A Hybrid RF-CART-SMOTE-GA Model for Early Warning in University Budget Management

Xu Qin
Jiangsu University Jingjiang College, Zhenjiang 212000, China
E-mail: 13815480090@163.com

Keywords: budget management, information warning, random forest, SMOTE, university finance

Received: June 20, 2025

With the advent of the big data era, university budget management faces heightened demands, necessitating effective responses to challenges arising from diversified funding sources and expanded scale. To address these challenges, this study employs the Random Forest algorithm as the foundation for a budget information early warning model. It combines the Classification and Regression Trees (CART) algorithm to optimize the decision tree structure, introduces weighted ensemble voting to enhance the classification process, and achieves data balancing and parameter optimization through Synthetic Minority Over-sampling Techniques and genetic algorithms. The model was validated using two datasets: the Integrated Postsecondary Education Data System (IPEDS) and the Higher Education Statistics Agency (HESA). Experimental results demonstrated that for the classification task during the budgeting phase, the model achieved a maximum classification accuracy of 94% on the HESA dataset, with a recall of 92.18%, an F1 score of 92.87%, and a sample balance rate of 81.64%, with the lowest budget information early warning error at 3.7%. Additionally, the average runtime was 0.04 seconds, and CPU utilization was only 24.17%, significantly outperforming models such as DBN, XGBoost, and VAE-TSAD. The research results demonstrate that the proposed model possesses high accuracy, real-time capability, and computational efficiency in early warning for university budget management information, providing reliable technical support for higher education financial decision-making.

Povzetek: Članek predstavi hibridni ansambelski model RF-CART-WIV-SMOTE-GA za zgodnje opozarjanje v univerzitetnem proračunskem upravljanju s pomočjo večnivojske optimizacije strukture.

1 Introduction

With the expansion of university scale, improvement of education level, and abundant educational resources, how to scientifically, reasonably, and efficiently manage budgets has become an important issue that university managers urgently need to solve [1]. Budget management is an important component of financial management in universities, which is related to the development of various undertakings and rational resource allocation, as well as the economic and educational benefits of universities [2]. The primary challenges currently facing university budget management include: (1) Strong data heterogeneity and low availability; (2) Time lag between budget execution and performance feedback; (3) Insufficient risk identification accuracy and delayed resource allocation response. Therefore, scholars at home and abroad have successively explored solutions. To further enhance the effectiveness of budget management information warning in universities, Teng et al. proposed a new warning method by combining data mining and deep learning models. The experimental results showed that the accuracy of the financial risk warning method was the highest, at 97.62%, which was 13.57% higher than that of traditional methods [3]. Valle-Cruz et al. proposed a new warning model that combined multi-layer perception algorithms to improve the clarity of allocation and expenditure in the budget execution process of universities. The experimental results showed that the model had high real-time and accuracy in financial budget information early warning of universities in the past three years [4]. Truong et al. argued that there were still difficulties in conducting timely financial management alerts in crowdsourced competition programs. For this reason, a budget management early warning strategy based on the HAIS algorithm was proposed. The experimental results showed that the warning time lag of this strategy was as low as 0.13 seconds, which was better than that of the traditional method [5]. Zheng et al. found that federated learning was highly effective in early warning of budget management information in universities, but at a higher cost. They proposed a fuzzy logic based federal learning budget information warning model. Experimental results showed that the effectiveness of the budget information early warning effectiveness far exceeded other existing methods [6].

Random Forest (RF) algorithm is an ensemble learning method that learns historical budget data by constructing a model to identify abnormal data that is inconsistent with the normal budget pattern, thereby achieving early warning of budget overruns, abnormal fund flows, and other situations [7]. To reduce the RF error between the revenue budget and expenditure budget of universities, Bénard et al. took the Sobol index to propose a new early warning method. The experimental results showed that the method could occur before the crisis to make effective judgments and early warnings, and the effectiveness was high [8]. Zhang et al. found that market economy fluctuations had a significant impact on budget management in universities. Therefore, a new budget warning algorithm was proposed by combining RF algorithm and logistic regression model. The experimental results showed that the new algorithm had the highest accuracy of 93.47% for teaching budget warning in three universities in mainland China [9]. Cao et al. proposed a new warning model for the risk warning level in the new university budget management based on RF and convolutional neural networks. The experimental results showed that the model had a high effectiveness of 96.38% in risk warning for campus construction budget and campus teaching budget [10]. Zeng found that there were still many risks in the current financial budget management of universities. Therefore, the RF and backpropagation neural network were combined for risk data mining. Finally, a new warning model was proposed. The experimental results showed that the new model had the highest prediction accuracy of 91.6% and the shortest prediction time of 7.65 s [11]. The comparison for each document is shown in Table 1.

As shown in Table 1, existing research primarily focuses on model structure innovation or feature optimization, yet significant shortcomings remain in data imbalance handling, classification stability, and computational efficiency. In contrast, the proposed RF-CART-WIV-SMOTE-GA model not only integrates the advantages of multiple algorithms but also achieves data balancing and optimization through the collaborative optimization of SMOTE and GA. achieving 94% accuracy and 3.7% Mean Absolute Error (MAE) on both the IPEDS and HESA datasets. It also demonstrates significant improvements in runtime and CPU utilization compared to existing State-of-the-art (SOTA) models.

In summary, many researchers at home and abroad have conducted in-depth exploration on budget information management in universities and proposed various processing methods. These methods can to some extent complete the warning work of budget information, but there are still challenges such as low warning accuracy and high warning cost. Therefore, the study introduces a multi-layer optimization mechanism based on the RF model to construct a high-precision early warning model for university budget management. Specifically, at the structural level, Classification and Regression Trees (CART) are employed to optimize the decision tree partitioning process, while Weighted Integration Voting (WIV) is adopted at the decision level to enhance classification robustness.

At the data level, the Synthetic Minority Over-sampling Technique (SMOTE) balances budget information samples. At the optimization level, a Genetic Algorithm (GA) is introduced for global parameter optimization. Unlike traditional bagging or boosting methods, WIV significantly reduces bias correlation among weak classifiers through its weight-adaptive adjustment strategy. This approach achieves higher classification accuracy and model stability in the imbalanced data environment of budget management.

This study aims to use a multi-algorithm fusion integrated optimization framework to accurately classify and detect abnormal budget management data, thereby improving the intelligence level and risk control capability of financial decision-making in universities. To this end, the following research questions are proposed:

Question 1: How to improve classification accuracy and anomaly detection accuracy in situations where budget information data is highly imbalanced?

Hypothesis 1: Combining SMOTE and GA-based data balancing and optimization mechanisms can significantly enhance the classification performance of RF models on imbalanced datasets.

Question 2: Compared with traditional ensemble learning methods (e.g., Bagging, Boosting), can WIV improve alert stability under multi-indicator constraints?

Hypothesis 2: Introducing WIV optimizes the decision fusion process through adaptive weighting mechanisms, thereby enhancing model stability and generalization performance.

Question 3: Can the proposed RF-CART-WIV-SMOTE-GA model outperform existing models (e.g., DBN, XGBoost, VAE-TSAD) across the four stages of university budget formulation, execution, adjustment, and evaluation?

Hypothesis 3: The proposed model will significantly outperform comparable SOTA methods in key metrics including accuracy, recall, and runtime.

In summary, the innovation of this research manifests in three key aspects: (1) Establishing a multi-layer ensemble learning framework that integrates structural optimization, decision enhancement, and data balancing; (2) Proposing a weighted strategy using WIV to replace traditional ensemble methods and enhance robustness; (3) Validating the model's high accuracy and real-time capability in university budget early warning through empirical testing on IPEDS and HESA datasets, thereby providing scalable intelligent decision support tools for university financial management.

Table 1: Comparison of existing state-of-the-art (SOTA) early warning models for higher education budget management

Author (Year)	Algorithm / Model	Dataset	Accuracy / MAE	Computational Cost / Time	Data Imbalance Handling	Main Limitation	Reference
Teng Y *et al.* (2023)	Data Mining + Deep Learning	University Financial Data	97.62%	—	None	High model complexity, limited generalization	[3]
Valle-Cruz D *et al.* (2022)	MLP	Government University Data	≈95%	Good real-time performance	None	Applicable only to execution stage	[4]
Truong V Q *et al.* (2023)	HAIS	Crowdsourcing Dataset	—	0.13 s	None	High computational cost, limited scope	[5]
Zheng Z *et al.* (2022)	Federated + Fuzzy Logic	Smart City Data	≈93%	High	Partial	Expensive computation, privacy constraints	[6]
Bénard C *et al.* (2022)	RF + Sobol Index	Financial Dataset	≈92%	Medium	None	Poor adaptability to nonlinear fluctuation	[8]
Zhang Z *et al.* (2022)	RF + Logistic Regression	3 Chinese Universities	93.47%	—	None	Insufficient hierarchical budget analysis	[9]
Cao Y *et al.* (2022)	RF + CNN	Campus Budget Data	96.38%	Medium	None	Complex structure, difficult tuning	[10]
Zeng H *et al.* (2022)	RF + BPNN	Financial Dataset	91.60%	7.65 s	None	Long training time, sensitive to data bias	[11]
DBN / XGBoost / VAE-TSAD (baseline)	Deep Belief / Boosting / Variational Autoencoder	IPEDS / HESA	89–91%	0.07–0.11 s	None	Unstable on imbalanced data	/
Proposed Model (RF-CART-WIV-SMOTE-GA)	Hybrid Ensemble	IPEDS / HESA	94% / MAE 3.7%	0.04 s	SMOTE-balanced	High accuracy, stable performance, efficient computation	/

