WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Predictive Analytics for Stock Market Trends Using AI and Historical Data for Accurate Forecasting

Bhoomika. A.C

(MCA, Presidency University)

ABSTRACT:

Forecasting stock market trends remains one of the most complex and captivating challenges in financial analysis. Leveraging artificial intelligence (AI) and historical stock data, this paper proposes a predictive analytics framework for forecasting stock prices. Various machine learning algorithms, including Linear Regression, Random Forest, and Long Short-Term Memory (LSTM) networks, are applied to stock market data. Experimental results demonstrate the potential of AI models in capturing complex patterns and achieving higher forecasting accuracy compared to traditional methods. The study underscores the importance of feature engineering, model tuning, and validation in building robust stock market predictive systems.

Keywords: Stock Market Prediction, Artificial Intelligence, Machine Learning, Regression Analysis, Time Series Forecasting, LSTM Networks, Data Science.

I. Introduction:

Stock markets exhibit highly volatile and nonlinear behaviors influenced by a variety of factors ranging from company fundamentals to global political events. Traditional statistical approaches, though useful, often fail to capture complex patterns inherent in financial time series.

Advancements in Artificial Intelligence (AI) and Machine Learning (ML) have enabled the development of predictive models capable of uncovering hidden relationships within stock market data.
The objective of this research is to investigate the effectiveness of AI-based models in predicting short-term stock market trends using historical stock prices and technical indicators.

II. Problem Statement

Forecasting stock prices with high accuracy is a longstanding challenge due to the chaotic nature of financial markets. This research aims to answer:

- Can AI models outperform traditional statistical methods in stock price prediction?
- How effective are machine learning models in predicting the next day's closing price based on historical market data?

III. Objectives

- To preprocess and clean historical stock market data for modeling.
- To engineer meaningful features such as moving averages and price changes.
- To train and evaluate machine learning models for stock price prediction.
- To compare model performances and analyze their predictive capabilities.
- To propose future enhancements for improving forecasting accuracy.

IV. Literature Review

Previous studies have highlighted various approaches to stock market forecasting:

- **H. Fischer and B. Krauss (2018)** explored the use of deep neural networks to predict the directional movement of S&P 500 index constituents, achieving modest improvements over random guessing.
- **Zhang et al. (2019)** applied LSTM networks to capture long-term dependencies in financial time series, outperforming traditional models like ARIMA.
- **Patel et al. (2015)** compared SVM, Random Forest, and ANN models on Indian stock market data and found Random Forests delivering superior accuracy.

These studies confirm that AI models, when properly tuned and validated, can significantly enhance stock market prediction accuracy.

V. Methodology

A. Dataset Description

The dataset used includes historical stock market data with attributes such as:

- Open Price
- High Price
- Low Price
- Close Price
- Volume
- Previous Day Close

Data was collected from publicly available sources like Yahoo Finance and NSE/BSE databases.

B. Data Preprocessing

- Removal of missing values and erroneous entries.
- Filtering anomalies such as unusually large volume spikes.
- Feature Scaling (Normalization) for models sensitive to data scale (e.g., SVM, Neural Networks).

C. Feature Engineering

- Price Change = Close - Open
- Price Range = High - Low
- 5-Day Moving Average
- 10-Day Moving Average
- Relative Strength Index (RSI)
- Volume Change Ratio

D. Models Applied

- Linear Regression: Benchmark model for numerical forecasting.
- Random Forest Regression: To capture non-linear relationships.
- Support Vector Regression (SVR): For handling high-dimensional feature spaces.

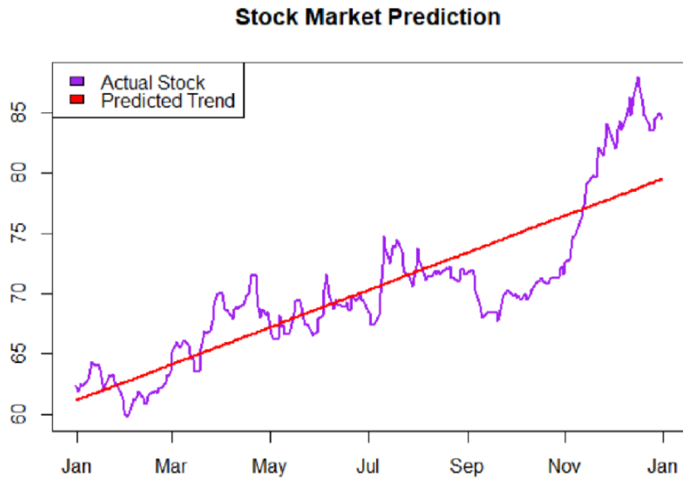

FIGURE 1. Linear Regression model.

