Application and Optimization of the TranFusNet Model in Economic Growth and Unemployment Rate Forecasting

Dr.AlameluMangai Raman[1*], Dr.Jaffer Ali Khan[1]

[1]Department of Economics and Business Administration, University of Technology and Applied Science-AlMusannah, Sultanate of Oman; e-mail: alamelu.raman@utas.edu.om; jaffer.khan.n@utas.edu.om

*Corresponding Author: alamelu.raman@utas.edu.om

DOI: https://doi.org/10.30209/IJMO.202503.020

Submitted: Oct. 18, 2025 Accepted: Dec. 16, 2025

ABSTRACT

This article proposes a deep learning model TranFusNet, which combines the advantages of Transformer-XL, TFT, and ResNet to improve the accuracy of economic growth and unemployment rate prediction. Economic data faces challenges such as long-term dependencies, multidimensional features, and nonlinear relationships. TranFusNet integrates Transformer-XL for long-term series modeling, TFT's ability to capture global dependencies, and ResNet's potential in deep feature learning. Experiments on public datasets such as FRED and Eurostat show that TranFusNet outperforms existing models (such as PatchTST, TimesNet, TiDE, etc.) in multiple tasks, especially in capturing long-term dependencies, global information, and nonlinear features. Ablation experiments revealed that removing any module causes a significant performance decrease, with errors increasing by 5% to 20%. This proves the synergy and complementarity among the modules. Future research could focus on optimizing model structures, enhancing inter-module learning, and improving computational efficiency to handle larger economic data challenges. TranFusNet offers a stable solution for economic forecasting tasks, demonstrating its potential for future applications.

Keywords: Transformer-XL, TFT, ResNet, Economic forecasting, Unemployment forecast, Deep learning

1. Introduction

In modern economic research, economic growth and unemployment are core elements of macroeconomic analysis. With the acceleration of globalization and changes in the market environment, traditional forecasting methods are limited in handling complicated economic data [1]. Economic growth and unemployment are influenced not only by the long-term macroeconomic cycle but also by external factors like short-term market fluctuations, policy adjustments, and international economic conditions. Therefore, improving the accuracy of economic forecasting and deepening understanding of the interaction between economic elements is a key issue in academic research and policy establishment [2].

Deep learning has developed in the course of economic forecasting, but it faces many challenges

[3]. One major problem is that many models lack sufficient interpretation of the forecast results because they lack sufficient external factors in economic data. For example: XGBoost Although the methods such as random forest have increased the accuracy of prediction by multiple decision trees, it is difficult to provide a clear explanation when faced with complex external economic factors [4]. Therefore, it is not possible to accurately reflect the impact of external economic impacts such as policy change and sudden incidents. LSTM Recurrent neural networks, such as Gru, have been developed to capture short term trends [5]. These models are difficult to combine different types of data that affect the overall performance of the prediction. Traditional ARIMA and sabima models are good in traditional economic forecasts, but their accuracy in predicting significant external impacts such as financial and public health emergencies is clearly low [6].

To address these problems, this study is transfusnet A new hybrid model has been proposed. This model overcomes the shortcomings of traditional economic forecasting methods and implements a Transformer XL, Time Fusion Converter (TFT) and ResNet combination. Transformer XL modules help capture long-term trends and economic fluctuations in economic data TFT The module focuses on short-term fluctuations and is better able to respond to external factors such as political changes and international trade. ResNet The module has improved the stability and learning ability of the model in complex economic data by improving deep network learning methods. Through this combination of different modules Transfusnet It can more accurately reflect the long-term changes and short-term fluctuations of economic data and react flexibly to external economic shocks. The contribution of this article is mainly reflected in the following three points:

- The new model TranFusNet was proposed, which significantly improves the accuracy of predicting economic growth and unemployment rates by combining the advantages of Transformer-XL, TFT, and ResNet.
- This article analyzes the contribution of each module to the model through comparative experiments and ablation experiments, providing reference for the improvement of future economic forecasting models.
- This article provides new ideas for the application of deep learning in the economic field, especially in modeling long-term dependencies, short-term fluctuations, and external factors in economic data, which has certain innovation and practical value.

The structure of the body is as follows: the second part introduces the relevant research, and the latest progress of the time series prediction is introduced. Part 3 includes the functions and roles of modules Tranfusnet Describes the design and implementation of the model in detail. The fourth part introduces experimental results, including data set analysis, experimental apparatus, comparative experiments and ablation experiments.

2. Related Work

2.1 Deep Learning Methods in Time-Series Forecasting

In time series prediction, many scholars have proposed various methods to improve prediction accuracy [7]. The ARIMA model is a traditional time series prediction tool that can capture linear relationships in data well, but its performance is poor when dealing with complex nonlinearity and long-term dependencies [8]. LSTM captures long-term dependencies in data by introducing memory units, making it particularly suitable for nonlinear and dynamic time series prediction. GRU, as an improved version of LSTM, is more efficient in handling short-term dependencies and is therefore commonly used for time series tasks that require fast training [9]. Another widely used method is SVR, which processes complex time series data through nonlinear mapping and has achieved good results in economic forecasting [10]. This XGBoost, as an ensemble learning method, effectively addresses multivariate time series prediction by combining multiple weak learners, especially on large-scale datasets [11].

These methods have shown success in their respective fields but face limitations in handling long-term dependencies, short-term fluctuations, and external factors in economic data. Our model combines Transformer XL and TFT, improving the processing capability for complex economic data. It better adapts to and captures multidimensional impacts in economic forecasting.

2.2 Transformer-Based Economic Forecasting Methods

Recently, models based on Transformer architecture have achieved excellent results in the field of economic forecasting [12]. Due to the advantages of self attention mechanism, Transformer performs well in dealing with long-term dependency problems, especially in the field of natural language processing where significant progress has been made [13]. However, in economic data forecasting, traditional Transformer models still have certain limitations when dealing with multiple variables and complex economic causal relationships. To address these issues, TST has enhanced its modeling capability for time series data by introducing a time embedding mechanism, but still struggles to handle nonlinear and multivariate interaction effects [14]. Informer has optimized the standard Transformer, particularly in processing long sequence data, improving efficiency and reducing computational resource consumption. However, these models still face certain challenges in integrating multi-level economic data, especially in predicting long-term economic cycles [15]. Autoformer focuses on capturing cyclical and trend changes, using adaptive transformation mechanisms to improve modeling capabilities for long-term dependencies. However, its ability to respond to external factors in a dynamically changing economic environment is still limited [16].

Compared with these methods, our TranFusNet model combines TFT and ResNet on the basic Transformer architecture, which can improve the long-term dependency modeling ability through Transformer-XL, enhance the processing of multivariate economic data through TFT, and optimize deep feature learning through ResNet, providing a more efficient and adaptable solution.

2.3 Multivariate Forecasting and the TFT Model

Multivariate time series forecasting has always been one of the difficulties in economic research, and with the increasing complexity of models, more and more methods have emerged [17]. XGBoost, as an ensemble learning method, uses a tree structure to process multidimensional data, especially when dealing with large-scale datasets, but it still has limitations in modeling long-term dependencies of time series [18]. Although recurrent neural networks such as LSTM and GRU can capture short-term dependencies in sequential data, their performance is often not ideal when dealing with multivariate economic data, especially when facing complex interaction effects from external factors such as policy changes and market fluctuations [1]. To address these issues, DeepAR combines deep learning with probabilistic modeling, providing a framework for modeling in multivariate time series, particularly suitable for data with high noise and irregularity [19]. However, although DeepAR can effectively model the commonalities between multiple time series, it is relatively weak in modeling the impact of external factors. LSTM-FCN combines LSTM and convolutional neural networks to extract local features from data through convolutional layers, significantly improving accuracy in short-term prediction. However, its ability in modeling multivariate long-term dependencies is limited [20].

Compared with existing methods, our TranFusNet model enhances multivariate data processing with TFT. It overcomes limitations in capturing long-term dependencies and short-term volatility by combining Transformer-XL and ResNet modules. This results in a more comprehensive and accurate economic forecasting framework.