2 Methods and materials

In response to the various risk issues in existing efficient budget management information, the study first uses the RF as a basis, introduces CART to adjust its decision tree structure, and adopts WIV to optimize its decision-making process. Finally, a budget management information classification model is constructed. Before model training, all raw data undergoes systematic preprocessing to ensure the accuracy and consistency of model inputs. Firstly, perform missing value detection on the IPEDS and HESA datasets. If the missing ratio is less than 5%, use the mean or mode estimate. If it exceeds 5%, use K-Nearest Neighbor Imputation (KNN Imputation) for repair. Next, Z-score normalization is applied to the continuous features, and a hot encoding process is used to classify the variables. Use the 3σ rule and interquartile range (IQR) method to detect and remove outliers. Finally, by ranking the importance of RF features, key features that have a significant impact on budget information classification are selected as model inputs. In addition, to successfully achieve classified information warning, the research introduces the SMOTE algorithm to adjust imbalanced data, and take the GA to strengthen the optimal warning solution. Finally, the research proposes a new university budget information management warning model (RF-SMOTE).

2.1 Construction of information classification model for university budget management based on RF

University budget management generally includes four stages: budget preparation, budget execution, budget control, and budget performance assessment [12]. Budget preparation is the foundation of university budget management. Budget execution is the process of putting the prepared budget into practice. Budget control is the supervision and adjustment of the budget execution process. Budget performance assessment evaluates the budget execution results. According to the description in the “Financial System of Higher Education Institutions”, Figure 1 displays the life cycle of budget management information in universities [13-14].

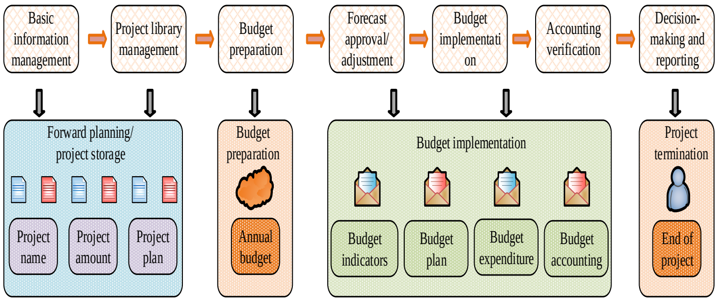


Figure 1: Schematic of the life cycle of budget management in higher education

As shown in Figure 1, when the basic management information is input into the system, it will undergo project library management, budget preparation, budget approval, budget adjustment, budget execution, and accounting verification before final settlement and reporting. Due to various reasons during the period, such as inadequate budget management, inability to dynamically meet budget requirements, inadequate connection between different levels of budget management, and incomplete budget systems, it is highly likely to result in inaccurate budgets, reduced budget management efficiency, and resource waste. To this end, the study introduces RF for multi-level early warning of university budget management information, in order to avoid the interference and intrusion of these abnormal information. Compared with other algorithms, RF algorithm can evaluate the importance of features and help identify key indicators that affect budget performance, providing a basis for budget preparation and adjustment [15]. Figure 2 displays the principle of RF.

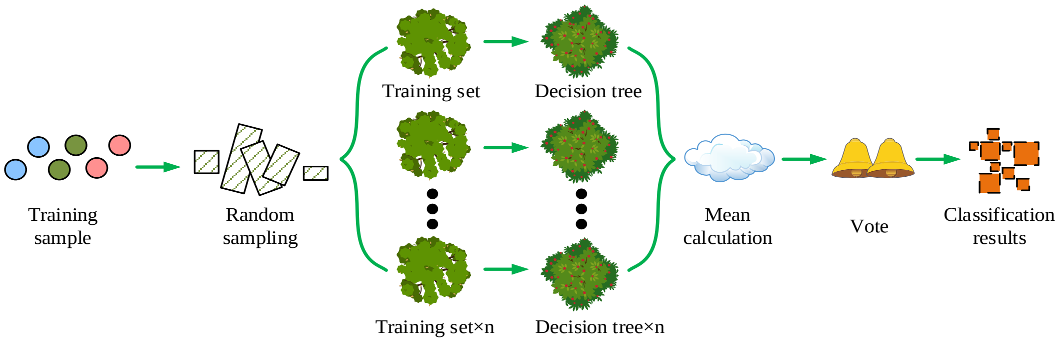


Figure 2: RF principle schematic

As shown in Figure 2, RF divides the original training samples into multiple sub samples after random sampling. When selecting segmentation nodes, randomly select some features to search for the best segmentation point, and then construct their own decision trees based on each sub sample. In general classification tasks, RF determines the final prediction result through majority voting, while in regression tasks, RF calculates the average of all decision tree prediction results to obtain the final prediction result [16]. RF needs to select the optimal splitting attribute based on indicators such as information gain and information gain rate. Assuming the initial sample set is D, the information entropy is shown in equation (1).

$$Entrpy(D) = -\sum_{i=1}^{c} P_i \log_2 P_i \quad (1)$$

In equation (1), P_i represents the proportion of class i in the sample set. $Entrpy(D)$ represents the information entropy of the original sample set. c represents the total number of samples. After introducing feature data A, the sample set is classified as information entropy, information gain, information value, and information gain rate, and their respective calculations are shown in equation (2).

$$\begin{cases} Entrpy(D_A) = -\sum_{i=1}^{k} \frac{|D_j|}{|D|} Entropy(D_A) \\ Gain(D,A) = Entropy(D) - Entropy(D_A) \\ SplitEntropy(D,A) = -\sum_{i=1}^{c} \frac{|D_i|}{|D|} \log_2 \frac{|D_i|}{|D|} \\ GainRation(D,A) = \frac{Gain(D,A)}{SplitEntropy(D,A)} \end{cases} \quad (2)$$

In equation (2), the set of values for the discrete feature A is denoted as $\{a_1,\dots,a_k\}$. Subdivisions based on the values of A yield subsets $D_1,\dots,D_k$, where $|D_j|$ represents the number of samples in the j-th subset and $|D|$ denotes the total number of samples. k represents the total number of parts that the set D is divided into. $Gain(D,A)$ represents the information gain of feature A in set D. $SplitEntropy(D,A)$ represents the information value of feature A in set D. $GainRation(D,A)$ represents the information gain rate of feature A in set D. Due to the random sampling method of RF, it is easy to cause strong correlation between adjacent decision trees, resulting in poor classification results of the final information data. For this purpose, the research takes CART to optimize the decision tree generation of RF in the study. Compared with other methods, CART is more suitable for RF because it is trained based on information entropy, avoiding more parameter settings and reducing the complexity of the RF optimization process [17]. The optimal index minimization node classification for CART is shown in equation (3).

$$Gini(D) = 1 - \sum_{i=1}^{c} P_i \quad (3)$$

In equation (3), $\sum_{i=1}^{c} P_i = 1$ is specified for $Gini_{split}(D) = \sum_{j=1}^{k} \frac{|D_j|}{|D|} Gini(D_j)$. When used to minimize node impurity, $argmin_A Gini_{split}(D)$ is selected. $Gini(D)$ represents the minimum node splitting index of the sample set D. At this point, the coefficient expression of the minimization index is shown in equation (4).

$$Gini_{split}(D) = \sum_{i=1}^{l} \frac{|D_i|}{|D|} Gini(D_i) \quad (4)$$

In equation (4), D_i represents the minimum index coefficient at the i-th set after partitioning according to l subsets. At this point, the optimal indicator for RF segmentation nodes is changed to information gain rate, which serves as the basis for node segmentation judgment. The information entropy, information gain, information value, and information gain rate of the improved RF are shown in equation (5).

$$\begin{cases} Entrpy(D_B|A) = -\sum_{i=1}^{l} \frac{|D_{Bi}|}{|D_B|} Entropy(D_B) \\ Gain(D_B|A) = Entropy(D_B) - Entropy(D_B|A) \\ SplitEntropy(D_B|A) = -\sum_{i=1}^{l} \frac{|D_{Bi}|}{|D_B|} \log_2 \frac{|D_{Bi}|}{|D_B|} \\ GainRation(D_B|A) = \frac{Gain(D_B|A)}{Gini_{Split}(D_B|A)} \end{cases} \quad (5)$$

In equation (5), D_B represents the B-th set. B denotes the B-th sub-sample set/tree obtained via bootstrap sampling, and D_B represents its training samples. Equation (5) is isomorphic to equation (2), but replaces the population D with D_B for the split metric on a single tree (or subset). In addition, to enhance the impact of attribute differences in different decision trees on the final data information classification and improve the effectiveness of management information classification, the final decision of RF is optimized by WIV, and the classifier is processed in a weighted manner. This process is shown in equation (6).

$$\omega_i = \sum_l P_{il} \Big/ \sum_{il} P_{il} \quad l = 1,2,3,\cdots,c \quad (6)$$

In equation (6), ω_i represents the weight of the i-th sample. P_{il} represents the probability of the i-th sample after the classifier splits the sample into l subsets. P_i denotes the normalization factor used for weighting (e.g., the prior proportion of class i in the training set or the validation set performance score), sharing the same sign but differing in meaning from the class probability in equation (1). The confidence calculation of the weighted classification results is shown in equation (7).

$$u_l(D) = \frac{1}{N} \sum_{i=1}^{N} \omega_i P_{il}(D) \quad (7)$$

In equation (7), $u_l(D)$ represents the confidence value of dividing the sample into l split subsets. A new classification model for university budget management information is proposed based on split improvement and voting improvement in RF. The process of this model is shown in Figure 3.

