Dean&Francis

ISSN 2959-6130

Exploring the Impact of Macroeconomic Factors on Credit Default Prediction Using Machine Learning and Neural Network Methods

Jierui Zhang[1, *]

[1] Department of Economics, Shenzhen University, Shenzhen, 518000, China

*Corresponding author: 2022041019@email.szu.edu.cn

Abstract:

Macroeconomics has a profound impact on the financial condition and loan behavior of individual borrowers. This article provides a specific analysis of the macroeconomic impact based on data from P2P lending platform LendingClub. Firstly, traditional machine learning models are used to train predictive data. By comparing the predictive performance of macro sensitive and non-sensitive micro features at similar scales, it is found that macro-sensitive micro features are more suitable for model training; Further inclusion of representative macro indicators leads to an improvement in performance, confirming the direct impact of macroeconomic factors on credit default predictions. Given the limitations of traditional machine learning models in dealing with feature multicollinearity and nonlinear relationship of the data, DNN (Deep Neural Network) models are introduced to optimize the performance. The research results indicate that at the feature selection level, macro-related features have a significant impact on credit default prediction. At the level of model evaluation, neural network models have better learning and default prediction performance compared to traditional machine learning models for imbalanced and nonlinear credit datasets in the real world.

Keywords: P2P lending; Credit default prediction; Machine learning; Macroeconomics

1. Introduction

After the 2008 financial crisis, due to the consideration of the actual cost of losses and future panic, major banks and financial institutions have shrunk their credit business scale, resulting in fewer channels and increasing difficulty for individuals and small and micro enterprises to obtain loans. Yet LendingClub, by virtue of its novel P2P loan model (referring to the financial model where individuals directly borrow funds through the Internet platform), precisely solves the pain point, providing a low threshold borrowing channel for the above specific groups to meet their needs.

The innovative platform has obvious advantages. Firstly, it directly connects borrowers and investors through Internet technology, reducing intermediate links and transaction costs, enabling borrowers to obtain loans at lower interest rates and investors to obtain higher returns. Moreover, compared to the strict and complex loan approval process of traditional financial institutions, this platform provides people with a new financing option. Its online approval process is relatively simple and fast, which can efficiently meet the funding needs of borrowers.

However, due to the audience's focus on individuals and small borrowers, the lending club platform also faces significant risks: the cash flow situation of individual investors is fragile and variable, and the platform's credit evaluation system may not cover all factors that affect borrowers' repayment ability, such as sudden changes in economic conditions, personal unexpected events, etc. Therefore, based on the difficulty in predicting individual default risks and the increasingly competitive financial market, it is of practical significance to focus on using new methods and some new features to improve the performance of default prediction models [1].

Macroeconomics is closely related to personal financial conditions and micro credit business. As mentioned earlier, the impact of macroeconomic emergencies or the deterioration of the macroeconomic environment may exacerbate the default risk of individual small borrowers from three aspects [2].

Firstly, direct damage to the income end leads to the disconnection of repayment sources. Shocks like the 2020 COVID-19 pandemic or cyclical economic weakness cause macro employment instability. Offline store closures and business shutdowns lead to salary cuts or layoffs and leave self-employed or salaried borrowers without main income, making loan repayment impossible. Secondly, the rising cost of living will indirectly increase the pressure of repayment. Economic decline may trigger inflation, pushing up expenses (e.g., food, energy). This reduces personal disposable income and funds for loan repayment. For instance, high inflation in economically unstable countries spikes daily costs—borrowers who once covered living expenses and repayments with monthly income now prioritize basic survival, failing to repay on time. Moreover, tightening credit policies increase borrowing costs, thereby increasing default risk. During recessions, some micro-lenders raise interest rates (e.g., from 10% to 15% annual rate). This significantly burdens borrowers with unstable income, easily leading to default.

From the above analysis, incorporating some representative macroeconomic indicators and some macro-sensitive micro features (such as DTI) into credit default prediction has practical significance.

Moreover, the selection of predicting models is equally important. The three main difficulties of using Lendingclub datasets to predict default are obvious: imbalanced data (low proportion of default records), multicollinearity between features (such as the correlation between loan interest rates and loan terms), and non-linear trends (such as the acceleration of default rate growth after DTI exceeds 50%). Compared to traditional linear regression models, machine learning and neural network models have significant adaptability for handling nonlinear features.

Furthermore, neural network models are superior to traditional machine learning models. The limitations of traditional machine learning model are shown below. Firstly, the issue of data imbalance still exists. For example, since non-default class data accounts for the vast majority (>85%), the random forest model will prioritize fitting the majority of non-default features, resulting in insufficient recognition of default class samples. Second, the issue of multicollinearity is still evident. For example, the GBDT model is prone to repeat the split due to feature redundancy, which can introduce noise and reduce the model's generalization ability. The third issue is that the capture of nonlinear features is not excellent enough. For example, GBDT is based on global considerations and tends to overlook the cumulative nonlinear effects of small-scale fluctuations.