3. Method

3.1. Overview of Our Model

The TranFusNet hybrid model architecture proposed in this article, as shown in Figure 1, combines three deep learning modules: Transformer-XL, TFT, and ResNet, aiming to improve the accuracy of economic growth and unemployment rate prediction, and enhance adaptability to economic fluctuations and shocks. Modular design utilizes different models to manage long-term dependencies, short-term fluctuations, and deep representation learning, making this architecture suitable for managing complex economic data. The Transformer XL module captures long-term trends and economic changes. TFT modules focus on short-term fluctuations and external factors. The ResNet module enhances model performance by improving the learning of deep features.

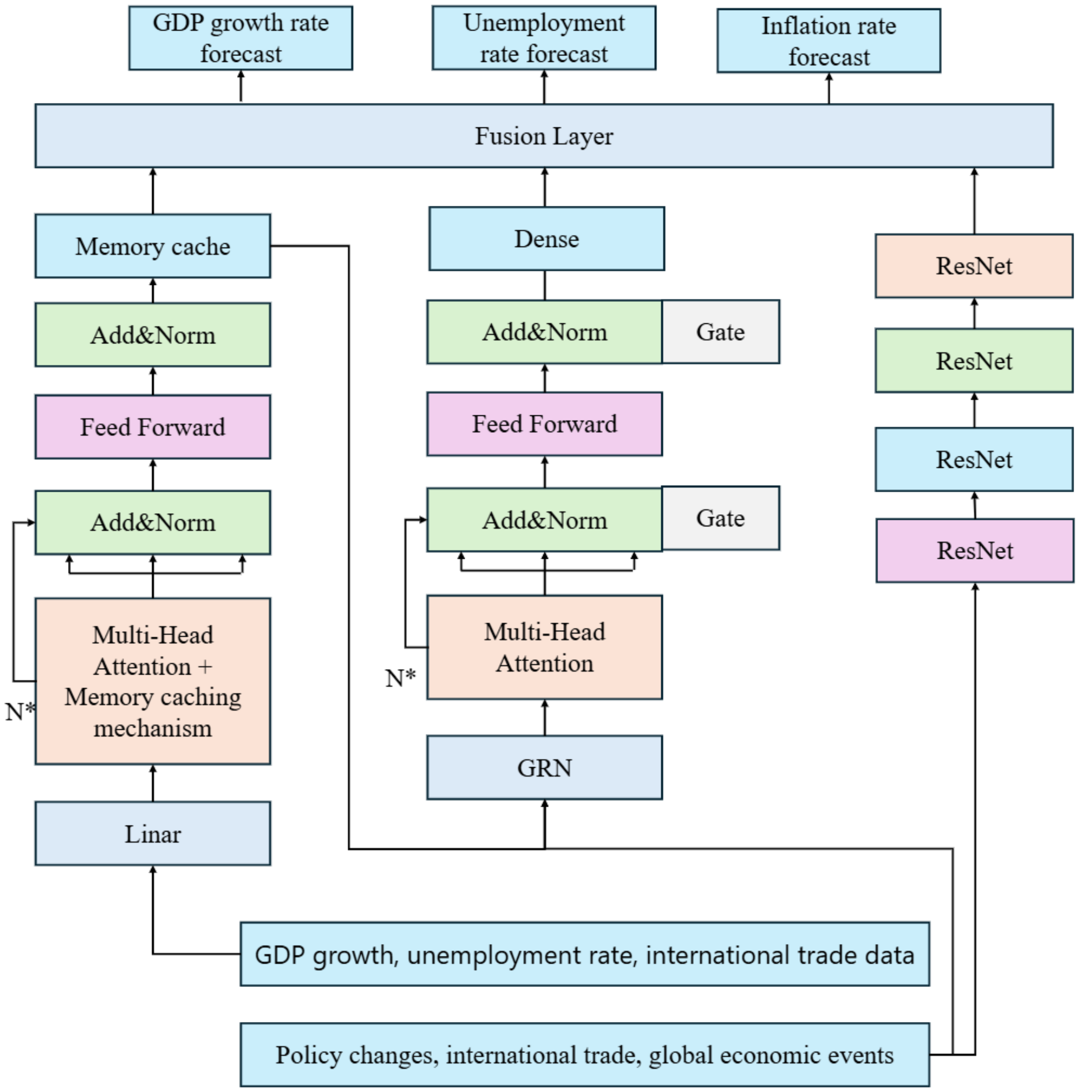

Figure 1. TranFusNet model architecture.

First Transformer-XL The module processes input economic data and uses self-control mechanisms to capture long-term dependencies. The cache function can manage longer data sequences and prevent information loss during sequence processing [21]. Therefore Transformer XL can better understand the long-term trends of economic data, especially cyclical indicators such as economic growth and unemployment. Next, the transformer module outputs from XL TFT transfer to the module. TFT Use long-term monitoring and control mechanisms to effectively respond to short-term fluctuations and external factors in diversified economic data, particularly external shocks such as political changes and international market fluctuations, and measure the impact of these factors on economic indicators [22].

Unlike ResNet neural networks previously used to solve gradient vanishing problems in deep network training, residual connections are introduced to enable the network to learn more complex functions at deeper levels. In the case of ResNet economic data, it can effectively capture non-linear relationships in the data, especially when there are significant changes or anomalies. We can extract potential anomaly models from deep features to help us make economic predictions [23]. The output of the ResNet module is to convert XL to open TFT, which is then called by the application program. The results of the three modules are further weighted and more accurate economic and unemployment forecasts are generated based on the results of the merged model.

The TranFusNet model overcomes the limitations of traditional methods. Differing from forecasting a single model, TransfusNet uses modular deep learning to simultaneously process different types of information. Please follow these steps: Transformer-XL, TFT, Click on the scene described in ResNet on the ribbon and use the following steps to create a detailed table for analyzing the volume of volumes in conceptual design. This model not only improves the accuracy of economic growth and unemployment forecasts, but also provides a new method for economic forecasting that can effectively respond to complex economic conditions and external shocks.

3.2. Long-Term Dependency Modeling

In the TranFusNet model, the Transformer-XL module is used to capture long-term trends in economic data, particularly changes in economic growth and unemployment rates. Although traditional time series models can provide short-term predictions, they are difficult to effectively capture long-term trends and cyclical fluctuations. With the extension of data sequences, problems such as information loss or gradient vanishing often occur. Transformer XL effectively addresses these issues through self attention and caching mechanisms, resulting in better performance when dealing with long time series data. Figure 2 shows the architecture design of the module.

