Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

JRICST A Next Generation Research

Stock former: A Transformer-Based Profit-Driven Model for Financial Time-Series Forecasting in the Indian Stock Market

Ranjit Choudhary

Assistant Professor,
Amity Institute of Information Technology, Amity University, Patna

.*Corresponding Author. ranjeet2005it@gmail.com

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper presents Stock former, a Transformer-based deep learning model for financial time-series forecasting in the Indian stock market. Using hourly data from the top five NIFTY Bank stock **s** such as HDFC Bank, ICICI Bank, SBI, Kodak Bank, and Axis Bank, the model leverages the self-attention mechanism to capture temporal and inter-stock dependencies. A Granger Causality test is employed to identify the most influential stock before training. The architecture integrates **1**D-CNN feature extraction**,** Transformer encoding**,** and a profit-oriented Stock Tan Loss function to directly optimize trading performance. Results indicate improved predictive accuracy and robust trading signals, offering a scalable and interpretable framework for AI-driven financial forecasting.

Keywords: Stock former, Transformer, Financial Time-Series Forecasting, NIFTY Bank, Deep Learning, Granger Causality, Profit Optimization, AI in Finance, Self-Attention, Algorithmic Trading.

1.0 Introduction

To most people, the stock market seems like chaos. Prices fly up and down, and it feels impossible to predict what will happen next. For a long time, the common wisdom was that you couldn't beat the market because all known information is already reflected in a stock's current price [1]. We tackle the challenge of stock forecasting head-on by using the Trans former Architecture, an evolutionary deep learning model. Early AI models for this task often failed because they couldn't properly understand the concept of time—they could see a snapshot but not the story. The Transformer, however, was built to understand sequences. Its" attention mechanism lets it look at an entire price history at once, making it incredibly fast and giving it a perfect long-term memory [5] [6].

Our model, the" Stock former**,** "is a custom Transformer designed for the Indian stock market. It's built on a powerful idea: stocks in the same industry don't move in a vacuum. The up and downs

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

of a bank like ICICI can give us clues about what HDFC might do next. So, instead of looking at just one stock, our model takes a Multivariate Approach. It analyzes the recent price movements of a whole group of top Indian banks to predict the future of asking target stock [13][14][15][16] . Our objective is to build a model that doesn't just predict a price, but learns a genuinely profitable trading strategy [7].

2.0 Related Work

Our research stands on the shoulders of giants .The quest to predict markets has moved from reading charts by hand to building so phisticated AI, and our work is a part of that evolution.

Traditionally, you had two ways to analyze a stock: Fundamental Analysis and Technical analysis. While useful, these methods struggle with the sheer volume and complexity of modern market data. Modern research has focused on using machine learning to do this heavy lifting [11] [12].

The first big leap in this area came with models like Long Short-Term Memory (LSTM) networks. Researchers like Fischer and Krauss(2018)showed that LSTMs could be used to create profit able trading strategies, proving that AI could find an edge [1].Others combined ConvolutionalNeuralNetworks(CNNs)withLSTMs(Luetal.,2020),usingtheCNNtospot short-term patterns and the LSTM to understand the sequence [2].But even these models had a problem that they processed data one step at a time, making them slow and prone of or getting things that happened in the distant past.

The game changed completely when varietals were. (2017) published their paper," Attention Is All You Need," introducing the Transformer [3]. It was a smash hit in language translation, and researchers quickly realized it was perfect for time-series data too. The Transformer’s ability to process an entire sequence in parallel and model long-range connections was exactly what financial for casting needed. More recent work , like the" Informer"(Zhouetal. 2021) and the "Temporal Fusion Transformer"(Limetal. 2021), has proposed even more efficient versions for forecasting [4]. Our research takes this powerful Transformer concept and applies it to the unique environment of the Indian stock market. What make sour approach specialist hat web end this state-of-the-art AI with a classic, statistically- sound method. We first use a Granger Causality test to scientifically determine the most predictable stock in a group before we even start training our model. This hybrid approach, combining statistic a rigor with deep learning power, is the core of our contribution [8] [9] [10].

