Journal of Engineering, Project, and Production Management
2026, 16(5), 2025-305

Developing A Stock Prediction Model Using Multi-Source Data and Temporal Encoding

Xu Ren
Associate Professor, School of Economics, Liaoning University of International Business and Economics, Dalian, 116000, China, E-mail: renxuxu@126.com

Project Management
Received December 5, 2025; revised January 21, 2026; accepted March 30, 2026
Available online June 17, 2026

Abstract: As a core component of the economic system, the stock market plays a crucial role in asset growth for investors, corporate financing decisions, and the stability of financial markets. Traditional prediction models face issues such as reliance on static weights for multi-source heterogeneous data integration and incomplete capture of temporal features, which limit their practical applicability. Therefore, this study proposes a stock prediction model based on multi-source data and temporal encoding. The model first constructs a multi-source data extraction algorithm using Empirical Mode Decomposition (EMD) and a Text Convolutional Neural Network (TextCNN). On this basis, the model combines Bidirectional Gated Recurrent Units and temporal encoding to predict stock prices. Experimental results show that the proposed model outperforms the comparison model in predicting A-share and H-share stocks. The coefficients of determination for Geely and Dongfeng Motor stocks are 0.925 and 0.918, respectively. All indicators outperform comparison models. The proposed model balances prediction accuracy and stability and provides a quantitative basis for investor strategy formulation, financial institutions risk management, and regulatory decision-making.

Keywords: Multi-source data, temporal encoding, stock prediction, bigru, time2vec.

DOI:10.32738/JEPPM-2025-305

1. Introduction

1.1. Research Background

The stock market, as a key component of the economic system, serves not only as an important platform for corporate financing but also as a core channel for investors asset allocation and wealth growth. Accurate prediction of stock prices is essential for investors to develop rational investment strategies, effectively avoid risks, and achieve stable asset growth. It also provides valuable reference for corporate financing decisions and strategic planning and plays an indispensable role in the stable operation of financial markets and the optimal allocation of resources (Patel et al., 2024; Htun et al., 2023). However, the mechanism underlying stock price formation is dynamic and highly complex. Its fluctuations are influenced by multiple factors, including macroeconomic factors, industry cycles, corporate fundamentals, market sentiment, regulatory policies, and the international environment. Traditional stock prediction models mainly rely on historical stock price data, capturing only linear features and simple temporal patterns, which limit their ability to fully reflect the complex economic relationships and market dynamics behind stock price fluctuations. With increasingly complex market environments and explosive growth in data, these models exhibit clear limitations, and their predictive accuracy often falls short of investors practical needs (Mukherjee et al., 2023). In summary, amid increasingly complex market conditions, expanding data dimensions, and rising volatility risks, traditional stock prediction models struggle to meet the actual demands of investors, enterprises, and regulatory authorities, underscoring the need for new prediction frameworks that integrate heterogeneous multi-source data and capture nonlinear temporal features.

1.2. Research Status

To address the dual challenges of multi-source data integration and nonlinear temporal modeling in traditional models, researchers have gradually explored new technical approaches. For example, the team led by Luo et al. (2024) proposed a hybrid model combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) with an attention mechanism to improve short-term stock correlation prediction accuracy. Their model reduces the Mean Absolute Error (MAE) by 8.3% and increases the Coefficient of Determination (R^2) by 0.09 compared with the optimal baseline, but it shows weaker adaptability in small-sample scenarios. Xing et al. (2024) proposed a stock prediction framework based on fuzzy graph learning to tackle the difficulty of modeling market momentum spillovers in traditional models. Their method reduces the mean absolute error by 7.2% and improves prediction accuracy by 5.8%, but performance decreases for highly volatile stocks. Tian et al. (2023) proposed a stock prediction framework based on graph-evolutionary recurrent units, using recurrent mechanisms to capture temporal features and adjusting stock correlation weights via an embedded graph evolution module. Their model reduces Mean Absolute Error (MAE) by 6.5%, but multi-step prediction performance deteriorates noticeably. Alam et al. (2024) used Long Short-Term Memory to capture temporal dependencies and employed Deep Neural Networks (DNN) to handle high-dimensional nonlinear features, achieving diversified stock prediction. Their model improves next-day limit-up prediction accuracy by 6.7% but shows limited tolerance to abrupt policy or market sentiment shocks. Li et al. (2024) proposed a framework combining hierarchical frequency decomposition and deep learning. In next-day closing price prediction, the mean absolute error decreases by 8.5% compared with the optimal baseline, and the average accuracy of stock rise/fall prediction improves by 5.7%. However, its computational efficiency is affected by the decomposition levels, and real-time prediction performance for large-scale stock portfolios still needs improvement.

1.3. Research Motivation and Method

Although current research methods have made progress in multi-source data fusion and time-series encoding, the integration of heterogeneous multi-source data is limited by a static weight allocation mechanism, which makes it difficult to dynamically adapt to the semantic features and correlation patterns of different data types. Moreover, it is prone to noise interference, leading to reduced feature extraction accuracy. At the same time, time-series encoding technology is insufficient for adapting to long-term dependencies and sudden fluctuations. It lacks an integrated modeling framework that can simultaneously account for long-term trends, short-term fluctuations, and potential reversal characteristics. Therefore, it is necessary to construct more efficient multi-source data extraction algorithms to achieve noise removal and modal de-mixing of structured time-series data, semantic quantification of unstructured text data, and design a time-series encoding module that adapts to the characteristics of the stock market, balancing the requirements for long-term trend modeling and short-term fluctuations, sudden reversal signal capture, and improving the model's prediction accuracy in complex market environments. Empirical Mode Decomposition (EMD) can separate high-frequency noise, trend components, and effective signals in stock trading data, providing clean base signals for feature extraction (Li et al., 2023a). Text Convolutional Neural Network (TextCNN) extracts semantic information and quantifies sentiment from news and comments (Shi et al., 2024). Graph Attention Network (GAT) provides graph-structured support for multi-source feature integration through node attention mechanisms (Li et al., 2023b). Bidirectional Gated Recurrent Unit (BiGRU) models accurately capture short-term trends in stock sequences using bidirectional gating (Imran et al., 2024). Time2Vec captures latent temporal patterns such as cycles and trends, providing representative temporal features for stock prediction (Wang and Shao, 2024). Transformer captures global trends and cross-time dependencies in stock price fluctuations (Selva et al., 2023). Therefore, this study combines EMD, TextCNN, and GAT to construct a multi-source data extraction algorithm and builds a stock prediction model using BiGRU, Time2Vec, and Transformer. This model aims to efficiently extract and integrate multi-source features and provide more accurate stock predictions, offering technical support for risk management and policy-making. The innovations of this study are: first, optimized EMD removes noise and redundant information from multi-source data, addressing interference in heterogeneous data integration. Second, combining TextCNN and GAT, multi-scale convolution captures local semantics of financial text, and graph attention aggregates structural features of stock networks, enabling cross-modal feature extraction of unstructured text and graph data. Finally, a hybrid temporal encoding model using BiGRU, Time2Vec, and Transformer balances local detail and long-term trend modeling, improving robustness and accuracy of stock prediction in complex markets.

1.4. Structure of the Paper

This study is divided into five sections. Section 1 introduces the research background and relevant literature, summarizes the limitations of existing stock prediction methods, and clarifies the research objectives. Section 2 details the design principles and overall process of the multi-source data extraction algorithm and the stock prediction model, fully presenting the technical approach. Section 3 compares the proposed method with baseline methods using MAR, R^2, and rise/fall prediction accuracy to verify model superiority. Section 4 analyzes experimental results in depth, discussing indicators in the context of related studies. Section 5 summarizes the findings and explores future research directions, including the efficiency of real-time prediction for large-scale portfolios and the integration of macroeconomic indicators and other heterogeneous data to further enhance prediction robustness.

