Pacific-Basin Finance Journal 89 (2025) 102612

Contents lists available at ScienceDirect

Pacific-Basin Finance Journal

journal homepage: www.elsevier.com/locate/pacfin

Farmers' credit risk evaluation with an explainable hybrid ensemble approach: A closer look in microfinance

Nana Chai [a,1], Mohammad Zoynul Abedin [b,1], Lian Yang [c,*], Baofeng Shi [d,e,**]

[a] School of Economics and Management, Tongji University, Shanghai 200092, Shanghai, China
[b] Department of Accounting and Finance, School of Management, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, UK
[c] College of Economics and Management, Shandong Agricultural University, Taian 271000, Shandong, China
[d] College of Economics and Management, Northwest A&F University, Yangling 712100, Shaanxi, China
[e] Research Center on Credit and Big Data Analytics, Northwest A&F University, Yangling 712100, Shaanxi, China

ARTICLE INFO

Keywords:
Farmers
Credit risk
Hybrid ensemble model
Interpretation

ABSTRACT

Artificial intelligence stimulates the vitality of microcredit by reshaping credit risk evaluation models, especially targeting the group of farmers. Therefore, the paper aims to establish a new interpretable hybrid ensemble model for evaluating the credit risk of microfinance for farmers, which is called ADASYN (Adaptive Synthetic Sampling)-LCE (Local Cascade Ensemble)-Shapash. It integrates the advantages of three ensemble models: bagging, boosting, and local cascading, including reducing model variance, reducing model bias, and simplifying complex problems by learning different parts of the training data. And it alleviates the problem of low generalization performance of traditional ensemble models caused by imbalanced loan data of farmers. Through the empirical analysis of the data of farmers' loans of China poverty alleviation agency "CHONGHO BRIDGE", it is found that its average rank is 2.1, which is better than other integrated models in the credit risk evaluation of farmers' microfinance. Finally, the global and local interpretation of our model is preliminarily explored.

1. Introduction

Microfinance is an important means to alleviate the financing constraints of vulnerable groups such as farmers, especially with the application of big data and artificial intelligence technologies in the microfinance industry, which has revitalized microcredit once again (Gao et al., 2022). The characteristics of the microfinance industry have also changed accordingly, such as the alleviation of information asymmetry between lenders and borrowers, further reduction of loan costs, greater diversification of loan products, and expansion of the coverage of loans. The main reason for the alleviation of information asymmetry is the increasingly improved credit data accumulation of farmers. Taking China, the country with the largest number of farmer groups, as an example, the construction of agricultural credit databases has been ongoing for 3 years (The Central People's Government of the People's Republic of China (CPGPRC), 2021), accumulating a large amount of data on farmers' operations, government affairs, tax information, etc. And the

* Corresponding author.
** Corresponding author: College of Economics and Management, Northwest A&F University, Yangling 712100, Shaanxi, China.
E-mail addresses: 23310234@tongji.edu.cn (N. Chai), m.z.abedin@swansea.ac.uk (M.Z. Abedin), yanglian@sdau.edu.cn (L. Yang), shibaofeng@nwsuaf.edu.cn (B. Shi).
[1] Nana Chai and Mohammad Zoynul Abedin have contributed equally to the paper as first authors.

https://doi.org/10.1016/j.pacfin.2024.102612
Received 13 June 2024; Received in revised form 7 November 2024; Accepted 23 November 2024
Available online 30 November 2024

establishment of the digital database led to the emergence of high-dimensional, low value density, and high complexity big data characteristics in farmers' credit data. This has increased the difficulty in selecting methods for assessing credit risks for microfinance to farmers (Dong et al., 2023).

Meanwhile, the simultaneous growth of the amount of microfinance issued to farmers and the corresponding balance of non-performing loans pose higher requirements for the accuracy of credit risk assessment models for microfinance to farmers. This is because a larger model discrimination error will result in a larger balance of non-performing loans, and as the amount of microfinance issued to farmers increases, the cost of microfinance for financial institutions becomes unsustainable, making it difficult for them to maintain sustainable development. This, in turn, worsens the long-term effectiveness of resolving farmers' financing constraints. Therefore, establishing a credit risk measurement model suitable for the characteristics of farmers microfinance is a huge challenge for microfinance institutions (Simumba et al., 2022).

Existing research has demonstrated significant improvements in the ensemble or hybrid credit risk model compared to weaker learners (Liu et al., 2023). The quality of constructing such models is closely tied to the diversity of base classifiers, with greater diversity leading to improved model generalization performance. Existing ensemble research has utilized two methods to enhance base classifier diversity. The first methods involve learning different parts of the training data, exemplified by the local cascading (LC) model (Carcillo et al., 2018; Liu et al., 2023). The second method involves altering the distribution characteristics of the training data, as seen Bagging and Boosting (Buchen and Wohlrabe, 2011; Sigrist and Hirnschall, 2019; Simsek et al., 2021; Wang et al., 2023). The second type of method finds it difficult to achieve a balance between predicting model performance, bias, and variance, and thus cannot simultaneously address issues of potential overfitting or underfitting on the training set. The first type of method, when combined with the second type, can effectively solve these problems. While existing research has focused on either the first or second method separately, there is limited exploration on combining these approaches to fully leverage their advantages and ensure base classifier diversity. Additionally, the lack of interpretability in such models hampers their application in microfinance for farmers.