2. Case Description

2.1 Data Preparation

Firstly, this article chose LendingClub data ranging from Jan 2012 to Dec 2018 at a scale of 188k pieces. In terms of feature selection, due to the mentioned influences that macro economy has on default prediction, this article incorporates some representative macroeconomic indicators and some macro-sensitive micro features (such as DTI) into credit default prediction.

2.2 Model Selection

Considering the three main limitations of the LendingClub datasets mentioned above, machine learning and neural network models are selected to train the data and perform default prediction.

Firstly, selected three traditional machine learning models: logistic regression model, random forest model, and gradient boosting decision tree model (GBDT). Logistic Regression is a simple, interpretable baseline for binary classification that models default probability using log-odds. While computationally efficient, its linearity and independence assumptions limit performance on credit data, resulting in low recall. Random Forest reduces overfitting

ISSN 2959-6130

by aggregating many decision trees built on bootstrapped samples and random features, capturing nonlinear patterns better than linear models [3]. GBDT builds trees sequentially to minimize residuals, handling complex nonlinearities (e.g., default rate accelerates when debt ratio > 70%) and interactions (e.g., term-interest rate-installment). It delivers high accuracy but requires careful tuning and more computation [4].

Then, the deep neural network model (DNN), with its deep hidden layer architecture, is more advantageous for addressing the mentioned problems. For multicollinearity, its multi-layer neurons can automatically reduce and recombine highly correlated features and dynamically allocate weights (assigning higher weights to effective features), eliminating redundant feature interference. To address data imbalance, it is possible to amplify the loss of minority class samples through custom loss functions, significantly improving minority class recognition rates. When dealing with nonlinearity, relying on nested multi-layer activation functions to construct complex mappings, the model can automatically achieve high-dimensional feature interaction, accurately capturing complex nonlinear laws such as non-monotonicity and weak interaction [5].

2.3 Performance Evaluation and Problem Solving

By comparing the predictive performances of traditional machine learning models using ROC curve and F1 score tools, the GBDT model was the most suitable for default prediction. However, even with parameter tuning, the model's recall improvement always comes at the cost of low precision, resulting in a low F1 score after comprehensive weighting. The main reason is still the incomplete resolution of the three main problems mentioned above. Certainly, the multicollinearity between features is also an important reason for the low performance of F1 score, especially that macroeconomic indicators at different levels may be influenced by the same economic event.

Therefore, this article introduces DNN models to optimize prediction, while using VIF test to remove the most significant features of multicollinearity. Finally, after introducing the more realistic evaluation metric WACC instead of the original metric F1 score, the default prediction ability of the model was further improved.

3. Analysis on the Problem

3.1 Feature Selection

Table 1. Screening and reasons for macro sensitive micro features (Picture credit: Original)

interest_rate	As a core tool of monetary policy, interest rates directly adjust with the economic cycle, significantly impacting borrowing costs and default risk.
term	Long-term loans face greater uncertainty during economic downturns, making the loan term an effective reflector of cyclical risk appetite.
annual_income	The borrower's income level is closely tied to the job market and is highly sensitive to fluctuations in the economic cycle.
dti (debt to income ratio)	This ratio directly measures repayment pressure arising from falling income or rising debt during an economic recession.
installment	The monthly payment, determined by interest rate and term, directly transmits changes in monetary policy, affecting household cash flow.
Fico_average	Economic recessions systematically increase default rates, leading to a decline in overall credit scores.
total high credit limit	Financial institutions adjust credit limits based on economic outlook, making this a key indicator of credit cycle tightness.
revolving balance	An abnormal increase in balance often signals that consumers are relying more on credit to maintain spending under economic stress.
bankcard utilization	A sharp rise in utilization is an early warning signal of personal financial liquidity stress and impending distress.
bankcard_open_to_buy	A reduction in available credit directly reflects a shrinking financial buffer and a decreased ability to withstand economic shocks.

Firstly, as shown in Table 1, choose micro features which are sensitive and related to the macro-economic environment. Their association with macroeconomics is explained in the second column of the table [6-8].

Table 2. Screening and reasons for macro insensitive micro features (Picture credit: Original)

grade	Risk rating assigned to borrowers by the platform; fixed at loan issuance.
sub_grade	Detailed breakdown of „grade"; also fixed at loan issuance.
home_ownership	Borrowers' long-term residential status (e.g., own, rent, mortgaged); rarely changes in the short term.
purpose	Reason for borrowing (e.g., debt consolidation, home improvement); usually fixed.
employment_length	How long the borrower has held their current job; cumulative indicator, unaffected by short-term economic fluctuations.
zip_code	Borrower's geographic identifier; static attribute.
state of residence	U.S. state where the borrower lives; geographic info, not influenced by economic cycles.
loan Amount	Principal amount applied for by the borrower; fixed at loan issuance.
application_type	Whether the loan is an individual or joint application; structural feature, fixed at application.

