Doi:doi.org/10.70693/cjst.vli.793 Vol. 1 No. 1(March and April, 2025)

# Crafting an Efficient Credit Risk Alert System: Assessment and Validation of the WOE-Enhanced Logistic Regression Model

Ran Meng-huan [1], Zhu Xiao-xin [1,*]

1. *Qingdao University of Technology, School of Business, Qingdao, Shandong, China, 266520*

**Abstract:** In the current development landscape of the credit industry, risk management faces a series of challenges. Although technological advancements have brought significant progress to this field, issues such as high labor costs and insufficient customer authentication persist, highlighting the urgent need to build efficient risk prediction models. Taking the GiveMeSomeCredit dataset on the Kaggle platform as an example, this study applies feature-engineering techniques to develop a debt default early warning model aimed at identifying potential credit risks in advance. By combining in-depth optimization of IV and WOE values, logistic regression models and an LR-WOE model were constructed. Comprehensive evaluation using metrics such as PSI, KS statistic, and AUC scores ensured the robustness of the models' risk prediction accuracy. The research findings reveal: (1) Family structure plays a crucial role in credit risk assessment. Applicants with 0 or 7 dependents exhibit a higher probability of default compared to the overall sample, while those with 6 or 8 dependents demonstrate relatively lower default risk. (2) The constructed LR-WOE model performed the best, indicating its effectiveness in distinguishing borrowers with different credit profiles and maintaining stable predictive performance across various thresholds. Integrating WOE transformation techniques with logistic regression models can help financial institutions assess credit risks more accurately and optimize risk management strategies.

**Keywords:** Debt default; Risk management; LR-WOE; Feature engineering

Personal debt default risk is a critical factor in financial market stability, directly impacting the asset quality and profitability prospects of banks as lending institutions. Amid fluctuations in the global economy and credit markets—particularly the rise in default rates due to unemployment and income declines following the COVID-19 pandemic—this risk has become increasingly prominent. In the face of potential default, banks not only confront asset impairment and potential credit losses but may also be compelled to restrict lending to consumers and businesses due to credit tightening, thereby limiting economic activity and growth. Moreover, an increase in personal debt defaults can negatively affect consumer psychology, leading to reduced spending willingness. This not only has long-term consequences for borrowers' creditworthiness but may also trigger a chain reaction at the societal level, exacerbating inequality and instability. Therefore, effective early warning systems and risk control measures for debt defaults are not only crucial for banks but also serve as essential safeguards for macroeconomic stability, consumer protection, and societal well-being. Additionally, by accurately assessing and monitoring debt risks, policymakers can formulate effective macroeconomic policies to stabilize financial markets and support long-term healthy economic development. Strengthening personal credit risk management, developing advanced early warning systems, and refining risk control strategies hold profound significance for banks, policymakers, and society at large—collectively forming the cornerstone of financial stability and economic prosperity.

In previous research, the logistic regression model has demonstrated its robust predictive capabilities and practical efficacy across multiple critical domains, including medical safety[1], and geological hazard prevention[2,3]. Notably, Costa e Silva, Eliana, et al. analyzed credit scoring data from a Portuguese financial institution using a logistic regression model to build a consumer loan default risk prediction model, the model achieved a default prediction accuracy of 89.79%, validating its effectiveness[4]. In the realm of financial credit risk management, logistic regression has been extensively and profoundly applied, substantially improving financial institutions' ability to assess borrower credit risk with precision. Moreover, it has refined credit decision-making processes, significantly optimizing the efficiency and effectiveness of risk management strategies. From Wu Yi and Pan Yuwen's proposal of using the pdC-RF algorithm for dimensionality reduction on data and applying WOE encoding to compare the performance of random forest, support vector machine, and logistic regression models[5]. To Dumitrescu, Elena, and others' suggestion of extracting rules from short-depth decision trees as predictors while maintaining the interpretability of the logistic regression model to improve model performance[6], credit assessment technology has progressively evolved from traditional methods toward more sophisticated algorithmic integration. Wei Ying and Hasan Hafnida's research highlights the advantages of applying the logistic regression model in bank credit risk management. Through empirical studies, they demonstrate that the model has high accuracy in credit risk measurement and significant potential for further development[7]. The research by Saha, Partha, et al. that the logistic regression scoring card performs well in automatically predicting the fraud risk level, risk impact, and ease of detection in bank loan processing[8]. Mao Yi et al. introduced Nadaray-Watson density estimation to address the nonlinear classification problem in logistic regression, and demonstrated that logistic regression has good model interpretability[9]. Meanwhile, Dzik-Walczak Aneta et al. demonstrated the high performance of hybrid logistic regression-neural network models in P2P lending credit scoring[10]. Shuqi Liang et al. (2022) expanded logistic regression's application by incorporating multi-dimensional data, leveraging WOE transformation and IV-based variable screening to offer novel perspectives on personal credit scoring and risk quantification[11]. Existing research not only underscores the widespread adoption of logistic regression in credit risk assessment but also traces the field's technological evolution—from standalone models to ensemble approaches, and from conventional evaluation to data-driven, complex analytics. It reflects the relentless innovation by scholars worldwide in refining risk assessment methodologies, equipping financial institutions with increasingly precise and efficient risk management tools.

