

An Approach Towards Developing Relationally Intelligent Multimodal Framework for Stock Movement Prediction (Student Abstract)

Manali Patel¹, Krupa Jariwala¹, Chiranjoy Chattopadhyay²

¹ Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India

² FLAME School of Computing and Data Sciences, Pune, Maharashtra, India

d21co004@coed.svnit.ac.in, knj@coed.svnit.ac.in, chiranjoy.chattopadhyay@flame.edu.in

Abstract

The dependency of stock prices on a multitude of factors makes the task of prediction exceedingly challenging. Given the volatile nature of stock data, it is imperative to integrate multiple sources of information to accurately encompass the various factors that influence market trends. To capture these complex dynamics, several multimodal methodologies have been proposed, integrating market data, technical indicators, and textual information. However, it is claimed that these coarse-grained information sources do not offer a holistic view of the market. Furthermore, these sources are stock-specific and do not elucidate the interconnections between various stocks. To address this deficiency, we propose a multimodal approach that incorporates this relational aspect alongside fine-grained information sources. The applicability of our framework is underscored by empirical results, which demonstrate the superiority of our approach.

Introduction

The advancement in AI has transformed the landscape of FinTech (Financial Technology) by enabling automated and data-driven decision-making that enhances the efficiency of stock market prediction. The stock market is influenced by numerous factors (Meher and Mishra 2025), which are reflected in different components: visual candlestick charts reflect price volatility, textual data like news headlines or social media reflect investor sentiment, technical indicators reflect price movements, and fundamental indicators reflect the economic health of a stock. Considering this, it becomes essential to incorporate these diverse sources of information in an efficient manner to build an efficient prediction model. Fig. 1 shows the various types of data sources for the stock market prediction task.

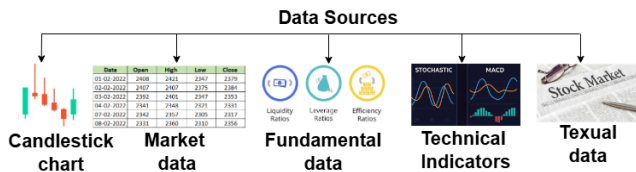


Figure 1: Types of stock market data sources.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

The approaches that consider diverse data sources have an advantage over the single-modality approaches. In the literature, various multimodal approaches have been proposed that aim to fuse heterogeneous modalities (Lee et al. 2025; Zong et al. 2025). However, recent models focus on a limited set of modalities and neglect the relational aspects among stocks that help to understand market dynamics in a comprehensive way. To fill this gap, we propose a relationally intelligent multimodal framework that integrates the interconnections between stocks along with the other data sources defined in Fig. 1 in a unified manner.

Problem Statement

Given historical data of the i^{th} modality at time instance t consisting of l previous data points denoted as $\mathcal{M}_t^i = \{x_{t-l+1}, \dots, x_{t-1}, x_t\}$, our objective is to encode and fuse such k modalities and predict the future price movement at time instance $t + 1$ as $y_{t+1}^{\hat{}}$ = $f(\mathcal{M}_t^{i:k})$. Here, $y_{t+1}^{\hat{}}$ represents the upward and downward movements, respectively. $f()$ is a user-defined function.

Multimodal Framework

From the literature survey, we identified two issues, i.e., 1) Collection of relational dataset and 2) Design of an efficient fusion technique to address noise and data alignment between multiple modalities. The input to the framework is multimodal input, and the output is the price movement (up or down). This framework is tailored to the Indian economy. We have considered three modalities, i.e., candlestick charts, technical indicators, and relational data. The reason for choosing these modalities is that candlestick charts present daily trading activity and account for market volatility. The technical indicators model temporal patterns, and relational data accounts for dependencies among stocks to identify sectoral patterns. Specifically, we consider the market data of the peer stocks belonging to the same sector as the target stock to capture the intra-sectoral trends efficiently. We propose three main modules, i.e., 1) **Data processing**, 2) **Encoder**, and 3) **Fusion module**.

The objective of the data processing module is to perform rolling window segmentation and encode each modality in a temporal sequence that allows the model to learn the effect of past values. The objective of the encoder module is

to extract features from each modality. We propose modular Convolutional Neural Networks (CNNs) to extract geometric features from candlestick charts. However, it further can be enhanced by incorporating market continuation and reversal patterns (Chen, Hu, and Xue 2024). The Temporal Convolutional Network (TCN) is used to encode the technical indicators. Moreover, a novel relational encoder is proposed that focuses on capturing dynamic intra-sector dependencies. The encoded features obtained from each modality are given to the fusion technique.