Secondly, as shown in Table 2, select feature sets with similar scales which are not influenced by the macro economy as comparison to verify the rationality of the main idea. If the prediction performance or macro-sensitive micro loan features is better than non-sensitive sets, the impact of macro economy can be seen as significant. These indicators mainly reflect relatively stable factors such as borrowers' personal background, loan product characteristics, and geography, and are less affected by short-term economic fluctuations and macroeconomic cycles.

Table 3: Screening and reasons for macro indicators (Picture credit: Original)

Indicator Class	Specific Indicator	Processing Method
Horizontal Indicators	unemployment rate	unemployment_rate_average unemployment_rate_max
Horizontal Indicators	cpi	- cpi_average
Horizontal Indicators	cci (consumer's confidence index)	cci_average cci_min (the lowest confidence value is used as a risk signal)
Change Rate Indicators	gdp	gdp_growth_average gdp_growth_standard_deviation
Change Rate Indicators	Industry production index	Industry_growth_average Industry_growth_standard_deviation
Policy Indicators	lpr (loan prime rate)	lpr_initial lpr_min_during_loan

Furthermore, as shown in Table 3, select representative macro indicators. The indicators cover leading, synchronous, follow-up indicators, as well as different economic sectors, providing a comprehensive portrayal of the macroeconomic environment [9].

3.2 Data preparation

Firstly, cope with micro features. Filter target micro characteristics, uniformly encode categorical data by one-hot encoding, retain only numerical values in the „term" column and convert them into numerical data. A step further, standardize numerical data to eliminate dimensionality [10].

Secondly, the processing of macro features is divided into the following steps: Obtain raw macro datasets from FRED; Linear interpolation of GDP to obtain monthly data from quarterly data; Merge both macro and micro data after using One Hot Encoding which convert non-numeric categorical features into recognizable numerical combinations by the models.

3.3 Model implementation and performance evaluation

Firstly, merely use the macro-sensitive micro features in the processed datasets to train the ML models. Compare the performance of Logistic Regression, Random Forest and GBDT models.

F1 score and ROC curve are used to evaluate the performance. In the situation where default matters, it is assumed that the recall ratio is an important indicator, since the true positive customers we ignored might give us significant losses. However, we cannot ignore the precision

ISSN 2959-6130

ratio, since refusing to many good borrowers makes us lose profits and Industry reputation. Therefore, F1-socre, which is the weighted average score of recall and precision, is of great value.