2. Materials and Methods

2.1. Construction of Multi-Source Data Extraction Algorithm

The stock market is not purely an economic-technical system, its operation is deeply embedded in social institutions, interpersonal networks, and policy environments. Traditional prediction models, by separating technical analysis from social factors, struggle to explain common phenomena such as irrational fluctuations, policy lags, and herd behavior. Institutional economics suggests that market rules directly determine the transparency of information and the efficiency of resource allocation. This perspective provides institutional-level logic for the design of the multi-source data extraction algorithm. For example, after the full implementation of the registration-based system in China's stock market in 2023, the disclosure requirements of listed companies shift from compliance review under the approval system to substantive disclosure under the registration system. The granularity of financial data and business risk information increases by more than 40%, but this also exacerbates information overload (Srivastava et al., 2023; Wu et al., 2023). EMD processes time series data, such as quarterly revenue and profit, in layers, removing noise from non-recurring gains and losses (Zhang et al., 2024). It decomposes signals into multiple Intrinsic Mode Functions (IMF) to obtain different functions, calculated as shown in Eq. (1).

$$H(t) = X(t) - M(t) \tag{1}$$

In Eq. (1), H is the IMF component obtained from decomposition, X represents the original time series data, M represents the local mean of the original data, and t represents the time variable. The extracted IMF components are then removed from the original data, and the decomposition process is repeated for the remaining components until the residual term becomes a monotonic function or contains only one extreme point, expressed in Eq. (2) (Liu et al., 2024).

$$R(t) = X(t) - \sum IMF \tag{2}$$

In Eq. (2), R represents the residual term, and $\sum IMF$ represents the sum of the extracted IMF functions. Although EMD can adaptively decompose non-stationary signals into intrinsic mode functions, it still suffers from mode mixing. Fast Independent Component Analysis (FastICA) is a blind source separation method based on the statistical independence of signals. It separates mixed IMF components in EMD decomposition by exploiting component independence (Zhao et al., 2024). Therefore, this study combines EMD and FastICA to construct a time series processing algorithm, EMD-FastICA, whose process is shown in Fig. 1.

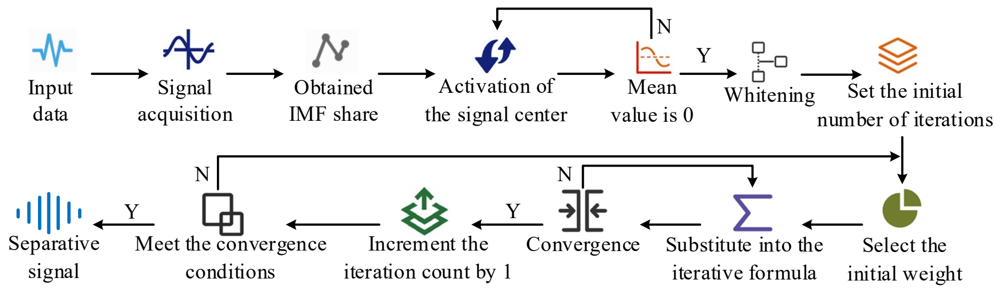


Fig. 1. EMD-FastICA algorithm operation process (Icon source from: iconpark.oceanengine.com)

As shown in Fig. 1, in the EMD-FastICA algorithm, the original data is first adaptively decomposed by EMD into multiple IMF components and residual terms, thereby preliminarily removing noise from non-recurring gains and losses. FastICA then separates the mixed IMF components by inverting a linear random-mixing matrix to obtain the observed signals, ultimately isolating the independent source signals. FastICA separates several independent source signals from multiple input signals by calculating them as shown in Eq. (3).

$$X'(t) = A(t) * S(t) + V(t) \tag{3}$$

In Eq. (3), X' represents the sampled signal, A represents the linear random mixing matrix, S represents the unknown source signal, and V represents noise. The goal is to iteratively optimize the process for recovering source signals from observed signals, as calculated in Eq. (4) (He et al., 2023).

$$Y(t) = W(t) * X(t) + V(t) \tag{4}$$

In Eq. (4), Y is the observed signal, and W represents the inverse of the linear random mixing matrix. EMD-FastICA can adaptively process structured time-series data related to stocks, producing pure signals free of noise or mode mixing. However, it is effective only for structured data such as financial or trading data and cannot process important unstructured textual data in the stock market, such as news reports, policy interpretations, and investor comments, which contain key information about policy trends, market sentiment, and public opinion that directly affect stock price fluctuations. TextCNN captures local semantic features of textual data through multi-scale convolutional kernels, converting unstructured text into quantifiable indicators. The information transformation process of TextCNN is shown in Fig. 2.

As shown in Fig. 2, after inputting stock-related unstructured text, TextCNN maps it to word vectors and constructs a word vector matrix that preserves semantic relationships. Multi-scale convolution captures local semantics of short phrases and long textual logic, generating multiple feature maps containing stock-related semantic information. Each feature map undergoes pooling to select key semantic features and reduce data redundancy. Finally, the pooled features are concatenated into a one-dimensional vector, outputting quantifiable textual features. The convolution operation is expressed in Eq. (5).

$$c_i = f(C \cdot W'_{i+h-1} + b) \tag{5}$$

In Eq. (5), c represents the feature, f represents the nonlinear activation function, C represents the convolution kernel, W' represents the word vector matrix, i indexes the convolution operation, h represents the sliding window size, and b represents the bias. The convolution kernel slides to extract features and generate feature maps, expressed in Eq. (6).

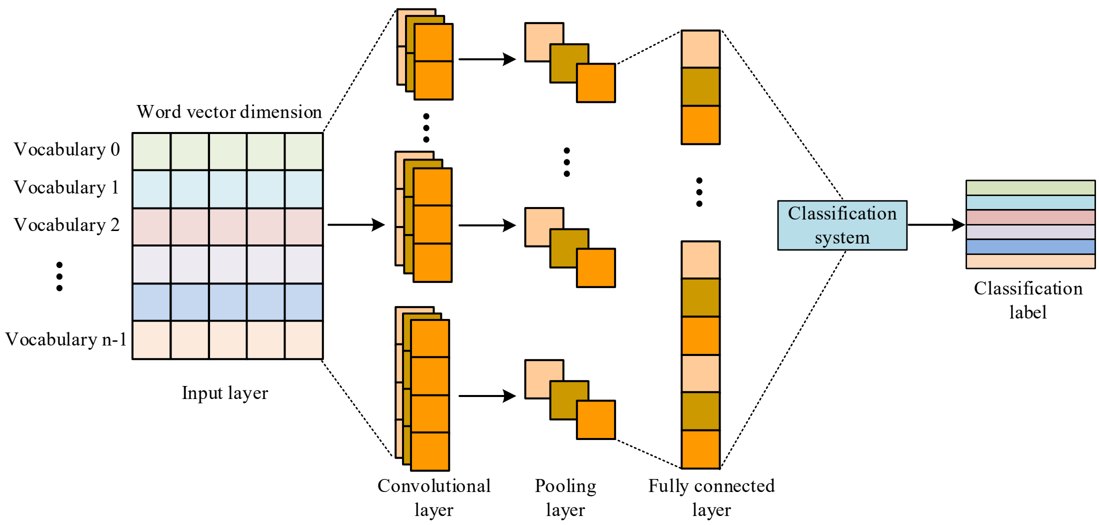


Fig. 2. Schematic diagram of the information conversion process of TextCNN

$$C' = [c_1, c_2, c_3, \ldots, c_{n-h+1}] \tag{6}$$

In Eq. (6), C' represents the final feature map, and n represents the total length of the input word vector sequence. Max pooling preserves key features and reduces dimensions by concatenating the maximum values, as shown in Eq. (7).

$$Z = [\tilde{C}_1, C_2, \tilde{C}_3, \ldots, \tilde{C}_n] \tag{7}$$

In Eq. (7), Z represents the new feature vector obtained by concatenating features extracted by multiple convolution kernels. The resulting feature vector is finally classified, calculated as shown in Eq. (8).

$$y = soft\max(W^x \cdot Z + b') \tag{8}$$

In Eq. (8), y is the output category probability, and W^x and b' are the weights and bias of the fully connected layer, respectively. EMD-FastICA and TextCNN extract features from some multi-source data in the stock market, but structured features of stock industry relationships, equity relations, and other graph-structured data remain difficult to cover. GAT aggregates structural features of the stock association network through node attention mechanisms, achieving cross-modal collaborative extraction of unstructured text, structured time series, and graph-structured features (Gheisari et al., 2023). Therefore, this study combines EMD-FastICA, TextCNN, and GAT to construct a multi-source data extraction algorithm, named EFTG, as shown in Fig. 3.