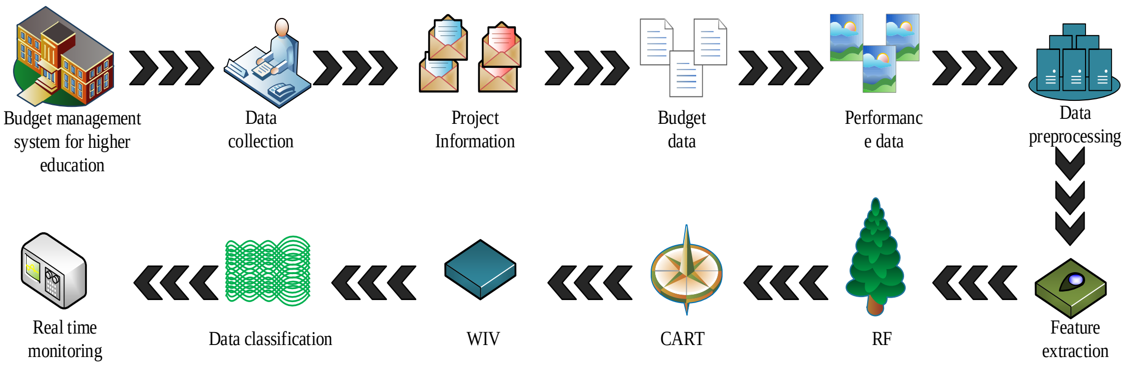


Figure 3: RF-CART-WIV information categorization model flow

As shown in Figure 3, firstly, all relevant data are collected from the university budget management system, including project information, budget data, performance data, etc. Secondly, perform data preprocessing to ensure the quality and consistency of the data. Then, extract useful features from the raw data, including amounts, project names, etc. After completion, the RF is used to evaluate the importance of each feature and select key features that have a significant impact on budget management. During this period, CART optimizes the decision tree structure of RF, and WIV optimizes the decision process of the classifier. Next, real-time data classification is performed after validation on the training set, and the classification results are monitored in real-time, with regular updates to the model.

2.2 Training and enhancement of early warning model with SMOTE algorithm fusion

After constructing a classification model for budget management information data in universities, there are still technical drawbacks in evaluating and making decisions based on different classification data. For example, noisy data after budget management information classification can cause overlapping classification results. To use information classification results for effective judgment and connection, and provide decision support for management, including adjusting budget allocation, optimizing resource allocation, or taking preventive measures, the study first investigates the sources of income and expenditure data that have the greatest impact on budget allocation. The sources of income and expenditure in universities are shown in Figure 4. To enhance the credibility and data support of Figure 4, the study further incorporates publicly available statistical data samples. Referencing the “2022 National Education Expenditure Implementation Statistics Bulletin” issued by the Ministry of Education and the Ministry of Finance, along with financial annual reports from selected universities, the following typical feature distributions are extracted: Fiscal allocations account for approximately 56% of total university revenue, institutional income accounts for about 27%, operational income accounts for roughly 9%, and other income accounts for approximately 8%. Correspondingly, educational expenditures constitute about 63% of total expenditures, personnel expenses account for 18%, project funding represents 12%, and operational expenditures make up 7%. This distribution pattern serves as the basis for setting input data proportions in the model and is reflected in the training sample structure of the subsequent RF-SMOTE model. This ensures that the income-expenditure relationship depicted in Figure 4 possesses practical statistical significance and representativeness.

As shown in Figure 4, the income includes government subsidy income, business income, operating income, and other income. In terms of expenditure, it can be roughly divided into education expenditure, health expenditure, employment security expenditure, personnel expenditure, project expenditure, and operating expenditure. According to the “Statistical Announcement on the Implementation of National Education Funds in 2022” released by the Ministry of Education, the National Bureau of Statistics, and the Ministry of Finance in 2022, in some cases, the financial budgets of universities have shown an imbalance where income is less than expenditure [18-19]. Therefore, to balance the input data of the warning model, the SMOTE algorithm is introduced. This algorithm balances the dataset by generating synthesized minority class samples, rather than simply copying minority class samples, thus avoiding overfitting. The new sample generated by SMOTE algorithm is shown in Figure 5.

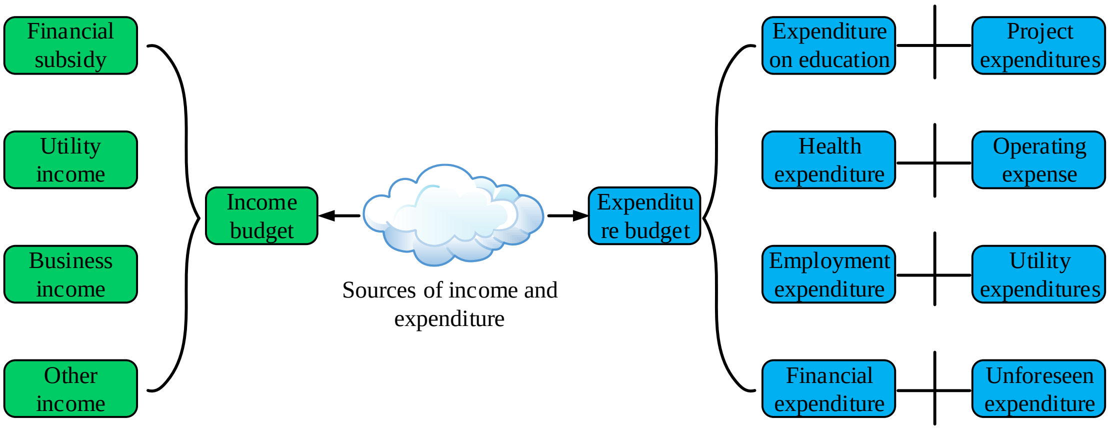


Figure 4: Demonstration of the sources of income and expenditure of higher education finances

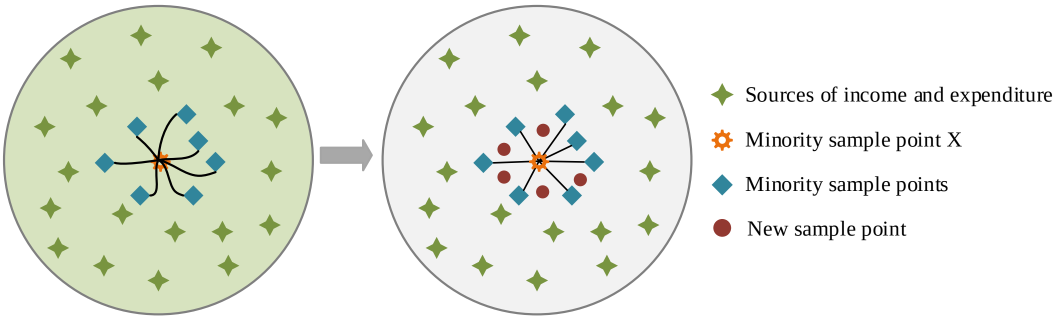


Figure 5: Schematic of SMOTE new sample generation

As shown in Figure 5, when there is an imbalance between financial income and expenditure data, in samples with a small amount of data, the nearest k neighboring samples of sample point x in the minority class can be found based on the Euclidean distance, and the upsampling magnification is set. For each selected neighbor, calculate the distance between the original sample and the neighbor, and generate a new composite sample from the random points on this line. The method for generating the new sample is linear interpolation between the original sample and the neighbor sample. The interpolation is shown in equation (8).

$$\begin{cases} Q_i = X + rand(0,1) \times (y_i - X) \\ i = 1, 2, ..., n \end{cases} \tag{8}$$

In equation (8), Q_i represents the generated interpolated sample points. X represents the data sample point in the minority class. For each minority class sample point X, its corresponding nearest neighbor sample is found. The i-th nearest neighbor sample is represented as y_i. Finally, the maximum value of i is n, which is the upsampling magnification. The imbalance in the dataset is calculated in equation (9).