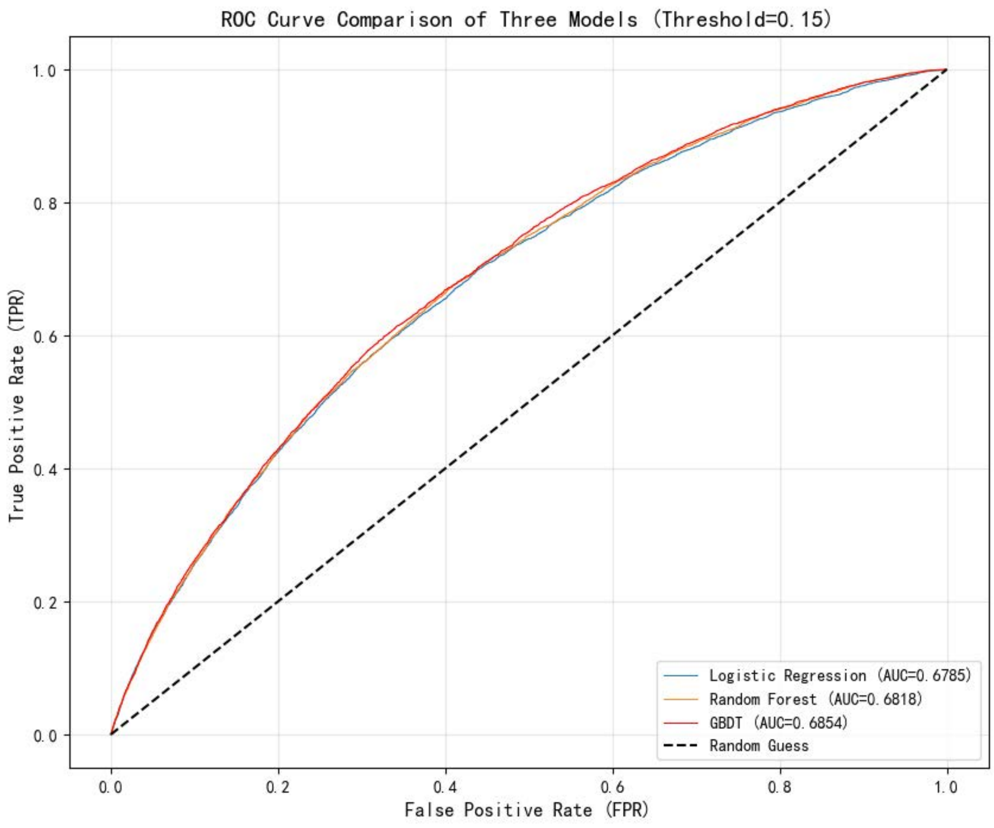

Figure 1: ROC curve comparison of three models (Picture credit: Original)

As shown in Figure 1 and Table 4, GBDT model has largest AUC and best ROC curve which is relatively obvious. At the same time, the GBDT model also has the highest F1 score as shown in the chart below, indicating that the GBDT model is the most suitable for default prediction.

Table 4: Performance comparison of Three Machine Learning Models (Picture credit: Original)

Model Type	F1 score	AUC
Logistic Regression	0.3490	0.6785
Random Forest	0.3515	0.6818
GBDT	0.3551	0.6854

Secondly, compare the performance of the model by separately using macro-sensitive micro features and micro features which are not sensitive to the macro economy. Only the most suitable model GBDT is applied in this comparison.

Table 5: Performance comparison of two types of features (Picture credit: Original)

Micro Feature Type	F1 score	AUC
Macro-sensitive	0.3551	0.6854
Macro-insensitive	0.3405	0.6746

As shown in Table 5, it can be observed that under the same model code and threshold, macroeconomic-related features are more powerful in predicting defaults, proving the rationality of the topic selection.

Furthermore, the model performance has been steadily improved by adding macro indicators into the trained sets.

Table 6: Performance comparison of GBDT model with and without macro indicators (Picture credit: Original)

Add macro indicators or not	F1 score	AUC
No	0.3551	0.6854
Yes	0.3573	0.6931

As shown in Table 6, it is clear that both F1 score and AUC are improved, demonstrating that the inclusion of macroeconomic variables brings practical benefits to credit default prediction.
However, the F1 score is still relatively low even using GBDT model and macro indicators plus macro-sensitive micro features. Particularly, the precision ratio is only 0.2431, demonstrating that under a strict threshold, the recall ratio improves at the cost of low precision. As the analysis deepens, it is believed that three issues—data imbalance (very small number of default samples exist in the trained sets), feature multicollinearity, and uncaptured data nonlinearities—reduce model precision via distinct ways. Data imbalance biases the model toward predicting the majority class (non-default): when predicting defaults, it generates excessive false positives, directly inflating the precision formula's denominator. Multicollinearity causes unstable parameter estimation and noise sensitivity, especially among macro indicators mentioned above, leading to biases that either raise false positives or lower true positives. Uncaptured nonlinearities trigger systematic misjudgments, also reducing true positives or increasing false positives.
Therefore, to better solve the problems above, VIF (Variance Inflation Factor) test is applied to examine and address the multicollinearity problem.

Table 7: VIF value of different features (Picture credit: Original)

Feature	VIF
industry_growth_std	1.461776e+04
gdp_growth_std	1.414363e+04
unemployment_rate_avg	1.098726e+03
cci_avg	7.263864e+01
gdp_growth_avg	5.193828e+01
industry_growth_avg	1.962247e+01
cpi_avg	8.761574e+00
lpr_initial	7.836945e+00
int_rate	2.099640e+00
Fico_avg	1.934716e+00
bc_util	1.932949e+00
bc_open_to_buy	1.912570e+00
tot_hi_cred_lim	1.659575e+00
annual_inc	1.546915e+00
term	1.478899e+00
revol_bal	1.448490e+00
installment	1.308278e+00
dti	1.155031e+00
lpr_min_during_loan	2.242590e-07
cci_min	4.579789e-08
unemployment_rate_max	4.266997e-08

As shown in Table 7, the VIF of all features are listed in a descending order. VIF>10 indicates severe multicol-

ISSN 2959-6130

linearity, therefore corresponding features (the top 6) are removed. The VIF test confirms the problem of multicollinearity in feature data, especially at the macroeconomic indicator level.

Table 8: Performance comparison of GBDT model with and without high VIF features (Picture credit: Original)

Remove features with high VIF value or not	F1 score	AUC
No	0.3573	0.6931
Yes	0.3597	0.6944

As shown in Table 8, after removing features with high VIF value, the performance of the model is further improved, indicating that the multicollinearity problem is significantly weakened.

