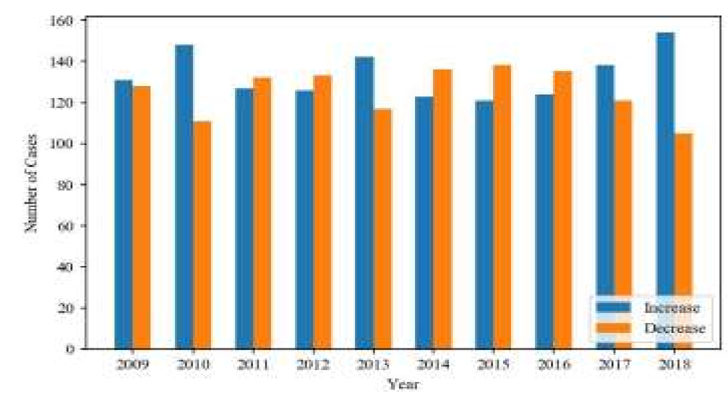


FIGURE 1. The number of increasing and decreasing cases (trading days) in each year for the diversified financials group.

Proposed System

In the proposed system, the system concentrates on comparing prediction performance of nine machine learning models (Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression and ANN) and two deep learning methods (RNN and LSTM) to predict stock market movement. Ten technical indicators are utilized as inputs to our models. The proposed

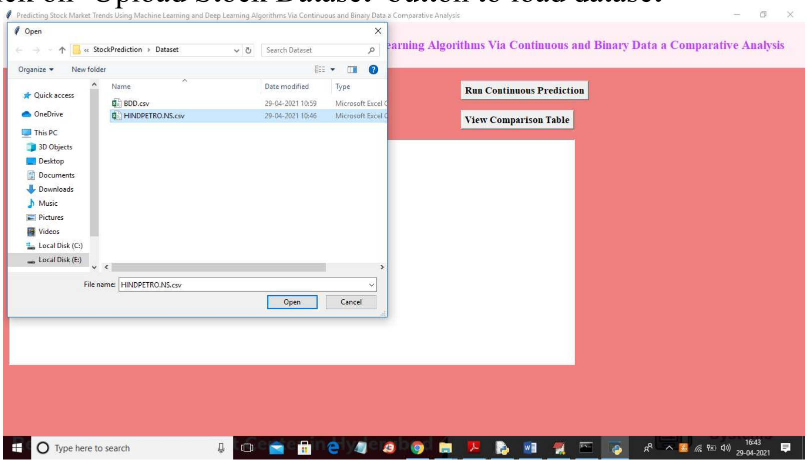
International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:5 Volume No.9 September - October – 2025
DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

study includes two different approaches for inputs, continuous data and binary data, to investigate the effect of preprocessing; the former uses stock trading data (open, close, high and low values) while the latter employs preprocessing step to convert continuous data to binary one. Each technical indicator has its specific possibility of up or down movement based on market inherent properties. The performance of the mentioned models is compared for the both approaches with three classification metrics, and the best tuning parameter for each model (except Naïve Bayes and Logistic Regression) is reported. All experimental tests are done with ten years of historical data of four stock market groups (petroleum, diversified financials, basic metals and non-metallic minerals), that are totally crucial for investors, from Tehran stock exchange. We believe that this study is a new research paper that incorporates multiple machine learning and deep learning methods to improve the prediction task of stock groups' trend and movement.

Results

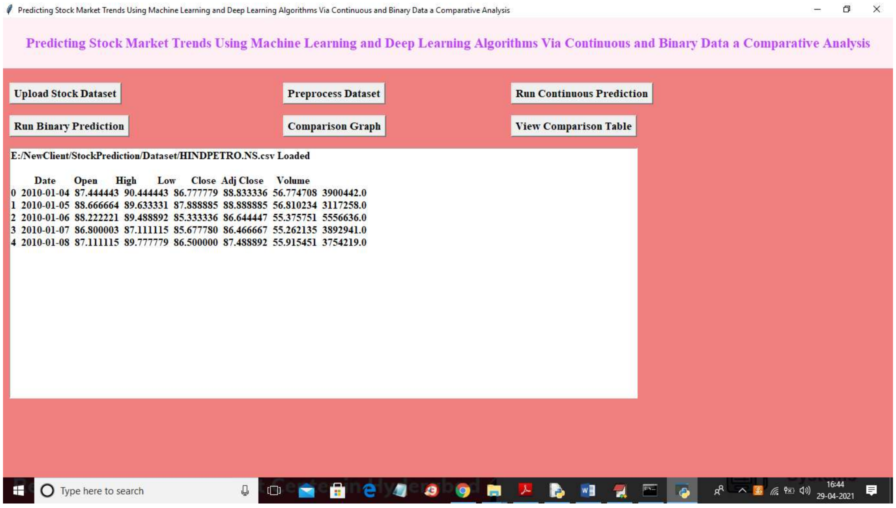
To run project double click on 'run.bat' file to get below screen

