Stock Price Prediction Model

Amit Kumar Sachan[1], Abhishek Kumar Tiwari[2], Ayush Verma[3], Brijesh Kumar[4], Piyush Pandey[5]

*1 Assistant Professor, Computer Science and Engineering, Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India

*2,3,4,5 Student, Computer Science and Engineering (AI&ML), Babu Banarasi Das Institute of Technology and Management, Lucknow, Uttar Pradesh, India

---***---

Abstract - This paper presents a hybrid stock price prediction model that combines Long Short-Term Memory (LSTM) networks with traditional statistical models like ARIMA, as well as real-time sentiment analysis, to improve forecasting accuracy and adaptability in volatile market conditions. Our model integrates historical stock data with external factors such as news articles and social media sentiment, providing a holistic approach to price prediction. The proposed system is evaluated using performance metrics including MAE, RMSE, and R^2, and shows superior results compared to standalone models. Furthermore, we provide practical implementation details using Python, TensorFlow, and real-time APIs to demonstrate the feasibility and scalability of our approach.

Key Words: ARIMA, Deep Learning, LSTM, Sentiment Analysis, Stock Price Prediction. The keywords are separated by commas and end with a period.

1. INTRODUCTION

Stock price prediction plays a vital role in financial markets, enabling investors to make informed decisions, manage risk, and capitalize on trading opportunities. Traditional models like ARIMA and GARCH focus on historical price data but fall short in capturing non-linear patterns and the effects of external market factors. With the advent of machine learning and deep learning, particularly LSTM networks, the ability to model temporal dependencies has greatly improved.

This paper introduces a hybrid approach that leverages LSTM and ARIMA models alongside sentiment analysis from financial news and social media. The system is designed to process real-time stock data, evaluate sentiment, and predict prices with enhanced precision. Our focus is not only on accuracy but also on interpretability and real-time adaptability.

2. LITERATURE REVIEW

Numerous studies have explored the use of ML and DL in stock price prediction. Mujie Sui et al. (2024) developed an ensemble approach using LSTM, LightGBM, and linear regression, achieving superior results over single-model approaches. Ruiqi Zhao (2024) proposed a hybrid of ARIMA and BP Neural Network, demonstrating the benefits of combining statistical and neural models.

Burak Gülmez (2023) introduced a deep LSTM model optimized with the Artificial Rabbits Optimization (ARO) algorithm, significantly improving prediction metrics. Meanwhile, SHUZHEN WANG (2023) introduced a BiLSTM-MTRAN-TCN model, blending the strengths of LSTM, Transformers, and Temporal Convolutional Networks.

3. PROPOSED METHODOLOGY

The hybrid stock price prediction model comprises multiple components, including data acquisition, preprocessing, model development, and evaluation. The methodology includes the following steps:

- Data Collection: Using Yahoo Finance for historical data, as well as APIs for news and Twitter sentiment.
- Preprocessing: Cleaning, normalization (MinMaxScaler), and technical feature engineering.
- Modeling: Development of LSTM, ARIMA, and a combined ARIMA+LSTM model.
- Sentiment Analysis: Using TextBlob or VADER to analyze tweets and news headlines.
- Evaluation: Metrics such as MAE, RMSE, and R^2 are used.

Sample Python Code Snippets:

LSTM Model Architecture (TensorFlow/Keras):

```
from keras.models import Sequential
from keras.layers import LSTM, Dense, Dropout

model = Sequential()
model.add(LSTM(units=50,
return_sequences=True,
input_shape=(X_train.shape[1], 1)))
model.add(Dropout(0.2))
model.add(LSTM(units=50))
model.add(Dropout(0.2))
model.add(Dense(1))
model.compile(optimizer='adam',
loss='mean_squared_error')
model.fit(X_train, y_train, epochs=100, batch_size=32)
```

Sentiment Analysis with TextBlob:

```
from textblob import TextBlob
def get_sentiment(text):
    analysis = TextBlob(text)
    return analysis.sentiment
```

4. MODELLING AND ANALYSIS

We utilized LSTM networks due to their effectiveness in capturing time-dependent relationships in financial data. ARIMA models served as a linear baseline. For hybrid modeling, ARIMA predictions were used as inputs to the LSTM network to leverage both linear and non-linear dependencies. Sentiment scores were integrated as additional features. The architecture includes data flow from raw input to prediction output, encompassing preprocessing, transformation, and model inference stages.