In addition, to make the performance of the model better and solve the three main problems, PCA method (Principal Component Analysis) is introduced. The core logic of PCA is to transform the original highly correlated features into mutually orthogonal (zero correlation) principal components through linear transformation. By applying the principle of maximizing variance, the default risk information dispersed among multiple related features is condensed into a few principal components, making it easier for the model to capture the differences between default samples [11].

For the problem of data imbalance, this method can decrease the number of model parameters, enable „small samples" to support more stable model training, and avoid overfitting caused by dimension redundancy. As targeted solution for multicollinearity problems, the method integrates and eliminates redundant features. For the problem of non-linear relationships in data, linear redundancy can be removed to highlight non-linear signals. Since the PCA method targets on reducing dimension, all of the 21 rather than 16 (VIF processed) features are used to synthesize new principal components.

By traversing the cumulative explanatory variance, AUC value (priority considering the imbalanced data feature), and F1 score under different principal component scores, the optimal number of principal components is obtained as 18 dimensions.

The Figure 2 below shows the process selecting best dimensions of principal components. In the left subgraph, when there are four principal components, the AUC value of the model rapidly increases and tends to flatten out, reaching its maximum value at the 18th principal component; In the right subgraph, when there are 14 principal components, the cumulative interpretable variance growth rate slows down, and tends to explain 100% of the data information when there are 18 principal components.

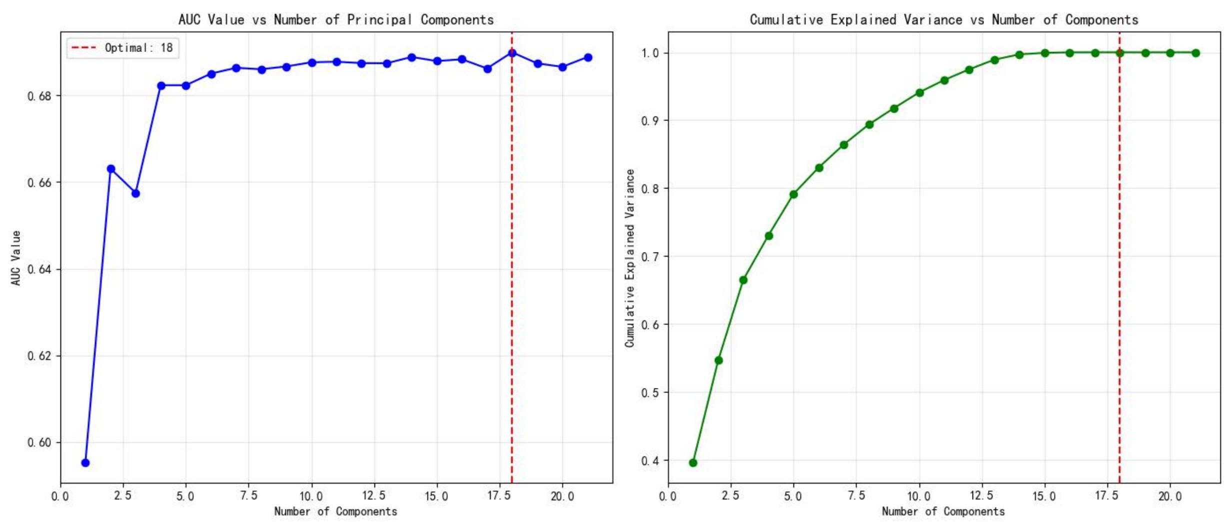

Figure 2: The selection of best dimensions of principal components (Picture credit: Original)

Yet, as shown in Table 9, the model performance after using 18-dimension principal components is not ideal, partly because the total dimensions are not high enough and the VIF method more precisely address the features with high multicollinearity as comparison.

Table 9: Performance comparison of GBDT model using VIF and PCA methods for optimization (Picture credit: Original)

Methods	F1 score	AUC
VIF processed	0.3597	0.6944
Apply PCA	0.3550	0.6899

Even though the model performance is steadily improved by measures above, the precision ratios is always low and the F1 score is equally not satisfying. Therefore, in the next step, neural network models and WACC evaluation standard are introduced based on retaining 16 feature data after VIF processing.

Firstly, explain the rationality of replacing F1 score with WACC as evaluation standard. The formula for WACC (Weighted Accuracy) is in the equation (1) to equation (3):

$$WACC = 0.25 \times Sensitivity + 0.75 \times Specificity \quad (1)$$

$$Sensitivity = TP / (TP + FN) \quad (2)$$

(measuring the ability to correctly identify bad loans, where TP is True Positive and FN is False Negative).

$$Specificity = TN / (TN + FP) \quad (3)$$

(measuring the ability to correctly identify good loans, where TN is True Negative and FP is False Positive).

