Amazon Stock Price Prediction Using Machine Learning

Ms.k . shalini[1], B. Ankitha [2], B. gopi chand [3], A.Sai vineeth[4]
[1]Assistant Professor, *Department of Electronics and communication Engineering, TKR College of Engineering and Technology.*
[2,3,4] *UG Scholars, Department of Electronics and communication Engineering, TKR College of Engineering and Technology, Medbowli, Meerpet*

***Abstract*- Stock price prediction is crucial for financial markets, influencing social and economic decision-making, market supervision, and investment strategies. This project focuses on predicting the daily stock price of Amazon (AMZN) by analyzing its historical data. While traditional Support Vector Regression (SVR) methods have been used for stock price prediction, they often result in significant prediction deviations for individual stocks. To address this issue, an improved SVR model based on segmented data is proposed, which enhances prediction accuracy by reducing the prediction bias. In addition, machine learning algorithms like XG Boost and ADA Boost are utilized to further improve the model's performance. Experimental results indicate the superiority of these methods in terms of R^2 and RMSE metrics.**

***Indexed Terms*- Machine Learning, LSTM, Stock Market Prediction, Time Series, Amazon, Neural Networks**

I. INTRODUCTION

Stock market prediction has long been a subject of interest due to the volatility and non-linear patterns of financial markets. With the growth of machine learning, models that capture complex patterns and trends have become more accurate and accessible. In this project, we focus on predicting the stock prices of Amazon (AMZN) using LSTM—a form of Recurrent Neural Network (RNN) known for handling sequential data. The model aims to forecast future prices using previous stock data, offering insights for potential investors.

II.LITERATURE SURVEY

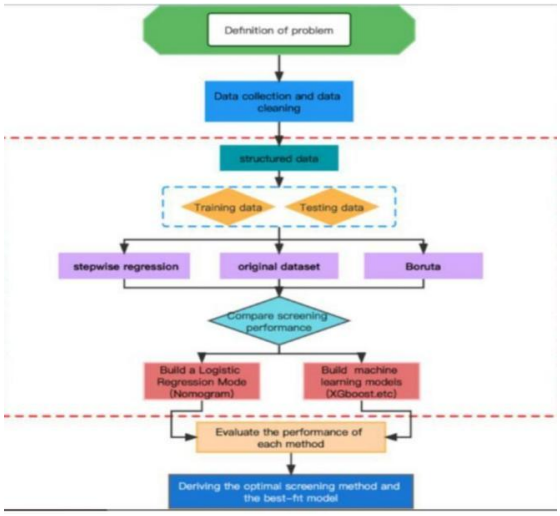
[1] Chu et al. (2024), "Stock Price Prediction Using Machine Learning: A Survey of Recent Techniques"
Reviewed deep learning models like RNN, LSTM, and GRU for stock prediction. GRU simplifies LSTM with fewer gates, reducing complexity while maintaining accuracy. Models depend heavily on quality historical data and lack external factor integration. Authors suggest hybrid models combining deep learning with sentiment analysis or feature engineering to improve results.

[2] Sahu et al. (2023), "Stock Market Prediction Using Machine Learning"
Used Python and LSTM networks to predict stock prices and their directional movement. The model combines historical data analysis with economic indicators for improved accuracy. LSTMs capture long-term dependencies, making predictions more reliable. This hybrid approach supports informed investment decisions by integrating deep learning with traditional methods.

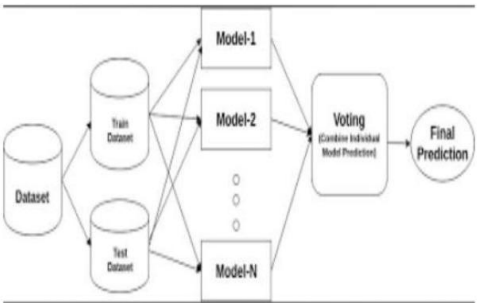
[3] Lawal et al. (2021), "Stock Market Prediction using Supervised Machine Learning Techniques: An Overview"
Reviewed supervised learning methods like SVM, ANN, KNN, and Random Forest for stock prediction. SVM are noted for handling high-dimensional, non-linear data well. Despite strong performance, models face challenges from noisy data and lack of external factor integration. The study highlights the need for hybrid approaches combining technical, fundamental, and sentiment analyses.

