International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Stock Market Price Prediction Using Machine Learning Algorithm

ANITHA C[1], BERNET ANTO M[2] , RISHABAVARTHINI M[3] , LOGESH S[4] , MRS. VISHNU PRIYA V V[5]

1,2,3,4 B.E, Student Department Computer Science And Engineering, Angel College Of Engineering And Technology, Tirupur, Tamil Nadu, India.

5 Asst. Professor, Department Computer Science And Engineering, Angel College Of Engineering And Technology, Tirupur, Tamil Nadu, India.

ABSTRACT:

Stock market trading prediction involves forecasting the future movements of stock prices or determining whether to buy, sell, or hold a particular stock. Predicting stock market movements is a challenging task due to the dynamic and complex nature of financial markets. Various methods, including traditional statistical models and machine learning algorithms, can be employed for stock market trading prediction. Using Convolutional Neural Networks (CNNs) for stock market price prediction involves applying deep learning techniques to analyses historical price data and capture patterns that can assist in forecasting future price movements. While CNNs are more commonly associated with image recognition tasks, they can also be adapted for time-series data like stock prices. Classifying stock market trading involves predicting whether the price of a stock will go up, down, or remain stable. Machine learning algorithms can be applied to this task. Here's a general guide on how you might approach classifying stock market trading using machine learning like SVM, Decision Tree.

INDEXTERMS: Data preparation, Model architecture, Model training, Evaluation & predictions.

I. INTRODUCTION:

In the ever-evolving world of finance, the ability to accurately predict stock market movements stands as a Holy Grail for investors, traders, and financial analysts alike. The dynamism and complexity inherent in the financial markets make this task exceedingly challenging, yet the potential rewards for successful predictions are substantial, ranging from significant financial gains to improved risk management and strategic investment planning. Traditional approaches to stock market prediction have primarily relied on statistical models and fundamental analysis, focusing on economic indicators, company performance metrics, and historical price trends. However, the advent of advanced computational techniques and the proliferation of data have paved the way for more sophisticated and nuanced approaches to understanding and forecasting market movements.

Among these advanced techniques, machine learning algorithms have emerged as particularly powerful tools for analyzing vast amounts of financial data and uncovering patterns that may not be immediately apparent to human analysts. Machine learning's ability to learn from data and improve over time offers a promising avenue for enhancing the accuracy and reliability of stock market predictions. Specifically, Convolutional Neural Networks (CNNs), traditionally celebrated for their prowess in image recognition and processing tasks, have been adapted to tackle the sequential nature of time-series data such as stock prices. By capturing temporal patterns and dependencies in historical price data, CNNs offer a novel approach to forecasting future stock movements.

Techniques such as Support Vector Machines (SVM), Decision Trees, and XGBoost have been employed to navigate the complexities of market data and provide actionable trading insights. These methods leverage historical data and a variety of features, including price movements, trading volumes, and technical indicators, to make informed predictions about future stock behavior.

This introduction sets the stage for a detailed exploration of using machine learning techniques, specifically CNNs and other algorithms like SVM, Decision Tree, and XGBoost, for stock market trading prediction. We will delve into the nuances of these methods, their application to financial data, and the challenges and opportunities they present in the quest to navigate the turbulent waters of the stock market.

II. LITERATURE SURVEY:

STOCK PRICE PREDICTION: SUPPORT VECTOR MACHINES (SVM) METHOD

Patel et al. (2015) used SVM along with other ML algorithms to predict the stock market.

Support Vector Machines (SVM) are supervised machine learning algorithms used for classification and regression tasks. In stock market prediction, SVM is typically used to classify whether the stock price will go up or down based on historical data and technical indicators.SVM works by finding the optimal hyperplane that separates data points of different classes (e.g., price up vs. price down).It performs well in high-dimensional spaces and is effective when the number of features exceeds the number of samples.SVM can handle non-linear data using kernel functions like the Radial Basis Function (RBF).Input features may include moving averages, volume, RSI, MACD, and other technical indicators.The model predicts future trends (e.g., bullish or bearish) based on historical patterns.Good generalization capability.Effective in cases with clear margin of separation.

ARTIFICIAL NEURAL NETWORKS (ANN) METHOD:

Kim (2003) applied ANN for stock prediction in the Korean stock market.