$$n = ROUND(IL) \tag{9}$$

In equation (9), IL denotes the imbalance degree. n_{maj} represents the number of majority class samples and $n_{\min}$ represents the number of minority class samples. Then, the imbalance degree IL is defined by equation (9) (e.g., commonly used $IL = n_{maj} / n_{\min}$ or $IL = (n_{maj} - n_{\min}) / n_{\min}$. IL represents the degree of imbalance. $ROUND(IL)$ represents the rounded value of IL. However, the simple SMOTE does not differentiate different samples when processing minority class samples, but uniformly generates new samples using the same method, which to some extent limits its applicability. Therefore, GA optimizes the sample selection in the study, which uses the roulette wheel selection method to make the newly synthesized samples more likely to cluster around high fitness samples [20]. By calculating the cumulative value of probabilities, the relative weight of each sample in the roulette wheel process is determined, as shown in equation (10).

$$cump(j) = \sum_{i=1}^{j} Q(i) \tag{10}$$

In equation (10), $Q(i)$ represents the probability of randomly selected minority sample points. At this point, the probability of selecting minority class samples is shown in equation (11).

$$Q(i) = \frac{k+1-f(x_i)}{\sum_{i=1}^{k}(k+1-f(x_i))} \tag{11}$$

In equation (11), $f(x_i)$ denotes the fitness of minority class samples, with a recommended range normalized to [0,1]. k represents the nearest neighbor parameter. Subsequently, the posEff function is used to calculate the generated sample distance, and the relative distance ratio is shown in equation (12).

$$u' = \left(\frac{d_1}{d_{xy}}\right) \times \left(\frac{d_2}{d_{xz}}\right) \tag{12}$$

In equation (12), d_{xy} and d_{xz} represent the average Euclidean distances between minority samples and majority samples, respectively. d_1 and d_2 respectively represent the distance between the sample and its nearest neighbors of the same class and the majority of its nearest neighbors. To further elucidate the specific implementation mechanism of the GA in the minority class sample selection process, this study explains its fitness function and roulette wheel selection. Firstly, to measure the representativeness and separability of minority class samples, the fitness function $f(x_i)$ is defined in equation (13).

$$f(x_i) = \alpha \cdot \frac{d_{intra}(x_i)}{d_{inter}(x_i)} + \beta \cdot (1-u_i') \tag{13}$$

In equation (13), $d_{intra}(x_i)$ denotes the average Euclidean distance between sample x_i and its class neighbors. $d_{inter}(x_i)$ denotes the average Euclidean distance between sample x_i and the nearest majority-class samples. u_i' denotes the relative distance ratio defined in equation (12). α and β are both weight coefficients. A smaller $f(x_i)$ value indicates that the sample resides in a dense region within its class and is distant from the majority-class boundary, making it more suitable for synthesis. Secondly, a roulette wheel mechanism selects samples based on their fitness distribution. The sample selection probability $Q(i)$ and cumulative probability $cump(j)$ are determined by equations (10)-(11), respectively, with the calculation shown in equation (14).

$$Q(i) = \frac{\varsigma+1-f(x_i)}{\sum(\varsigma+1-f(x_i))}, cump(j) = \sum_{i=1}^{j} Q(i) \tag{14}$$

In equation (14), ς represents the translation term used to prevent the denominator from becoming zero. The algorithm achieves probabilistic selection by generating a random number r ∈ [0,1] and locating the corresponding sample index within the cumulative probability interval. The roulette selection pseudocode is as follows:

```
Input: Minority sample set X_min, neighbor number k, population size N_pop,
    maximum generation G_max, fitness function f(x_i), crossover rate p_c, mutation rate p_m
Output: Optimized minority sample subset X_opt

Initialize population P with N_pop minority samples randomly selected from X_min
For generation g = 1 to G_max do
    Evaluate fitness f(x_i) for each individual in P using Eq.(f)
    Normalize fitness and compute selection probability Q(i) and cumulative prob cump(j)
    For i = 1 to N_pop do
        Generate random number r ∈ [0,1]
        Select parent p_i such that cump(j-1) < r ≤ cump(j)
    Apply crossover (probability p_c) and mutation (probability p_m) to generate new population P'
    Replace P ← P'
End For
Return individuals with minimum f(x_i) as X_opt
```

To improve the specificity of GA optimization, this study simultaneously monitored indicators such as classification accuracy, sample balance rate, and MAE during the iteration process of the genetic algorithm. These metrics serve as comprehensive performance evaluation criteria for sample selection schemes. By comparing the improvement margins of these metrics in each iteration, the GA dynamically identifies minority samples that significantly boosts overall model performance, thereby finely optimizing the SMOTE generation process. The hyperparameter table is shown in Table 2.

In summary, based on the previously proposed RF-CART-WIV information classification model for university budget management, a new warning model, RF-CART-WIV-SMOTE-GA, is proposed. The warning process is shown in Figure 6.

Table 2: Main hyperparameter settings for the proposed RF-CART-WIV-SMOTE-GA model

Module	Symbol / Parameter Name (in Formulas)	Description	Value / Determination Method	Formula Reference / Basis
RF / CART (Splitting & Voting)	B	Number of bootstrap sub-samples / trees ((D_B))	200	Matches (D_B) in Eq. (5); grid-search optimum balancing accuracy and runtime
	$\mathrm{Gain}(D,A)$, $\mathrm{SplitEntropy}(D,A)$, $\mathrm{GainRation}(D,A)$	Information gain, split entropy and gain ratio metrics	/	Eqs. (2) for information-based splits
	$\mathrm{Gini}(D), \mathrm{Ginisplit}(D)$	Gini impurity and weighted Gini for splitting	/	Eqs. (3), (4) for minimal impurity selection
Weighted Integration Voting	ω_i	Weight of the *i*-th sample or class	Computed automatically	Eq. (6)
	$u_i(D)$	Confidence value after weighted voting	Computed automatically	Eq. (7)
SMOTE	k	Number of nearest neighbors	5	Used for neighbor search and also in GA selection (Eq. 11 denominator range)
	$Q_i = X + rand(0,1) \times (v_i - X)$	Formula for synthetic sample generation	/	Eq. (8)
	IL	Imbalance Level indicator	Calculated per class	Eq. (9), drives sampling ratio
	$ROUND(IL)$	Rounded oversampling percentage	≈ 150 %	Derived from Eq. (9) after rounding
GA	$cump(j) = \sum_{i=1}^{j} Q(i)$	Cumulative probability in roulette selection	Computed	Eq. (10)
	$Q(i) = \dfrac{k+1-f(x_i)}{\sum (k+1-f(x_i))}$	Selection probability for minority sample (x_i)	Computed	Eq. (11)
	$f(x_i)$	Fitness function of minority sample (x_i)	Computed	Eq. (11)
	$u' = \left(\dfrac{d_1}{d_{xy}}\right) \cdot \left(\dfrac{d_2}{d_{xz}}\right)$	Relative distance ratio (posEff) for sample suitability adjustment	Computed	Eq. (12)

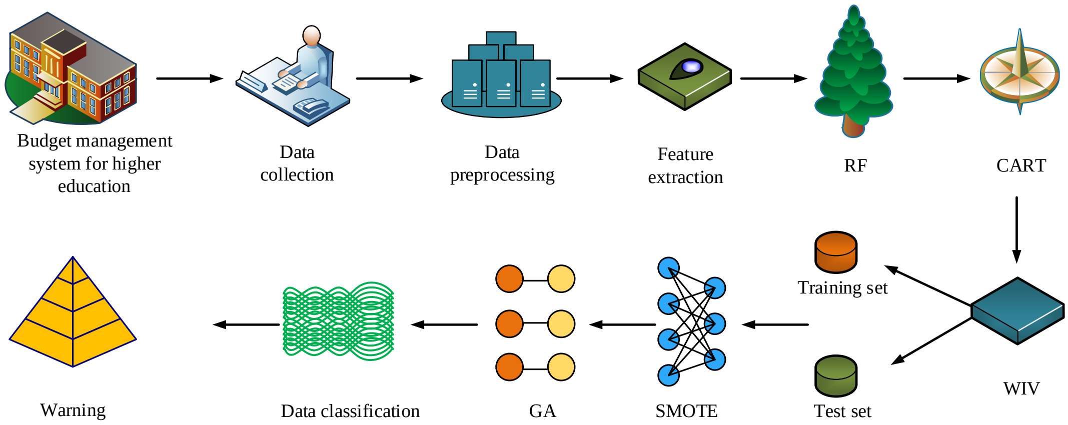


Figure 6: A novel early warning modeling process for budget management information in higher education institutions

As shown in Figure 6, firstly, collect all relevant data from the university budget management system. Secondly, perform data preprocessing to extract useful features such as quantity and project names. Thirdly, use RF algorithm to evaluate the importance of features, select key features, and introduce CART to optimize the decision tree structure of RF, while introducing WIV to optimize the classification process. Fourthly, divide the dataset into a training set and a testing set for model training and parameter adjustment. Fifthly, to handle imbalanced data, introduce the SMOTE algorithm to generate synthesized minority class samples, combined with genetic algorithm to optimize sample selection, improving the efficiency and accuracy of SMOTE. Sixthly, deploy the optimized RF-CART-WIV-SMOTE-GA model to the university budget management system. Perform real-time data classification and warning, continuously monitor and evaluate the effectiveness of the model, and regularly update it to adapt to new data and requirements. The algorithm pseudocode is as follows:

Algorithm 1: RF-CART-WIV-SMOTE-GA Budget Information Early Warning Model

---

Input: Original budget dataset D (IPEDS or HESA)

Output: Predicted budget category and early warning signal

// Data Pre-processing

Clean D; handle missing values (by mean/mode or KNN); standardize continuous features; encode categorical ones.