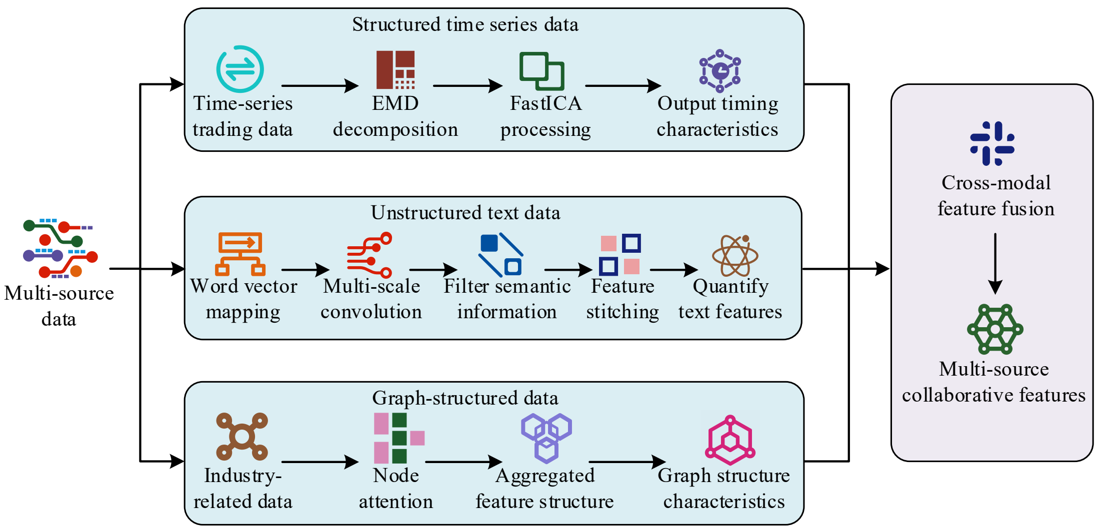


Fig. 3. EFTG algorithm operation flow chart (Icon source from: iconpark.oceanengine.com)

In Fig. 3, in the EFTG algorithm, after multiple data sources are input, the time series data is first decomposed using EMD to remove high-frequency noise. The decomposition of non-stationary time series data by EMD satisfies adaptive convergence, and the energy conservation error of the IMF components after decomposition is $\leq 10^{-4}$. Then, FastICA effectively compensates for the mode aliasing defect of EMD by iteratively optimizing negative entropy, while ensuring the stability of time-series feature extraction. For text data processing, the multi-scale convolution kernels of TextCNN reduce semantic noise interference through local feature aggregation and combine the node attention mechanism of GAT to suppress the weights of abnormal associated edges, thereby reducing the tolerance threshold of the feature vector after cross-modal fusion to data noise. These three types of features are then fused in a cross-modal layer, integrating temporal, textual, and graph features into a unified, semantically complete multi-source feature vector, providing high-quality input for the stock prediction model.

2.2. Stock Prediction Model based on Multi-Source Data and Temporal Encoding

The constructed EFTG algorithm performs cross-modal feature fusion across structured time-series, unstructured textual, and graph-structured data. It outputs unified and semantically complete multi-source feature vectors, providing comprehensive feature support for stock prediction. However, stock prices exhibit strong short-term fluctuations and complex long-term trends. Relying solely on multi-source feature extraction cannot achieve precise mapping from features to stock prices (Singh et al., 2023). In the stock market, many cyclical phenomena occur over time. For example, the opening effect refers to the phenomenon in which stock prices fluctuate more dramatically and with greater trading volume at the opening. Shortly after the market opens, participants digest and react to information from the previous trading day. The weekend effect occurs when stock performance on Fridays and Mondays differs from that on other weekdays. Investors may adjust positions on Friday to account for weekend uncertainty, and Monday prices may fluctuate in response to weekend news (Bathla et al., 2023). Time2Vec effectively captures these cyclical fluctuation patterns, integrating temporal information into the model to improve the understanding and prediction of stock price movements. Therefore, this study introduces Time2Vec temporal encoding, as shown in Fig. 4.

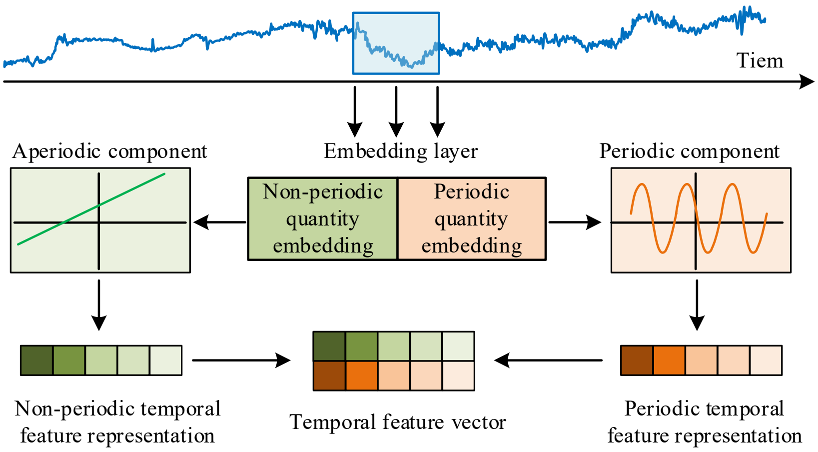


Fig. 4. Time2Vec timing encoding flow chart

As shown in Fig. 4, Time2Vec performs three-level processing: time decomposition, dual-component encoding, and feature fusion. It converts abstract temporal information into structured features suitable for stock prediction, accurately capturing market time-related patterns. After inputting stock trading timestamps, the algorithm first performs multi-granularity decomposition and then generates temporal features via dual-component encoding. Periodic components use trigonometric mapping to capture fixed patterns, while non-periodic components undergo linear transformation to extract dynamic trends, such as policy persistence or capital inflow decay. The encoded temporal feature vector can then be fused with the multi-source features output by EFTG, yielding enhanced temporal feature vectors that incorporate time semantics. The embedding equation of Time2Vec is shown in Eq. (9).

$$t2v(x)[i] = \begin{cases} \omega_{0h} x + \varphi_{0h}, i = 0 \\ \sin(\omega_{ih} x + \varphi_{ih}), 1 \le i \le k \end{cases} \tag{9}$$

In Eq. (9), $t2v$ represents the time vector embedding, x represents the input time scalar, i represents the i-th element of the embedding vector, ω_{0h} and φ_{0h} represent parameters for the linear component rate and phase shift, ω_{ih} and φ_{ih} represent angular frequency and phase shift parameters, h is a scaling parameter, and k is the number of periodic components. Time2Vec injects temporal semantics into multi-source features, but enhances features only for single time points and cannot model long-term temporal dependencies. Transformer dynamically calculates correlation weights between time points via self-attention, addressing long-range temporal dependencies, and enabling feature interactions among key temporal nodes (Xu et al., 2023). Therefore, this study combines Time2Vec and a Transformer to construct an improved network, Time2Vec-Trans, whose structure is shown in Fig. 5.

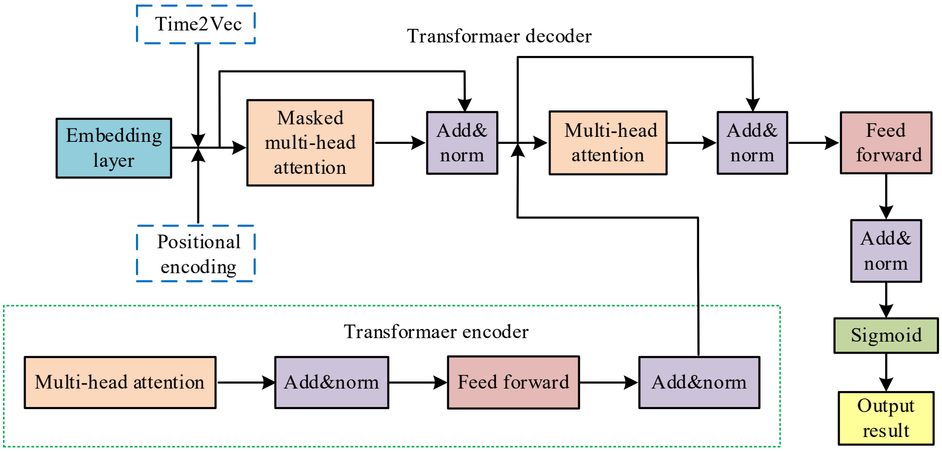


Fig. 5. Schematic diagram of the network structure of Time2Vec-Trans

As shown in Fig. 5, Time2Vec-Trans embeds positional encoding and Time2Vec temporal features at the input layer. In the encoder, multi-head attention computes correlation weights across different time nodes in parallel. Causal masking prevents future information leakage, and layer normalization standardizes each layer's output. Residual connections propagate features from the previous layer, jointly mitigating gradient vanishing during deep training and adapting to high-volatility market distributions. The output layer generates long temporal feature vectors with time semantics, providing global temporal support for downstream processing. Attention weight calculation is shown in Eq. (10).

$$\begin{cases} V = X'\tilde{W}^V \\ Q = X'\tilde{W}^Q \\ K = X'\tilde{W}^K \end{cases} \tag{10}$$

In Eq. (10), X' represents the input matrix, V, Q, and K represent value, query, and key vectors, and $\tilde{W}$ represents the weight matrix. Each head computes attention weights as shown in Eq. (11) (Xu et al., 2024).

$$A(Q,K,V) = Soft\max(\frac{QK^T}{\sqrt{d_K}})V \tag{11}$$

In Eq. (11), d_K represents the key vector dimension, and K^T represents the transposed key matrix. Similarity is scaled and normalized to obtain the weight function. Each head output is calculated and concatenated, then transformed to produce the final output, as shown in Eq. (12).

$$b_i = W^O[b_1^i;b_2^i] \tag{12}$$

In Eq. (12), b represents the attention head output, W^O represents the output transformation weights, $[b_1^i;b_2^i]$ represents concatenation, and i represents positions in the input sequence. Although Time2Vec-Trans captures long-term trends and temporal patterns, key signals may be diluted across long sequences, limiting its ability to adapt to intraday nonlinear reversals. BiGRU tracks short-term price evolution through its forward network and predicts potential reversals through its backward network, complementing Time2Vec-Trans. Therefore, this study combines Time2Vec-Trans and BiGRU to construct the stock price inference algorithm TTB, as shown in Fig. 6.

