Stock Movement Prediction

Tamlurkar Manoj
Information Technology
NITK SURATHKAL
Manglore, Karnataka, India
tamlurkarmanoj.221ai039@nitk.edu.in

Gandla Vikram
Information Technology
NITK SURATHKAL
Manglore, Karnataka, India
vikram.221ai044@nitk.edu.in

Duda Chethan Swaroop
Information Technology
NITK SURATHKAL
Manglore, Karnataka, India
chethanswaroop.221ai018@nitk.edu.in

Aourva Deshmukh Deepak
Information Technology
NITK SURATHKAL
Manglore, Karnataka, India
aourvadeshmukh.221ai017@nitk.edu.in

***Abstract*—While many approaches are found to predict stock movements, still many traders, financial analysts and institutional investors are striving to get new models with advanced methodologies for it. The traditional techniques are more often based on historical stock prices without considering the required insights from financial market news and changes within a day. The main objective of this paper is to inspect the stock movements using a multi-a deep learning framework that combines time series data, sentiment analysis, and real-time intraday stocks for appropriate forecasting. It uses several NLP techniques like Dynamic Knowledge Graphs for integrating the model with market sentiment and using Neural Networks Like GRU for time series forecasting of real-time intraday stocks. The proposed analysis outperforms traditional analysis in terms of integrating sentiment analysis and market data for a better understanding of real-time stock forecasting along with intraday insights. This implies that traders, financial analysts, and institutional investors should leverage advanced forecasting to improve their investment decisions and adapt it to other financial forecasting tasks.**

***Index Terms*—Deep Learning, GRU, Dynamic Knowledge Graph, Event Detection, Sentiment Analysis, Intraday Stock Prediction**

I. Introduction

Solely depending on historical stock prices will provide the insights in the dynamics of the stock market in today's world, a lot of factors influence the stock prices so incorporating the news data into the time-series model along with traditional features will capture the insights in the stock movement. This project aims to solve this problem by incorporating the news data sentiments scores along with the close and volume from the historical stock prices. This will help in grabbing the stock trends to help investors and short-term traders as in this project we are predicting intraday and longterm(30). The limitation of existing stock prediction models lies in their reliance on single-source data, usually price-based. To fully understand market dynamics, it is necessary to incorporate additional information streams such as news sentiment, market events, and even social media sentiment to improve the accuracy and responsiveness of stock predictions.The uniqueness of this project is implementing the dynamic-Knowledge graph using news data and using their sentiments to capture the stock market trends which will help the investors about the risks involved in that particular stock prediction.

Contributions of this project

1) We implemented a novel multi-modal deep learning approach to integrate different types of heterogeneous data like historical stock prices ,market sentiment to benefit the stock movement prediction.

2) We also explored on innovative aspects as event detection of companies.

3) We also reduced the gap between traditional models and practical trading applications by focusing on real-time intraday stock prediction.

4) We emphasized on performance improvement to highlight our model from baseline traditional techniques to demonstrate it's significance in evaluation metrics.

The paper is structured as follows: Section 2 reviews related works, Section 3 details the methodology, Section 4 discusses results and analysis, and Section 5 concludes with key findings and future directions

II. Literature Review

1. Hybrid Information Mixing Module for Stock Movement Prediction (2023)

Jooweon Choi et al. [1] introduced a hybrid model that combines price data with text data for stock prediction. The strengths lie in the novel integration of multiple data sources, which leads to better predictive accuracy. However, the paper highlights the need to evaluate the model on additional datasets and incorporate other external data sources like macroeconomic indicators.

2. Predicting Stock Price Movements in Volatile Markets: A Multi-Model Fusion Approach (2024)

Poojitha et al. [2] focused on combining sophisticated sentiment analysis with joint model training to improve prediction accuracy. Despite its innovation, it struggles with data overload and the quality of incoming data, especially during volatile market conditions. The authors recommend further refining models to better handle real-time data and adapt to market changes.