Artificial Neural Networks (ANN) are computational models inspired by the human brain. They consist of layers of interconnected "neurons" that process input data and learn patterns through training. ANNs are widely used in stock market prediction due to their ability to model complex, non-linear relationships.Composed of an input layer, one or more hidden layers, and an output layer.Learns by adjusting weights during training using algorithms like backpropagation.Can model relationships that are difficult to capture with traditional statistical methods.Inputs typically include historical stock data (Open, High, Low, Close, Volume) and technical indicators (RSI, MACD, etc.).Outputs can be future price values (regression) or movement direction (classification).Capable of modeling complex, non-linear patterns in financial data.Flexible and adaptable to different types of stock data.

RECURRENT NEURAL NETWORKS (RNN) AND LSTM METHOD:

Fischer & Krauss (2018) used LSTM networks for S&P 500 prediction.

Recurrent Neural Networks (RNN) are a type of neural network specifically designed for sequential data, such as time series. They have a memory mechanism that captures information from previous inputs, making them suitable for predicting stock prices based on historical data.Designed for time-dependent data.Maintains a hidden state that carries information from previous time steps.Struggles with long-term dependencies due to vanishing gradient issues.To overcome this, Long Short-Term Memory (LSTM) networks were developed.

LONG SHORT-TERM MEMORY (LSTM):

LSTM is a special type of RNN capable of learning long-term dependencies in sequences. It uses **gates** (input, forget, and output) to control the flow of information, making it highly effective for modeling stock price trends over time.Inputs: Historical price data (Open, High, Low, Close), trading volume, technical indicators.Output: Predicted stock price or trend.LSTM can capture complex patterns and trends that evolve over longer periods.Handles time-series data better than traditional models.Remembers both short- and long-term patterns in data.Superior performance in predicting trends in noisy and non-linear financial data.

RANDOM FOREST AND DECISION TREES:

Kara et al. (2011) compared ANN and Decision Trees for Istanbul Stock Exchange data.

DecisionTrees are supervised learning models used for classification and regression. They split data into branches based on feature values, creating a tree-like structure where each node represents a decision rule.Random Forest is an ensemble method that builds multiple decision trees and combines their outputs to improve prediction accuracy and reduce overfitting.Easy to understand and interpret.Can capture non-linear relationships.Prone to overfitting on noisy stock data.Builds multiple trees using random subsets of data and features.Averages the predictions (for regression) or takes the majority vote (for classification).More robust and accurate than a single decision tree.Input features often include historical prices and technical indicators (e.g., Moving Averages, RSI).Output can be future stock price (regression) or movement direction (classification: up/down).Handles large datasets and many features well.Less prone to overfitting than single trees. Useful for feature importance analysis.

lll. RELATEDWORK:

SUPPORTVECTOR MACHINES (SVM):

SVMs are widely used for classification and regression tasks. Patel et al. (2015) demonstrated that SVMs, combined with technical indicators like Moving Average and RSI, can effectively predict stock price movements. However, performance depends on careful parameter tuning.

ARTIFICIAL NEURAL NETWORKS (ANN):

ANN models are capable of learning non-linear relationships in historical price data. Kim (2003) applied ANN to the Korean stock market and achieved promising results. ANNs are flexible but prone to overfitting if not properly managed.

LONG SHORT-TERM MEMORY (LSTM) NETWORKS:

LSTM, a type of Recurrent Neural Network (RNN), is effective for time-series prediction due to its ability to remember long-term dependencies. Fischer and Krauss (2018) showed that LSTM models outperform traditional ML models on stock datasets like the S&P 500.

DECISION TREES AND RANDOM FORESTS:

These are ensemble learning methods known for their interpretability and robustness. Kara et al. (2011) found that Decision Trees, when used with technical indicators, can yield accurate predictions.

lV. EXISTING SYSTEM:

The existing system relies on traditional models like Linear Regression and basic classifiers such as Logistic Regression or Decision Trees. While these methods are easy to implement and interpret, they often fail to capture the complex, non-linear nature of stock market data. Simpler models may not perform well with noisy or non-stationary financial data. Their predictive accuracy is generally limited in volatile market conditions. As a result, these systems may provide suboptimal trading decisions.