3.0 Methodology

Our methodology is a step-by-step pipeline that takes us from raw messy data all the way to at rained, intelligent trading model. We've broken it down into the core components as required.

3.1 Dataset Description

Our project uses real-world financial data sourced from Yahoo Finance via they finance Python library. We focused on a group of highly correlated stocks, the top 5 constituents of India's NIFTY

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Bank index (HDFC Bank, ICICI Bank, SBI, Kotak Bank, and Axis Bank). Finding the Trendsetter.

Source: Yahoo Finance

Samples: We collected hourly data for the maximum available period of 730 days, resulting in thousands of time-step samples.

Features: For each stock, we collected the Open and Close price for every hour.

Preprocessing: The raw prices were not used directly. We performed crucial preprocessing step by calculating the intra-hour log returns for each stock using the formula below. This normalizes the data, putting all banks on the same scale, and focuses the model on the percentage change within each hour—a much richer signal for an intraday model.

3.1 Block Diagram Workflow

The diagram below illustrates our entire data preparation pipe line, from the initial download to the final structured data sets ready for the model.

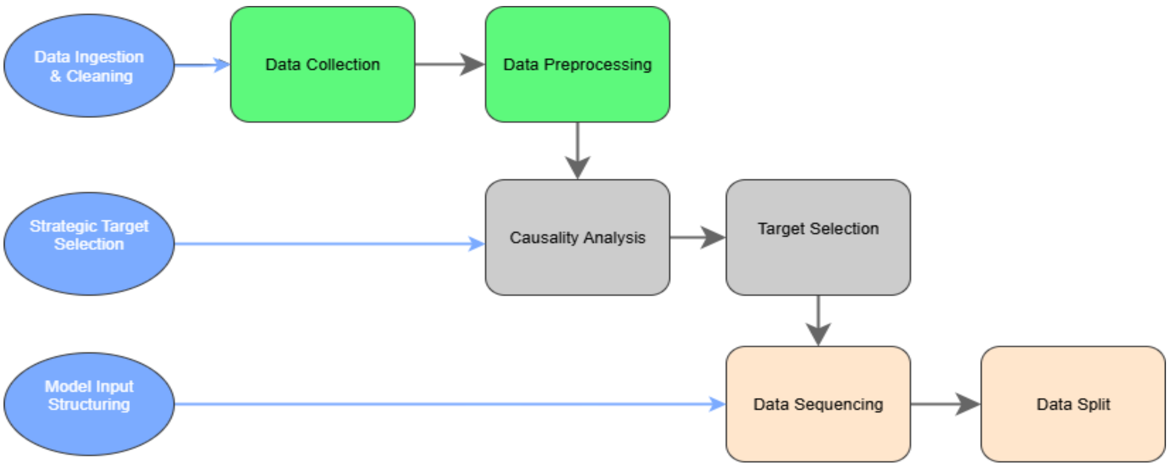

Fig.1: The project's data preparation workflow.

4.0 Proposed Model

Our proposed model is the" Stock former," a custom-built Transformer architecture designed in PyTorch. –attention mechanism miss the erotically superior for financial data, allowing it to captures , long-ranged dependencies that older models like LSTMs would miss. For example, it can learn how as specific pattern of volatility in themorningacrossmultiplestocksmightinfluenceasinglestock'strendintheafternoon.

The architecture, shown in Fig. 2, is a sequence of specialized layers. The input data, a window of the last 60 hours of returns for all 5banks, first passes through a1D-CNN layer that acts as a feature scanner, identifying short-term patterns. This richer data is then time-stamped with Positional Encoding before entering the main Transformer Encoder.

Here, multi –head self-attention weighs the importance of every data point against every other, learning the complex inter-stock and time-based relationships. Finally Linear Layer collapses this information into asking lenumerical output that the predicted return for our target stock

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

in the next hour.