WACC assigns a higher weight to specificity (measuring the ability to correctly identify good loans, associating with Type I errors, such as misjudging good loans as bad loans, which will result in loss of high-quality customers), which is three times higher than sensitivity (sensitivity measures the ability to correctly identify bad loans, associating with Type II errors, such as misjudging bad loans as good loans, which will result in losses due to default). This not only aligns with business priorities, but also ensures that hidden losses caused by Type I errors are avoided first, while indirectly reducing losses caused by Type II errors through high specificity. If the sensitivity weight is directly increased, it will lead to a surge in Type I errors, resulting in the loss of high-quality customers and even greater overall business losses. In contrast, the F1 score does not fully reflect the differences and priorities of error losses at the business level, so WACC is more suitable as the evaluation core for this scenario [11].

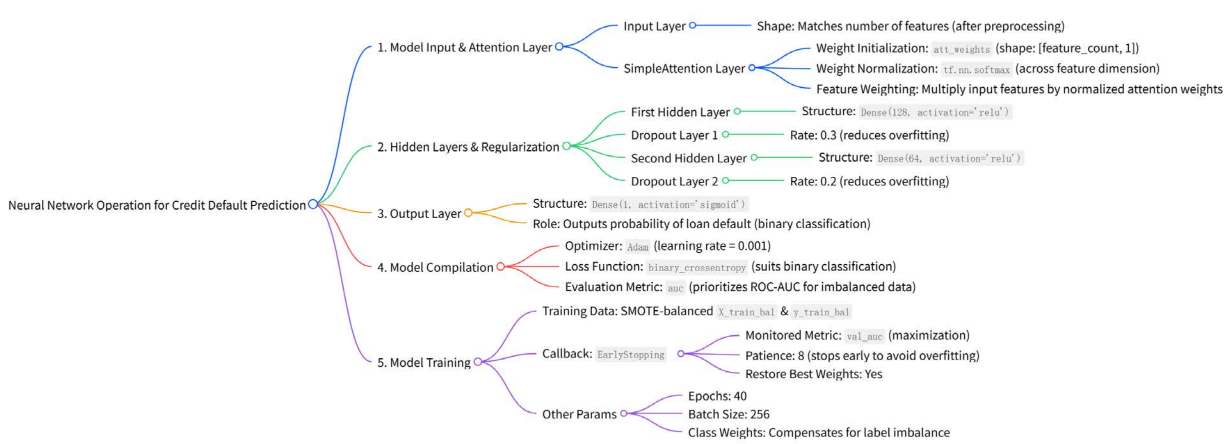

Figure 3: The mind map of DNN model parameter selection and process analysis (Picture credit: Original)

Secondly, the DNN model (Deep Neural Network) is applied with its advantages in solving the three main problems mentioned above. The Figure 3 above shows the comprehensive process of how the model works.

Here are the detailed explanations of methods and parameters chosen in the process. The Table 10 below shows the parameters chosen with detailed reasons; the Table 11 below shows the technical highlights during each step [12,13].

ISSN 2959-6130

Table 10: The detailed process analysis of the DNN model (Picture credit: Original)

Component	Description
Input Layer	Shaped to match the number of preprocessed features.
SimpleAttention Layer	Initializes weights with shape [feature_count, 1], normalizes via tf.nn.softmax for interpretability, and multiplies with input features to dynamically emphasize core risk factors [12].
Hidden Layers & Regularization	Consists of Dense (128, activation='relu'), Dropout(0.3), Dense(64, activation='relu'), and Dropout(0.2); learns complex nonlinear feature relationships and mitigates overfitting.
Output Layer	Dense (1, activation='sigmoid'); outputs loan default probabilities for binary classification.
Model Compilation	Uses Adam optimizer (learning rate=0.001), binary_crossentropy loss function, and AUC metric (robust to class imbalance).
Training Data	Utilizes SMOTE-balanced X_train_bal and y_train_bal datasets.
Training Callback	EarlyStopping monitors val_auc (maximization target), has patience=8, and restores best weights.
Training Parameters	40 epochs, batch size=256, and class weights to compensate for label imbalance.

Table 11: The technical highlights of each step of the DNN model (Picture credit: Original)

Step name	Technical highlights
SMOTE Oversampling + Class Weight and Penalty Mechanism	Oversample default (minority) samples via SMOTE(Synthetic Minority Oversampling Technique) to alleviate data imbalance; calculate class weights to emphasize default sample loss during training, enhancing the model's ability to recognize minority default cases [13].
Attention Mechanism	Construct a feature-level attention mechanism: initialize feature weights, normalize for interpretability, and dynamically weight input features to highlight core risk factors and reduce multicollinearity interference.
DNN Model Construction and Compilation	Use Adam optimizer (lr=0.001), binary_crossentropy loss, and AUC metric (insensitive to class imbalance, WACC-correlated).
DNN Model Training (Capturing Nonlinear Relationships)	Adopt early stopping (patience=8) for efficiency; training has three stages: early (basic patterns, rapid AUC rise), middle (nonlinear interactions, stable attention weights), late (avoid overfitting via early stopping).
Optimal Threshold Search	Traverse thresholds (0.1–0.91, step=0.01), calculate WACC from sensitivity/specificity per threshold, and retain the threshold with maximum WACC to optimize loan default prediction performance.