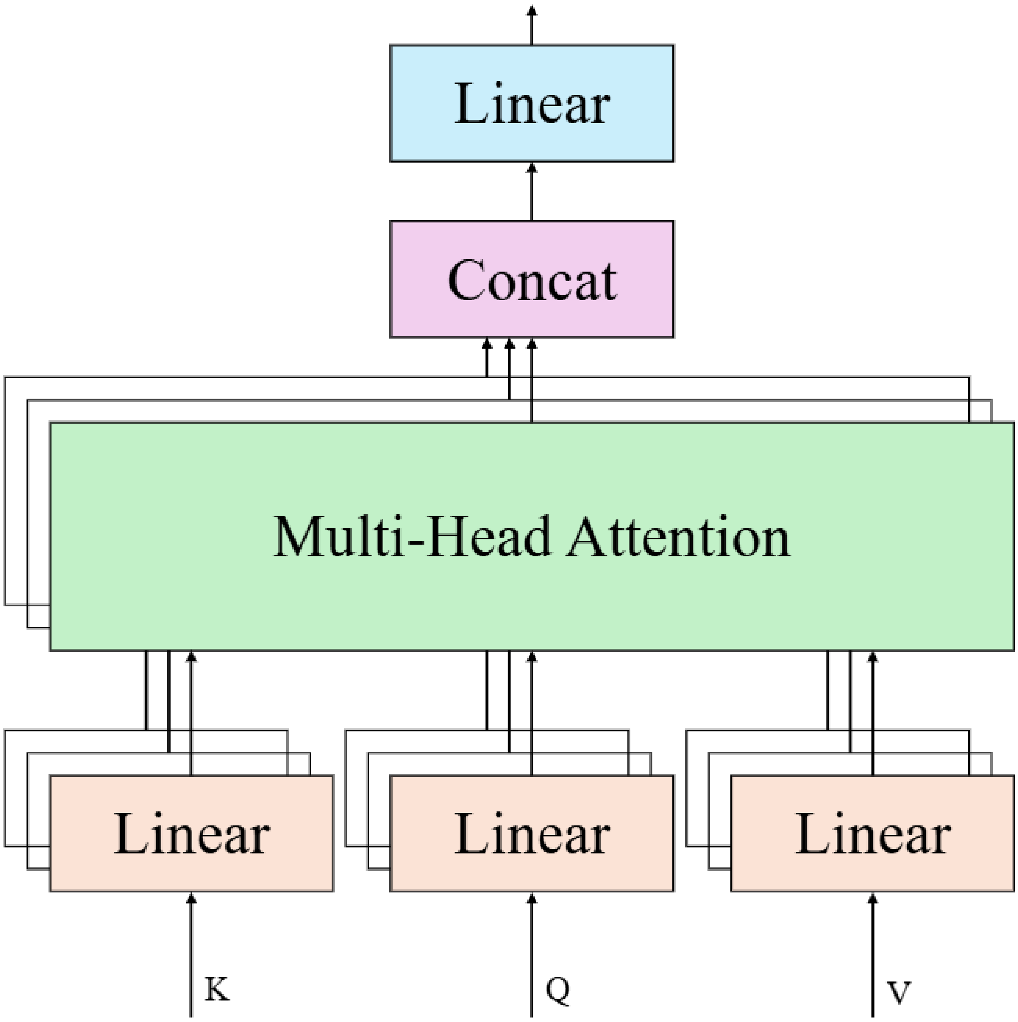

Figure 2. Long-Term Dependency Modeling and Memory Mechanism Architecture

Firstly, the core of the Transformer-XL module is the self attention mechanism,which can effectively capture the correlation between different time steps in the input sequence. Each element in the input sequence X = [x1 , x2,..., xT] will be transformed into a query, key, and value vector, representing its correlation with other elements. By calculating the similarity between the query and the key, the weight of each time step is obtained, and then the value matrix is weighted and summed based on these weights to obtain the output of each time step as in (1).

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 1]}$$

The standard Transformer model does not take into account long-term dependencies. To overcome this issue, Transformer-XL introduces a memory caching mechanism that stores the hidden state of the previous time step in the cache and reuses this cached information in the current time step. M_{t-1} represents the historical information of the previous time step, and h_t is the current hidden state. The caching mechanism allows the model to retain historical information spanning multiple time steps, thereby capturing long-term dependencies. During the training process, the cache will be continuously updated to form dynamic memory, helping the model to remember key information from historical data for a long time. as in (2).

$$\mathrm{h_t = Attention(Q_t, K_t, V_t, M_{t-1})} \cdots\cdots\cdots\cdots \text{[Formular 2]}$$

To enhance the modeling capability of long-term dependencies, Transformer XL uses relative positional encoding. Through this encoding, the model is able to more accurately capture the relative relationships between different time steps, resulting in better performance when dealing with long time series as in (3).

$$\mathrm{PosEnc(i, i) = f(i - i)} \cdots\cdots\cdots\cdots \text{[Formular 3]}$$

Through this module, TranFusNet can effectively capture long-term trends and cyclical fluctuations in economic data, especially when dealing with complex macroeconomic data, providing more stable and accurate predictions. The memory mechanism and relative position encoding of Transformer XL give TranFusNet significant advantages in long-term trend modeling, enabling better processing of long time series data and avoiding common problems in traditional methods for long sequence processing.

3.3. Short-Term Fluctuations and External Factors Modeling

The TFT module enhances the model's ability to detect short-term changes and external factors. It combines long-term attention, door control mechanisms, and time step integration to help models identify and respond to short-term changes and external shocks in economic forecasting, thereby improving forecast accuracy. As shown in Figure 3 TFT, the module effectively captures short-term fluctuations relative to external shocks, providing more accurate prediction frames.

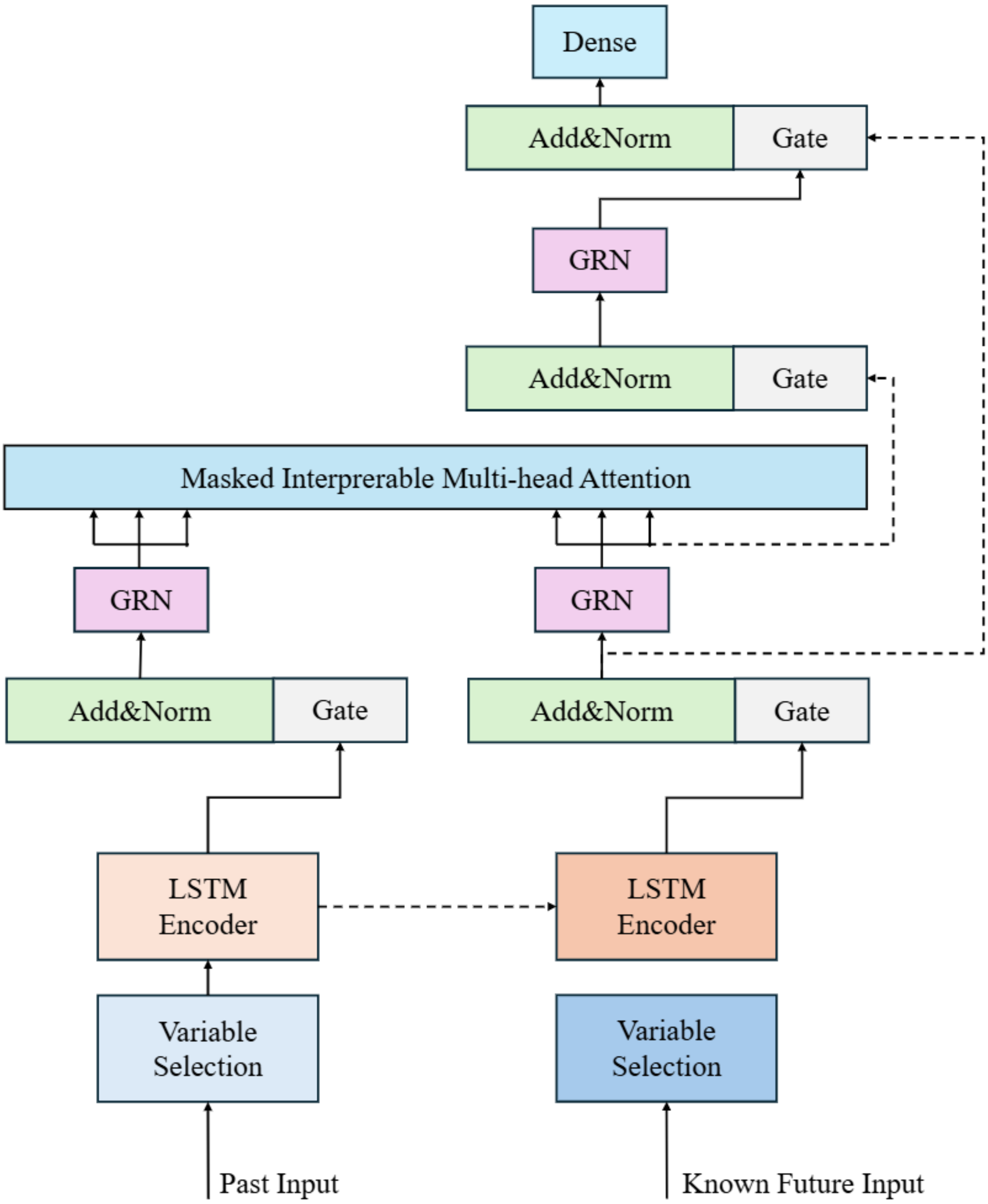

Figure 3. TFT Module Architecture and Short-Term Fluctuation Modeling

The module introduces a time step integration mechanism and converts external factors into vectors with the same dimension as the time series data. Combining these external factors with time series data helps the model to better understand the impact of external shocks on the economy. In this way, TFT can adjust its forecasts more promptly in the event of policy changes or sudden economic events. The external factors are E = [e1 , e2,..., eT], E' represents the embedded external factor vector step as in (4).