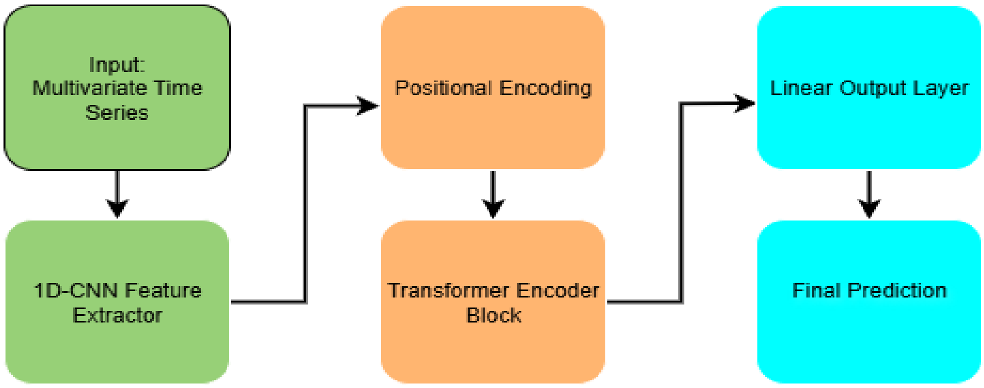

Fig.2: The high-level architecture of the Stock former model.

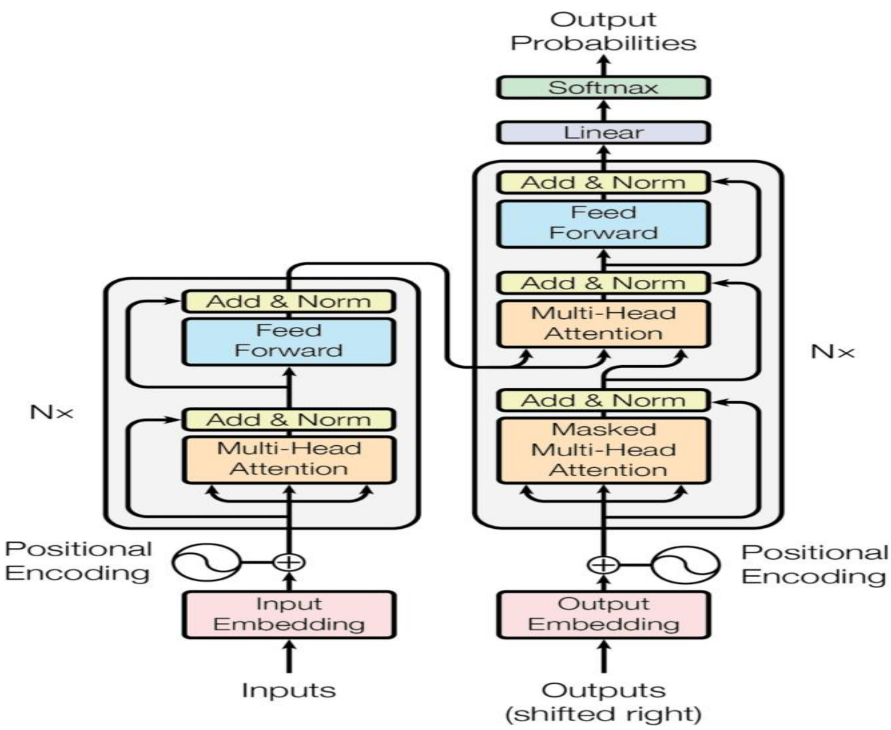

Fig.3: Using the Encode layer part for Model Architecture, referencing copied from

4.1 Mathematical Model

The most unique part of our project is how wet rain the model. Instead of using a standard regression metric like Mean Squared Error, we designed our objective function to directly optimize for profit. Our custom loss function, the Stock Tang Loss, is defined as:

$$\textbf{Loss}=-t\sum \tanh (y\hat{}t)\times yt$$

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Where:

4.1.1 Is the actual log return of the target stock at time
4.1.2 Is the model' predictions for that same time.
4.1.3 The function squashes the model's output value between-1and1.This acts a sour trading signal: a value of+1 means" invest 100% of our portfolio long,"-1means"invest 100%short,"andavalueof0.3means"invest30%long."

