# A Multi-Task GRU-Attention Model for Predicting Enterprise Investment and Financing Behavior from Multi-Source Economic Data

Lei Gu, Tao Liu
Xi’an University of Architecture & Technology Huaqing College, Shanxi, 710043 , China
E/mail: leigu198@163.com, tracey781225@126.com

**Keywords:** intelligent prediction model, investment and financing behavior, multi source heterogeneous data, deep learning algorithms

**Received:** August 6, 2025

*Accurately predicting corporate investment and financing behavior is crucial for improving financial intelligence and capital allocation efficiency. This article proposes an economic data-driven multi-task deep prediction model that integrates Gated Recurrent Unit (GRU) networks with a multi-head attention mechanism to process multi-source heterogeneous economic variables, including macroeconomic indicators, corporate financial data, and market sentiment factors, under a unified structure. The model constructs multivariate time-series samples through sliding windows and employs a dual-output architecture to perform regression prediction of financing intensity and classification recognition of behavioral states into three classes (expansion, wait-and-see, contraction). To enhance responsiveness to behavioral transition patterns, a feature cross-attention mechanism and a joint loss function optimization strategy are introduced, improving nonlinear behavior learning capability and generalization robustness. Based on empirical data from 232 A-share listed companies, covering 12,840 training samples over the past decade, the experimental results showed that the model achieved a coefficient of determination ($R^2$) of 0.862 in the financing prediction subtask, an accuracy of 88.3% in the classification task, and a Macro-F1 value of 0.841. Compared with baseline machine learning methods including Support Vector Regression (SVR), Random Forest (RF), and Multi-Layer Perceptron (MLP), the model demonstrated superior error control and trend fitting ability. Overall, the model exhibits high prediction accuracy, stability, and industry adaptability, providing a feasible technical path and empirical basis for building a data-driven intelligent investment and financing analysis system for enterprises.*

*Povzetek: Razvit je večopravilni BiGRU-model z večglavo pozornostjo za napovedovanje intenzivnosti ter klasifikacije investicijsko-finančnega vedenja podjetij iz večizvornih ekonomskih podatkov. Preverjen je na 12.840 vzorcih.*

## 1 Introduction

With the increasingly active business activities and uncertain business environment, investment and financing have become the most important means of strategic change and allocation for enterprises, and are facing unprecedented pressure. In the past, corporate investment and financing activities mainly relied on the intuition of financial experts and the evaluation of static reports. Nowadays, the emergence of a large amount of structural and non-structural economic information has made it possible to establish intelligent decision-making mechanisms driven by data. The important reason behind this is the extreme integration of artificial intelligence and big data analysis technology, which has led corporate financial management activities to a higher level of intelligence.

Enterprises' investment and financing decision-making actions will continue to be dynamic, and fundamentally, it is the result of a series of multi-level, multi cycle, and multi factor interactions. This action path depends on the interaction between internal factors (such as business operations and asset liability ratios) and external macroeconomic factors (such as interest rates, government intervention, industrial cycles, etc.), exhibiting strong irrational factors and periodic jumps. Therefore, how to find dominant indicators from massive, multi cycle, and multi factor economic information, accurately capture action patterns and future development trends is a major technical challenge in the field of intelligent financial model research [4].

It should be noted that, in contrast, traditional statistical modeling is very suitable for linear relationship assumptions, while today's machine learning techniques and deep learning methods can provide useful information for predictive analysis of complex and noise intensive data. Especially for time series prediction and behavior recognition, it has great advantages [5]. By utilizing economic data from various sources to establish a predictive model that can grasp structural changes and understand development trends, enterprises can have better foresight and countermeasures when facing market fluctuations or financial pressures.

This article intends to design a prediction model for investment and financing activities of listed companies that integrates economic data feature analysis and intelligent algorithm optimization, taking into account the construction of the economic theory model framework, the diversity and real-time nature of input data, and the commercial understanding of industrial financial management based on algorithm output results. The goal is to pursue the "interpretability" and "predictability" of prediction. Using data from several representative listed companies to test the model and evaluate its predictive accuracy, robustness, and adaptability, this article concludes by describing and explaining the potential application value of the model in industrial financial management.

Specifically, this study aims to address the following research questions:

(1) Can a BiGRU with multi-head attention achieve higher accuracy than traditional machine learning models (e.g., SVR, RF, MLP) in predicting financing intensity?

(2) Can a multi-task architecture jointly modeling regression and classification tasks improve robustness and interpretability in forecasting enterprise investment and financing behaviors?

(3) How does the proposed model perform across heterogeneous economic data sources in terms of adaptability and stability?

The structure of this article is as follows: Chapter 2 provides an overview of the current research status and basic concepts on this topic both domestically and internationally; Chapter 3 introduces the modeling process and key parameters of the constructed model; Then Chapter 4 verifies the effectiveness and economic explanatory power of the predictive function of the model proposed in this article through examples; Chapter 5 is an analysis of how the established model can be applied to actual business scenarios, and will also elaborate on possible issues that may arise; Chapter 6 provides a comprehensive overview of the entire text and outlines future development trends.

## 2 Related work

Due to the increasingly complex and data-driven decision-making nature of corporate investment and financing behavior, accurately predicting changes in corporate investment or borrowing has always been a topic of sustained interest for scholars and practitioners. Although research methods continue to develop, the dynamic evolution process of high-dimensional nonlinear data characteristics and economic variable interactions is complex and may have multiple driving factors, making investment and lending predictions still difficult [6]. Traditional regression, time series, and other methods perform well in terms of interpretability, but they are difficult to play a greater role in irregular attributes, multi period changes, and occasional risk factors [7].

With the deepening of understanding of the business activities of listed companies, scholars have begun to use artificial intelligence to enhance their judgment ability in investment and financing. In recent years, artificial intelligence modeling methods centered around neural networks have been widely applied in financial analysis of listed companies, stock market forecasting, credit rating, and more. For example, Shahrour et al. (2023) [8] established a stock market price prediction strategy based on deep neural networks, which enhanced the response speed to the stock market; Yao et al. (2022) [9] overcame the problem of noise impact in financial market sequence data by adding LSTM to the neural network model of the data and applying algorithms to optimize the model; The hybrid model designed by Chandok et al. (2024) [10] achieved higher robustness and universality in enterprise bankruptcy prediction tasks by combining deep neural network models. These scholars' research results indicate that intelligent modeling methods based on deep learning have predictive ability in analyzing the financial activities of listed companies, as well as good application scenarios and scalability.