As shown in Fig. 6, the TTB algorithm uses layer normalization in the Transformer to keep the output of each layer at a constant distribution. The residual connection effectively prevents gradient disappearance, ensuring the convergence of deep network training. The update and reset gates of BiGRU constrain the weights to [0, 1], effectively suppressing abnormal fluctuations in stock prices, preventing drastic oscillations in the output, and enhancing overall model stability. TTB captures long-term trends through temporal encoding and multi-head self-attention. The forward BiGRU unit traces historical price evolution, capturing high volatility and short-term trends caused by capital inflows, while the backward unit predicts short-term reversals, identifying policy corrections and pre-close price reversals. Forward and backward outputs are concatenated and fused with long-term features in the fusion layer, which outputs the prediction results. BiGRU consists of update, reset, and output gates. The update gate is calculated as shown in Eq. (13).

$$z(\hat{t}) = \hat{\sigma}(W_z[\hat{X}(\hat{t}), h(\hat{t}-1)] + b_z) \tag{13}$$

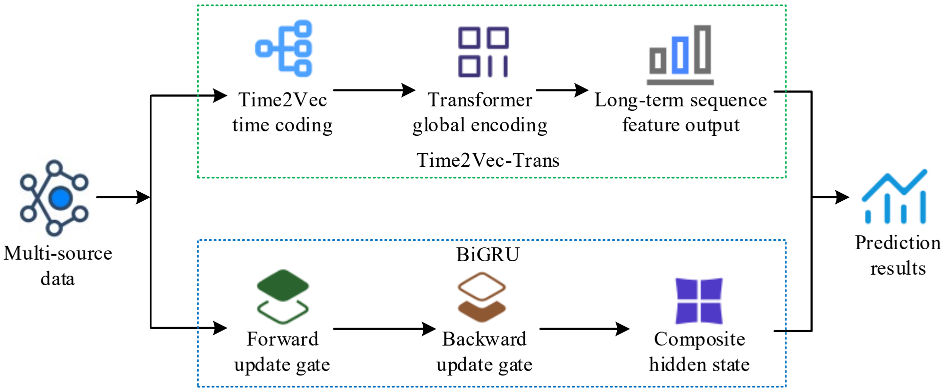


Fig. 6. Schematic diagram of the structure of the TTB algorithm

In Eq. (13), z represents the update gate output, $\hat{t}$ represents the time step, $\hat{\sigma}$ represents the activation function, W_z represents update gate weights, $\hat{X}$ represents input, h represents the hidden state, and b_z represents update gate bias. The reset gate is calculated as shown in Eq. (14).

$$r(\hat{t}) = \hat{\sigma}(W_r[\hat{X}(\hat{t}), h(\hat{t}-1) + b_r]) \tag{14}$$

In Eq. (14), r represents the reset gate output. The candidate hidden state is computed through selective memory storage, as shown in Eq. (15).

$$c'(\hat{t}) = \tanh(W_{c'}[\hat{X}(\hat{t}), r(t) \odot h(\hat{t}-1)] + b_{c'}) \tag{15}$$

In Eq. (15), c' represents the candidate's hidden state, and $\odot$ represents element-wise multiplication. The output gate decides output based on the memory cell, calculated as shown in Eq. (16).

$$h(\hat{t}) = z(t) \odot h(\hat{t}-1) + (1 - z(t)) \odot c'(t) \tag{16}$$

In Eq. (16), the update gate weights historical and candidate states to generate the new hidden state, retaining and updating information. In summary, this study combines EFTG and TTB to construct the final multi-source data temporal encoding stock prediction model, EFTG-TTB, whose structure is shown in Fig. 7.

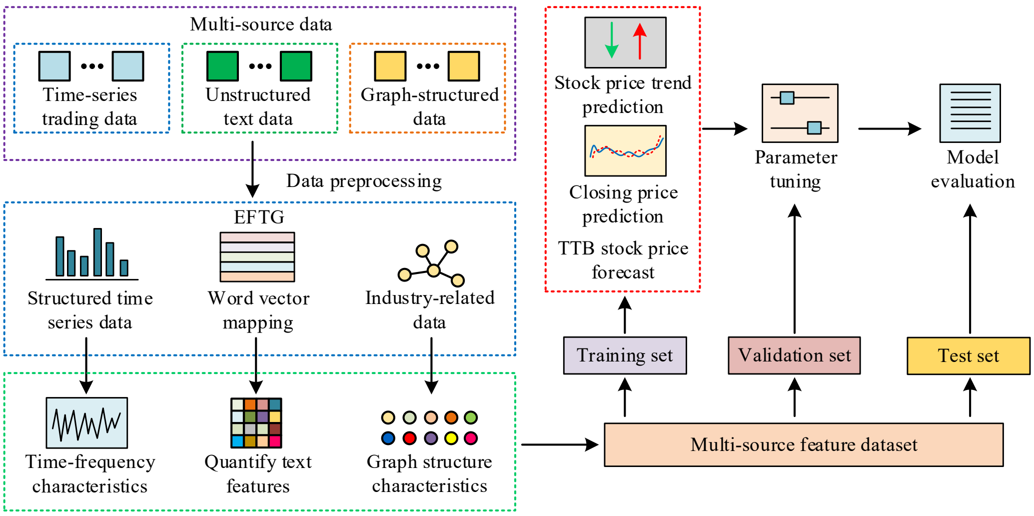


Fig. 7. EFTG-TTB model stock prediction model operation process

As shown in Fig. 7, EFTG-TTB processes multi-source heterogeneous data through EFTG, denoising time series, converting text into quantifiable semantic features, and dynamically calculating stock correlation weights to capture market structural dependencies. The unified, semantically complete multi-source feature vectors are input to the TTB prediction module. Time encoding captures periodic patterns and policy-driven trends, while multi-head attention captures long-term trends. BiGRU tracks price evolution and predicts short-term reversals, outputting short-term features. Long-term and short-term features are fused, anchoring macro trends and calibrating short-term fluctuations, ultimately predicting stock closing prices and rise/fall directions, providing technical support for investment strategy and risk management.

3. Results and Analysis

3.1. Performance Analysis of EFTG Multi-Source Data Extraction Algorithm

To verify the performance of the proposed EFTG multi-source data extraction algorithm, it was compared with the Variational Mode Decomposition (VMD) combined with FastText (FT) and Graph Convolutional Network (GCN) algorithm VMD-FT-GCN, the EMD combined with Bidirectional Encoder Representations from Transformers (BERT) and GAT algorithm EMD-BERT-GAT, and the Wavelet Transform (WT) combined with Deep Convolutional Neural Network (DCNN) and Graph Isomorphism Network (GIN) algorithm WT-DCNN-GIN. In the EFTG algorithm, the stopping criterion for EMD decomposition is set such that the residual term is a monotonic function, the maximum decomposition layer limit is 10 layers, the FastICA iteration times are set to 300 times, the convergence threshold is 1e-6, negative entropy is used as the independence criterion, and the learning rate is 0.001. The convolution kernel size of the TextCNN is set to three multi-scale combinations (2×128, 3×128, 4×128), and the number of convolution kernels for each size is 64. The pooling layer uses maximum pooling, and the pooling window size matches the convolution kernel size. The activation function is ReLU, the fully connected layer has 256 output units, and dropout is 0.3. The GAT network has 2 layers. The first layer has an output dimension of 128, and the second layer has an output dimension of 64. The number of attention heads is 8, and the multi-head attention output is fused using the concatenation mode. The activation function is LeakyReLU, and dropout is 0.2. When constructing the stock correlation network, the edge definition rules use a dual-determination standard. At the industry level, if two stocks belong to the same Shenwan first-level industry, or the upstream-downstream correlation of the industry chain is ≥0.7, then an industry correlation edge is established. At the equity level, if the direct shareholding ratio between listed companies is ≥5% or there is a consistent shareholder relationship, then an equity correlation edge is established. When constructing the adjacency matrix, edge weights are normalized by the correlation strength, and uncorrelated edges are assigned a weight of 0. The experiments ran on Windows 10 with Pytorch as the deep learning framework, Adam as the optimizer, an Intel Core i9-13900K@5.8GHz CPU, an NVIDIA RTX 3090 GPU (24GB), and 32GB memory. The data sources included financial reports of CSI 300 constituent stocks, financial news from Eastmoney, investor comments from Tonghuashun Stock Forum, and stock data from Tonghuashun iFinD. There are 825 time series records for a single stock, and a total of 247,500 records for 300 stocks. The text data includes 300,000 financial news articles and 200,000 investor comments, which are uniformly truncated to 512 characters per record for feature extraction. The graph structure data consists of the industry and equity association network for the 300 components of the Shanghai-Shenzhen 300 Index, with 300 nodes and 1,286 edges, and the adjacency matrix has dimension 300×300. The batch size for experiments is uniformly set to 32. The stock association network graph-structured data was first converted into adjacent matrices and visualized as graph images. The Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) of the four algorithms in processing graph-structured data were then tested. The results are shown in Fig. 8.

