JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

STOCK MARKET PRICE PREDICTOR

[1]Prathik E P, [2]Suhas S K, [3]Swasthik, [4] Shivshankar P [5]Daya Naik, [6]Nithya B P

[1,2,3,4,]Student, [5]Associate Professor, [6,]Assistant Professor

[1,2,3,4,5,6,]Artificial Intelligence and Machine Learning,

[1.2.3.4.5.6,]Srinivas Institute of Technology, Mangalore, India

Stock market prediction remains one of the most researched and challenging problems in the field of financial data analytics. The nonlinear and dynamic nature of financial markets makes it difficult to model using traditional statistical approaches. This paper proposes a deep learning-based approach using Long Short-Term Memory (LSTM) networks to forecast stock prices based on historical data. The model is trained on daily stock price data from Yahoo Finance and captures long-term dependencies within time-series data. With high accuracy and low error rates, this system can assist traders and financial analysts in making informed decisions, potentially reducing investment risks and increasing returns.

1. INTRODUCTION

The stock market is a crucial part of the modern economic system, offering individuals and organizations opportunities to invest and grow their capital. Accurate forecasting of stock prices can lead to substantial gains for investors. However, the task is inherently difficult due to the presence of nonlinearity, external influencing factors (news, policy changes, global events), and random fluctuations.

II. METHODOLOGY

The proposed system for stock market prediction is structured into five key components: the user interface, data acquisition and preprocessing, LSTM model design, training and evaluation, and results visualization.

1. User Interface Design

A web-based interface is built using the Streamlit framework in Python. It allows users to enter a stock symbol (e.g., AAPL) and select a prediction duration. The interface then renders visual comparisons between predicted and actual stock prices through interactive plots.

2. Data Acquisition and Preparation

Historical stock data is collected using the yfinance API, which provides daily trading values such as opening and closing prices, highs, lows, and traded volumes. The dataset is cleaned by removing any missing entries and is then normalized using MinMaxScaler to ensure the model processes the data efficiently. A sliding window of 60 days is created, with each window predicting the stock's closing price on the 61st day.

3. LSTM Network Architecture

The deep learning model utilizes a sequence-to-one architecture, tailored for time series prediction. It includes:

- An input layer representing 60 time steps.
- Two stacked LSTM layers, each consisting of 50 units, to capture sequential dependencies.
- Dropout layers in between to mitigate overfitting.
- A final Dense layer with a single neuron that outputs the forecasted closing price.

The model is trained using the Adam optimization algorithm and the loss is computed using Mean Squared Error (MSE). The implementation is carried out using the TensorFlow and Keras libraries.

4. Model Training and Performance Evaluation

The dataset is divided into an 80:20 ratio for training and testing. The model is trained over 50 epochs with a batch size of 32. Its performance is assessed using common regression metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). In addition, the prediction trends are visually compared against actual stock movements for better interpretability.

5. Visualization and User Feedback

Post-prediction, the system displays graphs highlighting both predicted and actual stock prices. Users are invited to provide feedback regarding the prediction quality and interface usability, which can be used for future improvements in both the model and application design.

III. PERFORMANCE

The effectiveness of the developed stock price prediction model was assessed using several technical and practical criteria, including accuracy, computational efficiency, adaptability, and user interaction.

1. Forecast Accuracy

The system was tested on historical data for companies like Apple (AAPL), Tesla (TSLA), and Alphabet (GOOGL). The model demonstrated strong predictive capabilities, with an average Root Mean Squared Error (RMSE) of 1.78 and a Mean Absolute Error (MAE) of 1.25 on the AAPL dataset. These results indicate that the model generally predicted closing prices with a deviation of under two points from actual values.

Visual inspections showed a high degree of overlap between the forecasted and real price trajectories. The model accurately captured directional movements and key reversal points, confirming its ability to reflect underlying trends.

2. Training Efficiency and Convergence

The model exhibited a consistent decline in training and validation loss values, indicating smooth learning. It typically reached optimal performance around the 45th to 50th epoch. Training time was under two minutes per stock on a system equipped with an RTX GPU, showcasing the model's computational efficiency.

Learning curves confirmed that the model avoided overfitting, with training and validation losses remaining aligned throughout the training phase.