3. Stock Market Prediction Using AI (2023)
Kanthimathi et al. [3] explored the use of machine learning algorithms for stock prediction, highlighting the efficacy of advanced analytics in improving model performance. Although the paper reports good results, challenges arise in terms of data quality, which impacts overall model reliability. The study suggests further exploration of ensemble methods and improved data diversity.

4. Multi-task Transformer for Stock Market Trend Prediction (2022)
Seyed Morteza Mirjebreili et al. [4] proposed a multi-task learning model using TimeVec embedding, which offers enhanced prediction accuracy. However, the complexity of the model leads to significant computational costs. Future work can focus on simplifying the model while maintaining its high performance.

5. Integrating Sentiment Analysis and Knowledge Graphs for Enhanced Stock Trend Prediction (2023)
Zhang [5] introduced the idea of combining sentiment analysis with knowledge graphs to boost prediction accuracy. While the approach yields promising results, it heavily depends on the availability of high-quality data, making real-time applications challenging.

6. Stock Movement Prediction with Social Sentiments and Interactional Data (2022)
Zhang [6] integrated social sentiment and interaction data for enhanced stock prediction. The study demonstrates improved prediction accuracy, but struggles with noise in social media data. The authors suggest refining sentiment analysis techniques to reduce noise and improve prediction reliability.

7. Cracking the Code: Sentiment Analysis for Moroccan Stock Market Forecasting (2023)
Sandeep et al. [7] focused on forecasting stock trends in the Moroccan stock market using sentiment analysis. Despite notable improvements in market understanding, limitations arise from the availability of accurate data. The study proposes incorporating advanced machine learning techniques to enhance prediction precision.

8. Public Sentiment Analysis in Twitter Data for Stock Price Movements (2020)
Bing et al. [8] achieved high predictive accuracy by analyzing Twitter sentiment for specific companies. However, the data quality is a concern due to the prevalence of bots and noise. Recommendations include the inclusion of more diverse data sources for better prediction accuracy.

9. Predictive Precision: LSTM-Based Analytics for Real-time Stock Market Visualization (2022)
Saravanan et al. [9] focused on using LSTM models for real-time stock market visualization and forecasting. While the model achieves high accuracy, it faces issues with overfitting and data subjectivity. Further work can focus on integrating more emotional categories into sentiment analysis for better predictions.

10. Improving Stock Prediction with roBERTa and LSTM (2021)
Poornima et al. [10] focused on sentiment analysis using social media is integrated with stock price forecasting. The approach improves accuracy but relies heavily on Twitter data, which may not fully represent market sentiment. Future work should incorporate financial reports and broader market data for enhanced predictions.

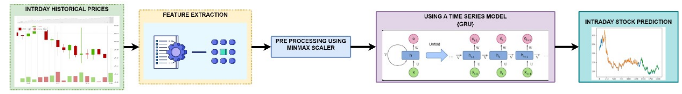

Fig. 1. Block Diagram For Intraday Stock Prediction

Outcomes of literature Survey
1. The combination of BERT for sentiment analysis and LSTM for stock forecasting shows an improvement in predictive accuracy.

2. LSTM-based models integrated with news data demonstrate strong performance in price prediction.

3. Incorporating additional sources like financial reports and social media sentiment can further enhance prediction reliability.

III. Methodology of the Project

We used different methodologies for long term prediction and intraday prediction

Methodology for Intraday predictions

In this Methodology shown in fig(1) we focus on the historical prices which are collected from the API yfinance and later. From this raw data we use only the feature(closing prices) and using minmax scalar we compresed the data to fit in the range between 0 to 1 which will help the time-series model(GRU) in performing it efficiently. Building GRU model:

First GRU Layer: This layer has 50 units, which are neurons designed to capture complex patterns in sequential data. The return sequences parameter is set to True, allowing this layer to output a sequence that can be fed into the next GRU layer.

Second GRU Layer: Another GRU layer with 50 units is used to further process the output from the first GRU layer. Here, return sequences is set to False, meaning that this layer will output only the last hidden state.