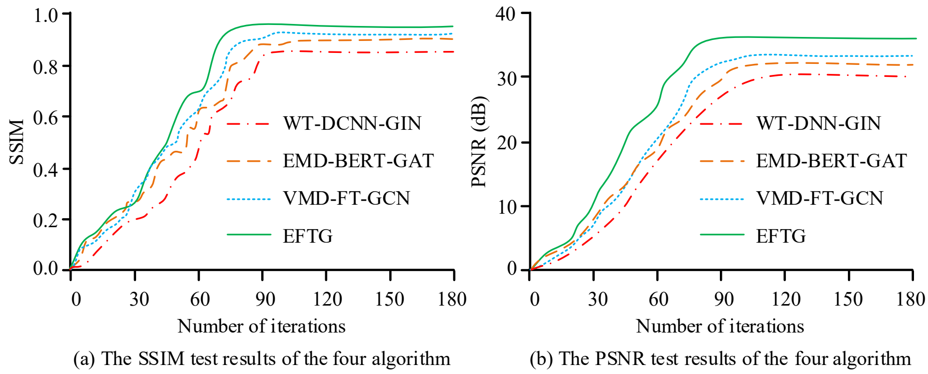


Fig. 8. Comparison of SSIM and PSNR test results

As shown in Fig. 8(a), the SSIM value of EFTG reached 0.955, which increased by 0.032 and 0.027 compared with VMD-FT-GCN and EMD-BERT-GAT, respectively. Fig. 8(b) shows that the PSNR of EFTG reached 36.79 dB, an improvement of 9.78% over EMD-BERT-GAT. In summary, EFTG caused less distortion to the original stock price trend when removing non-periodic and high-frequency noise, better preserving long-term price fluctuation patterns. Next, the text semantic classification accuracy of the four algorithms was compared with manual classification, and the Pearson correlation coefficient results are shown in Fig. 9.

Fig. 9(a) shows that EFTG achieved a high degree of fitting, with data points closely aligned with the fitted curve, and the correlation coefficient reached 0.9571. Figs. 9(b) and 9(c) show that the correlation coefficients of VMD-FT-GCN and EMD-BERT-GAT decreased by 0.0205 and 0.0144, respectively, compared with EFTG. Fig. 9(d) shows that WT-DCNN-GIN had a correlation coefficient of 0.9068, which was lower than that of the proposed model. These results indicated that EFTG extracted textual features that more accurately reflected policy trends and real-time sentiment, aligning better with actual market responses. The Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves for the four algorithms for industry node association extraction were then compared, as shown in Fig. 10.

Fig. 10(a) shows that the EFTG curve was closest to the top-left corner of the ROC curve, with an Area Under the Curve (AUC) of 0.914, higher than the comparison algorithms (0.902, 0.886, and 0.864), indicating better extraction performance. Fig. 10(b) shows that the PR curve of EFTG was closest to the top-right corner, significantly outperforming the other algorithms. This indicated that EFTG maintained high precision while achieving high recall, further validating its superior extraction performance. To further analyze EFTG performance, the runtime and memory usage of the four algorithms for multi-source data processing were tested. The running time and memory usage are the average values of 5 independent experiments. Each experiment was conducted after restarting the system in the same hardware environment to eliminate interference from background processes and ensure the reliability of the results. The results are shown in Table 1.

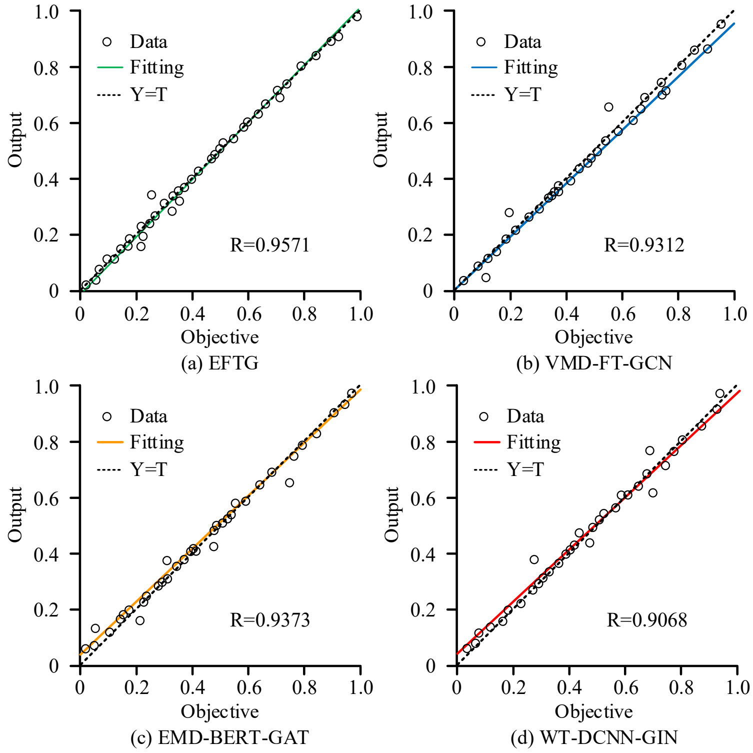


Fig. 9. Comparison of text semantic classification accuracy and manual classification

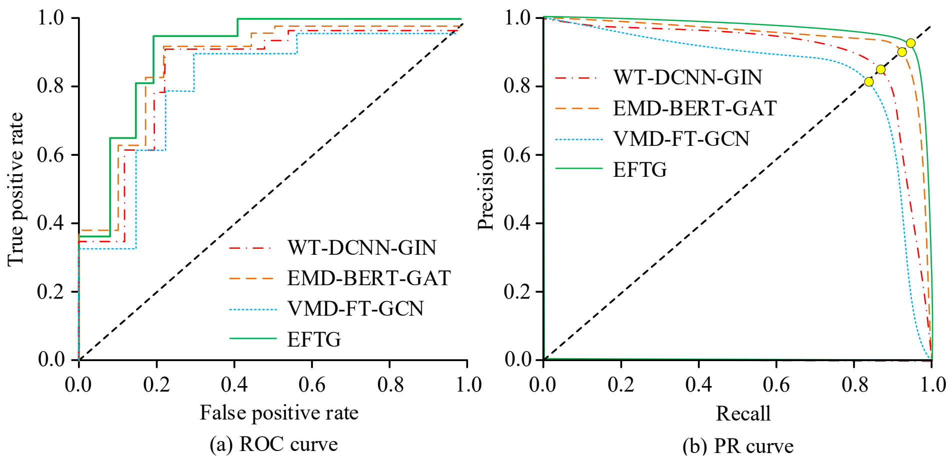


Fig. 10. Comparison of ROC curve and PR curve results

Table 1 shows that EFTG spent 13.2s to extract time-series data with 192.4 MB of memory usage, lower than VMD-FT-GCN and EMD-BERT-GAT. For text data processing, EFTG ran in 9.1s, slightly slower than VMD-FT-GCN at 7.5s, indicating a balance between processing speed and extraction accuracy. For graph-structured data, EFTG used 231.3 MB of memory, which did not significantly exceed that of other algorithms. These experiments demonstrated that EFTG achieved a good balance between algorithmic complexity and feature-extraction accuracy, meeting the real-time requirements of stock prediction.

3.2. Application Performance of EFTG-TTB Stock Prediction Model

After verifying the performance of the EFTG multi-source data extraction algorithm, the EFTG-TTB multi-source stock prediction model was evaluated and compared with the VMD-FT-GCN combined with Bidirectional Long Short-Term Memory Network (BiLSTM) model VFG-BiLSTM, the EMD-BERT-GAT combined with BiGRU model EBG-BiGRU, and the WT-DCNN-GIN combined with Temporal Convolutional Network (TCN) model WDG-TCN. In the EFTG-TTB model, the number of periodic components of Time2Vec is 16, and the parameter for the linear component change rate is

0.1. The number of layers of the Transformer encoder is 6, each layer has 8 attention heads, and the output dimension of multi-head attention is 512. The hidden layer dimension of the feedforward neural network is 2048, the activation function is GELU, and the dropout rate is 0.1. The hidden layer size of BiGRU is 128, and the number of layers is 2. The activation functions for the update and reset gates are Sigmoid, the activation function of the candidate hidden state is tanh, the dropout between layers is 0.2, and the recursive dropout is 0.1. The optimizer is AdamW, the initial learning rate is 0.001, the batch size is 32, the maximum number of training rounds is 300, and the early stopping strategy is based on the validation set loss, with patience set to 10. The experiments selected 200 stocks from both the A-share and H-share markets and collected multi-source data from January 1, 2020, to December 31, 2024, to construct the test dataset. The data were split into training and test sets in a 7:3 ratio. The four models were first tested on the prediction loss curves, and the results are shown in Fig. 11.