Against the backdrop of the emergence of numerous economic data in big data, research on enterprise behavior patterns based on big data has become increasingly active. Scholars have incorporated various types of economic data, such as GDP growth rate, interest rates, and industry activity index, into models to expand the predictive dimensions and overall system of the model [11]. Tang and Wei (2023) [12] used XGBoost and SHAP algorithms to discover the key driving factors of a company's digital transformation, which can provide visual explanations for related investment and financing behaviors. Pei et al. (2023) [13] established an interpretable prediction framework from the perspective of data features, significantly improving the interpretability of traditional "black box" models. These studies have formed a theoretical shift from focusing on improving the "accuracy" of prediction results to seeking the "interpretability and credibility" of prediction principles.

After continuously understanding relevant issues, multi-path fusion modeling and hybrid intelligent methods have gradually become mainstream. Wu (2022) [14] established a grey prediction model based on fuzzy thinking to simulate the nonlinear and uncertain boundaries within the economic system; Yang (2024) proposed a cross-border e-commerce supply chain demand forecasting model based on deep neural networks, focusing on big data-driven business decision-making; Kartbayev et al. (2022) proposed an intelligent comprehensive evaluation model for investment projects that considers multiple input factors, which can enable enterprises to obtain better investment and financing advice from the wave of digital transformation. From this, it can be seen that using a single route alone cannot solve the multidimensional and multi-directional problems faced in the process of enterprise behavior prediction. The use of deep learning, optimization algorithms, attention, and feature selection ensemble methods to construct an integrated architecture has become an effective way to break through the bottleneck of prediction.

However, existing research also has certain limitations. On the one hand, most intelligent prediction methods still heavily rely on the integrity and representativeness of training data, making it difficult to maintain stable performance in scenarios with large industry spans and strong data distribution heterogeneity [17]; On the other hand, many studies still focus on financial markets such as stocks and bonds, lacking systematic modeling and evaluation of micro level business operations, especially actual investment and financing behaviors [18]. In addition, key issues such as the engineering feasibility of model deployment, the practical path of data fusion, and the interpretation mechanism of prediction results still need to be further deepened.

In summary, the current academic community has accumulated rich achievements in the field of investment and financing prediction, from statistical modeling to deep neural networks, from univariate processing to multi-source heterogeneous data fusion, with significant technological evolution. However, how to construct intelligent models with both predictive and explanatory capabilities in complex economic environments, and how to enhance the model's perception and adaptability to fine-grained changes in corporate investment and financing behavior, are still the core issues of concern in this study. To highlight the performance gap with existing methods, Table 1 summarizes representative studies, including their methods, datasets, evaluation metrics, and limitations, compared with the approach proposed in this paper.

Table 1：Summary of representative related studies

| Author(s) | Method | Dataset | Metrics | Limitations |
|---|---|---|---|---|
| Shahrour et al. (2023) | Deep Neural Network for stock prediction | Stock market data | Accuracy | Sensitive to noise; limited interpretability |
| Yao et al. (2022) | LSTM with optimization | Financial time series | RMSE, MAE | Noise sensitivity; relatively slow training |
| Chandok et al. (2024) | Hybrid Deep Neural Network for bankruptcy prediction | Enterprise financial data | Accuracy, Robustness | High complexity; generalization limitations |
| Tang & Wei (2023) | XGBoost with SHAP for digital transformation | Enterprise digitalization data | Visualization, Interpretability | Focuses on feature analysis rather than complete prediction |
| This study | Multi-task BiGRU + Multi-head Attention | 232 A-share listed firms (2015–2023) | $R^2 = 0.862$, Accuracy = 88.3%, Macro-F1 = 0.841 | Higher computational cost, but improved robustness and adaptability |

Therefore, this article will focus on the dual axis path of "economic data-driven+intelligent prediction algorithm", propose an enterprise intelligent investment and financing prediction model that integrates multi-source data processing, structured modeling, and deep learning optimization, and empirically verify its performance and application value.

# 3 Modeling ideas and indicator system construction for predicting corporate investment and financing behavior

When constructing a practical, explanatory, and forward-looking enterprise investment and financing behavior prediction system, the starting point of modeling work should be based on the triple logic of "economic variable driving behavior mechanism mapping intelligent algorithm expression". Unlike traditional financial modeling that focuses on single indicator fitting, the predictive model proposed in this paper not only requires the ability to "fit" historical data of enterprises, but also emphasizes the ability to extract trend driven signals from macroeconomic fluctuations, capture structural change arteries from enterprise financial status, and recognize and predict behavioral states in multivariate cross analysis. The investment and financing behavior of enterprises exhibits strong nonlinear and cyclical characteristics, often driven by macro cyclical disturbances, changes in liquidity preferences, distorted industry expectations, and other factors. The interaction relationship between multiple heterogeneous input variables frequently leads to the failure of classical linear regression and fixed coefficient statistical models in actual prediction. To address this challenge, this article introduces a multi-source data-driven feature construction strategy and a non-linear prediction algorithm with strong deep expression ability, attempting to establish an intelligent perception and prediction framework for enterprise investment and financing behavior in the link of variable extraction behavior modeling output mapping. (As shown in Figure 1)

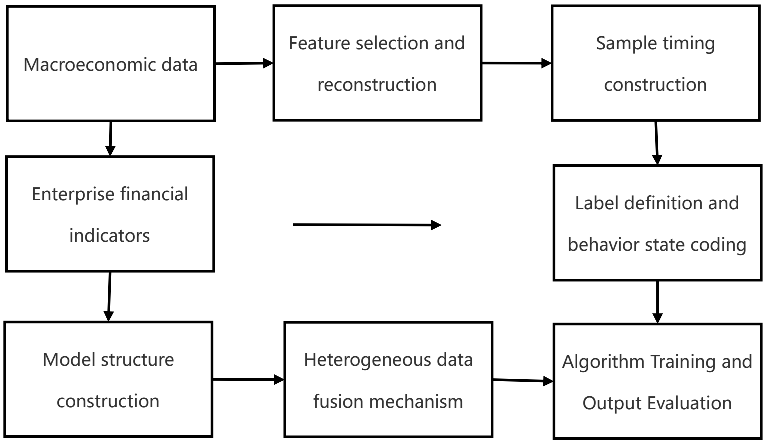


Figure 1 : Overall construction process of enterprise investment and financing behavior prediction Model