Dense Layer: A Dense layer with 25 units is added after the GRU layers. This layer fully connects each neuron to all neurons in the previous layer, allowing for a further transformation and refinement of the features extracted by the GRU layers.

Output Layer: The final Dense layer has only 1 unit, as it's responsible for producing the model's output—a single predicted stock price value.

This Model will help us in the Intraday prediction, where it will display the stock movement within the day so it will help the stock market traders to make meaning full decisions whether to buy or not.

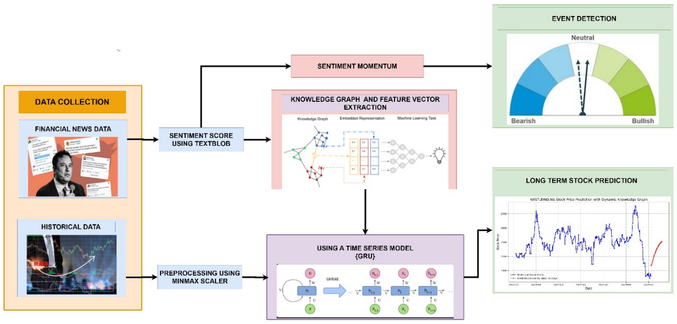

Fig. 2. Block Diagram For Long Term Stock Prediction

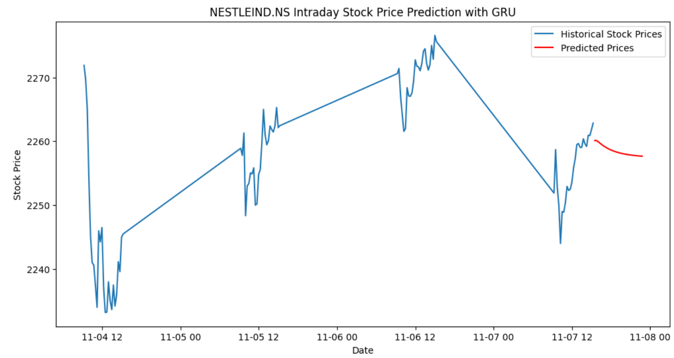

Fig. 3. Intraday Stock Prediction

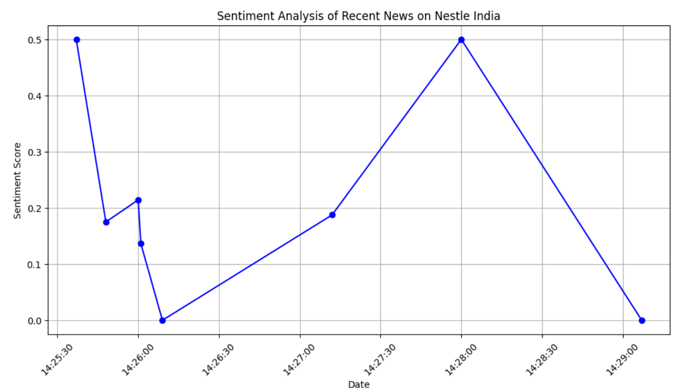

Fig. 4. Sentiment Score

Methodology for Long term predictions

Event detection through Sentiment momentum

Here in this methodology we collect news articles from yfinance. In these articles we take in a Dataframe containing at least the title. To this Dataframe we calculate the sentiment score using TextBlob, which ranges between the -1 to +1. This step leads to sentiment momentum calculation where we calculate the difference between the sentiment moment between to consecutive articles this helps in event detection of the stock whether its bullish or bearish. We calculate whether the stock event comes out to be bullish or bearish by the assigned threshold value. If the sentiment momentum is greater than 0.1 and the sentiment is positive, this indicating in the increasing positive sentiment and bullish. And if the momentum is less than -0.1 and the sentiment is negative then its bearish. And its neutral if neither conditions are met.