Table 1. Comparison of running time and memory usage results

Test content	Algorithm	Running time (s)	Memory usage (MB)
Time series data	EFTG	13.2	192.4
	VMD-FT-GCN	15.8	210.6
	EMD-BERT-GAT	14.5	203.2
	WT-DCNN-GIN	11.6	181.9
Text data	EFTG	9.1	163.7
	VMD-FT-GCN	7.5	132.5
	EMD-BERT-GAT	37.8	565.8
	WT-DCNN-GIN	12.3	192.1
Graph-structured data	EFTG	11.4	231.3
	VMD-FT-GCN	9.7	215.7
	EMD-BERT-GAT	11.6	235.9
	WT-DCNN-GIN	10.5	222.5

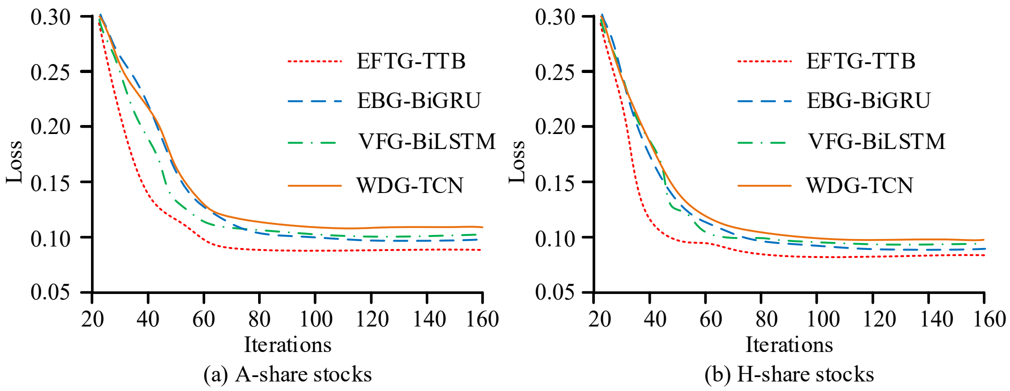


Fig. 11. Comparison of predicted loss curve results

Fig. 11(a) shows that, at 100 iterations, the loss of EFTG-TTB was 0.094, while those of EBG-BiGRU and VFG-BiLSTM were 0.113 and 0.125, respectively. Fig. 11(b) shows that, at 140 iterations, the EFTG-TTB loss was 0.092, which was 0.028 lower than that of WDG-TCN. These results indicated that EFTG-TTB predicted stock prices more quickly and accurately. Next, the relative prediction errors on the training and test sets were evaluated. The results are shown in Fig. 12.

Fig. 12(a) shows that EFTG-TTB achieved the lowest relative error on the test and training sets, at 2.31% and 2.82%, respectively. Fig. 12(b) shows that VFG-BiLSTM achieved the lowest relative errors of 4.24% and 4.95%. Fig. 12(c) shows that EBG-BiGRU achieved a minimum relative error of 4.12% on the test set. Comparing Fig. 12(d) with Fig. 12(a), WDG-TCN showed higher minimum relative errors than EFTG-TTB, with increases of 1.63% and 1.92% on the test and training sets, respectively. These results demonstrated that EFTG-TTB achieved more precise stock price predictions. To further illustrate model performance, predictions for Geely Automobile and Dongfeng Motor stocks were analyzed, as shown in Fig. 13.

Fig. 13(a) shows that for Geely Automobile, the EFTG-TTB prediction curve was the smoothest overall, with minimal errors at key points. On day 180, the predicted price differed from the actual price by only 0.08 yuan. The EBG-BiGRU and VFG-BiLSTM models had errors slightly larger than EFTG-TTB. Fig. 13(b) shows that for Dongfeng Motor, EFTG-TTB still had the lowest error, while the WDG-TCN prediction fluctuated more than the actual prices. Overall, EFTG-TTB provided more accurate predictions of actual stock prices, with superior fluctuation handling and greater robustness than the other models. Finally, the performance of the four models was tested using R2, MAE, and Root Mean Square Error

(RMSE). The results are shown in Table 2.

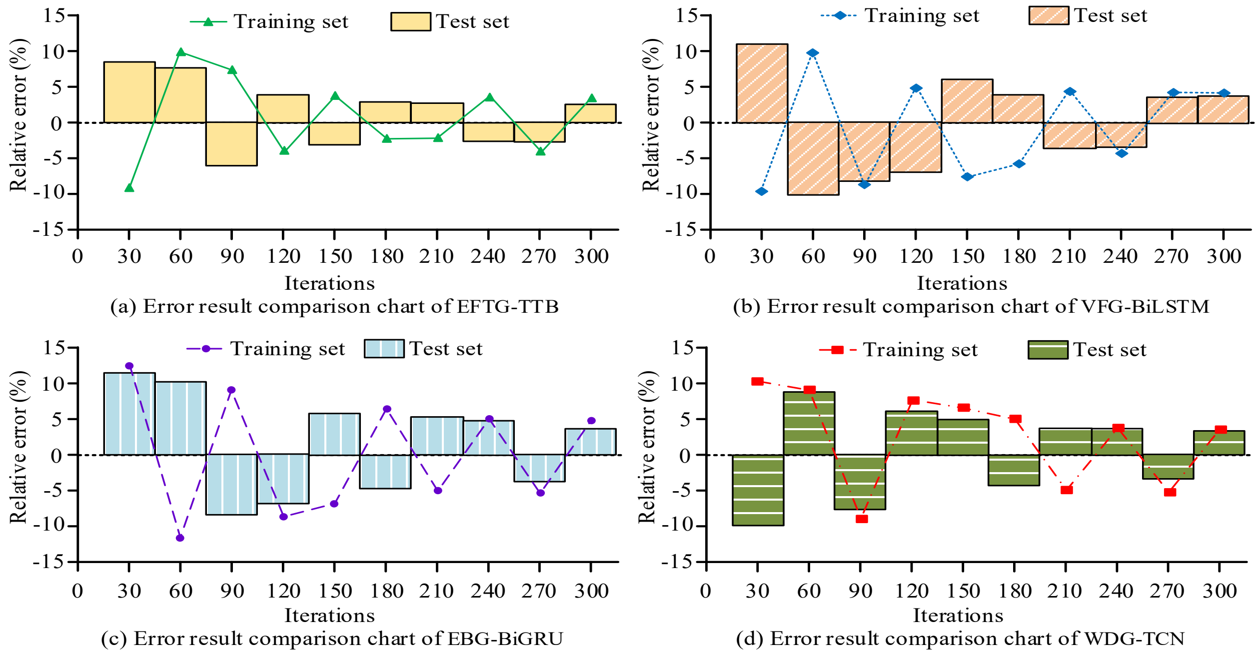
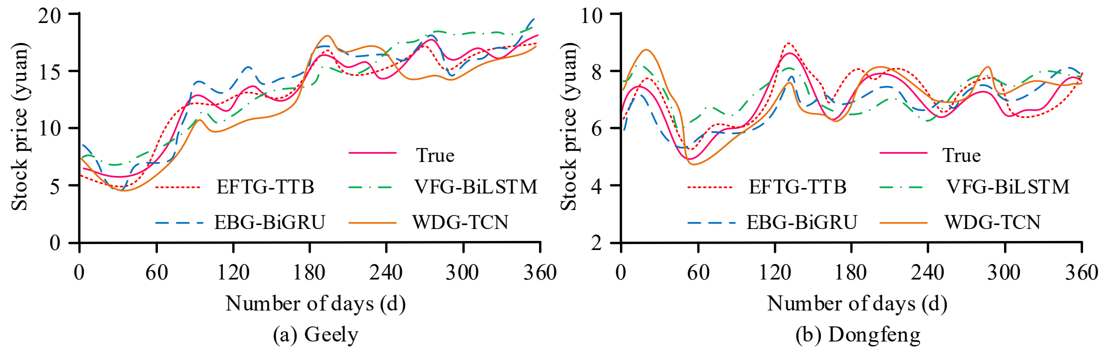


Fig. 12. Comparison of relative prediction errors of stock prices



(a) Dongfeng Motor (b) Different models.