Therefore, the objective of this study is to establish a new hybrid ensemble model (ADASYN-LCE-Shapash model) with interpretability that precisely assesses the credit risk of farmers. This model combines the advantages of LC, Bagging, and Boosting, while also maintaining model interpretability. And it is suitable for predicting credit risk in the context of imbalanced loan applicant dataset. Specifically, the paper employs LC to learn different parts of the training data and integrates Bagging and Boosting to enhance base classifier diversity and boost the model's generalization performance. ADASYN (Adaptive Synthetic Sampling) is introduced to address the issue of low predictive performance stemming from imbalanced data. Furthermore, we use Shapash to discuss the interpretability of ADASYN-LCE. Through empirical analysis of microfinance data from CHONGHO BRIDGE, the main findings are as follows. Firstly, the new model ADASYN (Adaptive Synthetic Sampling)-LCE (Local Cascade Ensemble) proposed in this study performs well in identifying default risks in microfinance for farmers. Secondly, ADASYN-LCE achieves an average rank of 2.25, making it the top-performing method for identifying credit risk among agricultural borrowers, with this performance advantage being statistically significant. Additionally, robustness tests conducted on German and Australian datasets demonstrate that ADASYN-LCE consistently outperforms other models, including ADASYN-XGBoost and ADASYN-LightGBM, which have corresponding average ranks of 1.75 and 1, respectively. The results from the Friedman test further confirm that ADASYN-LCE significantly excels in identifying default risks in microfinance for farmers. Overall, features such as marital status, regional academic output, and other non-financial features play a crucial role in the credit risk assessment of farmers' microfinance. Specifically, married or remarried farmers tend to default less frequently than their unmarried counterparts. This underscores the notion that, in contrast to typical non-vulnerable loan clients like publicly listed companies, financial features are not necessarily critical for assessing the credit risk of agricultural borrowers. Instead, non-financial features such as individual farmer features and macroeconomic conditions can significantly aid in identifying the credit risk associated with agricultural borrowers.

The contributions of this study are threefold. Firstly, this study designs a novel hybrid ensemble model with data balance, deviation variance minimization and interpretability, which is suitable for identifying default risks of microfinance to farmers. Secondly, the ADASYN-LCE model developed here acts as a reliable financial link between farmers and financial institutions by accurately pinpointing default risks in microfinance for farmers, thereby easing the financial constraints faced by farmers. Thirdly, the ADASYN-LCE model aids financial institutions in managing affordable default risk costs by identifying default risks in microfinance for farmers, aligning with the principles of inclusive finance and sustainable business practices. Particularly, it assists microfinance institutions in achieving sustainable development, as their unique social and financial dynamics necessitate more precise credit risk assessment models compared to commercial banks (Leif et al., 2014; Sun and Liang, 2021).

The remainder of this paper is organized as follows. Section 2 presents the Literature review. Section 3 describes the data. Section 4 represents empirical methodology. Section 5 reports the empirical results. Section 5 discusses the empirical results of our model. The last section concludes.

2. Literature review

Microfinance plays a crucial role in serving vulnerable populations such as farmers, particularly those facing poverty, and has evolved into a key policy tool for national economic development, especially in countries with significant poverty reduction and sustainable development objectives (Karlan and Zinman, 2011; Buera et al., 2021). This is highly related to its own inclusion of two major logics, social and financial, where the social logic refers to poverty alleviation, helping women, diverse products, affordable interest rates, etc., and the financial logic refers to the sustainable development of microfinance institutions (Leif et al., 2014; Sun and Liang, 2021). The integration of artificial intelligence technology in microfinance operations further enhances the realization of these dual social and financial logics (Benami and Carter, 2021; Dong et al., 2023). Specifically, the application of artificial intelligence technology in credit risk assessment for microfinance extended to farmers can facilitate increased access to credit for farmers and alleviate their financial constraints (Chai et al., 2023).