Although existing research has demonstrated the widespread use and application of logistic regression models in the credit industry, and has shown the evolution of technology from single models to integrated approaches, and from traditional assessments to data-driven methods, challenges still remain in the predictive capability of these models when adapting to changing economic environments or diverse customer groups. For instance, current models often struggle with generalization when handling high-dimensional data[12] or tend to overfit when facing imbalanced datasets[13]. Additionally, there are limitations in terms of interpretability and real-time prediction capabilities. In light of these issues, this study aims to improve the performance of logistic regression models in the field of debt default prediction by introducing new WOE encoding methods and model optimization techniques, thereby developing a more accurate, interpretable, and robust predictive model. This will help reduce default risks in the credit industry while providing new perspectives and tools for risk management in the credit sector.

## 1. Debt default model based on LR model

### 1.1 LR model

Logistic Regression is a binary classification model based on linear regression, where the input function values are mapped to a range through the sigmoid function. The study uses the sigmoid function to represent the probability of various credit card frauds, with $y$ as the binary dependent variable, indicating whether there has been a delinquent behavior of more than 90 days or longer overdue. $y=0$ represents no delinquent behavior, and $y=1$ represents delinquent behavior[14].

The binary logistic regression model can be expressed as :

$$P\left(y \mid X;\theta\right)=g\left(\theta^{T}X\right)=\frac{1}{1+e^{-\theta^{T}X}}\in\left(0,1\right), \tag{1}$$

Among them $y\in\{0,1\}$, $g\left(X\right)=\frac{1}{1+e^{-X}}$. The decision function is:

$$y=\begin{cases}0, P\left(y=1 \mid X\right)\leq 0.5\\ 1, P\left(y=1 \mid X>0.5\right)\end{cases}, \tag{2}$$

The parameters can be determined using maximum likelihood estimation $\theta$ [15].

### 1. 2 Model building

First, no preprocessing is done on the features, and no variable selection is performed on the model. The training data is directly input into the LR model, and the model's performance is tested using the test data to predict whether there has been a delinquent behavior of more than 90 days or longer overdue. Several important model evaluation metrics are used, including precision, recall, F1-score, and support. Precision refers to the proportion of samples predicted as the positive class that are actually of the positive class, with the positive class being the label 1, indicating high credit risk customers. Recall refers to the proportion of samples that are actually of the positive class and are correctly predicted as the positive class by the model. F1-score is the harmonic mean of precision and recall, comprehensively considering both the accuracy and the recall of the model. Support refers to the number of samples in each class. Accuracy refers to the proportion of correctly predicted samples out of the total number of samples. The results are shown in Table 1.

From Table 1, it can be seen that the precision is 0.59, indicating that the model has a certain level of error when predicting the positive class. The recall is 0.06, suggesting that the model has weak predictive power for the positive class, with many positive samples not being successfully identified. The F1-score is 0.12, indicating that the overall performance of the model is relatively low. The support for the positive class (label 1) is 1406, and the support for the negative class (label 0) is 19561. The accuracy is 0.93, indicating that the model has a relatively high proportion of correct predictions overall. The macro average and weighted average are results obtained by averaging the metrics across all classes. The macro average calculates the arithmetic mean of the metrics for each class, while the weighted average gives more weight to the class with a larger sample size, based on the support for each class. As observed from the results in Table 1, the model's predictive performance for positive class samples is significantly weaker than that for negative class samples. Despite the overall high accuracy, its ability to identify high-risk clients still requires further optimization. This indicates that the baseline model has weak predictive power for the target variable "whether there has been delinquent behavior of more than 90 days or longer overdue." Therefore, the next step is to consider preprocessing the features before inputting them into the binary logistic regression model for prediction. The ROC curve of the model is then obtained and shown in Fig. 1.