Fig. 13. Comparison of stock price prediction performance for Geely Automobile

Table 2 shows that for Geely Automobile, the coefficient of determination of EFTG-TTB reached 0.925, significantly higher than VFG-BiLSTM (0.886), EBG-BiGRU (0.863), and WDG-TCN (0.815), indicating the strongest explanatory power for stock price fluctuations. The MAE of EFTG-TTB was 0.86 yuan, which was 0.17 yuan lower than VFG-BiLSTM. For Dongfeng Motor, EFTG-TTB maintained leading performance, with a coefficient of determination of 0.918 and an RMSE of 0.96 yuan, lower than the 1.35 yuan of EBG-BiGRU. In summary, EFTG-TTB consistently achieved excellent prediction accuracy and stability across markets and was sensitive to stock fluctuations, effectively predicting stock prices. To verify the model's robustness in highly volatile, shock-prone market scenarios, the study conducted an impact stress test on the EFTG-TTB model. This included three types of scenarios: historical real impacts, hypothetical extreme and out-of-distribution tests (including three real events such as the spread of the COVID-19 pandemic in 2020 and the Russia-Ukraine conflict in 2022, two types of hypothetical extreme disturbances such as the increase in stamp duty and liquidity depletion, and out-of-distribution data from the 2018 Sino-US trade war). The study combined market characteristics from 2020 to 2024 and full-cycle data from 2018. After processing, such as Monte Carlo simulation to add extreme disturbances and remove noisy data, and using relative returns, maximum drawdown, and 95% confidence level Value at Risk (VaR) as evaluation indicators, the model parameters remained unchanged, and the test was completed with 187,500 records (300 stocks×25 features×25 daily windows). The test results are shown in Table 3.

As shown in Table 3, the EFTG-TTB model maintains strong predictive performance across all shock scenarios, with relative returns exceeding 1.05. It outperforms the average maximum drawdown of 35% for the Shenzhen 300 Index during the same period. In distribution-out scenarios, the model's performance is similar to that in historical shock scenarios, without a significant decline, confirming its robustness to highly volatile, sudden shocks and to deviations in data distribution.

Table 2. Comparison of test results of three comprehensive indicators

Stocks	Model	R^2	MAE (yuan)	RMSE (yuan)
Geely	EFTG-TTB	0.925	0.86	1.05
	VFG-BiLSTM	0.886	1.03	1.28
	EBG-BiGRU	0.863	1.16	1.45
	WDG-TCN	0.815	1.39	1.72
Dongfeng	EFTG-TTB	0.918	0.78	0.96
	VFG-BiLSTM	0.879	0.95	1.18
	EBG-BiGRU	0.855	1.08	1.35
	WDG-TCN	0.808	1.30	1.61

Table 3. Results of impact pressure test experiment

Test scenario	Relative yield rate (%)	Maximum drawdown (%)	VaR (95%)
Conventional scenario	1.23	8.7	2.3
Historical impact scenario	1.18	15.2	3.8
Assumed extreme scenario	1.05	22.6	5.1
Off-site scenarios	1.12	18.5	4.3

4. Discussion

The proposed EFTG multi-source data extraction algorithm demonstrated significant advantages in comprehensive stock feature extraction, providing high-quality input for subsequent prediction models. At the level of time series data processing, the adaptive hierarchical decomposition mechanism of EMD can precisely separate noise, such as non-recurring gains and losses, but a single EMD is prone to mode aliasing. FastICA, by exploring the statistical independence of signals, precisely compensates for this defect. This combined design is better aligned with the non-stationary, strong noise, and multi-modal superposition characteristics of stock time series data than the method of Li et al. (2023), which relies solely on clustering algorithms to integrate basic trading data. This design enables the processed signals to retain the long-term trend while avoiding local feature distortion. This is the core reason why SSIM and PSNR are superior to those of the VMD-FT-GCN and EMD-BERT-GAT algorithms. For processing graph-structured data, the node attention mechanism of GAT addresses the limitations of static weight allocation in traditional GCNs and is more dynamically adaptable than GCNs in the VMD-FT-GCN algorithm and GIN in the WT-DCNN-GIN algorithm. By dynamically calculating the weights of industry and equity correlations among stocks, the algorithm can adaptively capture changes in the market correlation network. This effectively alleviates the problem of the Shilpa and Shambhavi (2023) model's degradation in prediction performance for high-volatility stocks, as reflected in its AUC of 0.914, higher than the 0.902, 0.886, and 0.864 of the comparison algorithms, and in its PR curve, which is closer to the upper right corner. From the perspective of multi-source integration theory, the EFTG algorithm first purifies the single-modal features and then dynamically aggregates the logical relationships across multiple modalities. This breaks the fixed paradigm of presetting weights and feature concatenation in financial multi-source integration, demonstrating that, in scenarios with strong heterogeneity in financial data, prioritizing the resolution of single-modal noise and semantic distortion issues is key to improving integration performance. This logic can be transferred to other financial scenarios, such as credit assessment and quantitative trading.

In stock prediction tasks, the EFTG-TTB multi-source stock prediction model also exhibited outstanding overall performance. In the time series encoding and prediction module, the combination of Time2Vec-Trans and BiGRU provides dual reconstruction capability, capturing long-term trends and short-term fluctuations, which is more advantageous than a single time-series modeling method. Time2Vec injects temporal semantics through dual-component encoding, Transformer captures long-range dependencies, and BiGRU tracks short-term evolution and reversals. The synergy of these three elements is more sensitive to capturing short-term reversal signals in stock prices and more stable during convergence than the stock prediction model proposed by Huang et al. (2023), which uses a hybrid genetic algorithm and machine learning. Meanwhile, compared with the multi-scale time series feature convolution model proposed by Zhu et al. (2024) and the improved particle swarm optimization recurrent neural network proposed by Yang et al. (2023), the proposed model through Time2Vec injects temporal semantics, combines the multi-head self-attention mechanism of Transformer to solve the problem of long-range temporal dependencies, and realizes the dual reconstruction logic of macro trend anchoring and micro fluctuation calibration. This design enables the model to maintain high stability across different markets and stock types, with R2 reaching up to 0.925, significantly higher than those of comparison models such as VFG-BiLSTM and WDG-TCN, and both MAE and RMSE are lower. From the perspective of the temporal encoding methodology, the TTB algorithm constructs a three-dimensional framework for temporal semantic injection, global correlation capture, and local reversal modeling, thereby filling the gap in separating long-term dependency and short-term fluctuation modeling in financial time series prediction. It demonstrates strong periodicity, long memory, and sudden reversals in financial data. The hybrid time-series encoding should focus on the collaborative work of complementary modules rather than on optimizing a single module. This insight can be extended to time-dependent financial tasks such as futures price prediction and exchange rate fluctuation warning.

The contributions of this study were threefold: first, it proposed the EMD-FastICA combination algorithm, which simultaneously addressed noise interference and mode mixing in multi-source time-series data through adaptive layered decomposition and blind source separation, overcoming the limitations of traditional time-series processing that could only denoise without separating mixed components. Second, it constructed a cross-modal feature-extraction framework combining TextCNN and GAT, achieving semantic alignment and dynamic fusion of unstructured text, graph-structured data, and structured time-series data, thereby overcoming the static weighting and poor feature synergy in existing methods. Third, it designed the Time2Vec-Transformer-BiGRU hybrid time-series encoding module, which injected time semantics, captured long-range dependencies, and modeled short-term reversals, balancing long-term trends and short-term fluctuations in stock prediction and enhancing model robustness under highly volatile and shock-prone market scenarios, providing an integrated technical framework for precise multi-source time-series encoding in stock prediction.

5. Conclusion

To address the challenges of traditional stock prediction models in dynamically integrating multi-source heterogeneous data and comprehensively capturing nonlinear time-series features, this study proposed a stock prediction model based on multi-source data and time-series encoding. The study first constructed a multi-source data extraction algorithm to provide high-quality features with unified dimensions and complete semantics for subsequent prediction. On this basis, Time2Vec was employed to inject temporal semantics into the multi-source features, and Transformer was combined to balance the modeling of long-term trends and short-term fluctuations. Experimental results showed that the proposed EFTG-TTB model outperformed the comparison models in terms of prediction accuracy and stability. Although the proposed model performed well, it had not yet integrated broader heterogeneous data such as macroeconomic indicators. Future research will further incorporate distributed computing frameworks to enhance real-time processing capabilities for large-scale stock portfolios. In addition, integrating macroeconomic data and cross-market correlation features will improve the model's adaptability to multi-market linkage scenarios, further strengthening its practical value in complex financial environments and providing more scientific quantitative support for investor strategy formulation, financial risk management, and market supervision.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author used AI tools solely for language editing and readability improvement. The author reviewed and verified all content and takes full responsibility for the accuracy and integrity of the manuscript.