In above screen click on 'Upload Stock Dataset' button to load dataset



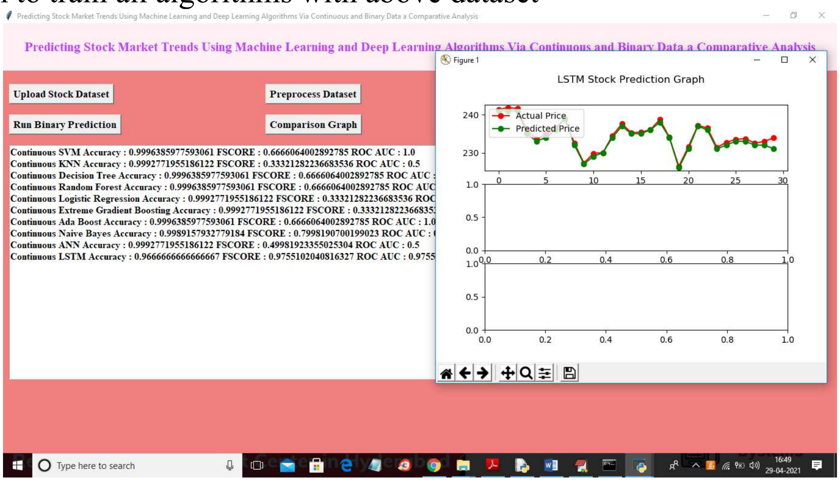
In above screen selecting and uploading "petrol" dataset and then click on 'Open' button to get below screen

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:5 Volume No.9 September - October – 2025
DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621



In above screen dataset loaded and dataset contains some missing values so to remove missing values and to split dataset into train and test part so click on 'Preprocess Dataset' button to get below screen

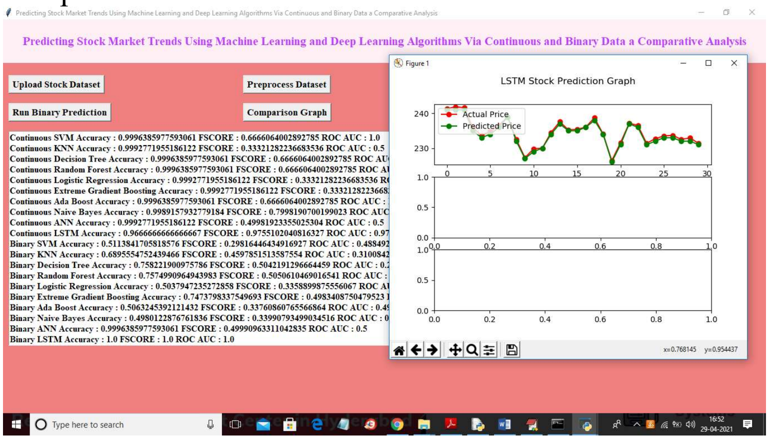
In above screen dataset contains total 2797 records and application using 2797 records for training and 30 records for testing and now train and test data is ready and now click on 'Run Continuous Prediction' button to train all algorithms with above dataset



In above screen in text area we can see accuracy, FSCORE and ROC_AUC values for all algorithms using continuous data and in above graph we can see x-axis represents number of days and y-axis represents stock price and red line represents actual price and green line represents predicted price and we can see there is close difference between actual and predicted so LSTM

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:5 Volume No.9 September - October – 2025
DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

performance is good and now click on 'Run Binary Prediction' button to convert dataset into binary values and then perform prediction



In above screen binary prediction also giving best result and in text area we can see LSTM accuracy is 1.0 which means 100% accurate. Now click on 'Comparison Graph' button to get graph between all algorithms

In above graph for continuous data ANN and LSTM is giving better result and now click on 'View Comparison Table' button to get below screen

Conclusion

The objective of this investigation was to predict stock market fluctuations using deep learning and machine learning algorithms. The dataset was constructed from ten years of historical records and ten technical features, and it was based on four stock market groups from the Tehran stock exchange: diversified financials, petroleum, non-metallic minerals, and fundamental metals. Additionally, nine machine learning models (Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression, and ANN) and two deep learning methods (RNN and LSTM) were implemented as predictors.. We considered two methods for supplying input values to models: continuous data and binary data. Additionally, we implemented three classification metrics for evaluation purposes. The efficacy of models was significantly enhanced when binary data was employed in place of continuous data, as demonstrated by our experimental work. In fact, our models were superior in both approaches due to the use of deep learning algorithms (RNN and LSTM).