Detect outliers using 3σ and IQR rules; remove extreme points.

// Feature Selection

Use Random Forest importance ranking to select top-K budget-related features.

// Model Training with CART and WIV

For each bootstrap subset D_B (b = 1 … B):

Train CART tree using GainRatio and Gini criteria (Eqs. 2–5);

Compute weighted vote ω_i and confidence u_i(D) (Eqs. 6–7).

Aggregate all trees using Weighted Integration Voting.

// Data Balancing via SMOTE + GA

For each minority class sample X:

Find k nearest neighbors v_i; generate synthetic sample Q_i = X + rand(0,1)·(v_i − X) (Eq. 8);

Compute imbalance level IL and ROUND(IL) (Eq. 9).

Use GA to optimize sample selection probability Q(i) and fitness f(x_i) (Eqs. 10–11);

Adjust by posEff distance ratio u' (Eq. 12).

// Model Integration and Prediction

Combine balanced dataset and trained RF-CART-WIV model;

Perform prediction and compute error metrics (Accuracy, Recall, F1 score, and MAE).

Return classified budget stage and early warning signal.

2.3 Model deployment and system architecture design

To validate the proposed model's integrability and applicability within university budget management information systems, the study further designs a system deployment architecture. The overall system adopts a four-layer structure comprising:

Data Collection Layer: Responsible for automatically extracting raw budget data from university financial databases and business management platforms, including project IDs, expenditure categories, amounts, and timestamps. It is connected to the system through ETL connection.

Algorithm Service Layer: Integrates the proposed RF-CART-WIV-SMOTE-GA model to provide budget information classification and risk warning services. This module is implemented using Python and the Scikit-learn framework, supports RESTful API calls, and can be deployed on Linux servers or university private cloud environments.

Application Interface Layer: Seamlessly connects with existing budget management systems via an API gateway. Users can invoke functional modules such as “Budget Anomaly Detection,” “Expenditure Trend Forecasting,” and “Early Warning Report Generation” through the system frontend, enabling real-time access to model outputs.

Visualization Layer: Presents model warning results and performance metrics through a web-based visual dashboard, supporting dynamic filtering and report exporting by department, project, and time dimension.

Furthermore, to ensure fairness and transparency in actual budget decision-making, the research specifically addresses data governance and algorithmic robustness during system design. Firstly, all input data undergoes anonymization and standardization during model training to prevent bias toward specific groups due to differences in department size or funding structure. Secondly, the system provides interpretable outputs in result presentation, including key feature contribution visualization and decision path visualization, to enhance the transparency and traceability of budget alerts. Simultaneously, to prevent false positives caused by external adversarial inputs (such as anomalous or misleading budget entries), the study introduces an anomaly detection mechanism based on Z-scores and IQR during data preprocessing. At the model level, a multi-round voting strategy using a RF sub-model is employed to mitigate noise interference.

3 Results

To comprehensively evaluate the model's performance in early warning for university budget management information, the study selected four core classification metrics: accuracy, precision, recall, and F1 score. Accuracy reflects the overall proportion of correct judgments made by the model. Precision measures the reliability of warning results. Recall indicates the model's sensitivity in identifying abnormal budgets. F1 score, as the harmonic mean of precision and recall, balances the detection capability and false alarm control. Additionally, to validate the model's feasibility in practical scenarios, MAE and sample balance rate are employed to measure the deviation of warning results and data balance.

3.1 Performance testing of university budget management information warning model

The research sets the CPU to an 8-core AMD Ryzen 7 4800H processor, GPU to NVIDIA GeForce, and memory to 64GB. The operating system is Ubuntu 20.04, and the experimental software includes Python 3.8 and the Scikit-learn library. The Integrated Postsecondary Education Data System (IPEDS) and the Higher Education Statistics Agency Finance Dataset (HESA) are used as the testing data sources. The above datasets have undergone data preprocessing, including data cleaning, data conversion, data integration, and data specification. These two datasets are divided into training and testing sets in an 8:2 ratio, used for ensemble training of the initial model. Both datasets exhibit significant category imbalance. Taking the budget execution phase as an example, the ratio of normal to abnormal samples is approximately 7:1 in the IPEDS dataset and 6:1 in the HESA dataset. Some sub-categories (such as special construction or research budget anomalies) account for less than 10% of the samples. Therefore, the SMOTE algorithm was introduced during model training to mitigate data bias, combined with GA optimization for sample selection to enhance the recognition capability of minority class. Additionally, hierarchical random segmentation was used to maintain consistency in the proportion of training and testing set categories and budget stages. The random seed was set to 42. All upscaling and parameter tuning were performed exclusively on the training set, with no resampling applied to the test set. To prevent data leakage from feature overlap, a dual-isolation strategy based on institution and time was implemented during data partitioning. Data from the same institution in the same year appeared exclusively in either the training or test set, with no cross-use. Cross-year samples were partitioned chronologically before stratified random sampling. This method ensures that the model cannot access any features originating from or appearing simultaneously with the training set during testing, thereby ensuring the independence and reliability of the evaluation results. The study first conducts ablation testing on the final model to verify the effectiveness of various modules. The test results are shown in Figure 7.

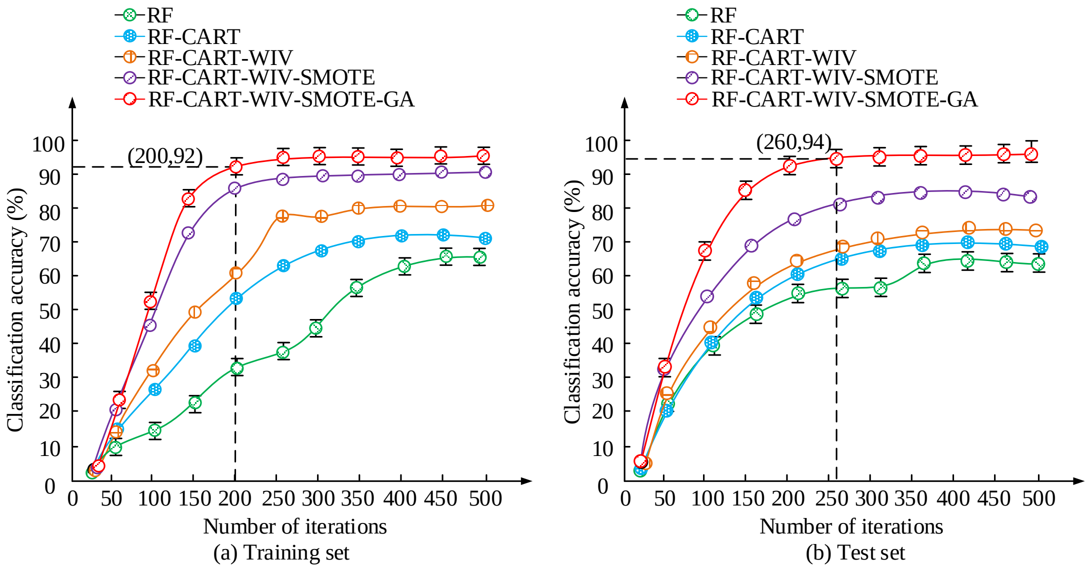


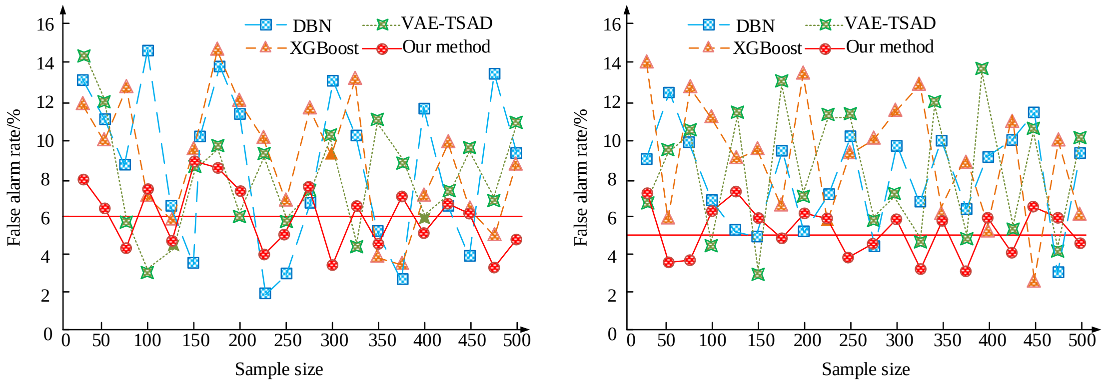
Figure 7: Ablation test results for a new budget information early warning model

Figure 7(a) shows the data classification test results of the new budget warning information model in the training set. Figure 7(b) shows the data classification test results of the new budget warning information model in the testing set. The horizontal axis represents model structure combinations, while the vertical axis shows average classification accuracy (unit: %). In the figure, “RF,” “CART,” “WIV,” “SMOTE,” and “GA” denote Random Forest, Classification and Regression Trees, Weighted Integrated Voting, Synthetic Minority Over-sampling Technique, and Genetic Algorithm modules, respectively. Budget data encompasses four categories: budget formulation expenditure items, fund disbursement records during implementation, budget adjustment projects, and annual evaluation indicators. From Figure 7(a), the budget information classification accuracy of the individual RF model gradually increased after training, with a maximum classification accuracy of 68%. After sequentially introducing CART, WIV, SMOTE, and GA, the budget information classification accuracy of the final warning model reached up to 92%. In Figure 7(b), similar to the training set, the introduced new modules in the testing set had a significant improvement effect on the final model, especially the RF-CART-WIV-SMOTE-GA model, which had a budget information classification accuracy of up to 94% and a minimum of 260 iterations. To further validate the independent contributions and stability impacts of each improvement module on model performance, multiple ablation experiments were conducted, with results shown in Table 3.