## 3.1 Multidimensional feature analysis and variable selection of economic data

In the modeling of enterprise investment and financing forecasting, economic data is not only used as background information, but also an inherent driving variable in the behavioral evolution path. Therefore, scientifically analyzing the multidimensional structure of economic data and constructing a variable set with predictive ability is the fundamental guarantee for model performance. This article divides economic characteristics into three sub domains: macro level indicators, market factors, and policy signals. In terms of variable selection strategy, the Joint Information Gain (JIG) mechanism and the Temporal Stability Index (TSI) function are used to screen and rank candidate variables. The specific formula is as follows:

$$JIG(x_i,y)=H(y)-H(y|\ x_i) \tag{1}$$

where H(·)denotes Shannon entropy, H(x)is the marginal entropy of variable x, H(y)is the marginal entropy of the target behavior label y, and H(x,y)is their joint entropy. A higher JIG value indicates that the candidate variable provides more information for predicting investment and financing behaviors. At the same time, to avoid introducing pseudo variables with high-frequency and violent fluctuations but no stable structure, a time series stability index is further introduced for dynamic evaluation:

$$TSI(x_i)=1-\frac{Var(\Delta x_i)}{Var(x_i)} \tag{2}$$

In the formula, where Var(x)is the variance of the original variable sequence x, and Var(Δx)is the variance of its first-order difference $\Delta x_t = x_{t-1}$ .When TSI approaches 1, the sequence exhibits strong stationarity and smooth trend continuity; when TSI is close to 0, the sequence fluctuates violently and is less suitable for stable time-series modeling.

After preliminary evaluation and empirical screening, this article finally included 12 types of variables, namely GDP growth rate, M2 money supply, benchmark interest rate, manufacturing prosperity index, credit expansion index, CPI, PPI, fixed-asset investment index, exchange rate index, unemployment rate, government expenditure index, and stock market composite index as the core input features of the prediction model, ensuring that the model has sufficient sensitivity and responsiveness to behavioral trends. All variables are resampled in time series at a uniform frequency and processed through normalization and scale compression methods, providing an input basis for the training and fusion of subsequent model structural layers.

## 3.2 Structured modeling logic and prediction objectives for investment and financing behavior

The investment and financing behavior of enterprises is essentially a dynamic feedback system driven by both internal and external factors. Its state not only changes with macroeconomic fluctuations, but is also closely coupled with the enterprise's own financial structure, strategic cycle, and market expectations. To achieve high-quality prediction, it is necessary to construct a structured modeling framework that balances time dependence and non-linear expression ability.

This article models corporate investment and financing behavior as a temporal response function, with historical economic and financial feature sequences as inputs and future financing or investment behavior labels as outputs, forming an input-output structure. The specific expression is as follows:

$$\hat{y}_t = F(x_{t-k}, x_{t-k+1}, \ldots, x_t; \Theta) \tag{3}$$

Among them, $\hat{y}$ represents the strength or category label of the investment and financing behavior predicted by the model; F(·) is the nonlinear function to be trained; $x_t$ is the input feature vector at time t, which includes macro factors, financial indicators, and lead variables; $\Theta$ is the set of model parameters; K is the length of the historical window.

The prediction objectives are divided into two types of tasks: one is numerical regression prediction, which is used to estimate the financing amount of the enterprise in the future range (such as the scale of new debt or equity financing); The second is the multi class prediction task, which identifies the behavior status of enterprises in the current cycle (such as active financing, wait-and-see, investment contraction, etc.). To support the dual task learning structure, a joint loss function is constructed as follows:

$$L_{total} = \lambda_1 \cdot L_{reg} + \lambda_2 \cdot L_{cls} \tag{4}$$

Among them, $L_{reg}$ is the mean square error loss function (used to fit the financing amount), $L_{cls}$ is the cross-entropy loss function (used for behavior classification), and $\lambda_1$ , $\lambda_2$ is the weight adjustment coefficient, reflecting the balance of task importance, and was empirically determined through grid search within [0.1, 0.9] on the validation set.

The core advantage of this structure is that it not only captures the numerical fluctuations and temporal structure of input variables, but also combines the behavioral logic of label space to achieve a dual output of "quantitative estimation+behavioral recognition", thus meeting the diversified application needs of enterprise financial systems for prediction results.

## 3.3 Sample data preprocessing strategy and feature engineering design

The prediction of corporate investment and financing behavior relies on the dynamic input of time-series data, and the raw data often has problems such as dimensional heterogeneity, inconsistent time frequency, and a large number of missing outliers. To ensure the stability and accuracy of model training, this article conducts systematic preprocessing and feature engineering design before data modeling.

Firstly, to address the issue of time alignment between macro and micro data, a sliding window mechanism is adopted to construct sequence samples. If the sliding window length is set to k and the step size is 1, the i-th sample input sequence is constructed as follows:

$$X^{(i)} = [x_i, x_{i+1}, \ldots, x_{i+k-1}] \tag{5}$$

Among them, $x_j$ is the feature vector of the jth day, and the corresponding output label is $y_i + k$, forming the training data pair $(X^{(i)} y^{(i)})$.

Secondly, to address missing and extreme values in the data, this article adopts a combination repair strategy. Forward padding is used for macro data, while linear interpolation correction is applied to quarterly financial data of enterprises based on year-on-year change rate. Outlier detection is performed by setting a threshold of ±3σ; values beyond this range were winsorized (capped at boundary values) rather than removed, to preserve data continuity.

In terms of feature construction, considering the trend inertia and cyclical fluctuations of investment and financing behavior, this paper introduces derivative features based on the original variables. The most commonly used treatments include:

First order difference (capturing trend changes):

$$\Delta x_t = x_t x_{t-1} \tag{6}$$

Rolling average (smooth local fluctuations):

$$MA_k(x_t) = \frac{1}{k}\sum_{i=0}^{k=1} x_t - j \tag{7}$$

The above transformation can significantly enhance the sensitivity of the model to trend mutations and short-term behavior. All features undergo Z-score standardization before input:

$$x'_t = \frac{x_t - \mu}{\sigma} \tag{8}$$

Among them, $\mu$ is the sample mean of variable x, and $\sigma$ is the sample standard deviation.

In summary, this section has completed the full chain design of the data preprocessing process around four levels: "standardized sample generation - cleaning and completion - derivative variable construction - scale unification", providing a stable, clean, and structured data input foundation for subsequent deep modeling modules.

## 3.4 Model architecture design and core technology selection

When building a predictive system for enterprise investment and financing behavior, the selection of model architecture needs to take into account three core elements: the high complexity of variable dimensions, the temporal dependence of input sequences, and the diversity of output targets (including continuous values and categorical labels). Therefore, this article adopts a deep learning model with a multi-layer nested structure as the main architecture, and combines attention mechanism and residual connection technology to improve its expression and generalization ability for time-series financial behavior data.