Building the Sentiment Enhanced Dynamic Knowledge Graph Using TextBLob

Here we fetch news articles, and Historical stock prices from the API yfinance. TextBlob is used to assign sentiment scores for each news headline polarity ranging from -1 to +1. As the news articles are retrieved from different time-zones we convert them to IST. After assigning the sentiment scores we then use MinMax Scaler to minimize the sentiment scores in the range 0-1. We then proceed to build the knowledge graph with the available sentiment information and the library NetworkX library. The nodes in the knowledge graph is either company name or publisher's name. The edges that connect to the nodes has sentiment weight it and the color represents whether it is negative or positive. Once the knowledge graph is build we extract the topic feature vector which has the nodes form the knowledge graph and its sentiment score.

Preparing the Stock Data and Sentiment Data

The Historical stock prices data is further scaled with the help of MinMax Scaler. We take Closing prices from the Historical stock prices and topic feature vector as features for our time-series model.

Creating the Time-series GRU

Here we set the sequence of 60 days, where the 60 days sequence of closing price and sentiment value is input for the GRU model. For each day in the 60 days sequence we target next day in the 30 days prediction for our output.The future prices are predicted through iterative prediction where each prediction is used as input for the subsequent time step allowing the model to generate a continuous forecast. The GRU has its first layer set to 50 units and return sequences, allowing it to connect to the next GRU layer. The Second GRU layer also has 50 units which captures the temporal dependencies in the data. The output layer of GRU and the knowledge graph features are concatenated. This layer passes through a final dense layer with one unit to produce the models output "The Predicted Stock Price"

IV. Results and Discussion

Visualization :

This section presents the results obtained from intraday and long-term stock predictions, using the GRU time-series model and sentiment analysis based on news data.

Intraday Predictions In the intraday model, we observed price fluctuations within the day. The GRU model captured the downward price movements, showing a dip of approximately 5–10 rupees. With the starting price at 2260 rupees, the model accurately reflected these short-term variations, aiding stock market traders in making timely decisions based on within-day trends. This is demonstrated in Fig. 2, Fig. 3, and Fig. 4, which

Title	Publisher	Published	Sentiment	Momentum	Event
Q3 2024 Grand Canyon Education Inc Earnings Call	Thomson Reuters	1970-01-01 05:30:01	0.5	-	Neutral
Metro by T-Mobile Drops Holiday Deals	Business Wire	1970-01-01 05:30:01	0.0	-0.5	Neutral
Argo's September 2024 Oil Production Update	Newsfile	1970-01-01 05:30:01	0.0	0.0	Neutral
Mexican economy chief says wants to sound out	Reuters	1970-01-01 05:30:01	0.2	0.2	Bullish
Tejon: Q3 Earnings Snapshot	Associated Press	1970-01-01 05:30:01	0.0	-0.2	Neutral
PayPay, Alipay+ expand e-wallet payment options	Electronic Payments	1970-01-01 05:30:01	0.0	0.0	Neutral
Dutch chipmaker NXP sees sales growth averaging	Reuters	1970-01-01 05:30:01	0.0	0.0	Neutral
Jefferies downgrades Palantir stock	Investing.com	1970-01-01 05:30:01	0.0	0.0	Neutral