As shown in Table 3, the baseline RF model achieved an average classification accuracy of only 68.21%. After incorporating CART structure optimization, the model's feature partitioning capability significantly improved, and the accuracy increased to 78.62%. When the WIV module was further introduced, the model's confidence in voting decisions increased, with all metrics improving by approximately 7%. Subsequently, the SMOTE algorithm improved sample balance, resulting in a significant increase in recall and F1 score, with an average accuracy of 90.31%. Finally, after incorporating GA optimization, the model achieved optimal overall performance, with average accuracy rising to 93.94% and standard deviation decreasing to 0.40, indicating greater stability in model results. Overall, the sequential integration of modules not only improves classification accuracy and recall, but also effectively reduces model uncertainty. This demonstrates that the proposed RF-CART-WIV-SMOTE-GA model outperforms other combinations in stability, generalization capability, and handling imbalanced data, validating its design rationality and practical application value. To validate the proposed model's effectiveness, three representative advanced models were selected for comparison: Deep Belief Network (DBN), which possesses strong multi-layer feature learning capabilities, Extreme Gradient Boosting (XGBoost), which demonstrates stable performance in structured data classification, and Variational Autoencoder-Based Time Series Anomaly Detection (VAE-TSAD), known for high accuracy in time series modeling and anomaly detection. All three models represent typical algorithmic frameworks in the field of budget information early warning, serving to comprehensively evaluate the performance advantages of the proposed model. The test results are shown in Figure 8.

Table 3: Results of ablation experiments with standard deviation and confidence interval

Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Std. (σ)	95% CI (±)
RF	68.21 ± 0.74	67.85 ± 0.69	68.13 ± 0.82	67.98 ± 0.77	0.76	±1.48
RF + CART	78.62 ± 0.65	78.15 ± 0.71	77.83 ± 0.67	77.98 ± 0.69	0.68	±1.31
RF + CART + WIV	85.74 ± 0.57	85.33 ± 0.62	85.26 ± 0.60	85.29 ± 0.59	0.59	±1.15
RF + CART + WIV + SMOTE	90.31 ± 0.49	89.96 ± 0.55	90.08 ± 0.53	90.02 ± 0.51	0.52	±1.02
RF + CART + WIV + SMOTE + GA (Proposed)	93.94 ± 0.38	93.57 ± 0.42	93.21 ± 0.40	93.39 ± 0.41	0.41	±0.79



(a) IPEDS (b) HESA

Figure 8: Comparison results of budget information warning error for different models

Figure 8(a) and Figure 8(b) respectively compare the early warning error performance of four models (DBN, XGBoost, VAE-TSAD, and the proposed RF-CART-WIV-SMOTE-GA model) on budget execution phase data from the IPEDS and HESA datasets. The horizontal axis represents test sample IDs, while the vertical axis shows the error rate (unit: %). The legends "DBN," "XGBoost," "VAE-TSAD," and "RF-CART-WIV-SMOTE-GA" represent the four models, respectively. Budget execution data originates from fund allocation and utilization records during the budget execution process of higher education institutions. The error rate is measured using the MAE. As shown in Figure 8(a), with the continuous increase of the sample size of test information, the four models exhibited certain stability fluctuations in early warning information processing. The budget information warning error rate of DBN and XGBoost models was as high as 15%, and as low as 2%. Although the VAE-TASD model had a smaller warning error compared with the previous two, its early data fluctuated greatly. The warning error fluctuation range of the proposed model was between 4% and 9%, and its performance was more stable. According to Figure 8(b), the data performance of the four models was similar to Figure 8(a). However, in terms of details, the warning error range of the proposed model was reduced to 3.7%-7.5%. The study conducts tests using Precision, Recall, F1 score, and Sample Equilibrium rate as indicators. Table 4 displays the test results.

Table 4: Multi-metric test results for four models

Data set	Model	Precision (P, %)	Recall (R, %)	F1 score (%)	Sample equilibrium rate (%)	95% CI (F1)	*p*-value (vs. Our model)
IPEDS	DBN	89.97 ± 0.42	87.62 ± 0.51	88.79 ± 0.48	75.34	[88.42, 89.16]	<0.01
	XGBoost	90.18 ± 0.36	88.87 ± 0.47	89.52 ± 0.41	74.21	[89.11, 89.93]	<0.01
	VAE-TSAD	91.23 ± 0.38	90.24 ± 0.43	90.73 ± 0.45	73.59	[90.29, 91.17]	0.013
	RF-CART-WIV-SMOTE-GA	93.57 ± 0.29	92.18 ± 0.33	92.87 ± 0.31	80.36	[92.55, 93.19]	/
HESA	DBN	90.98 ± 0.40	89.36 ± 0.44	90.17 ± 0.38	75.29	[89.80, 90.54]	<0.01
	XGBoost	89.14 ± 0.35	89.47 ± 0.41	89.31 ± 0.39	79.98	[88.91, 89.71]	<0.01

VAE-TSAD	91.33 ± 0.37	90.21 ± 0.42	90.77 ± 0.35	76.17	[90.39, 91.15]	0.016
RF-CART-WIV-SMOTE-GA	93.51 ± 0.31	91.46 ± 0.33	92.48 ± 0.28	81.64	[92.14, 92.82]	/

From Table 4, there were statistically significant differences in the main classification indicators of the four models in multiple experiments (10 independent runs). Taking the F1 score as an example, on the IPEDS dataset, the average F1 score of the proposed RF-CART-WIV-SMOTE GA model was 92.87%, the standard deviation was ±0.31%, and the 95% confidence interval was [92.55, 93.19]. It was significantly superior to VAE-TSAD (p=0.013) and other baseline models (p<0.01). On the HESA dataset, the F1 score of the proposed model was 92.48%, showing a statistical advantage in the significance test (p<0.05). In addition, the sample balance rate of the model was the highest in both sets (80.36% and 81.64%), with the improvement rates being 7.01% and 5.47% higher than those of other methods, respectively. These results indicate that the proposed model has statistically significant improvements in both classification accuracy and data balance.

3.2 Simulation testing of information warning model for university budget management

To verify the budget information management warning results of the proposed model in a real university environment, Chengdu Normal University is randomly selected as the test object. The school covers an area of 1,100 acres, which is a public full-time undergraduate institution sponsored and managed by the People's Government of Sichuan Province. The school offers 39 undergraduate majors, with over 17,000 students and more than 1,000 faculty members. The school adheres to the budget management philosophy of unified management and full responsibility integration. The study tests the information warning of the four stages of budget preparation, budget execution, budget adjustment, and budget assessment in the school. The test results are shown in Figure 9.

Figure 9(a) shows the absolute error test results of information management warning for budget preparation. Figure 9(b) shows the absolute error test results of information management warning for budget execution. Figure 9(c) shows the absolute error test results of information management warning for budget adjustment. Figure 9(d) shows the absolute error test results of information management warning for budget assessment. Figure 9 compares the absolute errors of budget information early warning for four models in the four stages of budget preparation, execution, adjustment and evaluation. The horizontal axis represents the budget stage, and the vertical axis represents the absolute error (unit: %). From Figure 9, compared with the first three links, the budget assessment had the smallest absolute error in early warning for each model due to clear indicators and less external interference. During the budget preparation stage, multiple factors such as revenue and expenditure balance, power and responsibility allocation, and performance orientation need to be comprehensively considered. The information complexity is the highest, and thus deviations are more likely to occur. Among the four models, the proposed model had the lowest early warning absolute errors in the four stages of budget preparation, execution, adjustment and evaluation, which were 4%, 2.5%, 5% and 3%, respectively. It is worth noting that this improvement in classification accuracy is not only reflected in the performance of the algorithm, but also plays a significant role in actual decision-making. It can identify potential overspending projects in advance during the budget preparation stage, provide early warning of abnormal funds during the execution stage, assist in reallocating funds during the adjustment stage, and improve the objectivity of performance evaluation during the evaluation stage. It can be seen from this that this model can effectively support the dynamic management and refined decision-making throughout the entire budget process. The VAE-TSAD model, which performs well, is compared with the proposed model for information confusion classification warning. The test results are shown in Figure 10.