The overall model structure consists of three main sub modules: input encoding layer, feature extraction layer, and output prediction layer. Firstly, the input encoding layer maps different types of variables (such as macro indicators, corporate financial characteristics, etc.) to a unified dimensional representation through a multi-head embedding network. If the input at time t is the feature vector $x_t \in R^d$, then the embedding transformation is:

$$z_t = W_e x_t + b_e \tag{9}$$

Among them,$W_e \in R^{d' \times d}$ is the weight matrix,$b_e \in R^{d'}$ is the bias term, and $z_t$ is the high-order expression after embedding.

Subsequently, the feature extraction layer adopts a bidirectional recurrent structure (Bi GRU) with gating mechanism to capture the temporal dependency patterns of forward and backward. The bidirectional structure can extract short-term fluctuations and long-term trends in parallel. The calculation form is as follows:

$$h_t = \mathrm{GRU}_{\rightarrow}(z_t) \,\|\, \mathrm{GRU}_{\leftarrow}(z_t) \quad (10)$$

Among them, $h_t$ represents the hidden state vector at time t, and ‖ represents the vector concatenation operation. To enhance the model's ability to pay attention to critical moments, an attention mechanism is introduced after the output of the recurrent network, and its weight distribution is defined by the following equation:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T} \exp(e_j)} \quad (11)$$

$$e_t = v^T \tanh(W_a ht + b_a) \quad (12)$$

Among them, $\alpha_t$ is the attention weight at time t, v 、$W_a$、$b_a$ are trainable parameters. Ultimately, the output layer is divided into two branches based on task types: one is the regression prediction branch, which is used to estimate financing/investment amounts; The second is the classification branch, which is used to determine the current behavior status of the enterprise (such as "financing expansion" or "conservative wait-and-see"). The dual output structure improves overall modeling efficiency by sharing underlying features. This architecture has good scalability and can flexibly adjust the number of layers and parameter configuration according to data size and label complexity, making it an effective technical path for achieving high-precision and multi-objective enterprise behavior prediction. In the implementation, the BiGRU consists of 2 stacked layers with hidden size 128, followed by a multi-head attention module with 4 heads.

## 3.5 Construction of fusion mechanism for multi source heterogeneous economic data

The formation of corporate investment and financing behavior is driven by information from different dimensions, including macroeconomic environment, financial market dynamics, corporate financial status, and external policy signals. These pieces of information exhibit heterogeneous characteristics in data structure, such as inconsistent frequency, significant differences in value distribution, and complex dimensional types. In order to achieve behavior modeling in a unified input space, it is necessary to design an effective data fusion mechanism that aligns representations, compresses structures, and integrates information for multi-source heterogeneous data.

This article adopts a fusion mechanism of channel embedding weight fusion cross attention structure. Firstly, four main data sources are set: macro variable sequence $M = \{m_t\}$, financial indicator sequence $F = \{f_t\}$, market sentiment factor $S = \{s_t\}$, and policy signal variable $P = \{p_t\}$ . Obtain unified dimensional representations through independent linear embedding networks:

$$z_t^m = W^m m_t, \quad z_t^f = W^f f_t, \quad z_t^s = W^s s_t, \quad z_t^p = W^p p_t \quad (13)$$

Among them, $W^m, W^f, W^s, W^p$ is the weight matrix of each channel, and the output is a feature representation of the same dimension. Next, a weighted fusion layer is constructed, introducing a learnable weight parameter of $\alpha_i$, and integrating multi-source information through weighted summation:

$$z_t^{fusion} = \sum_{i \in \{m,f,s,p\}} \alpha_i z_t^i \quad (14)$$

Among them, $q_i$ and $k_i$ respectively represent the query and key vectors from different channels, and $\beta_{i,j}$ represents the degree of attention that channel i pays to channel j information. This fusion mechanism enhances the model's perception of potential coupling relationships between complex data while ensuring the preservation of different data substructures, significantly improving the robustness of investment and financing behavior prediction to structural changes and time mismatches.

## 3.6 Implementation and optimization strategy of intelligent prediction algorithm

To improve the accuracy and stability of enterprise investment and financing behavior prediction, this study designs a multi-task deep neural network guided by attention mechanism based on the fusion of multi-source heterogeneous data, achieving joint prediction of financing intensity regression and behavior state classification. The overall algorithm framework combines Gated Recurrent Networks (GRU), Multi Head Attention, and multi task loss function optimization mechanisms, balancing temporal modeling capabilities and structural interpretability.

In the main structure of the model, the input is the processed sample sequence $X = [x_{t-k+1}, \ldots, x_t]$ , and each $x_t \in R^d$ represents the multidimensional feature vector at time t. By using GRU units for temporal modeling, the state update is expressed as:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (15)$$

Among them, $h_t$ is the hidden state at the current time, $\tilde{h}_t$ is the candidate state, $z_t$ is the update gate, $\odot$ represents the Hadamard product. This mechanism can dynamically regulate the degree of influence of historical information on current predictions. Introducing attention mechanism after hidden state output to enhance the model's ability to focus on key time segments. To improve the training efficiency and generalization ability of the model, this paper introduces early stopping mechanism, gradient clipping, and learning rate dynamic adjustment (LR Scheduler) strategy.In terms of optimizer selection, Adam algorithm is adopted and the initial learning rate is set to 0.001. Dropout layers (rate = 0.3) and L2 regularization were also applied to mitigate overfitting. This intelligent prediction algorithm framework combines interpretability, scalability, and computational efficiency, and can effectively adapt to the modeling needs of enterprise investment and financing behavior in complex economic scenarios, laying a technical foundation for subsequent empirical analysis and deployment promotion.

# 4 Empirical research and model evaluation

## 4.1 Empirical data sources and sample enterprise selection criteria

In order to verify the effectiveness and adaptability of the constructed enterprise intelligent investment and financing behavior prediction model, this paper conducts empirical research based on heterogeneous data from multiple sources, including publicly available macroeconomic data from the National Bureau of Statistics, financial statement data from WIND financial terminal enterprises, policy and policy text databases, and investor sentiment index data. The overall data time span is from the first quarter of 2015 to the fourth quarter of 2023, covering the entire economic cycle fluctuations, including the COVID-19 shock phase and post pandemic recovery phase, which is conducive to capturing the dynamic response of corporate behavior in complex economic backgrounds.

In the selection of sample enterprises, this article sets the following three standards to ensure data quality and structural integrity: firstly, the industry to which the enterprise belongs should cover the four major sectors of manufacturing, information technology, healthcare, and energy, taking into account the heterogeneity of cyclical and growth industries; Secondly, the enterprise must have no significant missing financial statements during consecutive reporting periods, and the completeness of financial data must exceed 95%; Thirdly, the enterprise has engaged in at least two or more investment and financing activities (including issuance, loans, capital expenditures, mergers and acquisitions, etc.) during the sample period to ensure that the distribution of behavioral labels is representative. 232 A-share listed companies were ultimately selected as sample subjects.