TABLE I
NEWS SENTIMENT SCORE AND EVENT PREDICTIONS

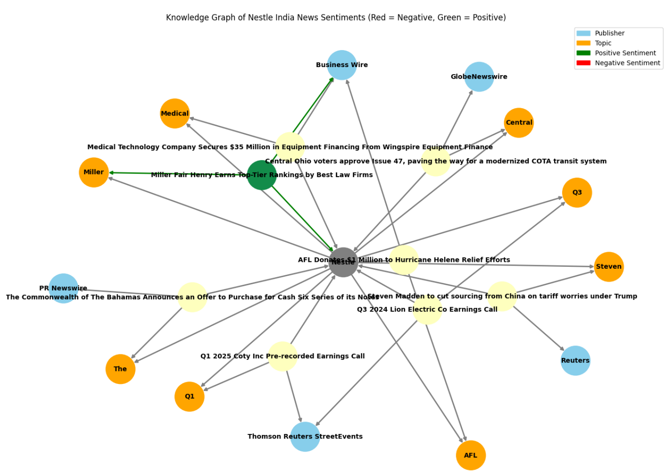

Fig. 5. Dynamic Knowledge Graph

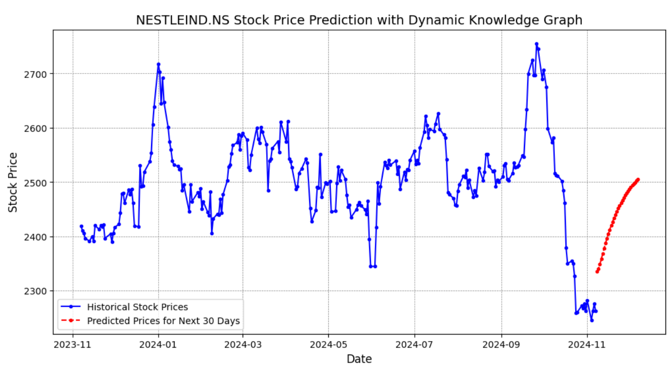

Fig. 6. Long Term Stock Prediction With Sentiment Analysis

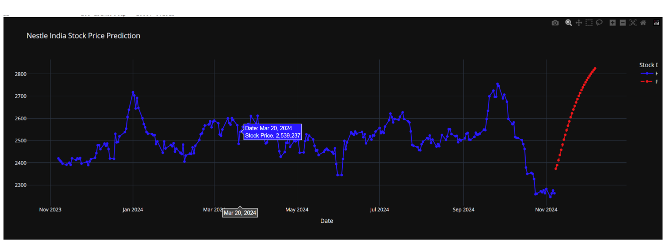

Fig. 7. Long Term Stock Prediction With only GRU

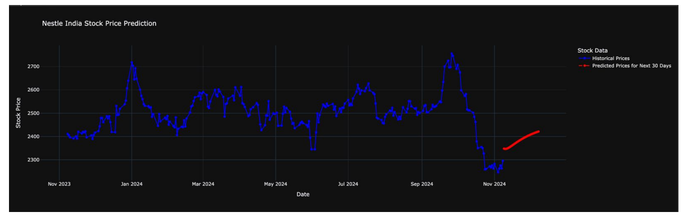

Fig. 8. Long Term Stock Prediction with LSTM

illustrate the intraday predictions, showing how effectively the model adapts to short-term price changes.

Long-Term Predictions For long-term predictions, our model combines sentiment momentum from news articles with time-series analysis. Over a 30-day period, the stock price demonstrated an increase of up to 200 rupees. This rise was in line with the cumulative bullish sentiment detected from news data. Fig. 5 shows this forecasted trend, where the predicted price increase reflects the model's effectiveness in capturing broader market trends. By integrating sentiment with time-series data, the model showed enhanced predictive accuracy for long-term price movements.

Sentiment Analysis and Event Detection The sentiment analysis on news headlines, as displayed in Table I, facilitated event detection impacting stock trends. For instance, articles with positive sentiment momentum and high sentiment scores aligned with bullish market behavior, while negative momentum indicated bearish tendencies. This analysis, shown in the "Event Prediction" column, provides additional insight into market sentiment and supports the model's long-term prediction capabilities.

Comparision for Evaluation Metrics

We have compared the results with knowlege graph and without knowledge graph along with LSTM and GRU

V. CONCLUSION

This project aimed to enhance stock prediction accuracy by combining GRU-based time-series modeling with sentiment analysis from news articles. The results indicate that the hybrid model successfully captures both short-term and long-term price trends. For intraday predictions, the model effectively mirrored within-day fluctuations, with observed price dips ranging from 5 to 10 rupees. This ability to

	Time	With Knowledge Graph	Without Knowledge Graph	Difference (With - Without)
0	0	2361.07	2284.69	76.3838
1	1	2374.38	2284.15	90.2339
2	2	2390.83	2282.99	107.837
3	3	2407.77	2281.69	126.083
4	4	2424.21	2280.44	143.775
5	5	2439.8	2279.29	160.508
6	6	2454.45	2278.27	176.185
7	7	2468.18	2277.36	190.823
8	8	2481.02	2276.54	204.474
9	9	2493.02	2275.82	217.199
10	10	2504.24	2275.18	229.058
11	11	2514.71	2274.6	240.107
12	12	2524.49	2274.09	250.398
13	13	2533.62	2273.64	259.982
14	14	2542.14	2273.23	268.906
15	15	2550.08	2272.87	277.211