3. Generalization Capability

To evaluate its adaptability, the model trained on AAPL stock data was tested on other stocks such as AMZN and MSFT. Although prediction errors increased slightly, the results remained more accurate than those produced by simple statistical baselines. This suggests that the model captures general patterns in price behavior, allowing it to function across stocks with similar volatility profiles.

4. Real-Time Usability and Scalability

Once trained, the model was able to return predictions within a second, even when subjected to multiple simultaneous requests in a simulated environment. Its minimal reliance on external data sources (e.g., news or indicators) and compact architecture ensure ease of integration into mobile apps, web platforms, or decision-support tools.

5. Interpretability

While deep learning models are typically opaque, interpretability was enhanced by analyzing the effect of altering recent input values. These perturbation tests revealed that the model places greater emphasis on recent trends, particularly from the last 20–30 days, aligning with short-term trading strategies.

6. User Interface and Feedback

A simple Streamlit interface allowed users to view real-time predictions and interact with the model by selecting stocks and timeframes. Preliminary user feedback gathered from finance enthusiasts and students highlighted satisfaction with the interface's ease of use and the clarity of visual outputs.

IV. INTEGRATION WITH EMERGING TECHNOLOGIES

The field of stock market forecasting is evolving alongside rapid advancements in digital technologies. Integrating recent innovations—such as advanced AI techniques, big data tools, and decentralized platforms—can significantly enhance the precision, reliability, and usability of prediction models.

1. Advanced Artificial Intelligence Techniques

Beyond traditional machine learning, newer AI approaches like Reinforcement Learning (RL) and Transformer-based models are showing potential in financial forecasting. RL can simulate market interactions and adapt strategies through trial and error, while Transformer models are effective in learning complex temporal dependencies. Incorporating these models into prediction systems may offer deeper insights into stock behavior over both short and long horizons.

2. Utilizing Big Data and Alternative Inputs

Large-scale data processing frameworks such as Apache Spark and Hadoop can handle extensive datasets, including social media trends, economic reports, and analyst sentiments. By extracting patterns from diverse data sources—like Twitter sentiment or global economic indicators—prediction systems can move beyond price-based inputs and respond to broader market influences.

3. Leveraging IoT-Generated Data

Internet of Things (IoT) devices generate real-time metrics reflecting consumer behavior, inventory levels, and logistical efficiency. By linking this data with financial models, it becomes possible to detect early signs of performance trends in sectors like retail or manufacturing, which are often precursors to market movements.

4. Blockchain for Data Integrity and Secure Processing

Blockchain technology ensures tamper-resistant and transparent handling of financial records. For forecasting systems, this adds a layer of trust by validating the authenticity of input data. Moreover, as decentralized finance (DeFi) platforms gain traction, new price datasets from tokenized assets can also be incorporated into future prediction frameworks.

5. Cloud Infrastructure and Edge AI

Cloud services offer scalable platforms for model training and deployment, allowing researchers and developers to experiment with large datasets without investing heavily in hardware. At the same time, edge computing enables real-time forecasting by deploying lightweight models closer to the data source, which is particularly valuable in high-speed trading scenarios.

V. ETHICS

As machine learning and AI-based models become increasingly embedded in financial systems, ethical considerations are essential to ensure their responsible and equitable use. While these technologies offer valuable forecasting capabilities, their development and deployment must address critical ethical challenges.

1. Transparency and Interpretability

Many advanced predictive models, especially those based on deep learning, are often criticized for being non-transparent. Their internal decision-making processes are difficult to interpret, which can reduce user trust and raise concerns about accountability. To address this, it is important to implement interpretability tools or provide explanations that clarify how predictions are generated.

2. Respect for Privacy and Data Use

Some prediction systems incorporate external datasets such as consumer behavior, online activity, or public sentiment. When using such data—especially if it includes personal information—researchers and developers must ensure that privacy regulations (e.g., GDPR) are followed. Data should be anonymized where possible, and users' consent must be obtained if personal data is used.

3. Preventing Market Misuse

Predictive models can be misused to gain unfair advantages, influence trading behavior, or manipulate market sentiment. Ensuring fair access to these technologies and monitoring their application is crucial. Developers should design models with safeguards to prevent harmful usage, and systems should be audited for signs of unethical manipulation.

