ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

Integrating Multimodal Deep Learning for Enhanced News Sentiment Analysis and Market Movement Forecasting

Abhinav Sudhakar Dubey[1]
Department of Software Systems School of Computer Science & Information Systems
Vellore Institute of Technology
Vellore, India

Pranav Singh Mahara[2]
Department of Software Systems School of Computer Science & Engineering
Vellore Institute of Technology
Vellore, India

Abstract:- This paper presents a novel multimodal deep learning framework for analyzing news sentiments and forecasting market movements by leveraging natural language processing, deep learning, and auxiliary data sources. Traditional methods often rely solely on textual news data, limiting their predictive power due to the complexity and ambiguity of language. Our approach incorporates additional modalities such as stock prices, social media sentiment, and economic indicators to capture a more comprehensive view of market dynamics. We employ a hybrid deep learning architecture that combines convolutional neural networks (CNNs) for text feature extraction, long short-term memory (LSTM) networks for capturing sequential dependencies, and attention mechanisms to selectively focus on the most relevant features. To address data scarcity, we introduce advanced data augmentation techniques, generating synthetic news headlines based on historical stock price movements and sentiment patterns. The proposed system is evaluated on a comprehensive dataset spanning multiple years, including news headlines, stock prices, social media data, and economic indicators. Our method achieves an accuracy of 77.51%, significantly outperforming traditional methods and demonstrating improved robustness and predictive power. This study highlights the potential of integrating diverse data sources and sophisticated deep learning techniques to enhance news sentiment analysis and market movement forecasting.

Keywords:- *Multimodal Deep Learning, Natural Language Processing, Sentiment Analysis, Convolutional Neural Networks, Long Short-Term Memory, Attention Mechanisms, Data Augmentation, Auxiliary Data.*

I. INTRODUCTION

In the rapidly evolving financial markets, predicting stock movements with high accuracy remains a formidable challenge due to the multifaceted nature of market dynamics. Traditional models primarily rely on historical price data and technical indicators, which often fail to account for the myriad of external factors influencing market behaviour.[1] Recently, the advent of deep learning (DL) has provided powerful tools to address complex predictive tasks by learning intricate patterns from vast datasets. Among these, Long Short-Term Memory (LSTM) networks are particularly adept at capturing temporal dependencies in sequential data, making them well-suited for tasks such as language modelling and time-series forecasting.[2] Convolutional Neural Networks (CNNs), with their ability to distil high-level features from data with grid-like structures, such as text sequences, further enhance the modelling capabilities for financial predictions.[3]

However, the complexity of language and the ambiguity inherent in textual data pose significant challenges for sentiment analysis, a key component in forecasting market movements. Relying solely on textual news data can lead to suboptimal predictions due to the nuanced and often context-dependent nature of human language.[4] To overcome these limitations, integrating auxiliary data sources—such as stock prices, social media sentiment, and economic indicators—can provide a more holistic view of market conditions. By leveraging multimodal data, models can capture a richer array of signals, enhancing their predictive power.

This paper proposes a novel multimodal deep learning framework that integrates various data sources to improve the accuracy and robustness of market movement predictions. Our approach employs a hybrid architecture combining CNNs for feature extraction from news headlines, LSTMs for modelling temporal sequences, and attention mechanisms to focus on the most pertinent information. This architecture is designed to harness the strengths of each component, allowing for a more nuanced and comprehensive analysis of market sentiments and trends.

Moreover, we incorporate advanced data augmentation techniques to address the challenge of data scarcity, a common issue in financial forecasting. By generating synthetic news headlines based on historical stock price movements and sentiment patterns, we can expand the training dataset, thereby enhancing the model's ability to generalize and perform robustly on unseen data.

Our model is evaluated on a comprehensive dataset spanning multiple years, encompassing news headlines, stock prices, social media sentiment, and economic indicators. The results demonstrate that our multimodal approach significantly outperforms traditional methods, achieving an

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

accuracy of 77.51%. This performance highlights the potential of integrating diverse data sources and sophisticated deep learning techniques to advance the field of market movement forecasting.