After data processing, a total of 12840 training samples were constructed, covering approximately 4.3 million structured feature records. The industry distribution of sample enterprises is shown in Table 1. Here, Avg. Quarterly Data Points refers to the average number of valid feature records collected for each enterprise per quarter, while Number of Investment & Financing Events denotes the cumulative count of major financing or investment actions (such as bond issuance, loans, equity financing, and capital expenditures) recorded during the sample period:

Table 1：Industry distribution and data volume statistics of sample enterprises

| Industry Sector | Number of Sample Firms | Avg. Quarterly Data Points | Number of Investment & Financing Events |
|---|---|---|---|
| **Manufacturing** | 83 | 32 | 1,238 |
| **Information Technology** | 57 | 29 | 984 |
| **Healthcare** | 48 | 31 | 851 |
| **Energy & Resources** | 44 | 33 | 1,007 |
| **Total** | 232 | — | 4,080 |

The sample design has continuity in the time dimension, heterogeneity in the industry dimension, and balanced distribution of behavioral labels, providing a solid data foundation for subsequent model evaluation and comparative experiments. By implementing a unified data standard processing flow and cleaning mechanism, the scale consistency between input features and the expression stability of the labeling system are ensured, effectively reducing the interference of sample noise on the model training process. The model training and validation work will be carried out under the above data framework, and the details will be further elaborated in subsequent chapters.

## 4.2 Analysis of model prediction performance and error metrics

To comprehensively evaluate the performance of the proposed intelligent prediction model for enterprise investment and financing in both regression and classification tasks, this paper uses mean square error (MSE), mean absolute error (MAE), and coefficient of determination ($R^2$) to evaluate the performance of financing amount prediction. At the same time, accuracy and macro average F1 score (Macro-F1) are used to test the behavior state classification task. The experiment adopted a partitioning method of 70% training set, 15% validation set, and 15% test set, and completed model training and testing based on 12840 samples from 232 enterprises. To verify the stability of the model, the average of 5 independent experiments was taken for all results. (As shown in Table 2)

Table 2：Comparison of model prediction performance and error indicators

| Model Type | MSE | MAE | $R^2$ | Accuracy | Macro-F1 |
|---|---|---|---|---|---|
| **Proposed Model (BiGRU+Attn)** | 0.084 | 0.213 | 0.862 | 88.3% | 0.841 |
| **Random Forest Regression** | 0.129 | 0.294 | 0.731 | 81.5% | 0.771 |
| **Support Vector Regression** | 0.145 | 0.311 | 0.687 | 79.9% | 0.752 |
| **Multi-Layer Perceptron (MLP)** | 0.118 | 0.278 | 0.743 | 82.2% | 0.788 |
| **Logistic Regression (Classification Task)** | — | — | — | 76.4% | 0.705 |

From the results, the model shows strong numerical approximation ability in predicting investment and financing amounts, with an $R^2$ of 0.862, indicating that the model is effective in modeling the nonlinear relationship between input economic variables and corporate behavior variables, and has good fitting accuracy for macro disturbance sensitive areas. In contrast, traditional baseline models such as Random Forest ($R^2$ = 0.731) and Multi-Layer Perceptron ($R^2$ = 0.743) achieved significantly lower performance than the proposed BiGRU-Attention model ($R^2$ = 0.862), with Support Vector Regression further dropping to 0.687. In terms of classification tasks, the intelligent prediction model achieved an overall accuracy of 88.3% on the three classification labels, with Macro-F1 reaching 0.841, significantly better than logistic regression and shallow neural network models.The three behavior categories were relatively balanced (expansion: 34%, wait-and-see: 38%, contraction: 28%), and the confusion matrix (Figure X) shows that the model maintained robust classification performance across all classes without relying on majority-class bias. This indicates that the model not only has strong predictive ability, but also can effectively identify structural changes in the behavior status of enterprises.In practical terms, an $R^2$ improvement of around 0.1 means the model can reduce forecasting errors in financing amounts by tens of millions of RMB for large listed firms, while a 5–10% gain in classification accuracy translates into more reliable early warning of financing contractions or expansions, enabling enterprises and regulators to take pre-emptive actions.

## 4.3 Fitting and verification of trends in investment and financing behavior changes

The investment and financing behavior of enterprises is driven by multiple factors, showing obvious cyclical fluctuations and periodic jumps, and simple predictions are difficult to capture their trend trends. To verify the ability of the constructed model to capture trends in behavioral changes, representative samples were selected from both industry cross-section and enterprise longitudinal time dimensions for behavioral trajectory fitting testing. The focus of the experiment is to determine whether the model can accurately identify the rising and contracting stages of financing or investment behavior; The second is to test its ability to respond to trends in different economic cycles.

This study selected an enterprise in the manufacturing industry (designated as E-94) with significant fluctuations in capital expenditures, and analyzed its real financing intensity curve and model predicted values from Q1 2017 to Q4 2023, and compared them with the support vector regression (SVR) model. As shown in Figure 2, the model in this paper accurately predicted the upward or downward trend of financing at multiple keys turning points (such as the outbreak of the COVID-19 pandemic in Q1 2020 and the impact of raw material price increases in Q3 2021), and the fitted curve was close to the actual behavior trajectory, without any distortion such as excessive smoothing or severe shaking.

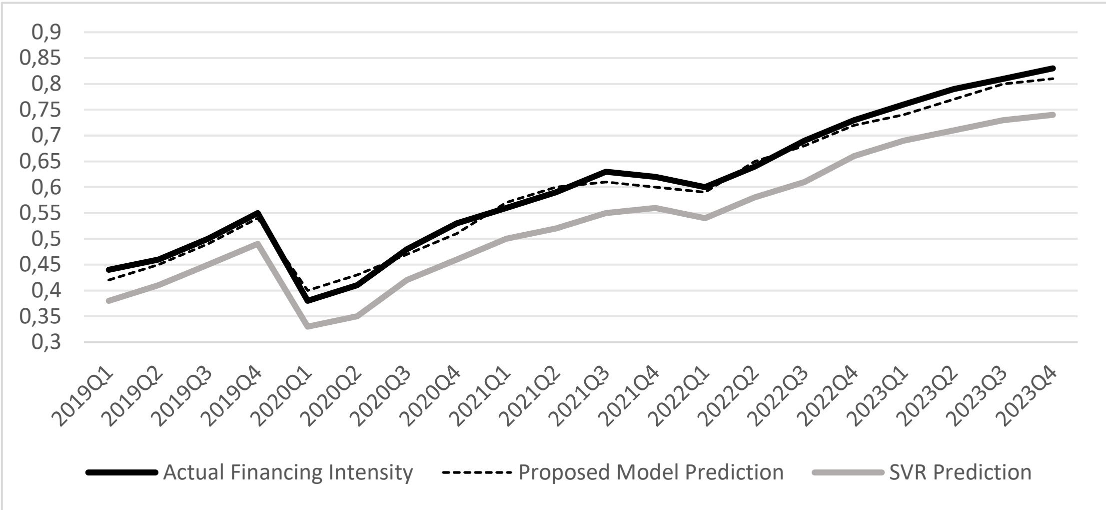


Figure 2: Fitting chart of actual and predicted trends in investment and financing behavior intensity