It is important to design a fusion technique that effectively combines different data sources. We propose a graph-based interactive fusion technique that focuses on the most informative source and models complex interdependencies across modalities. Fig. 2 depicts the general architecture of the discussed multimodal framework.

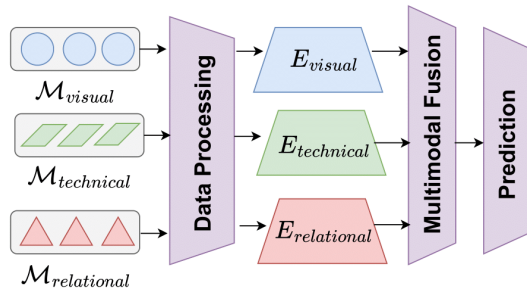


Figure 2: Multimodal Framework.

Results and Discussion

We have designed the multimodal framework for the Indian economy. Since there was no publicly available relational information for India, we curated a dataset consisting of the lists of peer stocks from official websites. Our approach has given superior performance compared to other approaches (Zhang, Ye, and Lai 2023; Shaban, Ashraf, and Slama 2024; Chandar 2022; Chen, Hu, and Xue 2024; Hoseinzade and Haratizadeh 2019). Fig. 3 presents the quantitative analysis on the accuracy parameter.

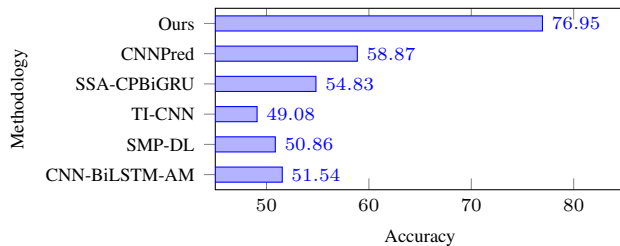


Figure 3: The comprehensive analysis between baseline and proposed model.

This framework is also generalized to other economies. We tested the proposed multimodal approach on US and Chinese stocks. It consistently achieves higher accuracy than the baselines in most stocks (See Fig. 4).

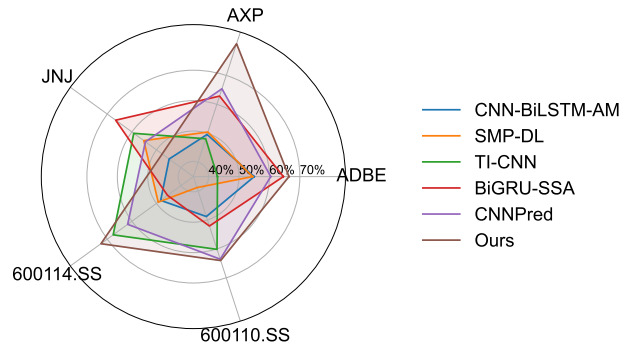


Figure 4: The comparison results for US and Chinese stocks.

Conclusions and Future Work

We proposed a multimodal relational framework that considers multiple sources of information, such as existing interconnections between the same sector sectors along with candlestick charts and technical indicators to predict the future price movement. In the future, we intend to include textual data, such as news headlines, to model the investor’s psychology. We will apply reinforcement learning methods to improve decision making in real-time market environments. We aim to create an interactive dashboard using our prediction model to offer insights like predicted movements, portfolio advice, and trading actions.

References

- Chandar, S. K. 2022. Convolutional neural network for stock trading using technical indicators. *Automated Software Engineering*, 29(1): 16.
- Chen, X.; Hu, W.; and Xue, L. 2024. Stock price prediction using candlestick patterns and sparrow search algorithm. *Electronics*, 13(4): 771.
- Hoseinzade, E.; and Haratizadeh, S. 2019. CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129: 273–285.
- Lee, K.; Kim, Y. J.; Kim, M.; and Kim, H. Y. 2025. Hierarchical multi-modal fusion architecture search for stock market forecasting. *Applied Soft Computing*, 113581.
- Meher, P.; and Mishra, R. K. 2025. Assessing the influence of factors affecting stock market: an ISM approach. *Qualitative Research in Financial Markets*.
- Shaban, W. M.; Ashraf, E.; and Slama, A. E. 2024. SMP-DL: a novel stock market prediction approach based on deep learning for effective trend forecasting. *Neural Computing and Applications*, 36(4): 1849–1873.
- Zhang, J.; Ye, L.; and Lai, Y. 2023. Stock Price Prediction Using CNN-BiLSTM-Attention Model. *Mathematics*, 11(9).
- Zong, C.; Wan, J.; Cascone, L.; and Zhou, H. 2025. Stock movement prediction with multimodal stable fusion via gated cross-attention mechanism. *Complex & Intelligent Systems*, 11(9): 396.