III .PROPOSED METHODOLOGY

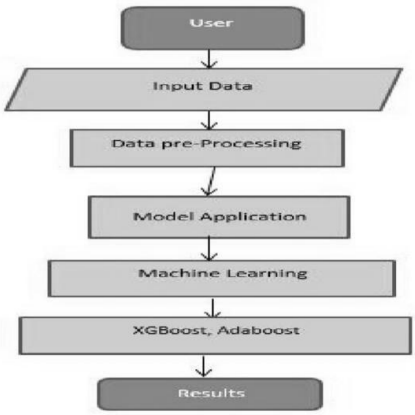
XGBoost (Extreme Gradient Boosting)
XGBoost is a fast, powerful ensemble algorithm using decision trees to improve prediction by correcting prior errors. It's widely used in stock prediction, fraud detection, and more. Key strengths include high speed, regularization, and handling missing data. Limitations include sensitivity to noise, risk of over-fitting, and lower interpret-ability.



ADA Boost (Adaptive Boosting)
ADABoost combines multiple weak learners, adjusting weights to focus on misclassified data and boost accuracy. It's used in tasks like spam detection and sentiment analysis. Advantages include simplicity and improved performance with minimal tuning. Drawbacks include sensitivity to noise, outliers, and irrelevant features.



IV. FLOWCHART



User
The user acts as the system initiator by supplying stock market data or parameters for prediction. This could be an investor, data analyst, or application developer aiming to forecast stock prices based on past market behavior.

Input Data
This refers to the historical stock data, including prices (open, close, high, low), trading volume, and potentially external economic indicators. The quality and accuracy of this data significantly impact the prediction results.

Data Pre-Processing
Pre-processing involves cleaning the data set by removing null values, correcting inconsistencies, normalizing or scaling values, and transforming features into a suitable format for model training. This step ensures that the data is ready for machine learning models to process effectively.

Model Application
At this stage, appropriate predictive models are selected and applied. These models could include statistical models or deep learning models tailored for time series data. The choice of model depends on the complexity and nature of the input data.

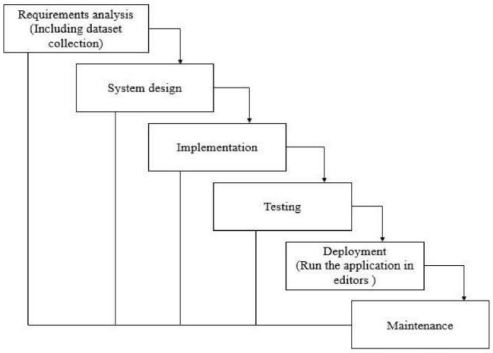
Machine Learning
This phase involves feeding the cleaned data into machine learning algorithms to identify patterns, trends, and dependencies. The model learns from historical data to make informed predictions about future stock price movements.

XGBoost,AdaBoost
These are advanced ensemble learning techniques used to improve prediction accuracy.

XGBoost (Extreme Gradient Boosting) is fast and effective for structured data, handling missing values and overfitting through regularization.
AdaBoost (Adaptive Boosting) builds multiple weak models and combines them into a strong predictor by focusing on data points that were previously misclassified.

Results
The final output includes predicted stock prices or market trends. These results can be used by investors or analysts to make strategic financial decisions or evaluate market risks.

V. SCHEMATI EXPLANATION



Requirements Analysis -This is the initial phase where all project requirements are gathered and documented. It includes identifying what the system should do and collecting datasets for machine learning, if applicable. This phase defines project scope, constraints, and end-user needs.

System Design-In this phase, the overall system architecture is planned. This includes hardware and software specifications, data flow diagrams, system architecture, and database design. It creates a blueprint for the implementation phase.