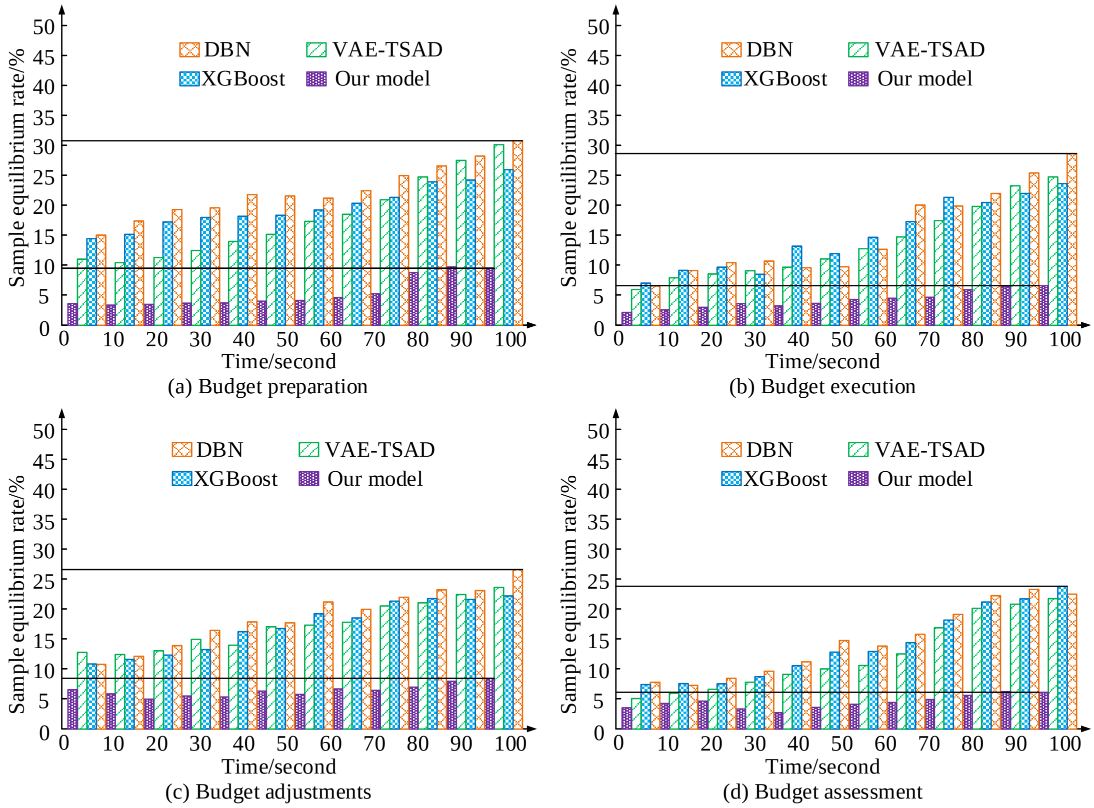


Figure 9: Results of different models of absolute errors in early warning for four types of budget link information

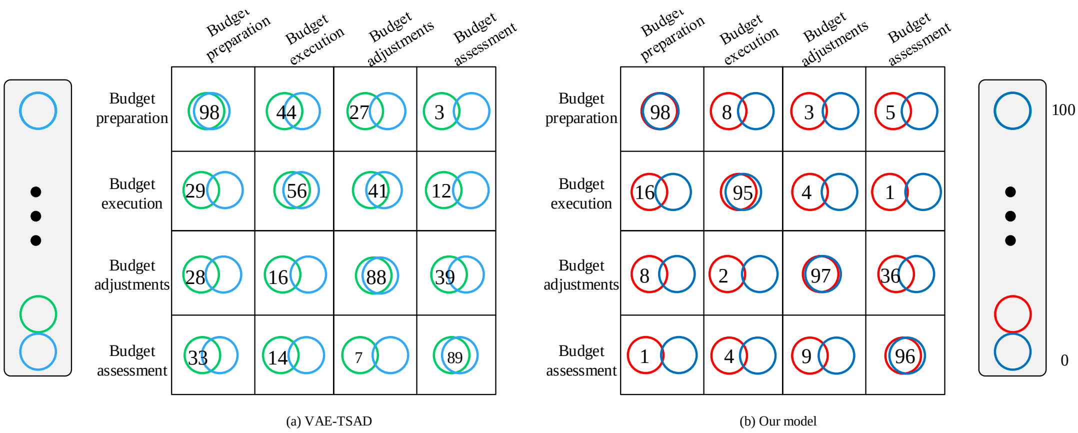


Figure 10: Information warning confusion matrix results for two types of warning models

Figure 10(a) shows the information warning classification results of the VAE-TASD model for four types of budget processes. Figure 10(b) shows the information warning classification results of the proposed model for four types of budget processes. The horizontal axis represents the predicted category, the vertical axis represents the true category, and the values in the matrix represent the classification accuracy at each stage. From Figure 10(a), the VAE-TASD model performed well in the correlation of information warning classification results for the four types of budget processes, but the warning results for some budget information were poor, such as confusion between budget adjustment and budget adjustment, and confusion between budget preparation and budget adjustment. The test classification results of the four types of budget stages are generally high, achieving over 90% effectiveness in information warning. Especially for budget execution and budget adjustment, the effectiveness of information classification warning for both reached 95%. To further verify the overall discriminative performance of the model, the ROC curves of four models on the HESA dataset were plotted, as shown in Figure 11. The curve shape and AUC value reflect the model's recognition ability under different discrimination thresholds.

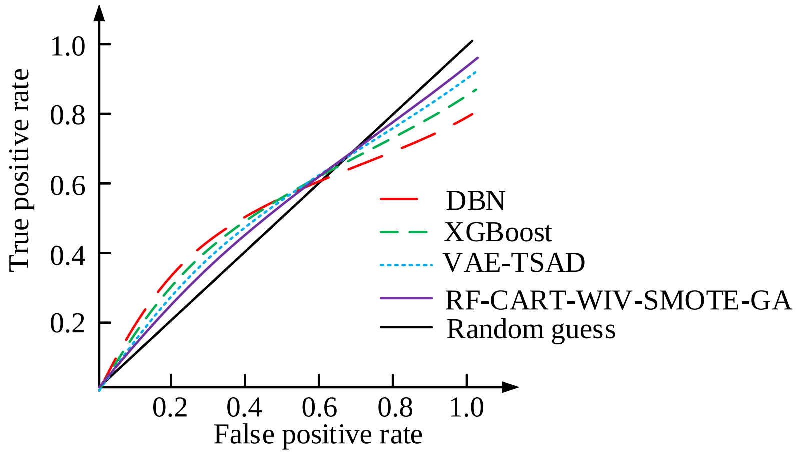


Figure 11: ROC curve comparison of four models on the HESA dataset

Figure 11 compares ROC curves of four models (DBN, XGBoost, VAE-TSAD, and RF-CART-WIV-SMOTE-GA) in the budget information classification task on the HESA dataset. The horizontal axis represents the False Positive Rate (FPR), and the vertical axis represents the True Positive Rate (TPR). The area under the curve (AUC) reflects the overall classification and discrimination ability of the model. From Figure 11, there were significant differences in the overall discriminative performance of the four models in the budget information classification task. The ROC curve of the RF-CART-WIV-SMOTE-GA model was always above that of other models, and its AUC value was the highest, reaching 0.962. This indicates that this model maintains a high TPR and a low FPR under different discriminant thresholds, and has a stronger comprehensive discriminant ability. In contrast, the AUC values of the VAE-TSAD, XGBoost and DBN models were 0.935, 0.913 and 0.887, respectively. The overall curves deviated from the ideal classifier line, indicating that their stability in identifying abnormal budget samples was relatively weak. The comprehensive results show that the proposed model not only has higher classification accuracy in budget information early warning, but also has better robustness and generalization performance. The test takes indicators such as warning success rate, warning time, CPU utilization, and resource utilization, as shown in Table 5.

Table 5: Multi-metric test results for different models

Dataset	Model	Early warning success rate (%)	Warning time (s)	CPU utilization (%)	Resource utilization (%)	95% CI (Success rate)	*p*-value (vs. Our model)
IPEDS	DBN	90.12 ± 0.43	0.08 ± 0.01	46.63 ± 1.52	77.62 ± 1.26	[89.69, 90.55]	<0.01
	XGBoost	91.33 ± 0.37	0.11 ± 0.02	37.89 ± 1.21	81.23 ± 1.08	[90.96, 91.70]	0.012
	VAE-TSAD	91.58 ± 0.39	0.07 ± 0.01	35.66 ± 1.09	84.79 ± 1.17	[91.20, 91.96]	0.015
	RF-CART-WIV-SMOTE-GA	93.87 ± 0.28	0.04 ± 0.01	24.17 ± 0.94	89.74 ± 0.83	[93.59, 94.15]	/
HESA	DBN	90.52 ± 0.45	0.13 ± 0.02	45.84 ± 1.48	79.24 ± 1.14	[90.07, 90.97]	<0.01
	XGBoost	91.77 ± 0.36	0.09 ± 0.01	43.11 ± 1.32	79.89 ± 1.02	[91.42, 92.12]	0.01
	VAE-TSAD	92.59 ± 0.33	0.08 ± 0.01	36.79 ± 1.18	86.21 ± 0.97	[92.27, 92.91]	0.018
	RF-CART-WIV-SMOTE-GA	94.11 ± 0.27	0.05 ± 0.01	27.52 ± 0.85	89.33 ± 0.91	[93.84, 94.38]	/

From Table 5, after 10 independent repeated experiments, the four models showed statistical differences in the early warning performance indicators. The average early warning success rates of the RF-CART-WIV-SMOTE-GA model on the IPEDS and HESA datasets were 93.87% and 94.11%, respectively, and their 95% confidence intervals were [93.59, 94.15] and [93.84, 94.38]. All were significantly higher than those of other models ($p<0.05$). Meanwhile, the average early warning time of this model was the shortest (0.04-0.05s), the CPU occupancy rate was the lowest (24.17% - 27.52%), and the resource utilization rate was the highest (about 89%), indicating that it had statistical advantages in terms of operational efficiency and energy consumption control. The overall results show that the proposed model has achieved significant improvements in both early warning accuracy and computational efficiency.