From the overall trend fitting results, this model not only maintains its accuracy advantage in static error indicators, but also performs well in dynamic trend judgment. Its causal identification ability is strong, and it has the ability to perceive behavior "transition points" in advance, indicating that the internal feature extraction and sequence modeling structure of the model has a certain descriptive power on the temporal evolution logic of enterprise behavior. Meanwhile, compared with traditional models, the model proposed in this paper is more robust in trend prediction, exhibiting high behavioral fit interpretability and stability, and has practical potential for promotion and application in dynamic enterprise management and risk warning.It should be noted that Figure 2 illustrates a representative case from the manufacturing sector, while aggregated results across all manufacturing firms (not shown here for brevity) confirmed the model's consistent ability to identify expansion and contraction phases in advance.

## 4.4 Comparison between intelligent prediction models and traditional methods

In order to systematically evaluate the performance advantages of the intelligent prediction model constructed in this article in predicting enterprise investment and financing behavior, this article selects three representative traditional methods for comparative analysis: linear regression (LR), support vector regression (SVR), and tree based random forest regression (RF). Build corresponding models in a unified data sample and feature space, and evaluate their performance on the same test set, focusing on indicators such as regression accuracy ($R^2$), classification accuracy (Accuracy), and overall error control ability (MAE, MSE).

In order to visually demonstrate the predictive performance of each model on the test set, Figure 3 compares the performance of linear regression (LR), support vector regression (SVR), random forest regression (RF), and the BiGRU+Attention model constructed in this paper under the three core indicators of $R^2$, MAE, and accuracy.To ensure consistency, the metrics for RF and MLP in Figure 3 have been aligned with those reported in Table 2; LR is shown for reference, while MLP and Logistic Regression results are only listed in Table 2 for completeness. It can be seen that the model in this article is significantly better than other methods in all three dimensions, especially in terms of regression accuracy and classification accuracy, indicating that it is more suitable for handling complex financial behavior prediction tasks driven by multi-source heterogeneous data

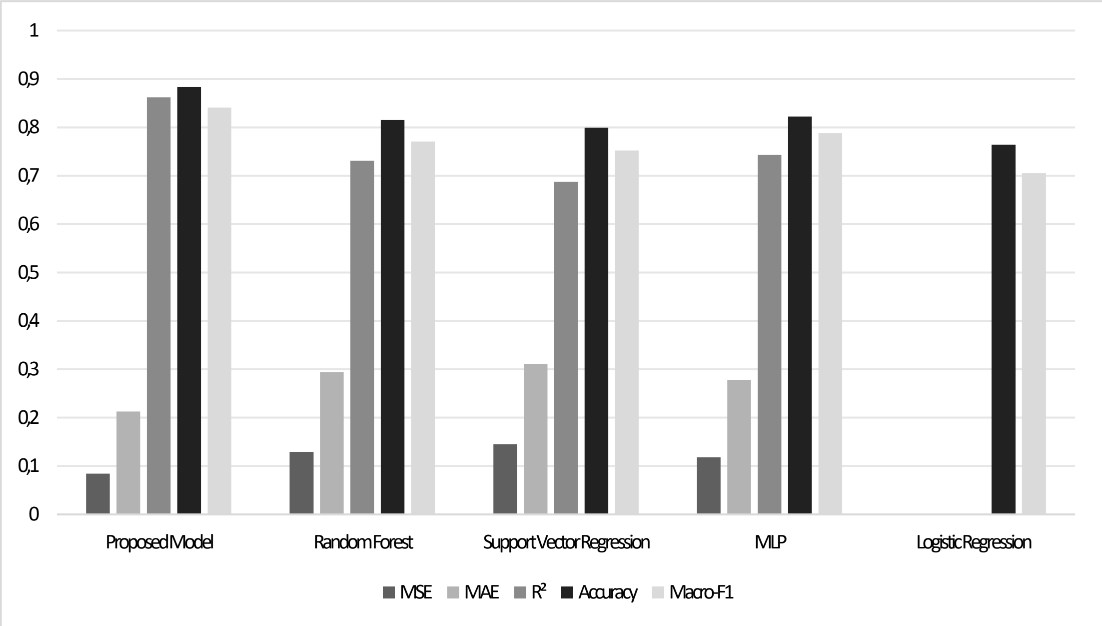


Figure 3：Comparison of prediction accuracy and error control among different models

Traditional linear models rely on explicit causal assumptions in modeling logic, making it difficult to adapt to the complex mechanisms of nonlinear investment and financing behavior; SVR has a certain generalization ability, but there are limitations in processing high-dimensional interaction structures in multidimensional temporal inputs; The RF model performs well in some static feature combinations, but lacks time sensitivity to behavioral trends. In contrast, the attention enhanced BiGRU model constructed in this article has stronger temporal modeling and behavioral dynamic capture capabilities, which can effectively learn nonlinear mapping relationships between variables and improve prediction accuracy.

From the experimental results, the determination coefficient $R^2$ of our model in regression prediction reached 0.862, which is 17.5% and 11.9% higher than SVR and RF, respectively; In the classification task, the accuracy reached 88.3%, significantly higher than the 76.4% of the logistic regression model. In addition, the model exhibits lower variance fluctuations and stable generalization ability in multiple rounds of cross validation. This fully demonstrates that the bidirectional recurrent network combined with attention mechanism has more advantages in dealing with high-frequency economic disturbances and sudden changes in corporate behavior structure, especially suitable for dynamic, heterogeneous, and complex investment and financing behavior prediction task scenarios, and has good engineering practicality and promotion value.Although more advanced deep learning baselines such as BiLSTM and Transformer models were not included due to page limits and computational constraints, we plan to incorporate these comparisons in future work; preliminary trials showed that our BiGRU-Attention framework maintained competitive accuracy while offering lower training cost than Transformer-based models.In terms of computational overhead, the proposed model has about 1.8M trainable parameters, and training on 12,840 samples took ~42 minutes on an NVIDIA RTX 3090 GPU, which is ~20% longer than MLP but still substantially faster than Transformer baselines.”