**Table 1 Performance Evaluation Table of the LR Model**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.94 | 1 | 0.97 | 19561 |
| 1 | 0.59 | 0.06 | 0.12 | 1406 |
| *accuracy* | | | 0.93 | 20967 |
| *macro* | 0.76 | 0.53 | 0.54 | 20967 |
| *weighted* | 0.91 | 0.93 | 0.91 | 20967 |

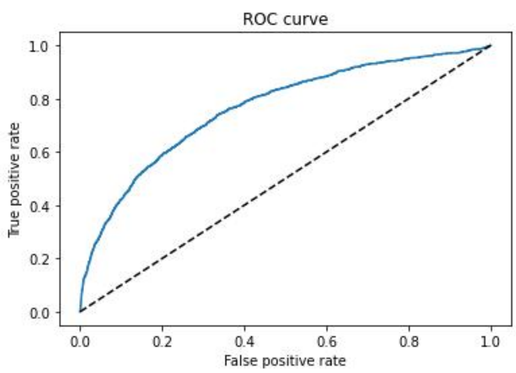


**Fig. 1 ROC Curve of the LR Model**

The AUC of the ROC curve is used to measure the performance of a machine learning model. AUC ranges from 0 to 1, and the larger the value, the better the performance of the machine learning classifier[14]. An AUC close to 1 indicates that the model has excellent performance and can effectively distinguish between positive and negative examples. An AUC close to 0.5 means that the model's performance is similar to random guessing, indicating almost no predictive capability. An AUC below 0.5 suggests that the model's performance is poor, even worse than random guessing. Fig. 1 shows that the AUC of the LR model is 0.78, indicating that the model performs reasonably well under the ROC curve, but further optimization is needed.

## 2. WOE adjustment of debt default model based on LR algorithm

### 2.1 WOE Coding

Binning and WOE encoding, based on machine learning methods, are used to mine sample features. First, binning is applied to discretize the data, followed by WOE encoding. Finally, the explanatory variables that are suitable for the model are selected. Binning is the process of converting continuous variables into discrete values, which involves grouping continuous data into several categories and then reclassifying the features based on those categories. After discretization, the within-group differences are small while the between-group differences are large, making the model more stable and reducing the risk of overfitting. After binning, the variables need to be encoded using WOE, which is a type of encoding for the original variables. To apply WOE encoding to a variable, it is necessary to perform an accurate and effective grouping of that variable[17]. The formula (3) is shown below:

$$WOE = \ln\left(\frac{Bad_i}{Bad_T} \Big/ \frac{Good_i}{Good_T}\right), \tag{3}$$

Formula (3) represents the difference between the “proportion of negative samples in the current group relative to all negative samples” and the “proportion of positive samples in the current group relative to all positive samples”.

Through the WOE formula, we found through observation that: in the same group, if the negative sample rate is low, the difference between different classes will become smaller, so the WOE value is smaller; conversely, if the negative sample rate is high, the difference between different classes will become larger, making the WOE value larger, which also means that the proportion of negative samples in the bin is higher. Through the WOE transformation, the feature values simultaneously represent the discriminative power between categories and their impact on the credit evaluation results. A positive WOE value indicates that the probability of an event in the group is higher than the probability of an event in the overall sample, which means that the group has a positive impact on the target variable; while a negative WOE value indicates that the probability of an event in the group is lower than the probability of an event in the overall sample, which means that the group has a negative impact on the target variable. The larger the absolute value of WOE, the stronger the predictive ability of the group for the target variable.

### 2.2 WOE encoding of data

Before calculating the WOE value, the IV value must first be calculated. The IV value is an indicator used to measure the predictive power of a variable, primarily for feature selection and variable screening. It evaluates the correlation between a variable and the target variable by calculating the information gain for each possible value of the variable. First, the WOE values are calculated for discrete variables. Among the existing features, there are 4 discrete features. Each unique value in these features is treated as a separate bin, and the corresponding WOE value is calculated for each. If bins with similar WOE values are found, they are merged. Next, the WOE encoding situation of three features is listed. The first feature is the number of family members, as shown in Fig. 2.

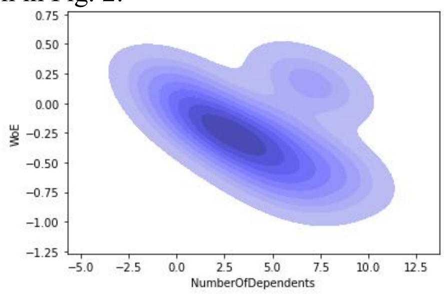


**Fig. 2 WOE Encoding Values Chart of Number of Dependents**

Fig. 2 is a 2D density plot of the number of family members and the corresponding WOE values, after excluding certain values. It can be observed that when the number of family members is 6 or 8, the default risk is very low. Only when the number of family members is 0 or 7, the WOE value is greater than 0, indicating that the probability of an event occurring in these two groups is higher than the overall sample's event occurrence probability. This means that these groups have a positive impact on the target variable. The WOE encoding results of the available credit ratio are further visualized and analyzed, as shown in Figure 3.