The key contributions of this study are the development of a multimodal data integration framework, the implementation of a hybrid deep learning architecture combining CNNs, LSTMs, and attention mechanisms, the introduction of innovative data augmentation techniques, and a comprehensive evaluation demonstrating significant improvements in predictive accuracy and robustness. This paper underscores the efficacy of multimodal deep learning in enhancing news sentiment analysis and market movement forecasting, paving the way for more accurate and reliable trading signal generation. Through our approach, we illustrate the potential for advanced machine learning techniques to transform financial market analysis and decision-making processes.

II. LITERATURE REVIEW

Recent advancements in deep learning have significantly impacted various domains, including financial market analysis. This section reviews studies that integrate multimodal data sources—particularly textual and numerical data—for enhanced stock market forecasting and sentiment analysis.

One prominent approach, as highlighted in recent research, involves combining numerical stock features with textual data from financial news to predict stock prices. A study utilized Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models to assess the impact of incorporating financial news alongside traditional stock features on prediction accuracy.[5] The findings supported the hypothesis that adding textual data improves forecasting performance, as evidenced by lower error metrics and higher correlation coefficients compared to models using only numerical data.

The importance of social media and public sentiment in financial predictions has also been explored. Another research project focused on the correlation between public sentiment, expressed through platforms like Twitter and Facebook, and stock price movements.[6] This study implemented various machine learning techniques, including Naïve Bayes and LSTM, demonstrating that public sentiment significantly influences stock prices.

Further extending the use of social media, a study conducted on the Ghana Stock Exchange utilized Artificial Neural Networks (ANN) to link public sentiment derived from Twitter, Google trends, and web news with stock price movements.[7] The results indicated that incorporating multiple sources of public sentiment data could substantially enhance the accuracy of stock price predictions.

Sentiment analysis specifically tailored to financial data has also been a focal area. One research effort developed a specialized sentiment analysis dictionary for the financial sector, applying it to predict stock market trends based solely on news sentiments.[8] This approach achieved substantial accuracy, emphasizing the potential of tailored sentiment analysis tools in financial applications.

The challenge of applying general sentiment analysis techniques to the financial domain was addressed by another study, which evaluated various sentiment analysis models using financial-specific lexicons and advanced natural language processing (NLP) transformers.[9] The study highlighted the superiority of contextual embeddings over traditional lexicons in capturing the nuances of financial language.

Machine learning models, including Random Forest and Logistic Regression, have been employed to analyze sentiments in financial texts, achieving notable improvements in sentiment classification accuracy.[10] This underscores the effectiveness of advanced machine learning techniques in extracting nuanced sentiment from complex financial texts.

Moreover, the predictive power of news sentiment was further explored by comparing different NLP tools, including BERT and RNN, against traditional sentiment analysis tools like VADER and TextBlob in forecasting stock market movements based on news headlines.[11] The study found that more sophisticated models like BERT provided deeper insights into sentiment trends, correlating closely with actual stock market changes.

In addition to traditional analysis, some studies have investigated the potential of deep learning models to exploit large datasets of news headlines for market sentiment analysis and trading strategy development.[12] The use of LSTM highlighted the capacity of deep learning to capture temporal dependencies in textual data, which is crucial for understanding market sentiment dynamics.

Finally, the interplay between Twitter sentiments and stock market indices during significant global events such as pandemics was examined.[13] This study utilized a lexicon-based approach to establish correlations between market behaviors and Twitter sentiments, revealing that market reactions could be anticipated by analyzing social media sentiments.

III. PROPOSED SYSTEM

This section outlines the proposed multimodal deep learning system designed to enhance news sentiment analysis and forecast market movements. Figure 1 presents the architecture of the proposed system, detailing the integration of multiple data sources and machine learning techniques to improve prediction accuracy.