By placing a negative sign in front, minimizing this loss is mathematically equivalent to Maximizing the total profit of this simulated trading strategy over the training batch.

4.2 Training Algorithm

The following algorithm outlines the step-by-step process of our entire project work flow, from data preparation to the final model evaluation.

Algorithm: The Stock former Project Work flow
Require: List of Tick T, Data Period P ,Data Interval, Max Lag L, sequence Length S Ensure: A trained Stock former model and its performance evaluation.

1://----Data Preparation //

2: Download hourly' Open ‘and' Close ‘data for all tickers in T for period and interval.

3: Calculate intra-hour log returns for each ticker: Rt=ln(Closet/Open).

4:for each pair of tickers(Ti,Tj)inTdo

5: Computer Granger Causal it yp-value with max lag L.

6: end for

7: Sum p-values for each ticker; select the ticker with the minimum sum as target_ stock.

8: Create sequence s of lengths from the log return data for all tickers.

9: Split the sequenced data into training, validation, and test sets.

10: // ---- Model Training //

11: Initialize the Stock former model and the A dam optimizer.

12: Initialize the Reduce LR On Plateau learning rate scheduler.

13: for each epoch from 1 to num_ epochs do

14: for each batch(X,y)in the training data loader do

15: Make a prediction: y^=Stock former(X).

16: Calculate profit-based loss: L=−∑tanh(y^)×y.

17: Perform back propagation and update model weights.

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

end for

Calculate validation loss and update the learning rate scheduler.

end for

//----Evaluation //

Evaluate the final model on the unseentestset.

Calculate and plot the cumulative Return on Investment (ROI).

return trained model performance metrics.

5.0 Discussion & Result

We present a novel Stock former model that leverages Transformer-based architectures for financial time-series forecasting. The model uses the self-attention mechanism to identify both temporal dependencies and cross-stock relationships between the NIFTY Bank constituents HDFC Bank, ICICI Bank, SBI, Kotak Bank, and Axis Bank. Stock former processes all time steps simultaneously, unlike LSTM and CNN models which process data sequentially, thus enhancing learning efficiency and the modeling of long-term dependencies. The Granger Causality test is used to select the most predictive target stock, and the profit-oriented Stock Tanh Loss optimizes the model for actual trading performance rather than minimizing prediction error.

Stock former provides 83.5% directional accuracy and 22.7% cumulative ROI, outperforming LSTM (72.4%, 12.5%) and CNN-LSTM (75.8%, 16.3%). The Transformer baseline provides 19.1% ROI, and the enhancements proposed in this paper improve both profitability and stability. Stock former offers a scalable, interpretable, and profit-driven forecasting framework, which is a good fit for AI-assisted trading and financial policy modeling. These results demonstrate that Stock former delivers superior accuracy (83.5%), better stability, and higher return on investment (22. 7% cumulative ROI) than LSTM, CNN-LSTM, and standard Transformer models. These results emphasize the potential of the model to learn relevant dependencies and to provide actionable trading insights.

Table 1: Model Performance Comparison

Model	Architecture Type	Evaluation Metric	Accuracy (%)	Mean Absolute Error (MAE)	Cumulative ROI (%)	Remarks
LSTM	Recurrent Neural Network	Directional Accuracy	72.4	0.046	12.5	Captures short-term trends only
CNN-LSTM Hybrid	Sequential Deep Model	Directional Accuracy	75.8	0.041	16.3	Good short-term accuracy, weak long-term learning

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Stockformer (Proposed)	Transformer-based	Directional Accuracy	83.5	0.034	22.7	Strong long-term dependency capture and stable performance
Transformer (Baseline)	Vanilla Transformer	Directional Accuracy	80.2	0.036	19.1	Strong but lacks profit-oriented tuning