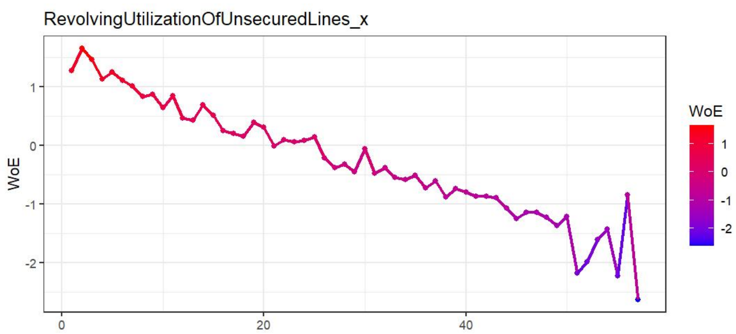


**Fig. 3 WOE Encoding Values Chart of Available Credit Ratio**

From the WOE calculation results of the available credit ratio shown in Fig. 3, it can be observed that the WOE value changes as the value increases. For example, in the lower range (such as below 0.02), the WOE is relatively high, while the WOE value starts to decrease once the value exceeds a certain threshold. The lower value range (e.g., below 0.5) is typically associated with higher WOE values, which may indicate that customers in this range are more likely to default compared to the overall sample. The higher value range (e.g., above 0.5) is typically associated with lower WOE values, suggesting that customers in this range are less likely to default compared to the overall sample. Next, the credit quantity is binned using single value bins, and the WOE values are calculated, as shown in Fig. 4.

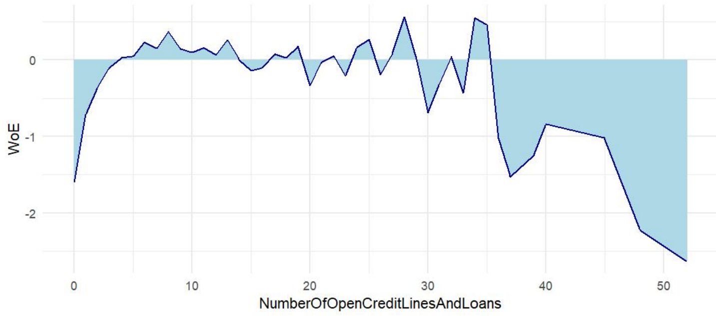


**Fig. 4 WOE Encoding Values Chart of Credit Quantity**

From Fig. 4, it can be seen that most of the WOE values are concentrated between -1 and 0.5. These values represent the relative differences in predictive power for default relative to the overall sample. When the credit quantity is relatively low, such as below 10, the corresponding WOE values are mostly positive, indicating that customers in this range are less likely to default compared to the overall sample.

**2.3 Establishment of LR-WOE model**

For the encoded continuous variables, manual binning is performed based on the frequency distribution of the features, while discrete variables are binned using single-value bins. The WOE values are then used to replace the corresponding features and input into the LR model for training, resulting in an LR-WOE model. The model is then used to predict the target values for the test set to determine whether there is any bad behavior, such as being overdue for more than 90 days or longer. The resulting confusion matrix of the model is shown in Fig. 5.

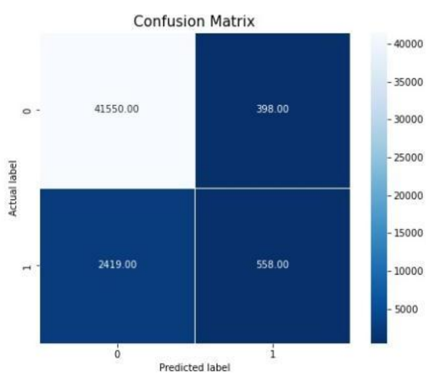


**Fig. 5 Confusion Matrix of the LR-WOE Model**

From the confusion matrix of the LR-WOE model, it can be seen that the number of samples that are actually negative class and correctly predicted as negative class by the model is 41,550. The number of samples that are actually negative class but incorrectly predicted as positive class by the model is 398. The number of samples that are actually positive class but incorrectly predicted as negative class by the model is 2,419. The number of samples that are actually positive class and correctly predicted as positive class by the model is 558. The evaluation table is shown in Table 2.

Table 2 indicates that the model performs well in predicting both the positive and negative classes, while balancing precision and recall. The ROC curve is then used to evaluate the performance of the

LR-WOE model, as shown in Fig. 6.

**Table 2 Evaluation Table of the LR-WOE Model**

| accuracy | precision | recall | f1-score |
|---|---|---|---|
| 0.937 | 0.991 | 0.945 | 0.967 |

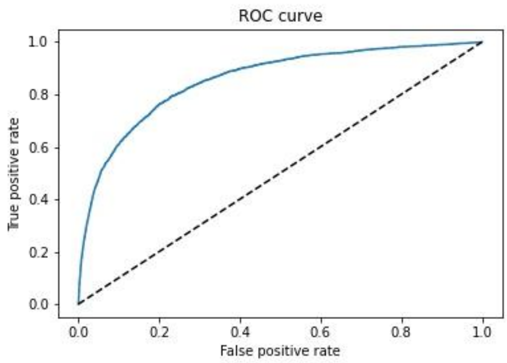


**Fig. 6 ROC Curve of the LR-WOE Model**

Fig. 6 shows that the AUC score of the model is 0.86, indicating that the model's performance is relatively good. The area under the ROC curve is much larger than the level of random guessing, which would have an AUC of 0.5. This suggests that the model is capable of distinguishing between positive and negative samples effectively, and its predictions are highly reliable.