References

Alam, K., Bhuiyan, M. H., Haque, I. U., Monir, M. F., and Ahmed, T. (2024). Enhancing stock market prediction: A robust LSTM-DNN model analysis on 26 real-life datasets. *IEEE Access*, 12, 122757-122768. https://doi.org/10.1109/ACCESS.2024.3434524

Bathla, G., Rani, R., and Aggarwal, H. (2023). Stocks of year 2020: Prediction of high variations in stock prices using LSTM. *Multimedia Tools and Applications*, 82(7), 9727-9743. https://doi.org/10.1007/s11042-022-12390-5

Gheisari, M., Hamidpour, H., Liu, Y., Saedi, P., Raza, A., Jalili, A., Rokhsati, H., and Amin, R. (2023). Data mining techniques for web mining: A survey. *Artificial Intelligence and Applications*, 1(1), 3-10. https://doi.org/10.47852/bonviewAIA2202290

He, Z., Tian, T., Wu, Q., and Jin, X. (2023). FTW-GAT: An FPGA-based accelerator for graph attention networks with ternary weights. *IEEE Transactions on Circuits and Systems II*, 70(11), 4211-4215. https://doi.org/10.1109/TCSII.2023.3280180

Htun, H., Biehl, M., and Petkov, N. (2023). Survey of feature selection and extraction techniques for stock market prediction. *Financial Innovation*, 9(1), 26-50. https://doi.org/10.1186/s40854-022-00441-7

Huang, J. Y., Tung, C. L., and Lin, W. Z. (2023). Using social network sentiment analysis and genetic algorithm to improve the stock prediction accuracy of the deep learning-based approach. *International Journal of Computational Intelligence Systems*, 16(1), 93-106. https://doi.org/10.1007/s44196-023-00276-9

Imran, H. A., Riaz, Q., Hussain, M., Tahir, H., and Arshad, R. (2024). Smart-wearable sensors and CNN-BiGRU model: A powerful combination for human activity recognition. *IEEE Sensors Journal*, 24(2), 1963-1974. https://doi.org/10.1109/JSEN.2023.3338264

Li, M., Zhu, Y., Shen, Y., and Angelova, M. (2023). Clustering-enhanced stock price prediction using deep learning. *World Wide Web*, 26(1), 207-232. https://doi.org/10.1007/s11280-021-01003-0

Li, Y., Chen, L., Sun, C., Liu, G., Chen, C., and Zhang, Y. (2024). Accurate stock price forecasting based on deep learning and hierarchical frequency decomposition. *IEEE Access*, 12, 49878-49894. https://doi.org/10.1109/ACCESS.2024.3384430

Li, Y., Li, J., Tu, P., Wang, H., and Wang, K. (2023). Gesture recognition based on EEMD and cosine Laplacian eigenmap. *IEEE Sensors Journal*, 23(14), 16332-16342. https://doi.org/10.1109/JSEN.2023.3279555

Li, Y., Lu, Y., Zhang, R., Ai, B., and Zhong, Z. (2023). Deep learning for energy efficient beamforming in MU-MISO networks: A GAT-based approach. *IEEE Wireless Communications Letters*, 12(7), 1264-1268. https://doi.org/10.1109/LWC.2023.3270361

Liu, J., Xiang, Z. Q., Ning, X. L., and Fang, J. C. (2024). GGA-EMD-based inversion method of spectrum reflectance template for celestial Doppler difference navigation. *IEEE Transactions on Aerospace and Electronic Systems*, 60(1), 658–674. https://doi.org/10.1109/TAES.2023.3327226

Luo, A., Zhong, L., Wang, J., Wang, Y., Li, S., and Tai, W. (2024). Short-term stock correlation forecasting based on CNN-BiLSTM enhanced by attention mechanism. *IEEE Access,* 12, 29617-29632. https://doi.org/10.1109/ACCESS.2024.3369419

Mukherjee, S., Sadhukhan, B., Sarkar, N., Roy, D., and De, S. (2023). Stock market prediction using deep learning algorithms. *CAAI Transactions on Intelligence Technology*, 8(1), 82–94. https://doi.org/10.1049/cit2.12059

Patel, M., Jariwala, K., and Chattopadhyay, C. (2024). A hybrid relational approach toward stock price prediction and profitability. *IEEE Transactions on Artificial Intelligence*, 5(11), 5844-5854. https://doi.org/10.1109/TAI.2024.3408129

Selva, J., Johansen, A. S., Escalera, S., Nasrollahi, K., Moeslund, T. B., and Clapes, A. (2023). Video transformers: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11), 12922-12943. https://doi.org/10.1109/TPAMI.2023.3243465

Shi, X., Hu, M., Ren, F., and Shi, P. (2024). A dual-ways feature fusion mechanism enhancing active learning based on TextCNN. *Intelligent Data Analysis*, 28(5), 1189-1211. https://doi.org/10.3233/IDA-230332

Shilpa, B. L., and Shambhavi, B. R. (2023). Combined deep learning classifiers for stock market prediction: Integrating stock price and news sentiments. *Kybernetes*, 52(3), 748-773. https://doi.org/10.1108/K-06-2021-0457

Singh, T., Kalra, R., Mishra, S., Satakshi, and Kumar, M. (2023). An efficient real-time stock prediction exploiting incremental learning and deep learning. *Evolutionary Systems*, 14(6), 919-937. https://doi.org/10.1007/s12530-022-09481-x

Srivastava, S., Pant, M., and Gupta, V. (2023). Analysis and prediction of Indian stock market: A machine-learning approach. *International Journal of System Assurance Engineering and Management*, 14(4), 1567-1585. https://doi.org/10.1007/s13198-023-01934-z

Tian, H., Zhang, X., Zheng, X., and Zeng, D. D. (2023). Learning dynamic dependencies with graph evolution recurrent unit for stock predictions. *IEEE Transactions on Systems*, Man, and Cybernetics: Systems, 53(11), 6705-6717. https://doi.org/10.1109/TSMC.2023.3284840

Wang, K. and Shao, B. (2024). Combined forecast of natural gas load based on Time2Vec-LSTM-TCN-Attention. *Nanjing Xinxi Gongcheng Daxue Xuebao*, 16(6), 801-809. https://doi.org/10.13878/j.cnki.jnuist.20240104002

Wu, J. M., T., Li, Z., Herencsar, N., Vo, B., and Lin, J. C., W. (2023). A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems*, 29(3), 1751-1770. https://doi.org/10.1007/s00530-021-00758-w

Xing, R., Cheng, R., Huang, J., Li, Q., and Zhao, J. (2024). Learning to understand the vague graph for stock prediction with momentum spillovers. *IEEE Transactions on Knowledge and Data Engineering*, 36(4), 1698-1712. https://doi.org/10.1109/TKDE.2023.3310592

Xu, J., Yuan, S., Shang, J., Wang, J., Yan, K., and Yang, Y. (2024). Spatiotemporal network based on GCN and BiGRU for seizure detection. *IEEE Journal of Biomedical and Health Informatics*, 28(4), 2037-2046. https://doi.org/10.1109/JBHI.2024.3349583

Xu, P., Zhu, X., and Clifton, D. A. (2023). Multimodal learning with transformers: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10), 12113-12132. https://doi.org/10.1109/TPAMI.2023.3275156

Yang, F., Chen, J., and Liu, Y. (2023). Improved and optimized recurrent neural network based on PSO and its application in stock price prediction. *Soft Computing*, 27(6), 3461-3476. https://doi.org/10.1007/s00500-021-06113-5

Zhang, K., Tang, L., Fan, W., He, Y., and Rao, J. (2024). Aging and echelon utilization status monitoring of power batteries based on improved EMD and PCA of acoustic emission. *IEEE Sensors Journal*, 24(8), 1214-12152. https://doi.org/10.1109/JSEN.2024.3370739

Zhao, H., Zhang, X., Chen, M., and Zhou, P. (2024). Online decomposition of surface electromyogram into individual motor unit activities using progressive FastICA peel-off. *IEEE Transactions on Biomedical Engineering*, 71(1), 160-170. https://doi.org/10.1109/TBME.2023.3294016

Zhu, W., Dai, W., Tang, C., Zhou, G., Liu, Z., and Zhao, Y. (2024). PMANet: A time series forecasting model for Chinese stock price prediction. *Scientific Reports*, 14(1), 18351-18367. https://doi.org/10.1038/s41598-024-69303-9

Xu Ren obtained her Ph.D. in Business Administration from the Lyceum of the Philippines University in 2023. Currently, she works as an Associate Professor at the School of Economics, Liaoning University of International Business and Economics. She has published more than 10 papers in various academic journals and has presented at international academic conferences. Her research interests include securities, digital finance and regional economy.