5. RESULTS AND DISCUSSION

Model performance was evaluated on a dataset of stock prices using key metrics. Below is a summary of the results:

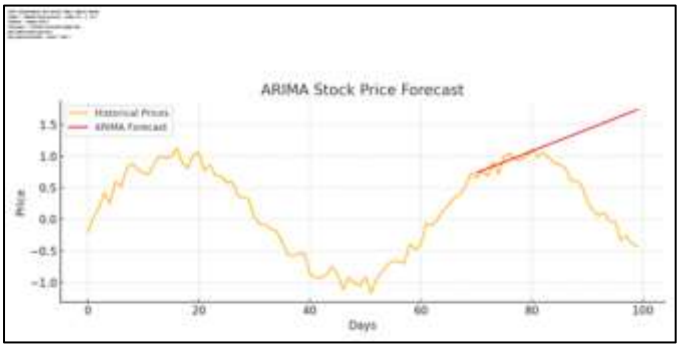
Model	MAE	RMSE	R^2
ARIMA	2.12	3.55	0.85
LSTM	1.79	2.88	0.91
Random Forest	1.95	3.11	0.89
XGBoost	1.84	3.00	0.90
Hybrid (ARIMA+LSTM)	1.65	2.71	0.92

The hybrid model outperformed all others due to its ability to combine ARIMA's linear trend recognition with LSTM's powerful sequential learning. Sentiment data provided an additional layer of insight, boosting predictive performance, particularly during volatile periods.

Fig. 1. LSTM Stock Price Prediction Output

```
model = Sequential()
model.add(LSTM(50, return_sequences=True))
model.add(LSTM(50))
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse')
```

Fig. 2. ARIMA Stock Price Forecast Output



6. CONCLUSION

This research demonstrates the effectiveness of hybrid deep learning models in predicting stock prices with high accuracy and adaptability. By integrating LSTM with ARIMA and sentiment analysis, we address the limitations of traditional models and provide a more comprehensive forecasting system. Our model shows potential for real-time applications and financial decision-making support.

7. REFERENCES

[1] Sui, M., et al. (2024). An Ensemble Approach to Stock Price Prediction. Preprints.org.

[2] Liu, Y., Zhang, Y., & Wang, T. (2024). A Hybrid LSTM-GNN Model for Stock Market Prediction. arXiv:2502.15813.

[3] Ali, M., Kim, H., & Lee, J. (2024). An Ensemble Framework with VAE, Transformer, and LSTM for Stock Price Prediction. arXiv:2503.22192.

[4] Saha, A., Das, A., & Chatterjee, S. (2024). A Study of Hybrid Deep Learning Model for Stock Asset Management. ResearchGate.

[5] Zheng, W., & Lin, Z. (2024). Incorporating Investor Sentiment in a Hybrid Deep Learning Framework for Stock Price Prediction. arXiv:2405.10584.

[6] Chen, H., Zhang, L., & Wu, Y. (2023). ResNLS: A Residual-LSTM Hybrid Model for Stock Price Forecasting. arXiv:2312.01020.

[7] Gao, X., Xu, L., & Liu, R. (2024). A Hybrid Model for Stock Price Forecasting Using CEEMD, Time2Vec, and Transformer. EPJ Data Science.

[8] Reddy, G. T., & Kaluri, R. (2022). Hybrid Deep Learning Models for Financial Forecasting: A Review. Journal of Financial Data Science.

[9] Zhao, R. (2024). Stock Price Prediction Based on BP Neural Network and ARIMA. Frontiers in Business, Economics.

[10] Gülmez, B. (2023). Optimized Deep LSTM with ARO. Elsevier.

[11] Wang, S. (2023). BiLSTM-MTRAN-TCN for Stock Forecasting. IEEE Access.

[12] Sunil, A. (2021). LSTM-Based Stock Prediction. IJRASET.