Disadvantages of Existing System

- Traditional statistical methods struggle to detect complex, non-linear patterns in stock market data.
- Many existing approaches require significant human effort for feature engineering.
- Traditional systems are not optimized for processing vast historical stock data.

- Some classical models may perform well on training data but poorly on unseen data.
- Existing systems may not be suitable for near real-time prediction and analysis.

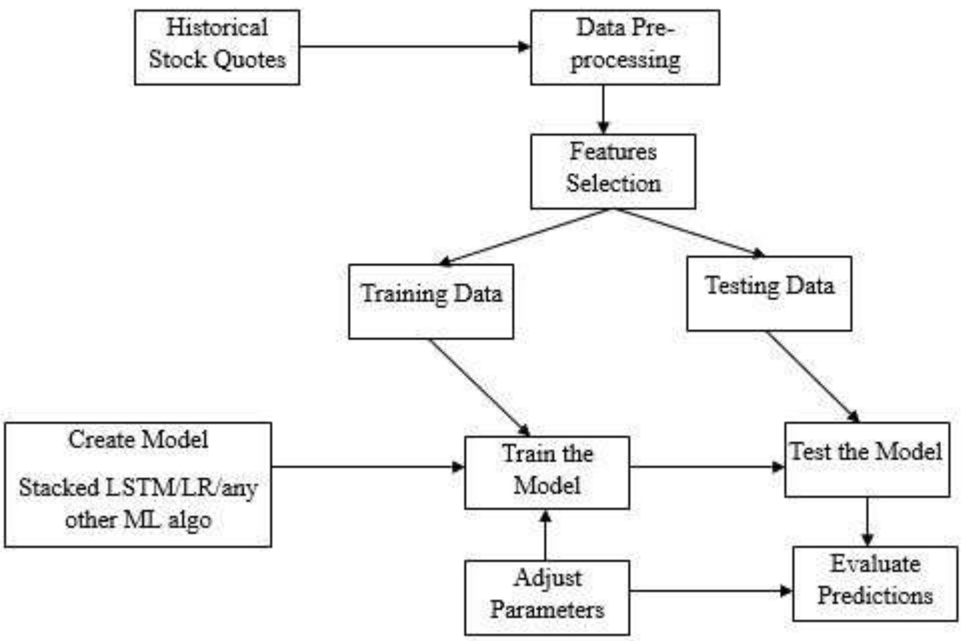
V. PROPOSED SYSTEM:

The proposed system enhances stock market prediction by using CNNs to detect temporal patterns in price data, treating time-series like one-dimensional images. It leverages CNN's ability to process complex patterns and multiple financial indicators for more accurate forecasting. Advanced models like XGBoost are used for classifying trading actions (buy, sell, hold) with high precision. These models handle high-dimensional, noisy financial data effectively. The approach improves prediction accuracy, scalability, and adaptability to market dynamics.

Advantages of Proposed System

- CNNs can automatically learn and extract important features from time-series data without manual intervention.
- Machine learning models like SVM, Decision Tree, and XGBoost offer better prediction accuracy over traditional methods.
- The system can capture complex and non-linear patterns in stock price movements.
- Can be trained on large datasets and adapted to different market conditions.
- The system can classify stock trends into multiple categories: up, down, or stable.

VI. ARCHITECTURE DIAGRAM:



VII. TECHNIQUES:

One of the most researched and difficult issues that affects so many academics and industry specialists from various departments of economics, business, mathematics, and computing science is stock price prediction. It is challenging to make accurate predictions about the stock market, primarily due to the stock time series near resemblance to random walks. This paper will examine various machine learning and artificial intelligence (AI) approaches to stock price prediction. This study gives a comprehensive assessment of 22 research publications that recommend various approaches, such as computation techniques, machine learning algorithms, performance metrics, and top journals. Research questions are used to guide the selection of the studies. As a result, this chosen researchis assisting in the discovery of ML methods and their dataset for stock market prediction. Due to its excellent performance and accuracy, LSTM (Long Short-Term Memory) was discovered to be the technique utilized the most commonly for stock price prediction. Numerous other methods, including CNN (Convolutional Neural Networks), RNN (Recurrent Neural Networks), SVM (Support Vector Machines), RF (Random Forests), and SVR (Support Vector Regressions), also produced encouraging prediction outcomes.