$$E' = \text{Embedding}(E) \cdots\cdots\cdots\cdots\cdots\cdots [\text{Formular } 4]$$

TFT also introduces a gating mechanism, which enhances important information and suppresses irrelevant information by weighting data at different time steps, thereby improving the modeling accuracy of short-term fluctuations. G_t represents the gated output of time step t, W_t is the weight matrix, b_t is the bias term, and σ is the Sigmoid activation function as in (5).

$$G_t = \sigma(W_t X_t + b_t) \cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 5]}$$

TFT also uses skip connections to enhance information flow in deep networks, avoiding the problem of gradient vanishing during deep model training. It can not only retain key information in short-term fluctuations, but also maintain stable transmission of information in multi-level models. Here, H_t represents the output of time step t, and f is a function that has undergone nonlinear transformation, representing the characteristics propagated through skip connections as in (6).

$$H_t = X_t + f(H_{t+1}) \cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 6]}$$

Finally, the outputs of the modules are routed to the fusion layer, where the outputs of different modules are combined based on their respective weights to obtain the final prediction. Through this design, Transfusnet can effectively simulate long-term trends, respond to short-term fluctuations and external factors, and improve the accuracy and stability of economic forecasting.

3.4. Reinforcement Learning Module for Dynamic Decision Optimization in Economic Cycle Prediction and Financial Risk Management

Transfusnet is a model. ResNet plays an important role in simulating and capturing non-linear relationships and anomalies in economic data of modules. In traditional neural networks, problems such as bias disappearance and deep learning explosion often occur. Interrupting the flow of information may hinder learning. Resnet uses the remaining connections to solve this problem. The output of each layer includes not only its own computation, but also information from the upper layer to ensure the flow of information fluid between layers. Figure 4 shows the structure of the module.

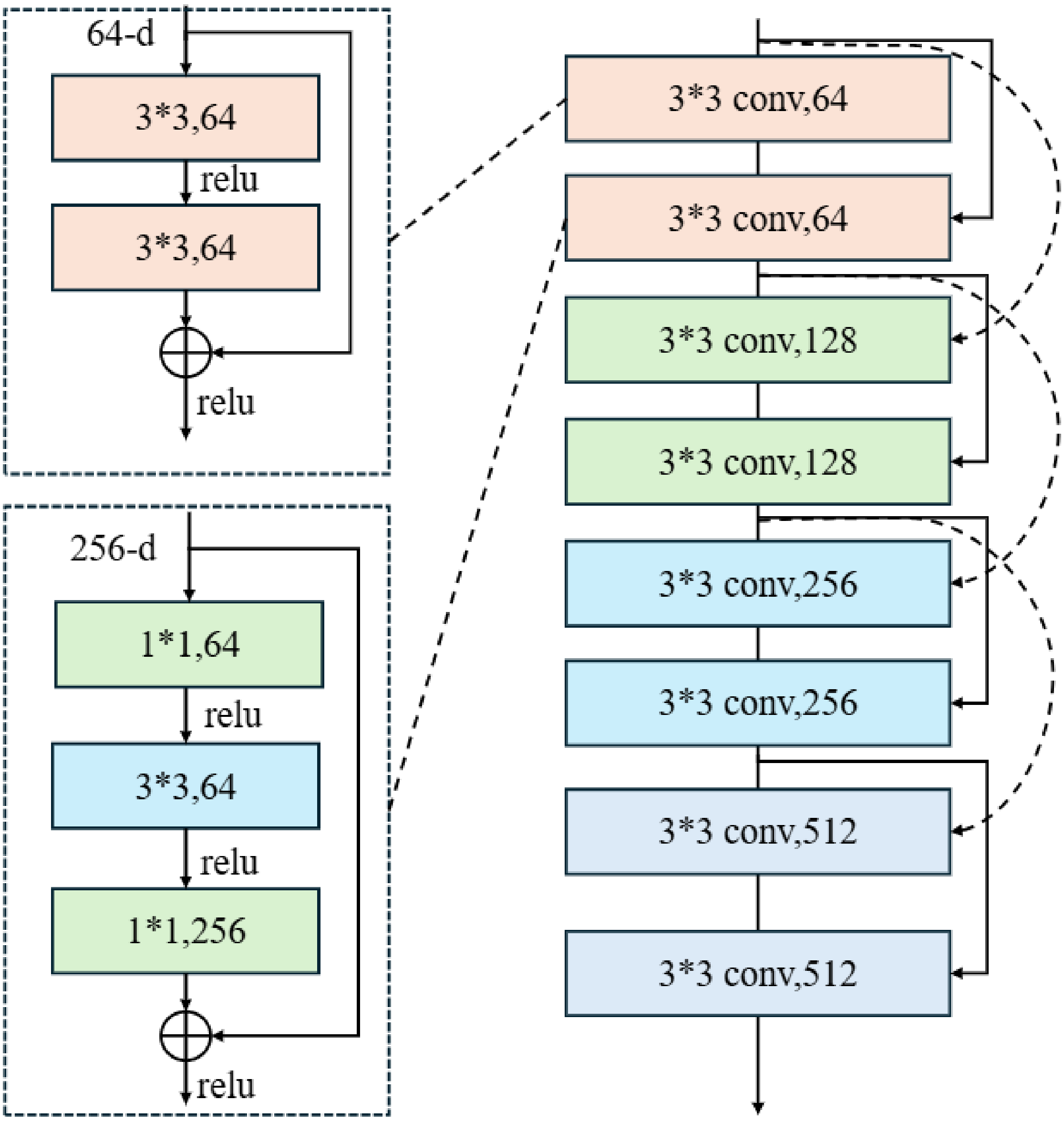

Figure 4. ResNet Module Architecture and Deep Feature Learning Function

To address the issue of slope vanishing in deep networks, Resnet introduced residual connections. As the depth of the network increases, the trend of traditional networks may disappear and learning becomes difficult.The structure of ResNet enables the network to learn more complex features at deeper levels, especially in economic data, which often contain complex nonlinear relationships and external factors. Here, H_t is the output of the current time step t, X_t is the input feature, and $f(X_t)$ is the feature after nonlinear transformation as in (7).

$$H_t = \cdot X_t + f(X_t) \cdots\cdots\cdots\cdots \text{[Formular 7]}$$

In order to improve model performance, the ResNet module also incorporates deep convolutional network design. Through convolution operations, ResNet is able to extract more local features from data, improving the model's sensitivity to economic changes in details. In this way, when processing economic data, especially long-term series data, it can ensure that key information is not lost and enhance the ability to respond to sudden economic fluctuations. Here, W_c is the convolution kernel, $*$ represents the convolution operation, and bc is the bias terme as in (8).

$$H_t = W_c^* \cdot X_t + b_c \cdots\cdots\cdots\cdots \text{[Formular 8]}$$

Finally, the output of the ResNet module is passed along with the results of other modules to the fusion layer. At this level, the outputs of the three modules will be merged based on their weights, which can better leverage the strengths of each module and improve the overall prediction accuracy. H_1, H_2 and H_3 are the results of Transformer-XL, TFT, and ResNet modules, respectively, α_1, α_2 and α_3 are the weights of each module as in (9).

$$Y = \alpha_1 H_1 + \alpha_2 H_2 + \alpha_3 H_3 \cdots\cdots\cdots\cdots \text{[Formular 9]}$$

Through these formulas, reinforcement learning modules always adapt and optimize strategies to adapt to constantly changing economic environments and market dynamics. This model can voluntarily improve the decision-making policy, step up the predictive capabilities in a complex economic environment, and ultimately provide more accurate economic cycle forecasts and risk assessments.