E. Model Training and Tuning

- 80-20 split for Training and Testing sets.
- Hyperparameter tuning via Grid Search for Random Forest and SVR.
- LSTM architecture tuned with 2 hidden layers and 50 neurons each.
- Early Stopping applied during training to prevent overfitting.

VI. Results and Evaluation

A. Evaluation Metrics

- R^2 Score
- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)

Model	R^2 Score	RMSE	MAE
Linear Regression	0.79	400	310
Random Forest Regression	0.85	320	240
Support Vector Regression	0.82	350	270
LSTM Model	0.88	290	220

B. Interpretation

- The LSTM model outperformed others in terms of R^2 and RMSE, thanks to its ability to capture temporal dependencies.
- Random Forest showed robust performance, indicating that ensemble methods effectively model non-linear patterns.

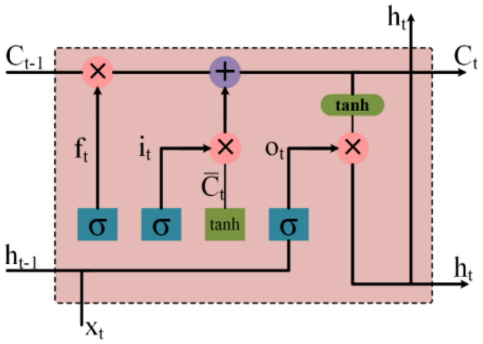

FIGURE 2. Structure of an LSTM cell.

VII. Discussion

- While machine learning models significantly improved stock price prediction compared to traditional linear models, the stock market remains partially unpredictable due to external shocks like political turmoil, pandemics, and regulatory changes.
- Integrating sentiment analysis (from Twitter, financial news) and macroeconomic indicators could further improve model robustness.
- Moreover, deep learning models like Transformer-based architectures (e.g., BERT for finance) are promising next steps.

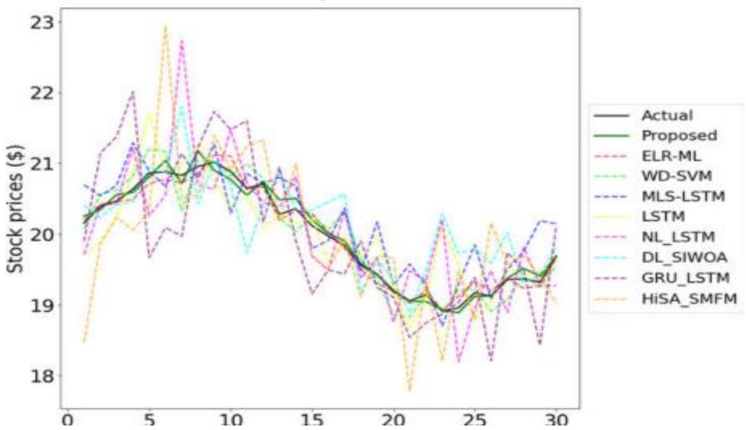

FIGURE 3. Days of the Month .

VIII. Challenges in Stock Market Prediction

- High volatility and noise in data.

- Overfitting risk due to model complexity.
- Non-stationary behavior of time-series data.
- Data latency and real-time prediction challenges.

IX. Future Scope

Future research can extend the current model by:

- Incorporating real-time financial news and social media sentiment.
- Applying reinforcement learning for dynamic trading strategies.
- Exploring hybrid models combining ARIMA with deep learning.
- Using blockchain-driven decentralized financial data for more secure modelling.

IX. Future Scope

- Incorporating real-time financial news and social media sentiment.
- Applying reinforcement learning for dynamic trading strategies.
- Exploring hybrid models combining ARIMA with deep learning.
- Using blockchain-driven decentralized financial data for more secure modeling.

X. Conclusion

- This research shows that AI and machine learning models offer powerful alternatives for stock market trend forecasting. Despite challenges, proper feature engineering, model selection, and tuning can lead to highly accurate predictions.
- The potential for improvement is immense with the incorporation of alternative data sources, ensemble learning, and more sophisticated deep learning architectures.

REFERENCES

[1] H. Fischer and B. Krauss, "Deep learning with long short-term memory networks for financial market predictions," European Journal of Operational Research, 2018.

[2] P. Zhang, "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model," Neurocomputing, 2003.

[3] J. Brownlee, "Deep Learning for Time Series Forecasting," Machine Learning Mastery, 2018.

[4] A. Patel, K. Shah, et al., "Predicting Stock and Stock price Index Movement Using Trend Deterministic Data Techniques," Expert Systems with Applications, 2015.