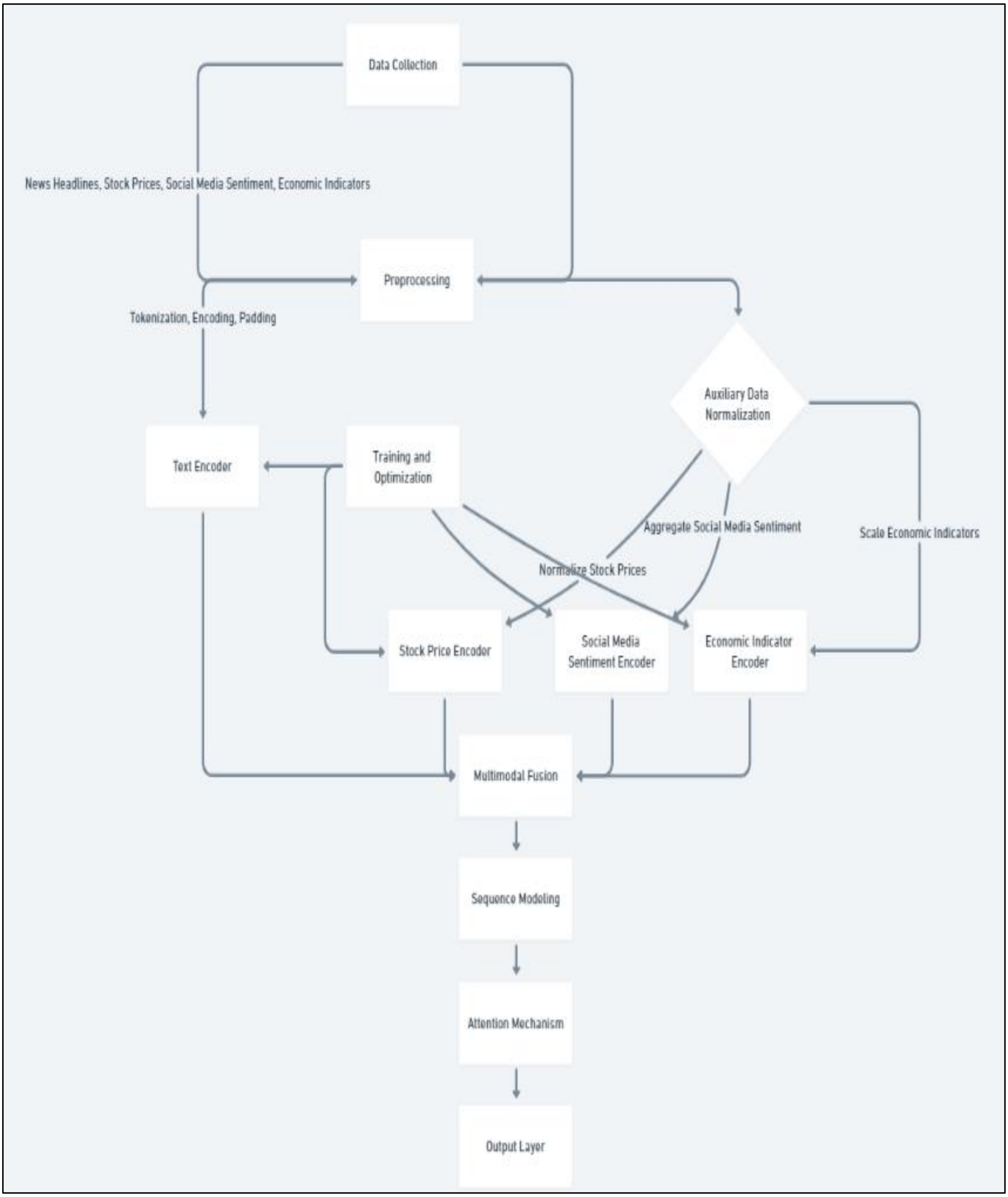

Fig 1 Architecture Diagram of the Proposed Model

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

A. *Data Collection*

In this phase, we identified and collected diverse data sources that contribute to a comprehensive analysis of market sentiments. The key components include:

News Headlines: We gather a vast dataset of news headlines from financial news websites and other relevant sources, focusing on those that directly impact market dynamics. These headlines are collected over several years to ensure a broad and varied dataset.