Finally, the Random Forest Model and GBDT model with same features are trained and evaluated by WACC and AUC as comparison. They represent the performance of traditional machine learning models under the new evaluation system of WACC.

Table 12: Performance comparison of different models under the WACC evaluation system (Picture credit: Original)

Model type	WACC	AUC
GBDT	0.7513	0.6944
Random Forest	0.7513	0.6893
DNN	0.7523	0.6917

As shown in Table 12, compared to the other two models traversing all thresholds, the DNN model performs best. Additionally, the difference in AUC between DNN and GBDT models is only 0.0027, which is a small fluctuation and not a significant difference in model performance. This gap may stem from small differences in model random initialization, training data partitioning, or GBDT's accidental adaptation to the current datasets, rather than the inherent disadvantage of DNN.

4. Suggestion

From the process of introducing the DNN model and WACC evaluation mentioned above, it can be seen that although this evaluation system more realistically reflects the good fitting performance of the model and confirms that the DNN model performs better than traditional machine learning models, which means it better solves the three main problems mentioned above, objectively speaking, the improvement is not significant and there is still room for further performance improvement.

Therefore, it is necessary to explore a step deeper to analyze the reason behind and suggest further improving methods. Since the data imbalance problem is solved by SMOTE method, which transforms the data size ratio of default from 0.15 to 0.5, and the multicollinearity problem is mostly eliminated by VIF method, the next step focuses on the examination and solution of nonlinear patterns among datasets.

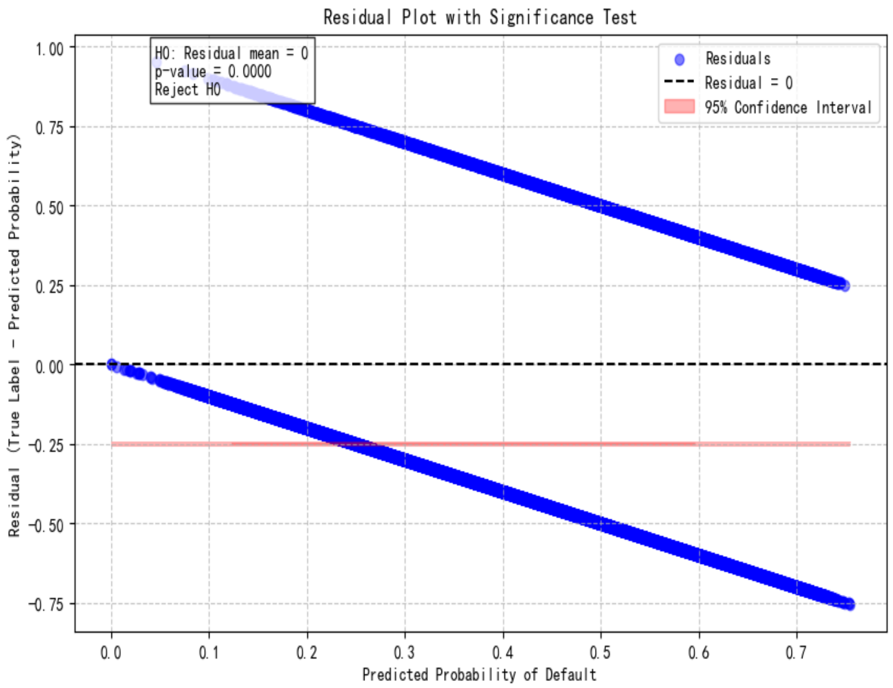

Figure 4: Residual plot with significance test (Picture credit: Original)

Table 13: Hypothesis testing statistical values (Picture credit: Original)

Statistic	Value
Residual Mean	-0.2479
Residual Std	0.3583
Sample Size	31822
95% CI (Lower)	-0.2518
95% CI (Upper)	-0.2439
Z-statistic	-123.3992
p-value	0.0000
H0 Test Result	Reject H0

The Figure 4 reflects the residual plot with significant test, and the Table 13 shows the specific hypothesis testing statistical values. In the above hypothesis test, the null hypothesis is that the mean residual is 0.

Based on the above images and charts, it can be seen that the residual mean of the DNN model is significantly different from zero. Although DNNs with theoretical nonlinear fitting capabilities were used (implemented through activation functions and multi-layer structures), the residual graph still exhibits systematic biases similar to linear models, rooted in the insufficient activation of the model's nonlinear ability: the network structure is simple, the

ISSN 2959-6130

static attention mechanism is essentially linear weighting, only linear feature standardization is performed without introducing nonlinear features, SMOTE overlaps with category weights and may ignore key samples, the early stopping strategy is conservative and the learning rate is fixed, resulting in the model falling into linear local optima.