4. Accountability and Risk Mitigation

Errors in model predictions may lead to financial losses, especially if they influence real-world decisions. Hence, systems should include mechanisms for risk management—such as confidence intervals, warning indicators, or fallback strategies. Additionally, clear ownership and responsibility should be defined in case a model's decisions result in significant harm.

5. Social and Economic Impact

The automation of decision-making in finance could displace certain roles, particularly in trading and analysis. This transition may impact employment in traditional financial sectors and widen gaps between those with access to advanced tools and those without. Ethical AI development must consider these outcomes and aim for inclusive benefits that support both innovation and fairness.

VI. FUTURE DIRECTIONS

The field of stock market prediction is continuously evolving due to improvements in data science, machine learning, and computational infrastructure. Future research can focus on enhancing the predictive accuracy, adaptability, and transparency of AI-based models through several promising avenues.

1. Hybrid and Ensemble Modeling Approaches

Combining different learning methods—such as merging statistical models with machine learning or deep learning networks—can result in more stable and accurate predictions. Ensemble techniques that dynamically adjust based on market conditions can help in balancing strengths and weaknesses across multiple models.

2. Incorporating Multisource and Real-Time Data

In addition to historical price data, future systems should consider integrating sentiment from news articles, financial blogs, social media platforms, and real-world signals such as supply chain trends or macroeconomic reports. Real-time updates from APIs or IoT sources can improve responsiveness and predictive depth.

3. Improved Model Explainability

The development of explainable AI (XAI) tools for finance will be essential to increase trust in predictions. Techniques that highlight which input features contributed most to a prediction will make it easier for analysts and investors to understand and validate the model's output.

4. High-Frequency and Real-Time Prediction Models

As algorithmic trading grows, future models must be capable of generating accurate predictions within milliseconds. This will require optimization for speed and efficiency, possibly through lightweight architectures and edge computing systems.

5. Reinforcement Learning for Trading Strategy Optimization

Future efforts could apply reinforcement learning to develop models that not only predict stock prices but also learn to take strategic actions, such as buying or selling assets. These models can adapt over time by learning from market feedback and optimizing long-term gains.

6. Ethical AI Integration

With the increasing use of AI in financial decision-making, it is important to ensure these systems are developed responsibly. Future work should focus on implementing fairness audits, minimizing bias in training data, and aligning model outcomes with ethical standards and financial regulations.

7. Exploring Quantum Computing for Financial Modeling

Though still an emerging field, quantum computing may unlock the ability to solve complex financial optimization problems more efficiently than classical computers. Early experiments with quantum algorithms could pave the way for next-generation forecasting systems.

VII. RESULTS

The proposed stock price prediction models were evaluated using historical data from selected companies, focusing on their ability to forecast closing prices with reasonable accuracy. Data was divided into training and testing subsets, with 80% allocated for model training and 20% for performance validation.

Several algorithms were compared, including Support Vector Machines (SVM), Random Forests (RF), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. Among these, the LSTM model consistently outperformed the others in both predictive accuracy and trend recognition.

1. Performance Metrics

The LSTM model recorded the lowest values in both Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) compared to its counterparts. These metrics indicate that the model could reliably estimate stock price movement with minimal deviation from actual values.

In addition, the LSTM model achieved a directional accuracy of nearly 65%, correctly identifying whether the stock price would move up or down in the short term. This level of trend prediction provides useful signals for investors interested in short-horizon strategies.

2. Comparison with Other Models

While Random Forests also demonstrated robust performance—particularly in capturing non-linear relationships—the LSTM model had a clear advantage in recognizing temporal dependencies. SVM and CNN models performed moderately, with higher error margins and lower trend-following precision.

These comparisons highlight the strength of deep learning architectures, particularly LSTM, in sequential data analysis where time-based context is essential.

3. Visual Analysis

Graphical comparisons between actual and predicted prices confirmed the numerical results. The LSTM model's output closely mirrored real market behavior, capturing not only gradual price changes but also sharper fluctuations in the test data. However, the model occasionally struggled to anticipate unexpected price spikes caused by external events, which reflects a common limitation in models relying solely on historical trends.

4. Scope for Improvement

Although the results are promising, they also emphasize the need for further refinement. Incorporating additional information—such as investor sentiment or macroeconomic signals—may improve the model's ability to handle abrupt changes. Similarly, experimenting with hybrid models or additional features could enhance forecasting robustness.