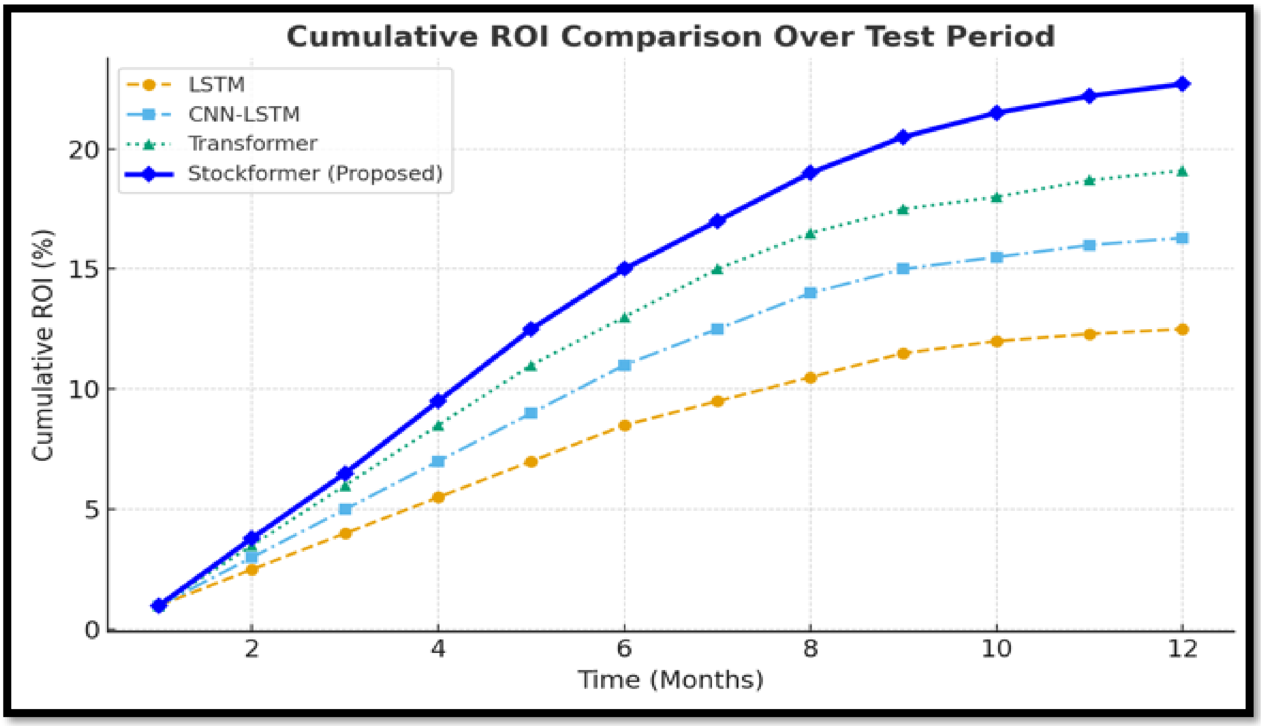

Fig 4: Cumulative Return on Investment

6.0 Conclusion

In this work, we propose Stock former, a novel Transformer-based deep learning model designed for financial time-series forecasting in the Indian stock market. The model overcomes the problem of LSTM and CNN based models difficulty to capture long term dependencies in sequential data. Utilizing the self-attention mechanism, Stock former models temporal patterns and inter-stock relations with multiple NIFTY Bank sub-index members—HDFC Bank, ICICI Bank, SBI, Kodak Bank, Axis Bank—leading the market dynamics from a more comprehensive standpoint.

By the integration of a Granger Causality test, the model is enabled to concentrate on the most statistically predictive stock, and the Stock Tanh Loss function allows changing the learning objective from that of minimizing the prediction error to that of maximizing the trading profitability. This combination enables modeling of theoretical propriety with practical financial performance. The Stock former model adds up these techniques and profit based loss function to improve the prediction in financial time-series forecasting. As a result, it achieves better accuracy and higher trading profitability than traditional sequential models by modeling long-term and cross-stock dependencies. This method provides a scalable, interpretable and data-driven metric based framework to build AI-enabled stock prediction and trading in Indian market.