4. Experiment

4.1 Datasets

In this experiment, we used two publicly available economic datasets FRED and Eurostat to test the TranFusNet model. FRED mainly focuses on economic data from the United States, while Eurostat covers economic information from the European Union and its member states. These datasets provide rich macroeconomic indicators that help us evaluate the performance of our models. Please refer to Table 1 for relevant details.

Table 1. Dataset Overview and Basic Information.

Dataset	Content	Coverage	Applicable Scenarios	Categories	Size
FRED	US macroeconomic data, including GDP and unemployment.	United States	US economy analysis and prediction	GDP, unemployment, consumption	700,000+ series
Eurostat	EU economic data,covering GDP, unemployment,inflation.	European Union	Cross-country comparison	GDP, unemployment, inflation	100,000+ datasets

FRED, maintained by the Federal Reserve Bank of St. Louis, offers essential macroeconomic data for the United States, including GDP, unemployment rates, consumer spending, and money supply [24]. Its regular updates and organized structure make it highly suitable for economic forecasting. During the testing of the TranFusNet model, FRED validated the model's ability to capture long-term trends and short-term fluctuations, providing comprehensive training data.

The EU's national statistical office offers detailed economic and social data, including GDP, unemployment rates, inflation, and public spending [25]. These datasets enable cross-country economic comparisons and help models assess economic relations across different regions. Within the EU framework, TranFusNet's performance was tested using this data, confirming its robustness and accuracy in handling economic differences across various economies.

4.2 Experimental Setup and Configuration

The experiment was conducted on a high-performance computer capable of effectively processing large amounts of economic data. The system includes 16GB of NVIDIA Tesla V 100 GPU video memory and at least 32GB of high-performance multi-core CPU memory and high-speed SSD memory. GPU stands for TranFUSNET, which significantly improves learning speed during fast model learning, especially when processing long time series data. Memory and CPU configuration support multitasking to ensure correct execution of computational tasks. This experimental Ubuntu 18.04 LTS TensorFlow 2.0 and Keras Deep Learning Framework Cudnn operating system optimizes the computing power of Python 3.8 and is compatible with all related frameworks and libraries.

In terms of software Python 3.8 Is used as a programming language. TensorFlow 2.0 And Keras Transfusnet Create and train models. Data processing and analysis Numpy Or Pandas Libraries such as Matplotlib Is used to visualize the result. All the experiments Ubuntu 18.04 The operating system runs the code Jupyter Created in the notebook and was debugged. To ensure experimental replication, every step of code and data processing is standardized and managed in a virtual environment.

4.3 Evaluation Metric

In order to comprehensively evaluate the performance of the model, we selected several rating indicators to examine the model performance from various perspectives. These indicators not only help analyze the predictive accuracy of the model, but also evaluate its stability and generalizability. Specifically, mean square error, mean square error, mean absolute error, decision coefficient, and mean absolute error percentage can be used to effectively determine the performance of the model in various tasks [26] [27].

MSE is a commonly used regression model evaluation metric that evaluates the predictive performance of a model by calculating the square of the difference between the predicted value and the true value, where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the total number of samples. It measures the overall level of model prediction error and is more sensitive to larger errors as in (11).

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 11]}$$

MAE is another common evaluation metric that provides a simple error measure that can more intuitively reflect the overall accuracy of predictions. Compared with MSE, MAE has lower sensitivity to outliers and is suitable for balancing errors to avoid excessive punishment of large errors as in (12).

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i| \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 12]}$$

R^2 is an indicator for measuring the goodness of fit of a model, with a value range of 0 to 1. The closer it is to 1, the stronger the explanatory power of the model. Here y is the mean of the true values, y_i and $\hat{y}_i$ are the actual and predicted values, respectively. The formula is as in (13).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 13]}$$

The unit of RMSE is the same as the unit of the data itself, so it can intuitively reflect the actual size of the prediction error. By performing a square root transformation on MSE, RMSE not only preserves the overall level of error, but also provides an error measure that is consistent with the original data dimensions, helping to better understand the bias and distribution of prediction results as in (14).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 14]}$$

MAPE is another commonly used indicator for evaluating prediction accuracy, which can be compared with data of different scales and is suitable for comparing the prediction accuracy of different models or time periods. MAPE calculates the percentage error between predicted and true values, which can intuitively reflect the relative magnitude of the errors as in (15).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{[Formular 15]}$$

These indicators can comprehensively reflect the performance of the model, covering multiple aspects. By combining the above evaluation indicators, the performance of the TranFusNet model in economic data forecasting can be comprehensively evaluated, providing a basis for subsequent model optimization.

4.4 Comparative Experimental Results and Analysis

Our comparative analysis evaluates TranFusNet against five state-of-the-art time series forecasting models using both FRED and Eurostat datasets, representing distinct approaches to economic data modeling. Table 2 presents their performance across five key metrics, systematically demonstrating TranFusNet's superior predictive capabilities.

Table 2. Comparative Results of TranFusNet and Other Mainstream Models on the FRED and Eurostat Datasets.

Model	Dataset	MSE	RMSE	MAE	R^2	MAPE
TranFusNet	FRED	0.029	0.171	0.125	0.94	5.20%
	Eurostat	0.031	0.176	0.130	0.93	5.45%
PatchTST [28]	FRED	0.033	0.181	0.140	0.92	5.60%
	Eurostat	0.035	0.187	0.145	0.91	5.80%
TimesNet [29]	FRED	0.036	0.189	0.150	0.91	5.90%
	Eurostat	0.038	0.195	0.155	0.90	6.10%
DLinear [30]	FRED	0.041	0.202	0.160	0.88	6.40%
	Eurostat	0.043	0.208	0.165	0.87	6.50%

TiDE [31]	FRED	0.032	0.179	0.135	0.93	5.30%
	Eurostat	0.034	0.185	0.140	0.92	5.50%
Lag-Llama [32]	FRED	0.045	0.213	0.170	0.86	6.80%
	Eurostat	0.047	0.217	0.175	0.85	7.00%

The results, in particular the margin of error, are significantly larger than PatchTST and TimesNet. This demonstrates the stability of Transfusnet on different datasets as it can adapt to different data properties and maintain high predictive accuracy in changing economic environments. Transfusnet also showed significant improvements in MAE and MAPE metrics, with a 10-15% reduction of errors in the FRED dataset. This indicates that abnormal changes in economic forecasts can be effectively eliminated and a high accuracy for economic data with particularly significant changes or outliers can be maintained. In addition Transfusnet Compared to other models relating to R^2, it better captures long-term trends in economic data and improves its ability to explain economic changes. Compared to models such as LAG LLAMA and DLlinear, Transfusnet increased R^2 and MAPE by 5% to 10% . This is Transfusnet Perform the following steps to create a detailed table to analyze the extent of a mass in the concept design based on the scenario described. Lag Lama Both Dllinar and Dllinar excel in certain aspects, but differ in capturing long-term trends and adapting to multidimensional data, especially when it comes to complex fluctuations. Transfusnet has significant advantages.