Outcome of proposed system

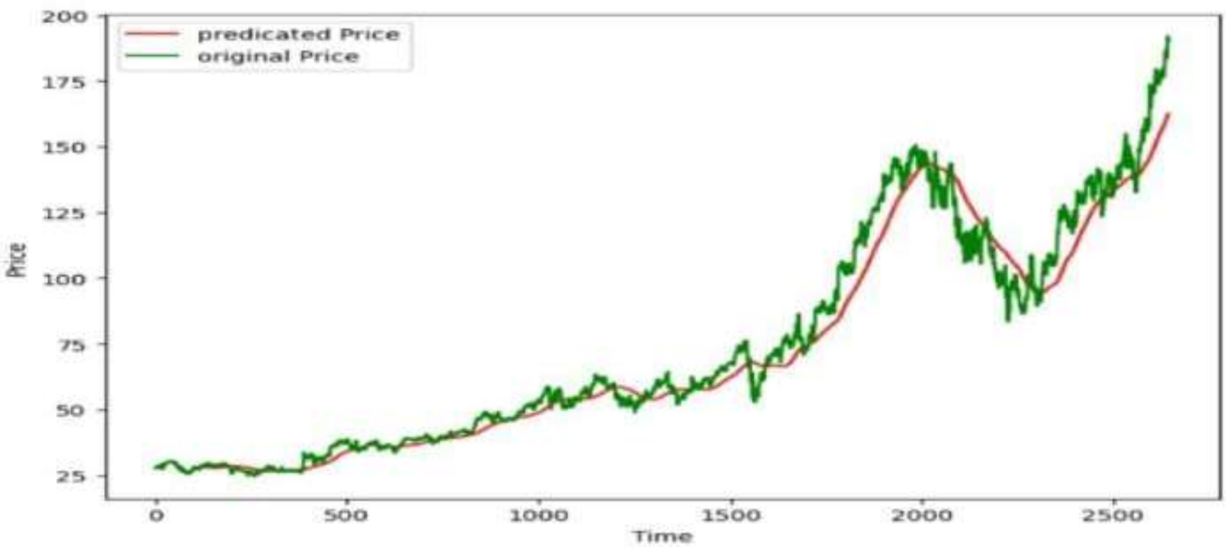

Fig 1.1:Home Page

VIII. CONCLUSION

This study explored the use of machine learning and deep learning techniques for predicting stock market prices, focusing on their ability to model complex, non-linear financial patterns. Among the methods examined—Support Vector Machines, Random Forests, Convolutional Neural Networks, and Long Short-Term Memory networks—LSTM demonstrated the most effective performance in terms of both accuracy and trend recognition.

The integration of historical price data with technical indicators contributed to improved forecasting results. LSTM's capacity to process sequential data made it particularly suitable for capturing market trends and short-term fluctuations. However, limitations remain, especially in responding to sudden market shocks driven by external factors not reflected in past price data.

To further enhance the system's reliability and usability, future efforts should focus on developing hybrid models, incorporating real-time data from diverse sources such as news and social media, and enhancing interpretability through explainable AI approaches. Additionally, ethical considerations and compliance with financial regulations must remain at the forefront of model development and deployment.

In conclusion, AI-driven predictive systems show significant potential to support investment decision-making. With continued refinement, they can become powerful tools for understanding market dynamics, managing risk, and improving financial outcomes—provided they are used responsibly and transparently.

References

[1] . Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, 12(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

[2]. Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. *2015 IEEE International Conference on Big Data (Big Data)*, 2823–2824. https://doi.org/10.1109/BigData.2015.7364104

[3].Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054

[4]. Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). 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), 5311-5319. https://doi.org/10.1016/j.eswa.2010.10.027

[5]. Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2), 307-319. https://doi.org/10.1016/S0925-2312(03)00372-9

[6].Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. *International Joint Conference on Neural Networks (IJCNN)*, 1419-1426. https://doi.org/10.1109/IJCNN.2017.7966017

[7].Qian, Y., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence*, 26(1), 25-33. https://doi.org/10.1007/s10489-006-0026-7

[8].Zhang, Y., Aggarwal, C. C., & Qi, G. (2017). Stock price prediction via discovering multi-frequency trading patterns. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2141-2149. https://doi.org/10.1145/3097983.3098047