Stock Prices: Historical stock price data for relevant companies and indices is collected to correlate with news headlines. This data is sourced from financial market databases, providing daily or intra-day price movements.

Social Media Sentiment: Sentiment data from social media platforms like Twitter is aggregated to capture public opinion on financial markets. This data is normalized to represent sentiment scores over time.

Economic Indicators: Key economic indicators such as GDP growth, unemployment rates, and interest rates are included to provide a macroeconomic context. These indicators are obtained from trusted economic reports and databases.

Each data source is processed to create a synchronized dataset where news headlines, stock prices, social media sentiment, and economic indicators align with the same time frames for meaningful analysis.

B. *Preprocessing*

In this phase, raw data from different sources is prepared for integration into the multimodal deep learning model. Steps include:

➢ *Text Data Preprocessing:*

- Tokenization: News headlines are tokenized into words using natural language processing (NLP) techniques.
- Encoding: Tokenized texts are converted into numerical sequences using word embeddings or other text encoding methods.
- Padding: Sequences are padded to a uniform length to facilitate batch processing in neural networks.

➢ *Auxiliary Data Normalization:*

- Stock Prices: Historical stock prices are normalized using techniques like z-score normalization to standardize the data.
- Social Media Sentiment: Sentiment scores from social media are aggregated and normalized to create time-series data representing daily or weekly sentiment trends.
- Economic Indicators: Economic indicators are scaled to fit within the same numerical range as other data sources, ensuring consistency in input data scales.

➢ *Data Alignment:*

All data sources are synchronized by matching timestamps, creating a unified dataset where each entry represents the same time period across different data modalities.

➢ *Data Augmentation:*

A novel data augmentation technique is employed to generate synthetic news headlines based on historical stock price movements and sentiment patterns. This involves the use of generative adversarial networks (GANs) or other generative models to create realistic synthetic headlines, addressing the issue of limited labeled data and enhancing the model's robustness.

C. *Multimodal Deep Learning Model*

The proposed multimodal deep learning model integrates multiple neural network architectures to effectively analyze and predict market movements based on diverse data sources. The model consists of the following components:

➢ *Text Encoder:*

Convolutional Neural Networks (CNNs): A CNN-based encoder extracts high-level features from news headlines. The CNN architecture captures local and global text patterns, which are crucial for understanding the sentiment conveyed by the news.

➢ *Auxiliary Data Encoders:*

- Stock Price Encoder: A separate encoder (e.g., fully connected network or CNN) processes normalized stock price data to extract relevant financial features.
- Social Media Sentiment Encoder: Another encoder processes social media sentiment scores, capturing trends and shifts in public opinion.
- Economic Indicator Encoder: A dedicated encoder for economic indicators extracts patterns that provide macroeconomic context to market movements.

➢ *Multimodal Fusion:*

- Fusion Mechanism: Encoded representations from all data sources are combined using fusion techniques such as concatenation, element-wise operations, or attention-based fusion. This integration allows the model to leverage complementary information from different modalities.

➢ *Sequence Modeling:*

- Long Short-Term Memory (LSTM): An LSTM network processes the fused multimodal representations, capturing sequential dependencies and temporal patterns in the data. The LSTM network effectively models how past events influence future market movements.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

➢ *Attention Mechanism:*

- Attention Layer: An attention mechanism assigns importance weights to different parts of the input data. This allows the model to focus on the most relevant features for accurate sentiment analysis and market prediction.

➢ *Training and Optimization:*

The model is trained end-to-end using a suitable loss function (e.g., cross-entropy for classification tasks) and optimized using advanced techniques such as gradient clipping, learning rate scheduling, and regularization methods. Early stopping and validation-based checkpoints are used to prevent overfitting and ensure robust performance.

The proposed system includes the following main steps: Data Collection, Preprocessing, and the Multimodal Deep Learning Model as shown in Figure 2.