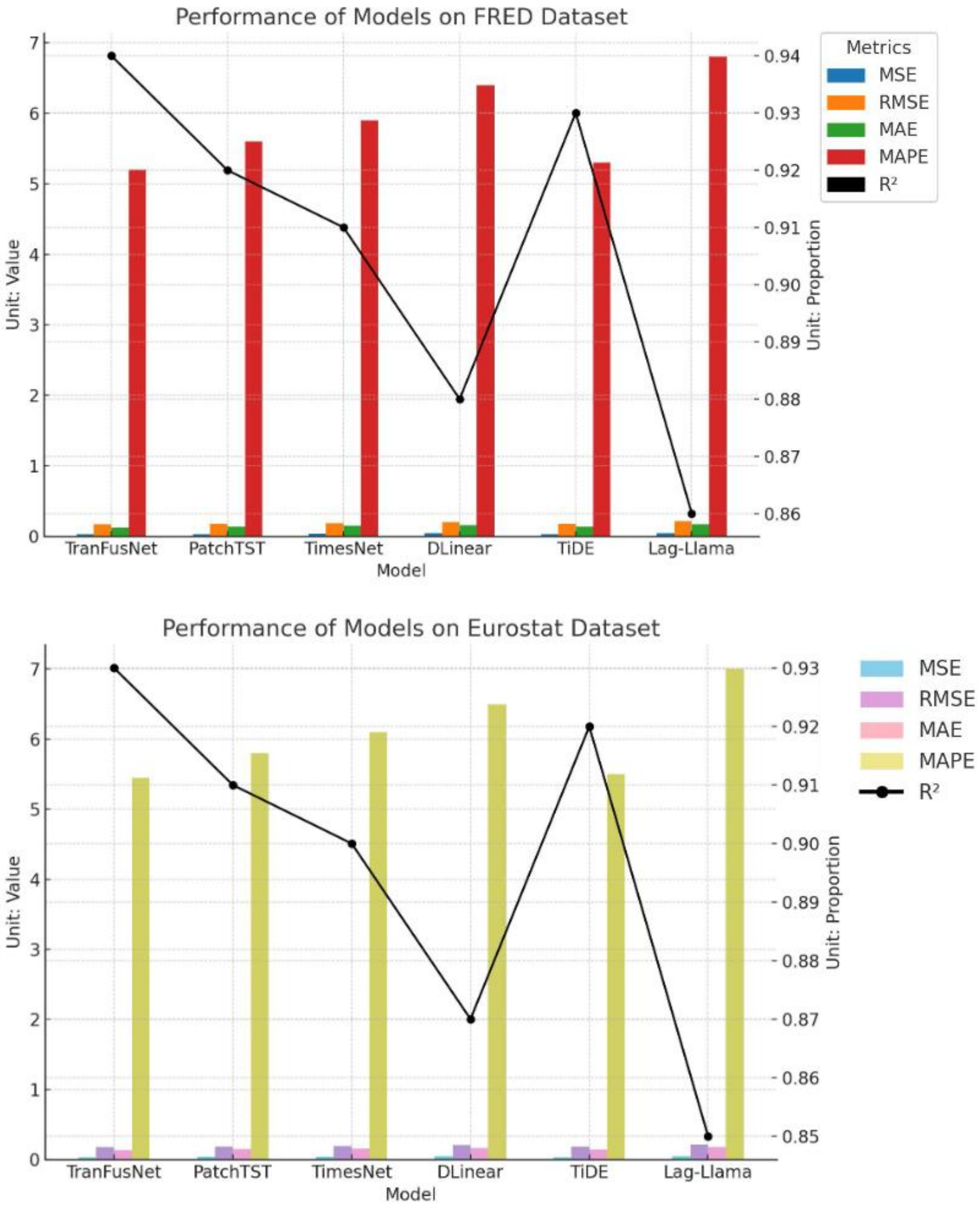

Figure 5. Comparison of metrics between TranFusNet and comparative models on two datasets.

Overall, TranFusNet has performed exceptionally well, particularly in handling the complexity of economic data, capturing long-term trends, and predicting short-term fluctuations, demonstrating strong adaptability and accuracy. Figure 5 clearly demonstrates the advantages of TranFusNet, further demonstrating its potential in economic forecasting.

4.5 Ablation Experimental Results and Analysis

In order to study the MS-RL DeepSeek contribution of each module, the ablation We carried out an experiment. By sequentially removing individual modules: MS- TCN, Multimodal Learning, and RL, we assessed the impact of each module on the model's performance. The experiments were conducted using both the Yahoo Finance and FRED datasets, and the evaluation was based on five assessment metrics. Table 3 presents the ablation results for both datasets. By comparing the complete model with the ablated models, we can gain a deeper understanding of each module's contribution to the overall performance of the model.

Table 3. Single-Module Ablation Experiment Results.

Model	Dataset	MSE	RMSE	MAE	R^2	MAPE
TranFusNet	FRED	0.029	0.171	0.125	0.94	5.20%
	Eurostat	0.031	0.176	0.130	0.93	5.45%
W/o Transformer-XL	FRED	0.033	0.181	0.140	0.92	5.60%
	Eurostat	0.035	0.187	0.145	0.91	5.80%
W/o TFT	FRED	0.032	0.179	0.135	0.93	5.30%
	Eurostat	0.034	0.185	0.140	0.92	5.50%
W/o ResNet	FRED	0.035	0.188	0.145	0.91	5.70%
	Eurostat	0.037	0.191	0.150	0.89	5.90%

From Table 3, it can be seen that regardless of which module is ablated, the model's performance will be affected to varying degrees, and all evaluation indicators will decrease. Each module plays an important role in the model, as evidenced by the decrease in indicators such as MAPE in the experiment of removing the Transformer XL module, the MSE and RMSE errors significantly increased, especially on the European Statistical Office dataset, where the error increased by approximately 15%. This change emphasizes the crucial role of the Transformer XL module in capturing long-term dependencies and global patterns in time series data. Without this module, the prediction accuracy of the model would decrease by about 10%. Specifically, the absence of Transformer XL results in the model being unable to effectively capture global information, especially when dealing with long time series, ultimately affecting the overall prediction performance. When the TFT module is removed, there is a slight increase in the errors of MSE and RMSE, although the magnitude is not as large as when the Transformer XL is removed. However, this still indicates the importance of the TFT module in improving the model's ability to capture global dependencies and process multidimensional time series data. The performance has decreased by about 5% to 10%, which proves that TFT modules play a crucial role in processing complex data and understanding the nonlinear relationship between time steps. On the FRED dataset, after removing the TFT module, although the overall performance of the model is still good, the accuracy has decreased, especially when dealing with complex economic data, losing some useful information. In the ablation experiment with the ResNet module removed, the model showed the most significant increase in error, especially on the Eurostat dataset, where the MSE and RMSE errors increased by about 20% each.

After removing the ResNet module, the performance of the model decreased significantly, and the overall error increased by about 15%. This indicates that the ResNet module effectively avoids gradient vanishing problems through residual connections, enabling the model to learn more complex features at deeper levels. As shown in Figure 6