Fig. 9. Comparision table for next 15 days with knowledge graph and without knowledge graph

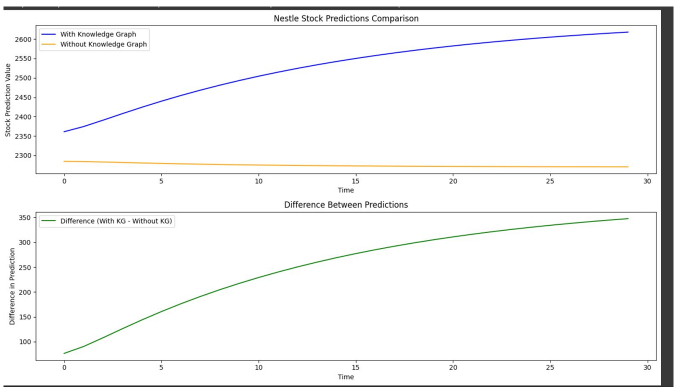

Fig. 10. Comparision Graph for next 30 days with knowledge graph and without knowledge graph

capture short-term price movements equips traders with actionable insights for day trading. The long-term predictions also proved promising, with the model forecasting a price increase of up to 200 rupees over a 30-day period. This trend aligns with the bullish sentiment detected in news data, underscoring the model's potential for anticipating broader market trends. The inclusion of sentiment analysis through TextBlob provided an additional layer of insight, particularly in detecting market events. The sentiment momentum analysis demonstrated a correlation between sentiment changes and stock price movements, with positive sentiment momentum often aligning with bullish trends and negative momentum indicating bearish tendencies. This suggests that sentiment-driven event detection can improve predictive accuracy by incorporating the influence of market sentiment. Despite the model's effectiveness, certain limitations must be acknowledged. The accuracy of sentiment scoring is contingent on the quality of news data, and sentiment analysis methods may introduce biases. Additionally, the model does not account for external economic or geopolitical events, which can impact stock prices independently of sentiment trends. In conclusion,

	Time	With GRU	With LSTM	Difference (With - Without)
0	0	2361.07	2347.82	13.2522
1	1	2374.38	2346.6	27.7837
2	2	2390.83	2347.25	43.5781
3	3	2407.77	2349.13	58.646
4	4	2424.21	2351.79	72.4238
5	5	2439.8	2354.93	84.8713
6	6	2454.45	2358.34	96.1094
7	7	2468.18	2361.89	106.285
8	8	2481.02	2365.49	115.531
9	9	2493.02	2369.06	123.956
10	10	2504.24	2372.59	131.648
11	11	2514.71	2376.03	138.679
12	12	2524.49	2379.38	145.108
13	13	2533.62	2382.63	150.987
14	14	2542.14	2385.78	156.359
15	15	2550.08	2388.82	161.265

Fig. 11. Comparision table for next 15 days with LSTM and GRU

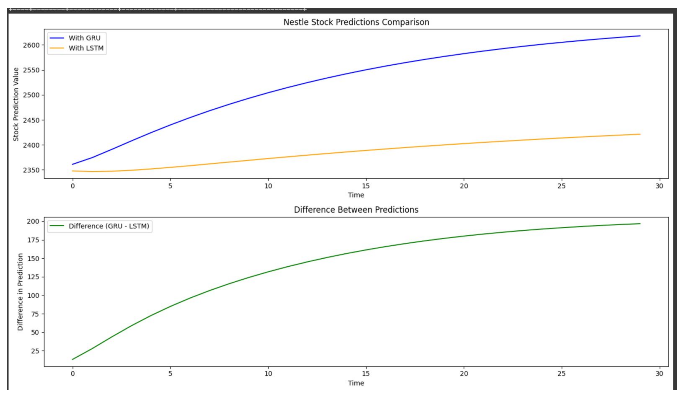

Fig. 12. Comparision Graph for next 30 days with LSTM and GRU

the combination of GRU time-series analysis with sentiment momentum provides a robust framework for stock prediction. Future improvements could include more advanced natural language processing techniques to enhance sentiment scoring accuracy, the incorporation of real-time data for timely event detection, and the addition of economic indicators to capture a broader range of market dynamics. These enhancements could increase prediction accuracy and provide a valuable tool for traders and investors looking to navigate both intraday and long-term market trends.