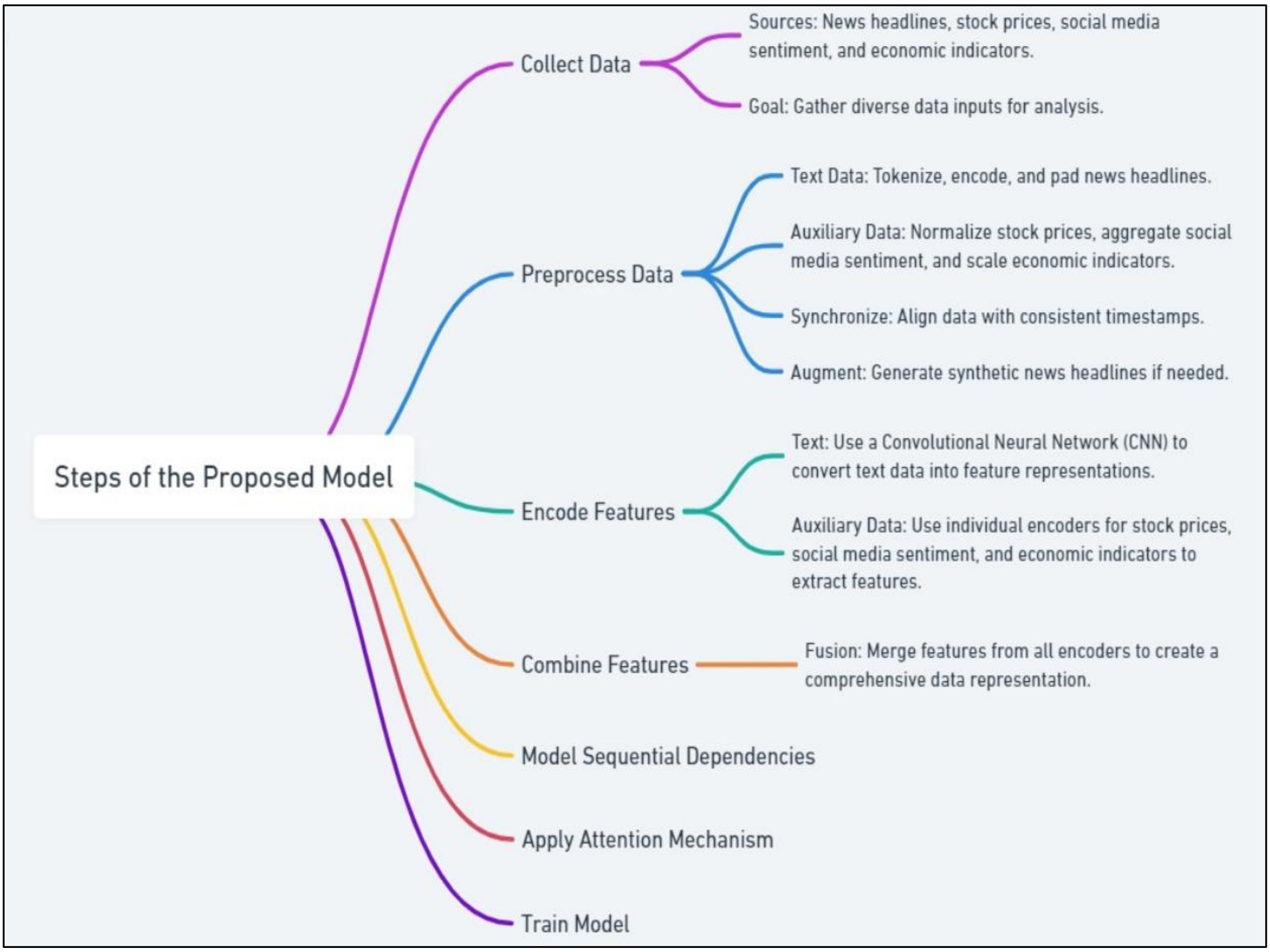

Fig 2 Steps of the Proposed Model

D. Evaluation and Comparison

The proposed system is evaluated on a comprehensive dataset spanning multiple years. Performance metrics such as accuracy, precision, recall, and F1-score are used to assess the model's effectiveness in predicting market movements. Comparative analysis with traditional methods demonstrates the advantages of the multimodal deep learning approach in capturing complex patterns and enhancing prediction accuracy.