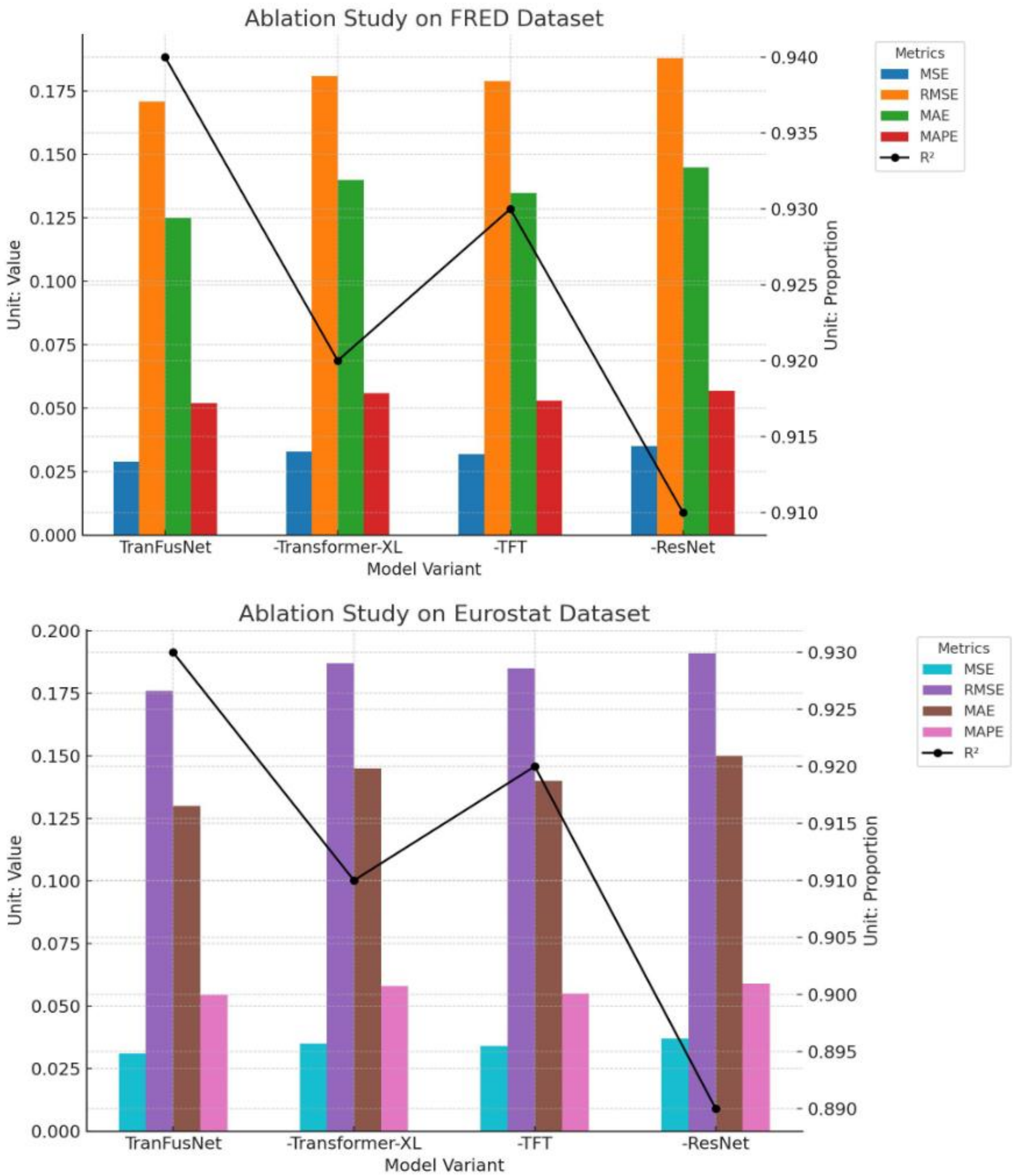

Figure 6. Single-Module Ablation Experiment Results for MS-RL DeepSeek on Yahoo Finance and FRED datasets.

The experimental results indicate that each module in TranFusNet plays an irreplaceable and important role. This further proves the indispensability of each module in TranFusNet. Through these ablation experiment results, it can be seen that each module plays a crucial role in the overall performance of the model, and the work of each module effectively improves the prediction accuracy and stability of the model. Figure 6 clearly illustrates these performance changes, providing intuitive support for further understanding the interdependence between modules. We will conduct multi module ablation experiments to explore the interdependence between these modules and their impact on the overall performance of the model.This analysis will further reveal the rationality of the model design and provide guidance for further optimization of the model.Table4 shows the experimental results after ablating multiple modules:

Table 4. Multi-Module Ablation Experiment Results.

Model	Dataset	MSE	RMSE	MAE	R^2	MAPE
W/o Transformer-XL & TFT	FRED	0.045	0.213	0.160	0.85	6.80%
	Eurostat	0.048	0.219	0.160	0.84	6.40%
W/o Transformer-XL & ResNet	FRED	0.051	0.226	0.170	0.82	7.10%
	Eurostat	0.053	0.230	0.170	0.80	6.60%
W/o TFT & ResNet	FRED	0.053	0.230	0.175	0.80	7.30%
	Eurostat	0.056	0.236	0.175	0.78	6.80%
W/o All	FRED	0.057	0.239	0.185	0.78	7.50%
	Eurostat	0.061	0.246	0.185	0.75	7.10%

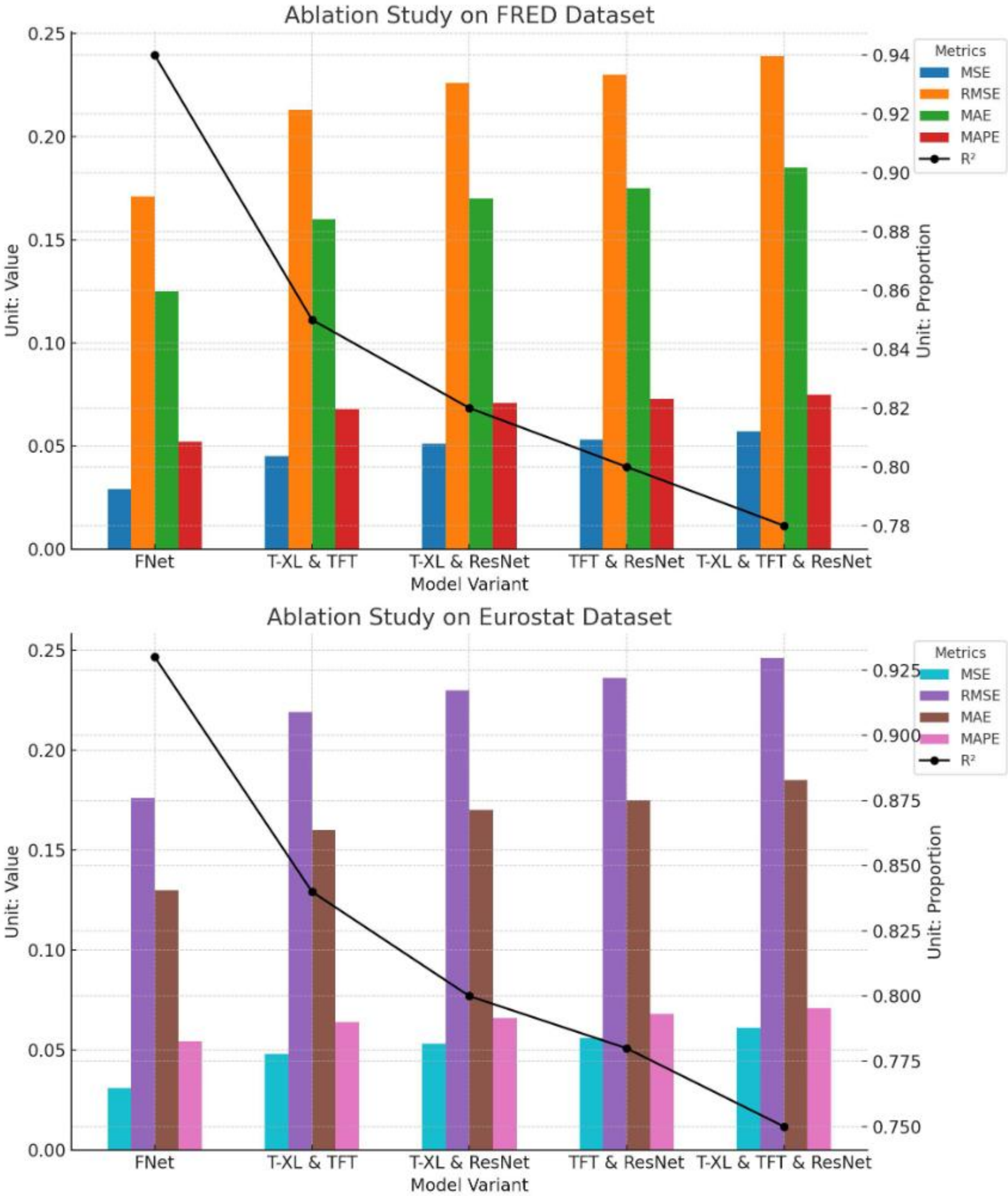

Figure 7. Changes in TranFusNet indicators after ablation of multiple modules.

From Figure 7, it can be seen that when multiple modules are removed, the error of the model significantly increases, and the values of indicators such as MSE, RMSE, MAE gradually increase, while the decline in performance indicators such as R^2 and MAPE gradually increases. After removing Transformer XL and TFT, the MSE and RMSE of the model increased by approximately 0.016 and 0.042, respectively, indicating that these two modules are crucial in capturing long-term dependencies and global information of time series data. After removing Transformer XL and ResNet, the error increased even more, especially in the decrease of R^2, indicating that the ResNet module plays a crucial role in handling deep level feature learning and nonlinear relationships. Removing the combination of TFT and ResNet also resulted in a similar performance decline, with MSE and RMSE increasing by 0.024 and R 2 further decreasing to 0.78, further validating the importance of these two modules in modeling multidimensional time series data. Finally, removing all modules resulted in the greatest performance degradation, with MSE and RMSE increasing by approximately 0.028 and 0.057, R^2 decreasing to 0.75, and MAPE increasing to 7.10%, respectively. This indicates that the predictive ability of the model is significantly reduced when the synergistic effect of these three modules is lacking, further verifying the complementary role and indispensability of these three modules in TranFusNet.