References

[1] Murphy, John J. Technical analysis of the financial markets: A comprehensive guide to trading methods and applications. Penguin, 1999.

DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

[2] Turner, Toni. A Beginner's Guide To Day Trading Online 2nd Edition. Simon and Schuster, 2007.
[3] Maqsood, Haider, et al. "A local and global event sentiment based efficient stock exchange forecasting using deep learning." International Journal of Information Management 50 (2020): 432-451.
[4] Long, Wen, Zhichen Lu, and Lingxiao Cui. "Deep learning-based feature engineering for stock price movement prediction." Knowledge-Based Systems 164 (2019): 163-173.
[5] Duarte, Juan Benjamin Duarte, Leonardo HernánTalero Sarmiento, and Katherine Julieth Sierra Juárez. "Evaluation of the effect of investor psychology on an artificial stock market through its degree of efficiency." Contaduría y Administración 62.4 (2017): 1361-1376.
[6] Lu, Ning. "A machine learning approach to automated trading." Boston, MA: Boston College Computer Science Senior Thesis (2016).
[7] Hassan, Md Rafiul, Baikunth Nath, and Michael Kirley. "A fusion model of HMM, ANN and GA for stock market forecasting." Expert systems with Applications 33.1 (2007): 171-180.
[8] Huang, Wei, Yoshiteru Nakamori, and Shou-Yang Wang. "Forecasting stock market movement direction with support vector machine." Computers & operations research 32.10 (2005): 2513-2522.gg
[9] Sun, Jie, and Hui Li. "Financial distress prediction using support vector machines: Ensemble vs. individual." Applied Soft Computing 12.8 (2012): 2254- 2265.
[10] Ou, Phichhang, and Hengshan Wang. "Prediction of stock market index movement by ten data mining techniques." Modern Applied Science 3.12 (2009): 28- 42.
[11] Liu, Fajiang, and Jun Wang. "Fluctuation prediction of stock market index by Legendre neural network with random time strength function." Neurocomputing 83 (2012): 12-21.
[12] Tsai, Chih-Fong, et al. "Predicting stock returns by classifier ensembles." Applied Soft Computing 11.2 (2011): 2452-2459.
[13] AraúJo, Ricardo De A., and Tiago AE Ferreira. "A morphological-rank-linear evolutionary method for stock market prediction." Information Sciences 237 (2013): 3-17.
[14] Ballings, Michel, et al. "Evaluating multiple classifiers for stock price direction prediction." Expert Systems with Applications 42.20 (2015): 7046-7056.
[15] Basak, Suryoday, et al. "Predicting the direction of stock market prices using tree-based classifiers." The North American Journal of Economics and Finance 47 (2019): 552-567.
[16] Weng, Bin, et al. "Macroeconomic indicators alone can predict the monthly closing price of major US indices: Insights from artificial intelligence, time-series analysis and hybrid models." Applied Soft Computing 71 (2018): 685-697.
[17] Long, Jiawei, et al. "An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market." Applied Soft Computing (2020): 106205.
[18] Rekha, G., et al. "Prediction of Stock Market Using Neural Network Strategies." Journal of Computational and Theoretical Nanoscience 16.5-6 (2019): 2333-2336.
[19] Pang, Xiongwen, et al. "An innovative neural network approach for stock market prediction." The Journal of Supercomputing (2018): 1-21.
[20] Kelotra, A. and P. Pandey, Stock Market Prediction Using Optimized Deep-ConvLSTM Model. Big Data, 2020. 8(1): p. 5-24.
[21] Baek, Yujin, and Ha Young Kim. "ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module." Expert Systems with Applications 113 (2018): 457-480.
[22] Chung, H. and K.-s. Shin, Genetic algorithm-optimized long short-term memory network for stock market prediction. Sustainability, 2018. 10(10): p. 3765.
[23] Kara, Yakup, MelekAcarBoyacioglu, and ÖmerKaanBaykan. "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange." Expert systems with Applications 38.5 (2011): 5311-5319.

DOI:10.46647/ijetms.2025.v09i05.009 ISSN: 2581-4621

[24] Patel, Jigar, et al. "Predicting stock market index using fusion of machine learning techniques." Expert Systems with Applications 42.4 (2015): 2162-2172.
[25] Patel, Jigar, et al. "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques." Expert systems with applications 42.1 (2015): 259-268