**2.4 Debt Default LR-WOE Model after Feature Screening**

To understand the relationships between the features, a correlation analysis is performed on all features except the target variable. The resulting heatmap is shown in Fig. 7.

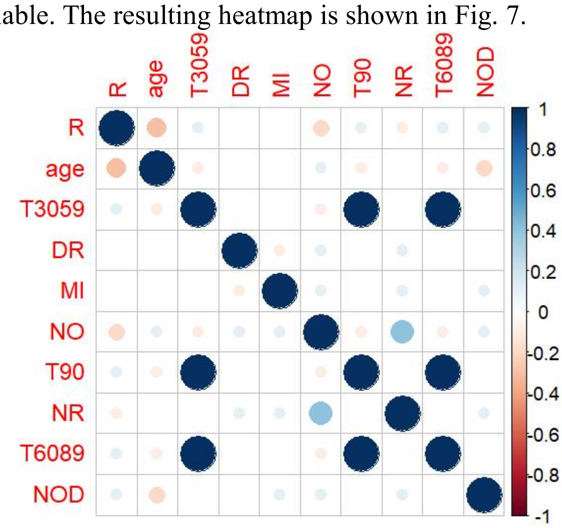


**Fig. 7 Feature Heatmap**

From Fig. 7, it can be seen that the features "Number of overdue 30-59 days," "Number of overdue 60-89 days," and "Number of overdue 90 days" have a high correlation. Therefore, two of these features are selected for removal, and the "Number of overdue 30-59 days" feature is retained. The remaining features after the deletion are then used to train the LR-WOE model, and the resulting confusion matrix of the model is shown in Fig. 8.

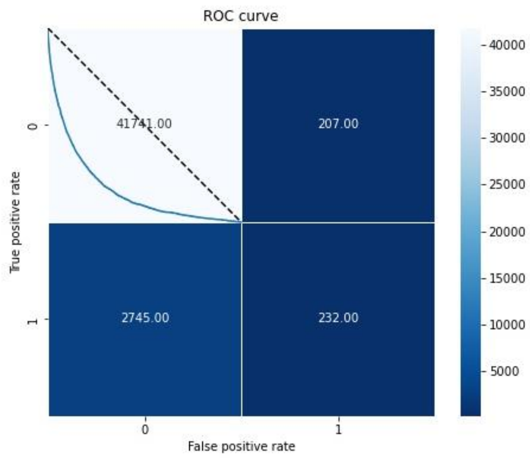


**Fig. 8 Confusion Matrix and ROC Curve of the LR-WOE Model After Feature Selection**

From Fig. 8, it can be seen that the model correctly identified 232 positive cases and correctly excluded 41,741 negative cases. However, it incorrectly predicted 207 negative cases as positive and 2,745 positive cases as negative. The evaluation results of the feature-selected LR-WOE model are presented in Table 3.

**Table 1 Evaluation Table of the LR-WOE Model After Feature Selection**

| accuracy | precision | recall | f1_score |
|---|---|---|---|
| 0.934 | 0.995 | 0.938 | 0.966 |

Table 3 shows several important evaluation metrics of the model. The accuracy is 0.934, the precision is 0.995, the recall is 0.938, the F1 score is 0.966, and the AUC score calculated based on the model's ROC is 0.83.

## 3. Model Evaluation

To further compare the models, this paper introduces a new evaluation metric—CAP (Comparative Assessment Parameter). This metric aims to provide a more comprehensive and detailed perspective to measure and compare the subtle differences in prediction performance and the overall performance of different models. The rating system performance indicator derived from the CAP curve is the accuracy AR （$-1\leq AR\leq 1$）, which is given by the ratio of two areas: one area, Q, is the area bounded by the ideal performance curve and the positive diagonal of the unit square, representing that the ideal performance is better than random performance; the other area, R, is the area bounded by the observed CAP curve and the positive diagonal, indicating that the observed performance is better than random performance. The ratio of these two areas, R/Q, indicates how well the observed performance compares to the ideal performance[18]. The LR-WOE model and the feature-selected LR-WOE model are further evaluated by adding the CAP metric. The results are shown in Fig. 9 and 10.