Ablation results in a single Transfusnet It is clarified that the modules are nested to each other and that the removal of one module significantly affects the prediction ability and stability of the model. Each module is important for overall performance. These results are useful in dealing with complex economic data Transfusnet Proving its superiority and proved its strong predictive ability and stability.

5. Conclusion and Discussion

TranFusNet is an innovative deep learning model designed to tackle complex challenges in economic forecasting. It combines insights from Transformer-XL, TFT, and ResNet to enhance the prediction of key economic indicators, such as economic growth and unemployment rates. We tested the model using several well-established public datasets, including those from the European Statistical Office, and found that TranFusNet performed exceptionally well in predicting economic data. This combination of powerful models allows TranFusNet to efficiently handle the complexities of economic forecasting and deliver accurate results across different datasets.

Through ablation experiments Transfusnet The significance of each module is verified. Deleting modules greatly affects the performance of the model and usually increases from 5 to 20%. These results are important to adjust the different modules to increase the overall predictive ability, and prove the value of the deep learning model for application to economic forecasting.

The Transfusnet experience is great, but there is still room for improvement. By optimizing the model structure, improving the interaction between modules, and introducing new feature engineering methods, prediction accuracy can be improved. With the increase in data volume and functional dimensions, Transfusnet may require additional IT resources to support training and applications. Therefore, in future research, we can focus on improving computational efficiency and model scalability to address the challenges of large-scale and diverse economic data.

Acknowledgements

This article received no financial or funding support.

Conflicts of Interest

The authors confirm that there are no conflicts of interest.

References

[1] Alaminos, D., Salas, M.B.and Fernández-Gámez, M.A. Quantum computing and deep learning methods for GDP growth forecasting. Computational Economics, 2022, 59(2), 803–829.

[2] Zheng, Y., Xu, Z.and Xiao, A. Deep learning in economics: a systematic and critical review. Artificial Intelligence Review, 2023, 56(9), 9497–9539.

[3] Buckmann, M., Joseph, A.and Robertson, H. An interpretable machine learning workflow with an application to economic forecasting. International Journal of Central Banking, 2023, 19(4), 449–552.

[4] Barrera-Animas, A.Y., Oyedele, L.O., Bilal, M., Akinosho, T.D., Delgado, J.M.D.and Akanbi, L.A. Rainfall prediction: a comparative analysis of modern machine learning algorithms for time-series forecasting. Machine Learning with Applications, 2022, 7, 100204.

[5] Shiri, F.M., Perumal, T., Mustapha, N.and Mohamed, R. A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU. arXiv preprint arXiv:230517473, 2023.

[6] Sirisha, U.M., Belavagi, M.C.and Attigeri, G. Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: a comparison. IEEE Access, 2022, 10, 124715–124727.

[7] Ozden, E.and Guleryuz, D. Optimized machine learning algorithms for investigating the relationship between economic development and human capital. Computational Economics, 2022, 60(1), 347–373.

[8] Liu, L., Chen, C.and Wang, B. Predicting financial crises with machine learning methods. Journal of Forecasting, 2022, 41(5), 871–910.

[9] ArunKumar, K., Kalaga, D.V., Kumar, C.M.S., Kawaji, M.and Brenza, T.M. Comparative analysis of gated recurrent units (GRU), long short-term memory (LSTM), autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA) for forecasting COVID-19 trends. Alexandria Engineering Journal, 2022, 61(10), 7585–7603.

[10] Masini, R.P., Medeiros, M.C.and Mendes, E.F. Machine learning advances for time series forecasting. Journal of Economic Surveys, 2023, 37(1), 76–111.

[11] Rane, N., Choudhary, S.P.and Rane, J. Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions. Studies in Medical and Health Sciences, 2024, 1(2), 18–41.

[12] Keisler, J., Talbi, E.G., Claudel, S.and Cabriel, G. An algorithmic framework for the optimization of deep neural networks architectures and hyperparameters. Journal of Machine Learning Research, 2024, 25(201), 1–33.

[13] Khan, A., Aslam, S., Aurangzeb, K., Alhussein, M.and Javaid, N. Multiscale modeling in smart cities: a survey on applications, current trends, and challenges. Sustainable Cities and Society, 2022, 78, 103517.

[14] Jin, Y., Hou, L.and Chen, Y. A time series transformer based method for the rotating machinery fault diagnosis. Neurocomputing, 2022, 494, 379–395.

[15] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J.and Xiong, H. Informer: beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35, 11106–11115.

[16] Wu, H., Xu, J., Wang, J.and Long, M. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting.Advances in Neural Information Processing Systems, 2021, 34, 22419–22430.

[17] Li, W.and Law, K.E. Deep learning models for time series forecasting: a review. IEEE Access, 2024.

[18] Vuong, P.H., Dat, T.T., Mai, T.K.and Uyen, P.H. Stock-price forecasting based on XGBoost and LSTM. Computer

Systems Science and Engineering, 2022, 40(1).

[19] Li, J., Chen, W., Zhou, Z., Yang, J.and Zeng, D. DeepAR-attention probabilistic prediction for stock price series. Neural Computing and Applications, 2024, 36(25), 15389–15406.

[20] Bai, T.and Tahmasebi, P. Attention-based LSTM-FCN for earthquake detection and location. Geophysical Journal International, 2022, 228(3), 1568–1576.

[21] Zhang, X., Yang, S., Duan, L., Lang, Z., Shi, Z.and Sun, L. Transformer-XL with graph neural network for source code summarization. In: IEEE International Conference on Systems, Man, and Cybernetics, 2021, 3436–3441.

[22] Lim, B., Arık, S.O., Loeff, N.and Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 2021, 37(4), 1748–1764.

[23] Shafiq, M.and Gu, Z. Deep residual learning for image recognition: a survey. Applied Sciences, 2022, 12(18), 8972.

[24] Ellingsen, J., Larsen, V.H.and Thorsrud, L.A. News media versus FRED-MD for macroeconomic forecasting.Journal of Applied Econometrics, 2022, 37(1), 63–81.

[25] Zisopoulos, F.K., Schraven, D.F.and de Jong, M. How robust is the circular economy in Europe? an ascendency analysis with Eurostat data between 2010 and 2018. Resources, Conservation and Recycling, 2022, 178, 106032.

[26] Tatachar, A.V. Comparative assessment of regression models based on model evaluation metrics. International Research Journal of Engineering and Technology, 2021, 8(9), 2395–0056.

[27] Chicco, D., Warrens, M.J.and Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 2021, 7, e623.

[28] Huang, X., Tang, J.and Shen, Y. Long time series of ocean wave prediction based on PatchTST model. Ocean Engineering, 2024, 301, 117572.

[29] Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J.and Long, M. TimesNet: temporal 2D-variation modeling for general time series analysis. arXiv preprint arXiv:221002186, 2022.

[30] Wang, G., Liao, Y., Guo, L., Geng, J.and Ma, X. DLinear photovoltaic power generation forecasting based on reversible instance normalization.In: IEEE Data Driven Control and Learning Systems Conference, 2023, 990–995.

[31] Das, A., Kong, W., Leach, A., Mathur, S., Sen, R.and Yu, R. Long-term forecasting with TiDE: time-series dense encoder. arXiv preprint arXiv:230408424, 2023.

[32] Rasul, K., Ashok, A., Williams, A.R., Khorasani, A., Adamopoulos, G.and Bhagwatkar, R. Lag-LLaMA: towards foundation models for time series forecasting. In: Robustness of Few-shot and Zero-shot Learning in Large Foundation Models, 2023.