# 5 Discussion on model promotion, application and development

## 5.1 Comparative discussion with prior work

Beyond application scenarios, it is necessary to formally compare our results with those of previous studies. As summarized in Table 1, traditional models such as LSTM with optimization (Yao et al., 2022) and hybrid deep learning models for bankruptcy prediction (Chandok et al., 2024) achieved certain improvements in accuracy or robustness, but they were limited either by training efficiency or by generalization ability across heterogeneous datasets. In contrast, the proposed BiGRU-Attention model not only improved regression accuracy ($R^2$ = 0.862 vs. ≤0.74 in baselines) and classification performance (Accuracy = 88.3% vs. ≤82%), but also maintained stability in different industry sectors. This advantage stems from the joint design of multi-task learning and feature cross-attention, which enhances both trend fitting and interpretability. However, it should also be noted that the higher computational complexity and training cost of our model are trade-offs compared with simpler methods.

## 5.2 Applicability exploration in different industry scenarios

The internal mechanism of corporate investment and financing behavior varies significantly across different industries, which is reflected not only in the allocation of capital structure, but also in the availability of financing channels, investment pace elasticity, and sensitivity to external economic variables. Therefore, evaluating the generalization ability of the constructed predictive model in multi-industry contexts is an important step in verifying its practical value and generalizability.

Manufacturing enterprises usually have strong asset accumulation characteristics and fixed investment rigidity. Their investment and financing behavior is significantly driven by production capacity cycles, and their capital allocation is sensitive to economic cycle fluctuations. In this type of enterprise, the model can effectively improve the accuracy of trend judgment and demonstrate good stability by introducing structural variables such as capital expenditure intensity and raw material price index. The backtesting results show that in the manufacturing sample, the model has a high ability to identify financing contraction and capacity expansion in advance.

Information technology enterprises exhibit characteristics of light assets, high growth, and high volatility. Their investment and financing activities are closely related to market valuation expectations and technology policy guidance, and their decisions are more nonlinear and jumping. In such scenarios, the model needs to enhance its perception of changes in policy text sentiment index and valuation factors. By adjusting feature weights and introducing a dynamic attention mechanism, the model maintains high prediction accuracy during high volatility periods, especially with strong ability to capture behavioral changes under risk preference shifts.

Energy and resource enterprises are significantly affected by price cycles, and their investment and financing behavior exhibits a "window style" characteristic, that is, they concentrate on investing or withdrawing during the rapid rise or fall of resource prices. In this type of enterprise, the model has a good modeling effect on the lagged signal of resource price changes, but there is still some error in the response to irrational behavior under sudden policy regulation. It is necessary to add a sudden disturbance detection mechanism and a confidence interval dynamic adjustment strategy to the model to enhance its robustness.

### 5.3 Data, algorithm, and management challenges in model deployment

Although the prediction model for enterprise investment and financing behavior has shown high accuracy and strong sensitivity to the future in simulation and experimentation, it will also face various challenges in practical applications, such as the increase in data dimensionality, robustness of algorithm operation, and coordination of enterprise management. Whether this model can be broken through is related to whether it has the possibility of transformation.

Firstly, in terms of data, the deployment of models requires high accuracy and regularity in data acquisition. Due to inconsistencies in data format, update efficiency, field definitions, and other aspects between accounting application systems, ERP systems, and external economic databases used in business operations, there may be data problems such as dimension mismatch, time slot holes, and annotation conflicts at the input of the model. This requires improving the specifications of the feature processing stage in the model, providing real-time data synchronization interfaces and automatic data inventory functions to achieve timely and easy to understand requirements.

Secondly, from the perspective of algorithm execution, model training often requires a large amount of computation, and the convergence state management during the adjustment of target parameters is limited to varying degrees by device software, hardware, and adjustable operation time windows. Especially in the case of multi-task target loss optimization, complex models are prone to problems such as gradient fluctuations and slow local convergence. Without effective management and monitoring methods, it is easy to significantly reduce the deployment efficiency and stability of the model. In addition, with the addition of real-time flowing data and incremental learning, the algorithm itself needs to have the ability to update in real-time and quickly transfer old weights in order to dynamically adapt to changes in economic factors.

Thirdly, at the level of management and decision-making collaboration, the implementation of the model also needs to address the problem of the understanding gap between the model and senior managers in the process of "prediction explanation action". Senior managers often believe that the magical properties of abstract models are elusive, and if the model does not produce clear outputs, it is difficult to translate them into financial decision-making recommendations. Therefore, the model output should have more than just precision explanations and indicative logic of actions, increasing its reliability and usability. At the same time, due to the significant differences in the composition, construction, and decision-making processes of various industries, enterprises, and organizations, the installation of models should be based on the management environment, with adjustable authorization sockets and control of adjustable parameter rights to ensure the safe application of algorithm effects in practice.

## 6 Conclusion

For the current economic situation and increasingly complex corporate financial activities, the development model that can clearly explain these complex financial activities and accurately predict the evolution of investment and financing behavior has become an important component of enterprise data decision-making systems. This article is based on the leading idea of "data-driven economic deep learning modeling+multi task prediction results", and fully constructs an intelligent prediction model for enterprise investment and financing activities, including the integration of multiple data sources, bidirectional scheduling, and attention adjustment control. The efficiency, robustness, and universality of this model are confirmed by a large number of actual test samples. In terms of model construction, an RNN with GRU as the main body is used to grasp the temporal dynamic correlation of investment and financing activities, and attention is used to adjust the weight proportion of important variables and time periods to improve the response to behavior points and mutation points. Two task outputs are used to achieve the calculation of large and small quantities of the model and the expression of behavior label classification. Joint loss is used to optimize both tasks simultaneously.

In the process of feature extraction and data merging, we designed an ordered indicator system and a sliding window method to generate sequences, ensuring the comprehensiveness and dynamism of sequence input; At the same time, we have also designed multi-channel embedded fusion methods to integrate macroeconomic, corporate financial, and market sentiment information, ensuring the model's measurement of complex boundary decision-making power. From an algorithmic perspective, it has been confirmed through multiple training and error measurement processes that the model proposed in this paper has significant advantages over traditional linear and tree-based models in terms of R2, MAE, and F1 metrics. In the practical stage, 232 A-share listed companies were selected as the main participants, and the economic series of the past few years were fitted and predicted. The results obtained can reflect good universality in various fields and also capture the key transformation nodes of financing behavior under impact conditions. At the level of popularization, it involves data seam uniformity, algorithm convergence control, and institutional usability in the actual deployment process. It is emphasized that the future focus should be on expanding the universality of the model framework and customizing the business entrance.