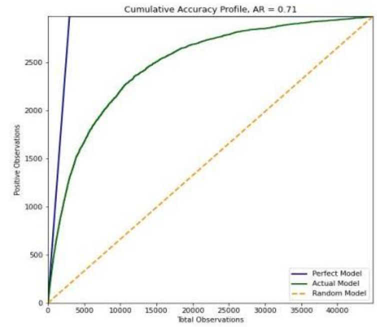


**Fig. 9 CAP Chart of the LR-WOE Model**

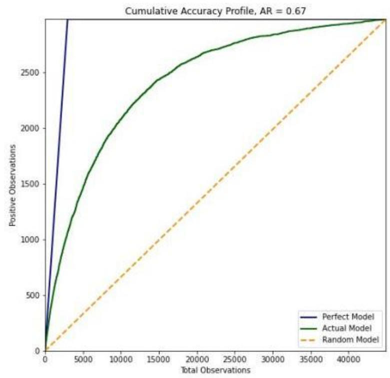


**Fig. 10 CAP of the LR-WOE Model After Feature Selection**

From Fig. 9 and 10, it can be seen that the AR score of the LR-WOE model is 0.71, while the AR score of the feature-selected LR-WOE model is approximately 0.67. Since the AR index closer to 1 indicates higher similarity, the AR score of the LR-WOE model is closer to 1, meaning that this model is more similar to the real situation.

Based on the comparison results, the LR-WOE model is selected as the final model. The resulting logistic regression model is:

$$p = sigmoid\left(x^T \beta + b\right), \tag{4}$$

in:

$$\beta = \begin{pmatrix} -0.496299193 \\ -0.419656597 \\ -0.522785723 \\ -0.884935759 \\ -0.219174254 \\ 0.112226311 \\ -0.521525077 \\ -0.592248789 \\ -0.373283042 \\ -0.271610416 \end{pmatrix}, b = -2.65221902 .$$

Finally, an indicator called the PSI value, or Population Stability Index, is introduced to assess the stability of the model across different sample sets. The PSI is an indicator used in credit scoring and risk management to measure the stability of model predictions. It reflects the difference between the distribution of validation samples across different score ranges and the distribution of the modeling samples. The stability test requires a comparison between the actual distribution and the expected distribution. During the modeling process, the training samples are typically used as the expected distribution, while the test or validation samples serve as the actual distribution. The PSI is calculated based on the difference between these two distributions and is used to evaluate the stability of the model across different sample sets. The PSI chart for the LR-WOE model is shown in Fig. 11.

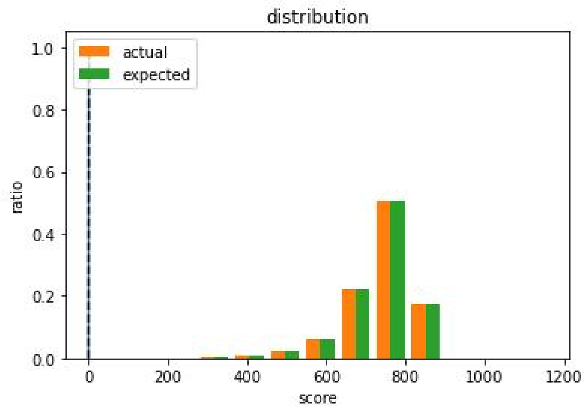


**Fig . 11 PSI Chart of LR-WOE Model**

As shown in Fig. 11, the distribution of the training set and the test set are generally consistent. The PSI value is approximately 0.0008, indicating that the distribution of the validation samples and the modeling samples across score ranges is relatively stable, with a small difference between the two. This is because a small difference in the PSI curve suggests that the model has a strong ability to adapt to new data, maintaining good performance stability. The model demonstrates good generalization ability across different data sets[19].

WOE encoding can significantly enhance the predictive power of features, optimize model stability, and simplify feature engineering through automation, collectively contributing to the outstanding performance of the LR-WOE model in credit risk assessment. The reasons why the LR-WOE model performs the best are as follows: First, WOE encoding transforms features by measuring the default rate of each category relative to the total sample, making each feature value more predictive. It effectively differentiates between samples from different categories, especially when there are large distribution differences between categories. This can more clearly reflect those differences, thereby improving the model's predictive ability. Second, the logistic regression model is sensitive to the distribution and scale of the input data. WOE encoding reduces the impact of outliers on the model and transforms feature values into similar scales, balancing the contribution of different features to the model, which in turn improves model stability. Third, the logistic regression model assumes a linear relationship between features and the target variable. WOE encoding makes the relationship between feature values and the target variable closer to linear, thus improving the fit of the logistic regression model. At the same time, WOE encoding reduces non-linear complexity by transforming the non-linear relationships of categorical features into linear ones, making it easier for the model to capture these relationships. Fourth, WOE encoding automatically converts categorical variables into a format that can be directly used by the logistic regression model, reducing the need for complex feature engineering steps and avoiding inconsistencies that may arise in feature processing.