VIII. PROJECT DISCRIPTION:

DATA PREPARATION:

Data Collection: Gather historical stock price data, including features such as open, high, low, and close prices, as well as trading volume.Data Normalization: Normalize the data to ensure all features are on a similar scale. This step is crucial for the performance of neural networks.Time Series Segmentation: Divide the time series data into sequences or windows. Each window comprises a set of consecutive data points, and these can be used as input sequences for the CNN.

MODEL ARCHITECTURE:

Input Layer: The input layer of the CNN would take in the segmented time series data. Convolutional Layers: Apply one or more convolutional layers to capture local patterns and relationships within the input sequence. The filters in these layers act as feature detectors.Pooling Layers: Use pooling layers (e.g., max pooling) to down sample the spatial dimensions and reduce the computational load. This helps in retaining important features while discarding less relevant information.Flattening Layer: Flatten the output from the convolutional and pooling layers into a one-dimensional vector.Dense (Fully Connected) Layers: Connect the flattened vector to one or more dense layers for higher-level feature learning. The final layer may consist of a single neuron for regression or multiple neurons for classification (e.g., predicting price movement categories).

TRAINING:

Loss Function: Define a suitable loss function for regression tasks, such as mean squared error.Optimization Algorithm: Use an optimizer (e.g., Adam) to minimize the loss function during training.Training Data: Split the data into training and validation sets. Train the CNN on historical data, adjusting weights during backpropagation.

EVALUATION AND PREDICTION:

Validation Set: Assess the model's performance on a validation set to ensure it generalizes well to unseen data.Testing Set: Evaluate the model on a separate testing set to simulate real-world performance.Prediction: Use the trained CNN to make predictions on data.

IX. CONCLUSION AND FEATURES:

In conclusion, measuring the accuracy of the different algorithms, we found that the most suitable algorithm for predicting the market price of a stock based on various data points from the historical data is the logistic regression algorithm. The algorithm will be a great asset for brokers and investors for investing money in the stock market since it is trained on a huge collection of historical data and has been chosen after being tested on a sample data. The project demonstrates the machine learning model to predict the stock value with more accuracy as compared to previously implemented machine learning models. This project signifies a significant advancement in leveraging machine learning for stock market prediction. By surpassing the performance of previously implemented models, logistic regression underscores its effectiveness in providing actionable insights for investment decision-making.

REFERENCE:

1. Egeli B, Ozturan M, Badur B, “Stock market prediction using artificial neural networks”, Third Hawaii international conference on business, Honoluluu Hawaii, 2003.
2. Mruga Gurjar1, Parth Naik2, Gururaj Mujumdar3, Prof. Tejaswita Vaidya4, "stock market prediction using ANN": International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 03 (Mar-2018)p-ISSN: 2395-0072
3. Gaurav Kshirsagar, Rukshad Amaria, Mohit Chandel, Shantanu Kakade, “ Stock Market Prediction using Artificial Neural Networks”: Volume 6, Issue 3, March 2016 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering, Research Paper.
4. Risul Islam Rasel, Nasrin Sultana and Nasimul Hasan, Financial Instability Analysis using ANN and Feature Selection Technique: Application to Stock Market Price Prediction- IEEE 2016 [5] PhayungMeesad, “Predicting Stock Market Price Using Support Vector Regression”, IEEE 2014
5. Kizilaslan R, Freund S., Oztekin, Iseri A., “A data analytic approach to forecasting daily stock returns in an emerging market”, European Journal of Operational Research, 253(2016), , 2016. ISSN 2278-6856
6. Zhang, R.; Yuan, Z.; Shao, X. A new combined CNN-RNN model for sector stock price analysis. In Proceedings of the IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 23–27 July 2018; pp. 546–551.
7. Singh, K.; Booma, P.M.; Eaganathan, U. E-Commerce System for Sale Prediction Using Machine Learning Technique. *J. Phys. Conf. Ser.* **2020**, *1712*, 012042.
8. Deng, C.; Liu, Y. A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 9969357.
9. Agrawal, Manish, Piyush Kumar Shukla, Rajit Nair, Anand Nayyar, and Mehedi Masud. 2022. Stock Prediction Based on Technical Indicators Using Deep Learning Model. *Computers, Materials & Continua* 70: 287–304.