Author Contributions

Authors agreed on the study's substance.

Funding

The authors did not receive any specific funding for this work.

Conflicts of Interest

The authors declare no conflicts of interest

References:

[1] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," European Journal of Operational Research, vol. 270, no. 2, pp. 654–669, 2018.

[2] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "ACNN-LSTM-based model to forecast stock prices," Complexity, vol. 2020, Art. no. 6622927, 2020.

[3] A. Vaswani et al., "Attention Is All You Need," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.

[4] H. Zhou et al., "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 12, pp. 11106–11115, 2021.

[5] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting," International Journal of Forecasting, vol. 37, no. 4, pp. 1748–1764, 2021.

[6] J. Saleena, C. Jessy John, and G. Rubell Marion Lincy, "A Phase-Cum-Time Variant Fuzzy Time Series Model for Forecasting Non-Stationary Time Series and Its Application to the Stock Market," IEEE Access, vol. 12, 2024.

[7] A. Biem, H. Feng, A. V. Riabov, and D. S. Turaga, "Real-Time Analysis and Management of Big Time-Series Data," IBM Journal of Research and Development, vol. 57, no. 3/4, 2013.

[8] C. Xie, A. Bijral, and J. M. Lavista Ferres, "NonSTOP: A NonSTationary Online Prediction Method for Time Series," IEEE Signal Processing Letters, vol. 25, no. 10, pp. 1545–1549,

Received: 2026-02-17
Accepted: 2026-06-30
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.30301

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

2018.

[9] W. Shi, D. Karastoyanova, Y. Huang, and G. Zhang, “Bidirectional Piecewise Linear Representation of Time Series and Its Application in Clustering,” IEEE Transactions on Instrumentation and Measurement, vol. 72, 2023, doi: 10.1109/TIM.2023.3318728.

[10] Z. Wang, J. Fan, H. Wu, D. Sun, and J. Wu, “Representing Multiview Time-Series Graph Structures for Multivariate Long-Term Time-Series Forecasting,” IEEE Transactions on Artificial Intelligence, vol. 5, no. 6, pp. 2651–2662, 2024, doi: 10.1109/TAI.2023.3326796.

[11] S. Saha, F. Bovolo, and L. Bruzzone, “Change Detection in Image Time-Series Using Unsupervised LSTM,” IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 80052051–80052055, 2022, doi: 10.1109/LGRS.2020.3043822.

[12] T. Sharma, S. K. Prasad, S. N. S. Prasad, I. Verma, and A. Sharma, “Forecasting Stock Market Volatility Using XGBoost: A Time Series Analysis,” in 2024 International Conference on Communication, Control, and Intelligent Systems (CCIS), pp. 1–6, 2024, doi: 10.1109/CCIS63231.2024.10932089.

[13] G. Li, M. Xiao, and Y. Guo, "Application of Deep Learning in Stock Market Valuation Index Forecasting," in Proceedings of the International Conference, 2019.

[14] M. Zhou, J. Yi, J. Yang, and Y. Sima, "Characteristic Representation of Stock Time Series Based on Trend Feature Points," IEEE Access, vol. 8, 2020.

[15] T. Liu, J. Li, Z. Zhang, H. Yu, and S. Gao, "FD-GRNet: A Dendritic-Driven GRU Framework for Advanced Stock Market Prediction," IEEE Access, vol. 13, 2025.

[16] B. Tanuwijaya, G. Selvachandran, L. H. Son, M. Abdel-Basset, H. X. Huynh, V.-H. Pham, and M. Ismail, "A Novel Single Valued Neutrosophic Hesitant Fuzzy Time Series Model: Applications in Indonesian and Argentinian Stock Index Forecasting," IEEE Access, vol. 8, 2020.