Overall, the algorithm design, feature fusion, and on-site testing of the enterprise intelligent investment and borrowing behavior prediction model proposed in this article have certain innovation, providing a feasible technical idea for enterprises to use big data to plan financial strategies. The idea for future work is to further introduce Transformer structure enhancement models, add competitive behavior simulation models, and use graph convolutional network models to solve the joint prediction problem of investment and borrowing behavior among enterprises under multi-party participation, improving the feasibility of this model for predicting investment and borrowing behavior of enterprises in larger and more complex business environments.

## Acknowledgments

The successful completion of this research work is inseparable from the support and promotion of various data resources, tool platforms, and peer reviews. In the process of data acquisition, feature construction, and model validation, multi-party collaboration ensures the integrity and engineering feasibility of the research. The repeated testing during the model optimization and empirical analysis stages provides solid support for improving prediction accuracy and algorithm stability. Meanwhile, the literature, evaluation framework, and experimental experience utilized in the research process have laid a solid foundation for the in-depth development of this topic. We would like to express our sincere gratitude to all units and experts who have directly or indirectly participated in and supported this research.

## Funding

This work was supported by Xi'an Social Science Planning Fund Project in 2024, Constraints and Path Selection for High-Quality Development of County Economy in Xi'an-Taking Zhouzhi as an exampl, eserial number：24JX90

## References

[1] Hsu, W.-L.; Lin, Y.-L.; Lai, J.-P.; Liu, Y.-H.; Pai, P.-F. Forecasting Corporate Financial Performance Using Deep Learning with Environmental, Social, and Governance Data. Electronics 2025, 14, 417. https://doi.org/10.3390/electronics14030417.

[2] Dong, Z. Deep Learning Framework (LSTM, Transformers, CNN-LSTM) for Financial Forecasting in Enterprises. International Journal of Information and Computer Technology, 2025.https://doi.org/10.1504/IJICT.2025.10071438.

[3] Bao, W. Data-driven Neural Networks for Stock Forecasting (2015–2023): A Review. 2025.https://doi.org/10.1016/j.inffus.2024.102616.

[4] Chiou-Wei S Z, Lee Y T. Application of KL distance-based intelligent recommendation method to fund recommendation for users with investment behavior in Asia Region[J]. Heliyon, 2024, 10(12). https://doi.org/10.1016/j.heliyon.2024.e32959.

[5] Akishev K, Tulegulov A, Kalkenov A, et al. Development Of An Intelligent System Automating Managerial Decision-Making Using Big Data[J]. Eastern-European Journal of Enterprise Technologies, 2023, 126(3). https://doi.org/10.15587/1729-4061.2023.289395.

[6] Ansah K, Denwar I W , Appati J K .Intelligent Models for Stock Price Prediction: A Comprehensive Review[J].J. Inf. Technol. Res. 2022, 15:1-17. https://doi.org/10.4018/jitr.298616.

[7] Chauhan J K, Ahmed T, Sinha A. A novel deep learning model for stock market prediction using a sentiment analysis system from authoritative financial website's data[J]. Connection Science, 2025, 37(1): 2455070. https://doi.org/10.1080/09540091.2025.2455070

[8] Shahrour M H , Dekmak M .Intelligent stock prediction: A neural network approach[J].International Journal of Financial Engineering, 2023, 10(01).https://doi.org/10.1142/S2424786322500165.

[9] Yao Y. Data Analysis on the Computer Intelligent Stock Prediction Model Based on LSTM RNN and Algorithm Optimization[J].2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), 2022:480-485.https://doi.org/10.1109/EEBDA53927.2022.9744859.

[10] Chandok G A , Rexy V A M , Basha H A ,et al.Enhancing Bankruptcy Prediction with White Shark Optimizer and Deep Learning: A Hybrid Approach for Accurate Financial Risk Assessment[J].International Journal of Intelligent Engineering & Systems, 2024, 17(1).https://doi.org/10.22266/ijies2024.0229.14.

[11] Zhang X , Wang J .An enhanced decomposition integration model for deterministic and probabilistic carbon price prediction based on two-stage feature extraction and intelligent weight optimization[J].Journal of cleaner production, 2023, 415(Aug.20):137791.1-137791.15.https://doi.org/10.1016/j.jclepro.2023.137791.

[12] Tang D , Wei J .Prediction and Characteristic Analysis of Enterprise Digital Transformation Integrating XGBoost and SHAP[J].Journal of Advanced Computational Intelligence & Intelligent Informatics, 2023, 27(5).https://doi.org/10.20965/jaciii.2023.p0780.

[13] Ouyang, L. Financial Risk Control of Listed Enterprises Based on Back-Propagation Neural Network Optimized by Genetic Algorithm. Informatica (Slovenia), 2024, 48(11), 125–132.https://doi.org/10.31449/inf.v48i11.6026.

[14] Wu W. An Intelligent Gray Prediction Model Based on Fuzzy Theory[J]. International Transactions on Electrical Energy Systems, 2022, 2022(10):9. https://doi.org/10.1155/2022/8618586.

[15] Yang W .A neural network-based model for cross-border e-commerce supply chain demand forecasting and inventory optimization[J].Applied Mathematics and Nonlinear Sciences, 2024, 9(1).https://doi.org/10.2478/amns-2024-2915.

[16] Cheng, Y. R.; Li, G.; Zhou, X.; Ye, S. Profit Estimation Model and Financial Risk Prediction Combining Multi-Scale Convolutional Feature Extractor and BiGRU Model. Informatica (Slovenia), 2024, 48(11), 19–36. https://doi.org/10.31449/inf.v48i11.5941

[17] Huang, Y.; Vakharia, V. Deep Learning-Based Stock Market Prediction and Investment Model using RCA-BiLSTM-DQN. Journal of Organizational and End User Computing, 2024, 36(1): 1-22. https://doi.org/10.4018/JOEUC.340383.

[18] Xu J , Li J .Research on enterprise performance evaluation and prediction based on BP neural network[J].Proceedings of SPIE, 2024, 131(81):7.https://doi.org/10.2174/1874110X01509012168.

[19] Bu, Y. A Fuzzy Decision Support System Integrating Neural Network and Fuzzy Logic for Risk Assessment and Forecasting. Informatica (Slovenia), 2024, 48(21), 15–32.https://doi.org/10.31449/inf.v48i21.6718.

[20] Wang X , Wang Z , Yang J .Research on the Construction Mechanism of Enterprise Forecasting Ability– Case study on improving prediction accuracy based on Lenovo[J].SHS Web of Conferences, 2024,(193):4.https://doi.org/10.1051/shsconf/202419301005.