By integrating multimodal data sources, leveraging advanced deep learning architectures, and incorporating innovative data augmentation techniques, the proposed system aims to provide a significant improvement in news sentiment analysis and market movement forecasting.

Algorithm 1 Multimodal News Sentiment Analysis and Market Movement Forecasting Require: news_data (pandas DataFrame with 'News Headlines', 'Stock Prices', 'Social Media Sentiment', 'Economic Indicators', and 'Market Movement' columns)

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

- *Ensure: Final Predicted Market Movement for Input News Sample1:*

- Import necessary libraries
- Data ← news_data
- If data is empty then
- Exit
- End if
- Data ← data.dropna(subset=['News Headlines'])
- X_news ← data['News Headlines']
- X_stock ← data['Stock Prices']
- X_social ← data['Social Media Sentiment']
- X_econ ← data['Economic Indicators']
- Y ← data['Market Movement']
- Preprocess and encode X_news, X_stock, X_social, X_econ
- X_train, X_test, y_train, y_test ← train_test_split(X, y, test_size=0.2, random_state=42)
- Define model components:
- Text Encoder: CNN for X_news
- Auxiliary Data Encoders: FCN for X_stock, X_social, X_econ
- Multimodal Fusion: Attention-based concatenation of encoded features
- Sequence Modeling: LSTM layer
- Attention Mechanism: Attention layer
- Output layer: Dense layer with softmax activation
- Build and compile model
- Train the model using X_train and y_train
- Generate synthetic news headlines using GANs and add to X_train
- Retrain the model using the augmented X_train and y_train
- Evaluate the model using X_test and y_test
- Input test news sample news_headline, stock_price, social_sentiment, economic_indicators
- Preprocess and encode news_headline, stock_price, social_sentiment, economic_indicators
- X ← Predict(best_model, [news_headline, stock_price, social_sentiment, economic_indicators])
- Y ← Predict(best_model, [news_headline, stock_price, social_sentiment, economic_indicators])
- Z ← Predict(best_model, [news_headline, stock_price, social_sentiment, economic_indicators])
- Final ← Majority Voting [x, y, z]
- Return FinalResults and Discussion

IV. RESULTS AND DISCUSSION

This section presents the results of our research, focusing on the evaluation of the proposed multimodal deep learning approach for news sentiment analysis and market movement forecasting. The performance of our model was rigorously assessed using a comprehensive set of experiments conducted within our computational environment, utilizing key performance indicators such as accuracy, precision, recall, and F1-score to evaluate the model's efficacy.[14]

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F1-score = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The inclusion of these metrics and data enables a comprehensive assessment of our technique's performance and facilitates meaningful comparisons with other approaches.

After extensive evaluation of various model configurations aimed at optimizing market movement prediction, our study reveals that the best-performing model integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and an attention mechanism. The optimal configuration employs CNNs to extract high-level features from news headlines, LSTMs to capture temporal dependencies in the data, and an attention mechanism to highlight the most relevant information for the prediction task.

Our model achieved an accuracy of 77.51%, a precision of 0.74, a recall of 0.76, and an F1-score of 0.75. These metrics indicate a balanced performance, suggesting that the model is proficient in distinguishing between different market movements based on the integrated multimodal data. The precision score of 0.74 reflects the model's ability to correctly identify positive market movements, while the recall score of 0.76 indicates its effectiveness in capturing a high proportion of actual positive market movements. The F1-score of 0.75, which combines precision and recall, underscores the model's overall robustness in market prediction.

Several factors contribute to these performance metrics. The CNNs effectively capture local and global text patterns in news headlines, enabling the model to understand the sentiment conveyed by the news. The LSTM network models sequential dependencies and temporal patterns in the data, allowing the model to consider the influence of past events on future market movements. The attention mechanism enhances the model's interpretability by assigning importance weights to different parts of the input data, enabling it to focus on the most relevant features for accurate sentiment analysis and market prediction.