4 Discussion

The RF-CART-WIV-SMOTE-GA model demonstrated significant advantages in budget information early warning. Its average classification accuracy could reach up to 94%, with an AUC value of 0.962, showing stable statistical significance in multiple independent experiments ($p<0.05$). Compared with the SOTA methods listed in Table 1 and Table 2, this model had outstanding performance in handling the imbalance of budget data and improving the real-time performance of early warnings. Especially compared with the DBN model, the RF framework can more effectively capture the nonlinear relationships between features. Through the structural optimization of CART and WIV, the model reduces overfitting and improves the discrimination accuracy in the small sample budget stage, thereby maintaining a high recall rate and sample balance rate even on imbalanced datasets. Further analysis indicates that the SMOTE-GA is one of the key reasons for the performance improvement. Compared with the uniform oversampling method of the common SMOTE, GA optimizes the sample selection process, making the synthetic samples closer to the high-fitness region, thereby avoiding noise diffusion and sample redundancy, and improving the data quality and model generalization ability. Meanwhile, although VAE-TSAD has a relatively good reconstruction ability under some conditions, it is relatively sensitive to input noise and distribution drift, resulting in large performance fluctuations during the budget execution or adjustment stage. Overall, the RF-CART-WIV-SMOTE-GA model is superior to existing methods in terms of stability and interpretability, providing a more practical solution for the intelligent early warning of university budget management.

5 Conclusion

The information asymmetry and low transparency in university budget management have always been a challenge in budget management. Based on the RF classification model, a budget information classification model was constructed by combining CART and WIV, and optimized by SMOTE and GA. Finally, a new university budget information management warning model was proposed. The experimental results showed that the budget information classification accuracy of the designed warning model reached 94%, with a minimum of 260 iterations. Compared with the DBN, XGBoost, and VAE-TASD models, the new model had a minimum warning error range of 3.7%-7.5%. Compared with more similar advanced methods, the maximum reduction was 7.5%. The highest P-value was 93.57%, the highest R-value was 92.18%, the highest F1 score was 92.87%, and the highest sample equilibrium rate reached 81.64%. Among them, the sample equilibrium rate has increased by 7.01% compared with the top three models. After testing the four types of budget information warning at Chengdu Normal University, the new model had the lowest absolute error in warning for the four stages of budget preparation, budget execution, budget adjustment, and budget assessment, which were 4%, 2.5%, 5%, and 3%, respectively. Especially in budget execution and budget assessment, the effectiveness of their early warning exceeded 95%. The optimal values for the success rate, warning time, CPU utilization, and resource utilization of this new model were 94.11%, 0.04 seconds, 24.17%, and 89.74%, respectively. In summary, the proposed warning model can effectively enhance the information warning capability of university budget management, with significant warning accuracy and real-time performance. However, the research still has certain limitations. The model training is mainly based on historical financial data, and its adaptability to policy changes, macroeconomic fluctuations and inter-institutional differences still needs further verification. In future research, it can be considered to introduce cross-institutional multi-source heterogeneous data to enhance the model's adaptability to different educational systems and financial structures. Meanwhile, by integrating causal reasoning and explainable learning frameworks, the cause analysis and prediction mechanism of budget anomalies are further explored, thereby achieving dynamic, transparent and intelligent support for the entire process of budget management in colleges and universities.

References

[1] Kumar S, Sharma D, Rao S, Lim W M, Mangla S K. Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. Annals of Operations Research, 2025, 345(2): 1061-1104. DOI:10.1007/s10479-021-04410-8

[2] Lei RD, Berto F, Hu C, Lu ZH, Yan XF. Early-warning signal recognition methods in flawed sandstone subjected to uniaxial compression. Fatigue Fract Eng M. 2023; 46(3): 955-974. DOI:10.1111/ffe.13911

[3] Teng YS, Zhang J, Sun T. Data-driven decision-making model based on artificial intelligence in

higher education system of colleges and universities. Expert Syst. 2023; 40(4): e12820. DOI:10.1111/exsy.12820
[4] Valle-Cruz D, Fernandez-Cortez V, Gil-Garcia JR. From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation. Gov Inform Q. 2022; 39(2): 101644. DOI:10.1016/j.giq.2021.101644
[5] Truong NVQ, Dinh LC, Stein S, Tran-Thanh L, Jennings NR. Efficient and adaptive incentive selection for crowdsourcing contests. Appl Intell. 2023; 53(8): 9204-9234. DOI:10.1007/s10489-022-03593-2
[6] Zheng ZH, Zhou YZ, Sun YL, Wang Z, Liu BY, Li KQ. Applications of federated learning in smart cities: recent advances, taxonomy, and open challenges. Connect Sci. 2022; 34(1): 1-28. DOI:10.1080/09540091.2021.1936455
[7] Swinney CJ, Woods JC. K-means clustering approach to UAS classification via graphical signal representation of radio frequency signals for air traffic early warning. IEEE T Intell Transp. 2022; 23(12): 24957-24965. DOI:10.1109/TITS.2022.3202011
[8] Bénard C, Da Veiga S, Scornet E. Mean decrease accuracy for random forests: inconsistency, and a practical solution via the Sobol-MDA. Biometrika. 2022; 109(4): 881-900. DOI:10.1093/biomet/asac017
[9] Zhang ZS, Xiao Y, Fu ZT, Zhong KY, Niu HY. A study on early warnings of financial crisis of Chinese listed companies based on DEA–SVM model. Mathematics-Basel. 2022; 10(12): 2142. DOI:10.3390/math10122142
[10] Cao YL, Shao Y, Zhang HX. Study on early warning of E-commerce enterprise financial risk based on deep learning algorithm. Electron Commer Res. 2022; 22(1): 21-36. DOI:10.1007/s10660-020-09454-9
[11] Zeng H. Influences of mobile edge computing-based service preloading on the early-warning of financial risks. J Supercomput. 2022; 78(9): 11621-11639. DOI:10.1007/s11227-022-04329-2
[12] Ludkovski M. Statistical machine learning for quantitative finance. Annual Review of Statistics and Its Application, 2023, 10(1): 271-295. DOI:10.1146/annurev-statistics-032921-042409
[13] Alkan N, Kahraman C. Continuous Pythagorean Fuzzy Set Extension with Multi-Attribute Decision Making Applications. Informatica, 2025, 36(2): 241-283. DOI: DOI: https://doi.org/10.15388/25-INFOR58
[14] Anderer M, Li F. Hierarchical forecasting with a top-down alignment of independent-level forecasts. Int J Forecasting. 2022; 38(4): 1405-1414. DOI:10.1016/j.ijforecast.2021.12.015
[15] Chen ZQ, Han M, Wu HX, Li MH, Zhang XL. A multi-level weighted concept drift detection method. J Supercomput. 2023; 79(5): 5154-5180. DOI:10.1007/s11227-022-04864-y
[16] Hoang D, Wiegratz K. Machine learning methods in finance: Recent applications and prospects. European Financial Management, 2023, 29(5): 1657-1701. DOI: 10.1111/eufm.12408
[17] Aldhyani THH, Alzahrani A. Framework for predicting and modeling stock market prices based on deep learning algorithms. Electronics-Switz. 2022; 11(19): 3149. DOI:10.3390/electronics11193149
[18] George A S. Securing the future of finance: how AI, Blockchain, and machine learning safeguard emerging Neobank technology against evolving cyber threats. Partners Universal Innovative Research Publication, 2023, 1(1): 54-66. DOI:10.5281/zenodo.10001735
[19] Liu LBA, Chen C, Wang B. Predicting financial crises with machine learning methods. J Forecasting. 2022; 41(5): 871-910. DOI: 10.1002/for.2840
[20] Shi HT, Zhang HP, Huo DA, Yu S, Su J, Xie YC, Li W, Ma LH, Chen HQ, Sun YQ. An ultrasensitive microsensor based on impedance analysis for oil condition monitoring. IEEE T Ind Electron. 2022; 69(7): 7441-7450. DOI:10.1109/TIE.2021.3100982