## 4. Discussion

Traditional feature selection methods predominantly rely on significance or linear correlations, which often fail to effectively capture nonlinear feature associations in credit risk scenarios. To address this limitation, this study proposes a dynamic feature screening framework by synergistically integrating Information Value (IV) and Weight of Evidence (WOE) encoding. First, to better reflect the relationship between features and target variables in highly imbalanced credit datasets, IV is employed to quantify the predictive power of features for default risk, thereby filtering out high-IV features. Continuous variables are discretized into bins, and WOE values are computed to transform complex nonlinear relationships in raw features into linearly separable numerical representations. Subsequently, bins with similar WOE values are merged to mitigate overfitting risks and enhance feature generalization capabilities. This framework ensures that the feature engineering process achieves higher interpretability and robustness. This research innovatively combines the Logistic Regression model with WOE encoding to construct an LR-WOE integrated model. While conventional model evaluations focus primarily on predictive accuracy, stability and business adaptability are equally critical in credit scenarios. Therefore, this study innovatively constructs a triadic evaluation system encompassing "predictive performance, distribution stability, and risk coverage capability." It not only validates the model's predictive performance but also assesses its practical value (e.g., PSI ensuring long-term stability, AR value guiding risk stratification). This approach addresses the challenge of "high AUC but low usability" in traditional evaluations, providing financial institutions with an end-to-end model optimization pathway. Results demonstrate that the LR-WOE model achieves an AUC of 0.86, an AR of 0.71, and a remarkably low PSI of 0.0008, significantly outperforming conventional LR models and confirming its superior generalization ability. The proposed "feature optimization + multi-metric validation" framework not only enhances prediction accuracy but also delivers a highly interpretable and stable solution for credit risk assessment.

Based on the above, this study presents the following suggestions:

First, banks should introduce and apply machine learning and artificial intelligence technologies. These models can handle large-scale datasets and extract complex patterns, thereby improving the accuracy of default prediction. Banks should also invest in data governance and quality control, establishing effective data cleaning and integration processes to ensure that the data used is highly accurate and consistent. Additionally, as the characteristics of credit risk change over time and with economic conditions, banks need to regularly update and retrain models to ensure they reflect the latest economic conditions and customer behavior. Finally, models based on data should focus on transparency, ensuring that the decision-making process and predictive results are interpretable by approval personnel, to enhance trust and acceptance. Banks should also provide explanatory analyses of the models to help credit approval personnel understand the basis and results of the predictions.

Second, the government should formulate and promote transparent and fair risk assessment methods for financial institutions, ensuring that there is no discrimination or unfairness in the credit approval process. For example, standardized model review procedures should be established, requiring banks to disclose basic information about the models and evaluation methods to ensure their fairness. Additionally, policy support and funding can encourage fintech companies to innovate in credit approval and risk management. The government should also encourage financial institutions and educational institutions to conduct financial education to enhance the public's understanding of credit management and risk assessment. Through education and training, consumers can learn about the factors influencing credit scores and how to improve their credit records, thereby increasing their financial literacy and risk awareness. Lastly, relevant laws and regulations should be formulated to protect consumers' rights during credit applications and risk assessments. A robust complaint and appeal mechanism should be established to ensure that consumers receive effective redress and support when facing erroneous assessments.

Third, individuals should regularly check their credit reports to ensure the accuracy of all information. If errors or omissions are found, they should promptly contact the credit reporting agencies to correct them. Regular checks can help individuals identify potential credit issues and take corrective action in advance. Secondly, making timely repayments is key to building and maintaining a good credit record. Individuals should create a reasonable budget to ensure that all loans and credit card bills are repaid on time, thus improving their credit score and reducing the risk of default. Finally, individuals should reasonably control their credit card usage and avoid excessive borrowing. It is recommended to keep the credit card balance below 30% of the credit limit to reduce the credit utilization rate, thereby improving the credit score. Frequently applying for new credit cards or loans can lead to hard inquiries, which may negatively affect the credit score. Therefore, individuals should be cautious about applying for credit products and only do so when necessary.

## 5. Conclusions

Personal credit, as a "double-edged sword" in economic activities, not only greatly stimulates consumer vitality and improves people's living standards but also plays an indispensable role in driving economic growth. However, this financial tool may also lead to social inequality and increase personal financial pressure, highlighting the potential risks and challenges of personal credit. It is precisely this duality that makes research into the risk management of personal credit particularly important, aiming to find effective ways to balance its positive effects with potential risks. Although logistic regression models are widely used in the field of economics due to their high prediction efficiency and speed, especially when applied to large-scale datasets, their algorithms may perform inadequately when handling complex feature data. This study, through WOE encoding and model fitting of the input features, found that features processed with WOE encoding performed better after being mapped through the sigmoid function compared to untreated features. In other words, the improved LR-WOE model outperformed the traditional LR model in terms of prediction accuracy and risk assessment. This finding not only validates the effectiveness of logistic regression models in credit risk evaluation but also significantly enhances the model's predictive performance through innovative feature engineering methods. It provides financial institutions with a more precise risk assessment tool, contributing to the long-term stability and sustainable development of the financial market.