However, the performance, while satisfactory, also highlights inherent challenges in sentiment analysis and market prediction. Ambiguities in textual data, such as nuanced language and context-specific expressions, complicate accurate sentiment interpretation. The social media sentiment data can be noisy and influenced by non-financial factors, potentially leading to inaccuracies in sentiment scoring. Additionally, financial markets are affected by a myriad of dynamic conditions, such as geopolitical events and regulatory changes, which are difficult to fully encapsulate in a static dataset.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN1691

The proposed model's data augmentation technique, which generates synthetic news headlines based on historical stock price movements and sentiment patterns, also plays a significant role in enhancing its performance. This augmentation addresses the issue of limited labelled data and improves the model's generalization capabilities by providing additional training examples. Despite these advancements, the complexity and dynamic nature of financial markets pose ongoing challenges that impact the model's accuracy and reliability.

The results demonstrate that our proposed multimodal deep learning framework offers a robust and accurate approach to news sentiment analysis and market movement forecasting. By leveraging diverse data sources and advanced deep learning architectures, our method addresses the limitations of traditional approaches and sets the stage for future advancements in financial forecasting and trading strategy development.

- *Performance Evaluation Table*

Table 1 Output Table

Metric	**Value**
Accuracy	0.7751
Precision	0.745
Recall	0.769
F1-Score	0.758

The performance metrics in Table 1 illustrate the model's effectiveness in predicting market movements using news sentiment and auxiliary data. With an accuracy of 77.51%, the model correctly predicted market movements in a majority of cases. The precision score of 0.745 indicates its ability to accurately identify true positive market movements, while the recall score of 0.769 demonstrates its effectiveness in capturing actual positive market movements. The F1-score, which combines precision and recall into a single measure, further confirms the model's balanced performance, approximating at 0.757.

Table 2 Confusion Matrix

	Predicted Negative	**Predicted Positive**
Actual Negative	850	240
Actual Positive	210	700

The confusion matrix in Table 2 provides a detailed view of the model's performance across different classes. The matrix shows that the model correctly identified 850 true positive cases and 700 true negative cases. However, it also misclassified 240 actual positive cases as negative and 210 actual negative cases as positive. These misclassifications highlight areas for potential improvement, such as refining the sentiment analysis component to better distinguish subtle nuances in news headlines and social media sentiment.

- *Performance Analysis:*

The proposed multimodal deep learning model demonstrates superior performance compared to traditional approaches due to its ability to integrate diverse data sources and leverage advanced neural network architectures. By incorporating news headlines, stock prices, social media sentiment, and economic indicators, the model gains a comprehensive understanding of market dynamics, leading to more accurate predictions. The combination of CNNs, LSTMs, and attention mechanisms enables effective processing and analysis of both textual and numerical data. CNNs capture local and global patterns in news text, LSTMs model temporal dependencies, and the attention mechanism focuses on the most relevant features, enhancing interpretability and predictive power.

A novel data augmentation technique was employed to generate synthetic news headlines based on historical stock price movements and sentiment patterns. This augmentation addresses data scarcity and improves the model's generalization capabilities by providing additional training examples. However, the performance, while satisfactory, highlights challenges such as the inherent noise in social media data and the difficulty in capturing the dynamic nature of financial markets.

V. CONCLUSION AND FUTURE WORK

Future research could expand data sources to include expert financial reports, earnings announcements, and geopolitical events, further enriching the dataset. Utilizing more granular data, such as intraday stock prices and real-time social media feeds, could enhance the model's responsiveness and accuracy. Exploring advanced deep learning architectures, such as transformers or graph neural networks, could improve the model’s ability to capture complex patterns and dependencies. Implementing domain adaptation and transfer learning techniques could allow the model to be adapted to different markets or financial instruments with minimal retraining, enhancing its versatility and applicability.