References

[1] J. Choi, S. Yoo, X. Zhou, and Y. Kim, "Hybrid Information Mixing Module for Stock Movement Prediction," *IEEE Access*, vol. 11, pp. 28781–28790, 2023, doi: 10.1109/ACCESS.2023.3258695.

[2] K. Poojitha, L. V. Krishna Rao, K. T. Dhilip Kumar, K. C. Chandra, and M. Varsha, "Predicting Stock Price Movements in Volatile Markets: A Multi-Model Fusion Approach," in *Proceedings of the 2024 Third International Conference on Distributed Computing and Electrical*

Circuits and Electronics (ICDCECE), 2024, pp. 1–6, doi: 10.1109/ICDCECE60827.2024.10548212.

[3] N. Kanthimathi, N. Saranya, C. I. Hermina, and M. S. Varma, "Stock Market Prediction Using AI," in *Proceedings of the 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, 2023, pp. 1–5, doi: 10.1109/ViTECoN58111.2023.10157327.

[4] S. M. Mirjebreili, A. Solouki, H. Soltanalizadeh, and M. Sabokrou, "Multi-Task Transformer for Stock Market Trend Prediction," in *Proceedings of the 2022 12th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2022, pp. 101–105, doi: 10.1109/ICCKE57176.2022.9960122.

[5] Z. Zhang, "Integrating Sentiment Analysis and Knowledge Graphs for Enhanced Stock Trend Prediction-A Deep Learning Approach," in *Proceedings of the 2023 3rd International Signal Processing, Communications and Engineering Management Conference (ISPCEM)*, 2023, pp. 16–20, doi: 10.1109/ISPCEM60569.2023.00009.

[6] Z. Zhang, "Integrating Sentiment Analysis and Knowledge Graphs for Enhanced Stock Trend Prediction-A Deep Learning Approach," in *Proceedings of the 2023 3rd International Signal Processing, Communications and Engineering Management Conference (ISPCEM)*, 2023, pp. 16–20, doi: 10.1109/ISPCEM60569.2023.00009.

[7] U. Sandeep, T. V. Vardhan, J. A. Mayan, B. Yaswanth, and A. R. Doni, "Cracking the Code: Unleashing the Power of Sentiment Analysis & ML for Moroccan Stock Market Forecasting," in *Proceedings of the 2023 8th International Conference on Communication and Electronics Systems (ICCES)*, 2023, pp. 1274–1278, doi: 10.1109/ICCES57224.2023.10192697.

[8] L. Bing, K. C. C. Chan, and C. Ou, "Public Sentiment Analysis in Twitter Data for Prediction of a Company's Stock Price Movements," in *Proceedings of the 2014 IEEE 11th International Conference on e-Business Engineering*, 2014, pp. 232–239, doi: 10.1109/ICEBE.2014.47.

[9] V. Saravanan, A. Pramod, L. Poudel, and P. Paramasivam, "Predictive Precision: LSTM-Based Analytics for Real-time Stock Market Visualization," in *Proceedings of the 2023 IEEE International Conference on Big Data (BigData)*, 2023, pp. 1–6, doi: 10.1109/BigData59044.2023.10443091.

[10] N. Poornima, D. Abilash, and M. Theodaniel, "Improvising the Stock Prediction by Integrating with roBERTa and LSTM," in *Proceedings of the 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)*, 2023, pp. 1–8, doi: 10.1109/IConSCEPT57958.2023.10169904.