**References**


[1] Zhang AQ, Fu HY, Wang JJ, et al. Establishing a nomogram to predict refracture after percutaneous kyphoplasty by logistic regression [J]. Frontiers in Neuroinformatics, 2023, 17.

[2] Li HC, Shao YM, Jiang HL, et al. Research on machine learning method for disaster prediction caused by heavy rainfall in heilongjiang province [J]. Journal of Catastrophology, 2024, 39(03): 60-65.

[3] Goel A, Gorse D. A comparison of deep and shallow models for the detection of induced seismicity [J]. Geophysical Prospecting, 2024, 72(01): 285-297.

[4] Silva ECE, Lopes IC, Correia A, et al. A logistic regression model for consumer default risk [J]. Journal of Applied Statistics, 2020, 47(13-15): 2879-2894.

[5] Wu Y, Pan YW. Application analysis of credit scoring of financial institutions based on machine learning model [J]. Complexity, 2021.

[6] Dumitrescu E, Hué S, Hurlin, C, et al. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects [J]. European Journal of Operational Research, 2022, 297(03): 1178-1192.

[7] Wei Y, Hasan H. Application of logical regression function model in credit business of commercial banks [J]. Applied Mathematics and Nonlinear Sciences, 7(01): 513-522.

[8] Saha P, Bose I, Mahanti A. A knowledge based scheme for risk assessment in loan processing by banks [J]. Decision Support Systems, 84: 78-88.

[9] Mao Y, Chen WL, Guo BL, et al. A novel logistic regression model based on density estimation [J]. Acta Automatica Sinica, 40(01): 62-72.

[10] Aneta WD, Mateusz H. An implementation of ensemble methods, logistic regression, and neural network for default prediction in peer-to-peer lending [J]. Zbornik radova Ekonomskog fakulteta u Rijeci : č asopis za ekonomsku teoriju ipraksu , 2021, 39(1): 163-197.

[11] Liang S, Tan L. Research on the construction of personal credit score model based on WOE analysis and logistics model [J]. Academic Journal of Business Management, 2022, 4(10): 38-42.

[12] Yao G, Hu XJ, Wang GX. A novel ensemble feature selection method by integrating multiple ranking information combined with an SVM ensemble model for enterprise credit risk prediction in the supply chain [J]. Expert Systems with Applications, 2022, 200.

[13] Niu K, Zhang ZM, Liu Y, et al. Resampling ensemble model based on data distribution for imbalanced credit risk evaluation in P2P lending [J]. Information Sciences, 2020, 536:120-134.

[14] Hwang RC, Chu CK, Yu KH. Predicting the loss given default distribution with the zero-inflated censored beta-mixture regression that allows probability masses and bimodality [J]. Journal of Financial Services Research, 2021, 59(03): 143-172.

[15] Zhang H, Chen QG, Shen QJ. Applied condition of maximum likelihood estimated method in logistic regression for case - control study [J]. Chinese Journal of Health Statistics, 2006, 23(03): 206-208.

[16] Jiang JY, Tao Q, Gao QK, et al. Dual coordinate descent method for solving AUC optimization problem [J]. Journal of Software, 2014, 25(10): 2282-2292.

[17] Zhou YQ, Han DQ, Yang Y. A research on the influence of evidence distance selection on combination of conflict evidences [J]. Journal of Xi'an Jiaotong University, 2018, 52(06): 1-8.
[18] Irwin JR, Irwin C T. Appraising credit ratings: does the cap fit better than the roc? [J]. International Journal of Finance Economics, 2013, 18(4): 396-408.
[19] Barddal JP, Loezer L, Enembreck F, et al. Lessons learned from data stream classification applied to credit scoring [J]. Expert Systems with Applications, 2020, 162.

**Funding project:** National Natural Science Foundation of China (72204130), Higher Education Outstanding Youth Innovation Team Program (2023RW028), Social Science Planning Project of Shandong Province (25CLJJ36), Social Science Planning Project of Qingdao (QDSKL2401132)

**[1]First author's profile:Ran Meng-Huan (2002-), female, Master's degree in progress, School of Business, Qingdao University of Science and Technology, research direction: risk management (data mining direction).Email: 2867648576@qq.com**

***Corresponding author: Zhu Xiao-xin (1988-), female, associate professor, main research: supply chain risk management; public safety and emergency management. E-mail: zhuxx0813@hotmail.com**