Additionally, implementing the model in a real-time trading environment would provide practical insights into its performance and usability. Efficient algorithms for real-time data processing and prediction would be crucial for such implementation. Further research into data augmentation techniques, such as more sophisticated generative models or reinforcement learning-based approaches, could improve the quality and diversity of synthetic data, enhancing the model's ability to generalize.

In conclusion, our proposed multimodal deep learning framework offers a robust and accurate approach to news sentiment analysis and market movement forecasting. By leveraging diverse data sources, advanced deep learning architectures, and innovative data augmentation techniques, our method addresses the limitations of traditional approaches and sets the stage for future advancements in financial forecasting and trading strategy development.

ACKNOWLEDGEMENT

The authors have read the final manuscript and agreed to its publication. Abhinav Sudhakar Dubey and Pranav Singh Mahara are the principal authors and have contributed equally towards the study.

REFERENCES

[1]. Henrique, Bruno Miranda, Vinicius Amorim Sobreiro, and Herbert Kimura. "Literature review: Machine learning techniques applied to financial market prediction." Expert Systems with Applications 124 (2019): 226-251.

[2]. Yu, Yong, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. "A review of recurrent neural networks: LSTM cells and network architectures." Neural computation 31, no. 7 (2019): 1235-1270.

[3]. Wu, Di, Ziyu Wang, and Weichao Zhao. "XLNet-CNN-GRU dual-channel aspect-level review text sentiment classification method." Multimedia Tools and Applications 83, no. 2 (2024): 5871-5892.

[4]. Zhang, Xulang. "Disentangling syntactics, semantics, and pragmatics in natural language processing." (2024).

[5]. Dahal, Keshab Raj, Nawa Raj Pokhrel, Santosh Gaire, Sharad Mahatara, Rajendra P. Joshi, Ankrit Gupta, Huta R. Banjade, and Jeorge Joshi. "A comparative study on effect of news sentiment on stock price prediction with deep learning architecture." Plos one 18, no. 4 (2023): e0284695.

[6]. Mehta, Pooja, Sharnil Pandya, and Ketan Kotecha. "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning." PeerJ Computer Science 7 (2021): e476.

[7]. Nti, Isaac Kofi, Adebayo Felix Adekoya, and Benjamin Asubam Weyori. "Predicting stock market price movement using sentiment analysis: Evidence from Ghana." Applied Computer Systems 25, no. 1 (2020): 33-42.

[8]. Shah, Dev, Haruna Isah, and Farhana Zulkernine. "Predicting the effects of news sentiments on the stock market." In 2018 IEEE International Conference on Big Data (Big Data), pp. 4705-4708. IEEE, 2018.

[9]. Mishev, Kostadin, Ana Gjorgjevikj, Irena Vodenska, Lubomir T. Chitkushev, and Dimitar Trajanov. "Evaluation of sentiment analysis in finance: from lexicons to transformers." IEEE access 8 (2020): 131662-131682.

[10]. Shahapur, S. S., A. Koralli, G. Chippalakatti, M. M. Balikai, D. Mudalagi, R. Dias, S. Devali, and K. Wajantari. "Discovering Untapped Potential in Finance Markets Using NLP-Driven Sentiment Analysis." Indian Journal of Science and Technology 17, no. 21 (2024): 2240-2249.

[11]. Nemes, László, and Attila Kiss. "Prediction of stock values changes using sentiment analysis of stock news headlines." Journal of Information and Telecommunication 5, no. 3 (2021): 375-394.

[12]. Vicari, Mattia, and Mauro Gaspari. "Analysis of news sentiments using natural language processing and deep learning." AI & society 36, no. 3 (2021): 931-937.

[13]. Dhingra, Barkha, Shallu Batra, Vaibhav Aggarwal, Mahender Yadav, and Pankaj Kumar. "Stock market volatility: a systematic review." Journal of Modelling in Management 19, no. 3 (2024): 925-952.

[14]. Rath, Swarnalata, Nilima R. Das, and Binod K. Pattanayak. "An Analytic Review on Stock Market Price Prediction using Machine Learning and Deep Learning Techniques." Recent Patents on Engineering 18, no. 2 (2024): 88-104.