To address this, it is necessary to strengthen the nonlinear structure of the network (increasing depth width, replacing dynamic attention, adding skip connections), enhance the nonlinearity of input features (nonlinear transformation, constructing interactive features), optimize training strategies (extending early stopping patience, using learning rate scheduling, retaining extreme samples), so that DNN can fully leverage its nonlinear advantages and capture complex correlation patterns in data.

5. Conclusion

Through a series of data filtering and optimization measures of model and method mentioned above, this study believes that macroeconomics has a profound impact on loan default prediction, which is reflected not only in the use of direct macroeconomic indicators as training data, but also in the use of macro sensitive micro indicators as part of the datasets.

In addition, in response to the three core issues with the Lendingclub datasets and corresponding features: multicollinearity between features, non-linear relationships between data, and imbalanced data samples, this study continuously reduces the impact of these issues on the model through feature screening, model optimization, and targeted method implementation. At the same time, a more realistic evaluation method, WACC, was used to assess the performance of various models, and it is ultimately concluded that the DNN model is the optimal choice for predicting loan defaults in this type of scenario.

However, there is still room for improvement in solving nonlinear problems, and in the future, the study will continue to focus on better utilizing neural network models and more appropriate parameter settings to solve this problem.

From a practical perspective, this study has practical value. For financial institutions, the “macro+micro” dual dimensional data fusion approach, targeted data governance plan, and optimized DNN model provided in this study can help institutions more accurately capture the linkage risk between macroeconomic fluctuations and micro borrower characteristics. The method effectively reduces the probability of misjudgment in credit approval, reduces default losses, and optimizes the allocation efficiency of credit resources towards low-risk and high return areas, thereby improving the sustainability of credit business.

For the borrower group, a more scientific risk assessment system can avoid “misjudgments” caused by the limitations of traditional models, allowing borrowers with good qualifications but slightly affected by short-term macroeconomic fluctuations to obtain more fair credit opportunities and promote the inclusiveness of the credit market.

6. References

[1] Muslim, M. A., Nikmah, T. L., Pertiwi, D. a. A., Subhan, N., Jumanto, N., Dasril, Y., & Iswanto, N. (2023). New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning. Intelligent Systems With Applications, 18, 200204.

[2] Baltazar, J., Reis, J., & Amorim, M. (2020). Sustainable economies: Using a macro-economic model to predict how the default rate is affected under economic stress scenarios. Sustainable Futures, 2, 100011.

[3] Zhang Liying, Yang Ruojin. Application Research on Machine Learning-Based Personal Loan Default Prediction Models[J]. Financial Regulatory Research, 2022, (06): 46-59. DOI: 10.13490/j.cnki.frr.2022.06.002

[4] Cao Yujia. Research on the Causes of Personal Credit Customer Default and Prediction Schemes in Commercial Banks[J]. Finance and Economics, 2022, (26): 51-53. DOI: 10.19887/j.cnki.cn11-4098/f.2022.26.007

[5] Kawano, K., Kutsuna, T., & Sano, K. (2025). Minimal sufficient views: A DNN model making predictions with more evidence has higher accuracy. Neural Networks, 190, 107610.

[6] Eggertsson, G. B., & Krugman, P. (2012). Debt, Deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach*. The Quarterly Journal of Economics, 127(3), 1469–1513.

[7] Avery, R. B., Calem, P. S., & Canner, G. B. (2004). Credit report accuracy and access to credit. Federal Reserve Bulletin, 90(3), 0.

[8] Stiglitz, J. E., & Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. American Economic Review, 71(3), 393–410.

[9] Ganong, P., & Noel, P. (2022). Why do Borrowers Default on Mortgages? The Quarterly Journal of Economics, 138(2), 1001–1065.

[10] Duan, J. (2019). Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction. Journal of the Franklin Institute, 356(8), 4716–4731.

[11] Bayraci, S., & Susuz, O. (2019). A Deep Neural Network (DNN) based classification model in application to loan default prediction. DOAJ (DOAJ: Directory of Open Access Journals).

[12] Zandi, S., Korangi, K., Óskarsdóttir, M., Mues, C., & Bravo, C. (2024). Attention-based dynamic multilayer graph neural networks for loan default prediction. European Journal of Operational Research.

[13] Di, S., Wang, Y., Yang, D., Liu, Y., Zhang, J., & Zheng, W. (2025). SMOTE-enhanced XGBoost for rapid seismic damage assessment of bridge portfolios. Soil Dynamics and Earthquake Engineering, 199, 109712.