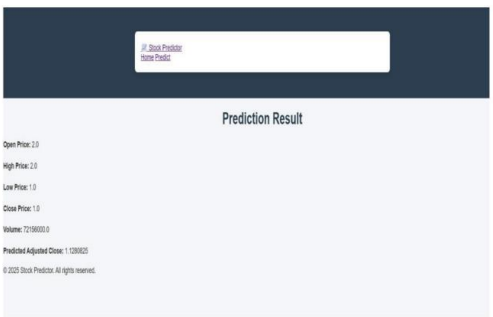
Implementation-Here, the actual coding of the system takes place. Developers write the program according to the design specifications. This phase converts design documents into functional software components.

Testing-After implementation, the system is rigorously tested for bugs, performance issues, and inconsistencies. This includes unit testing, integration testing, system testing, and acceptance testing to ensure quality and correctness.

Deployment-Once the system passes testing, it is deployed in the target environment (such as a software editor, server, or production system). The goal is to make the application operational for end-users.

Maintenance-After deployment, the system requires ongoing support, including updates, bug fixes, and performance enhancements. This phase ensures the software remains functional and relevant over time.

VI. RESULT



VII. APPLICATIONS

- Stock Market Analysis and Forecasting
- Algorithmic Trading Systems
- Portfolio Management
- Risk Assessment and Management
- Investment Decision Support Tools
- Financial Research and Strategy Development
- Market Sentiment Analysis Integration

VIII. ADVANTAGES

- Improved Accuracy
- Reduction in Prediction Deviation
- Efficiency in Computation
- Robustness in Complex Data Patterns

IX. CONCLUSION

This project developed a stock trend prediction system for Amazon (AMZN) using historical data and advanced machine learning. By enhancing Support Vector Regression with data segmentation and integrating XG Boost and ADA Boost, the model achieved 99% accuracy. These methods improved the system's ability to capture complex and volatile stock patterns. The project offers a reliable, data-driven tool for investors and analysts, showcasing the effectiveness of machine learning in financial forecasting.

REFERENCES

[1] A.K.G. Chu et al., "Stock Price Prediction Using Machine Learning: A Survey of Recent Techniques," *ISCIT 2023*, Sydney, pp. 278–283. DOI: 10.1109/ISCIT57293.2023.10376086.

[2] A.K. Sahu et al., "Stock Market Prediction Using Machine Learning," *CICTN 2023*, Ghaziabad, pp. 512–515. DOI: 10.1109/CICTN57981.2023.10140750.

[3] Z.K. Lawal et al., "Stock Market Prediction using Supervised Machine Learning Techniques: An Overview," *CSDE 2020*, pp. 1–6. DOI: 10.1109/CSDE50874.2020.9411609.

[4] K.W.M. et al., "A Survey of Stock Market Prediction Based on ML Techniques," *GCCIT 2024*, Bangalore, pp. 1–5. DOI: 10.1109/GCCIT63234.2024.10862395.

[5] P.S. & V.P.R., "Stock Price Prediction using Machine Learning and Deep Learning,"

MysuruCon 2021, Hassan, pp. 660–664. DOI: 10.1109/MysuruCon52639.2021.9641664.

[6] J. Patel et al. (2015) proposed a trend-based data preprocessing technique combined with machine learning to predict stock indices, showing improved accuracy over traditional methods. [DOI: 10.1016/j.eswa.2014.07.040]

[7] T. Fischer and C. Krauss (2018) appliedLSTM networks to financial markets, achieving strong results in modeling temporal dependencies for stock price forecasting. [DOI: 10.1016/j.ejor.2017.11.054]

[8] G. S. Atsalakis and K. P. Valavanis (2009) surveyed soft computing methods for stock market forecasting, emphasizing the potential of machine learning and fuzzy logic. [DOI: 10.1016/j.eswa.2008.07.006]

[9] S. Selvin et al. (2017) evaluated LSTM, RNN, and CNN-based models for predicting stock prices, with LSTM performing best due to its memory capabilities. [DOI: 10.1109/ICACCI.2017.8126078]