A good credit risk evaluation model, serving as a financial intermediary technology, ensures that microfinance institutions secure full loan repayment irrespective of the borrowers' productivity levels or collateral availability (Bai et al., 2019). Regarding credit risk evaluation models, the traditional classification is between econometric statistical models and machine learning models (Li et al., 2016). This study focuses on the analysis of machine learning models based on the current characteristics of agricultural loan data. In terms of the composition diversity of machine learning model, credit risk evaluation models are composed of single weak learners or multiple weak learners. Models comprised of multiple weak learners are typically categorized into ensemble models and hybrid models, which have gained popularity recently, as depicted in Table 1. In ensemble models, they are further divided into Local cascading, Bagging, and Boosting models. Local cascading, a decision tree-based classification algorithm, leverages the divide and conquer strategy (Carcillo et al., 2018). Bagging and Boosting models focus on altering the data distribution features to achieve accurate assessment of credit risk for loan customers (Buchen and Wohlrabe, 2011; Sigrist and Hirnschall, 2019; Simsek et al., 2021; Wang et al., 2023). However, a limitation of existing ensemble models lies in their reliance on homogeneous weak learners, prompting the emergence of hybrid models. These models consist of heterogeneous weak learners capable of addressing various challenges in credit risk assessment processes, such as handling imbalanced credit evaluation data, selecting credit evaluation features, and computing credit scores (Ezgi and Selma, 2016). Hybrid model variations include linear hybrids, nonlinear hybrids, and hybrid ensemble models (Koutanaei et al., 2015). Hybrid ensemble models typically combine ensemble models with other models, with ensemble models primarily utilized for credit risk assessment of loan clients while other models are employed for credit evaluation feature selection or data preprocessing (Marqués et al., 2012; Jadhav et al., 2018; Xiao et al., 2020; Xiao et al., 2021; Shen et al., 2021; Yi et al., 2023; Yang et al., 2024). This type of model is flexible in combination, suitable for solving various tasks, better able to capture complex patterns in data, enhance model predictive performance, and improve model robustness. Hence, a well-designed hybrid ensemble model is imperative for the effective performance of the credit risk evaluation model in microfinance.

The black box feature of traditional machine learning conflicts with the transparency requirement that the credit scoring model under Basel II regulations (Zhang and Yu, 2024). To address this issue, the existing research focuses on enhancing interpretability through interpretable models or model-independent approaches that can elucidate the intricate workings of the hidden layers within black-box machine learning models (Hayashi, 2016; Dumitrescu et al., 2022). LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) are classical interpretable methods of model-independent approaches (Chen et al., 2024), which need to be matched with tree models to explain the black box of tree models, or in other words, the interpretable realization of tree models needs to be matched with model-independent approaches unrelated to tree models (Liu et al., 2023; Caigny et al., 2024).

Existing research has proposed a variety of machine learning models in credit risk evaluation, especially ensemble models and hybrid models have performed well. However, existing ensemble models struggle to simultaneously address both the variance and bias of model predictions. They typically focus on reducing either variance or bias alone, overlooking the challenges posed by imbalanced data and the interpretability of the model. To fill this gap, this study introduces a novel hybrid ensemble model comprising a new ensemble model and an imbalanced data processing model. This innovative ensemble model integrates bagging, boosting, and local cascading techniques to overcome the limitations of existing ensemble models in effectively reducing both bias and variance in credit risk predictions for microfinance. Importantly, our model prioritizes transparency, ensuring a clear understanding of its decision-making processes.

3. Data

3.1. Data source

This study leverages exclusive data from CHONGHO BRIDGE, a poverty alleviation organization in China. The name "CHONGHO" is a phonetic rendering of "Zhonghe," derived from the traditional Confucian principle of "Zhonghe Weiyu." The institution's mission

Table 1
Classification of credit risk evaluation models based on machine learning.

(1) Model type	(2) Definition	(3) Literature	(4) Attribute
Weak learner	A machine learning algorithm with slightly better performance than random guess	Xu et al., 2009; Medina-Olivares et al., 2022; Liu et al., 2022a; Altman et al., 2023	Single; Weak robustness
Ensemble model	Multiple combined individual models (Bagging, Boosting, etc.)	Buchen and Wohlrabe, 2011; Carcillo et al., 2018; Papouskova and Hajek, 2019; Sigrist and Hirnschall, 2019; Xiao et al., 2020; Simsek et al., 2021; Liu et al., 2023; Wang et al., 2023	Multiple; Homogeneous
Hybrid model	Multiple different combinations of single models (linear hybrid models, nonlinear hybrid models, hybrid ensemble models, etc.)	Lee et al., 2002; Akkoc, 2012; Marqués et al., 2012; Koutanaei et al., 2015; Jadhav et al., 2018; Trivedi, 2020; Xiao et al., 2020; Xiao et al., 2021; Shen et al., 2021; Machado and Karray, 2022; Wang et al., 2024; Teng et al., 2024; Yang et al., 2024	Multiple; Heterogeneous; Diversity

Note: column (1) shows three different types of credit risk evaluation models; Columns (2) to (4) respectively describe the definitions, source literature, and basic characteristic attributes of three different models.

is to address the increasing financial requirements of individuals in rural regions through inclusive and sustainable microfinance services, aligning with the concept of "Zhonghe." The term "BRIDGE" signifies the institution's commitment to bridging the gap between urban and rural areas, wealth disparities, and diverse demographics.

This institution focuses on serving the last hectometer of rural areas, facilitating the integration of farmers and small-scale enterprises into the process of agricultural and rural modernization, and promoting rural revitalization and common prosperity. The institution's microfinance customers are spread across more than 100,000 villages in 21 provinces nationwide. Among them, 77.56 % are farmers, 72.25 % are female customers, and 18.61 % belong to ethnic minorities, representing 48 different ethnic groups. Additionally, 37.56 % of the customers are over 45 years old, and 79.80 % have junior high school education or below (Chongho Bridge (CB), 2023). Therefore, this dataset effectively captures the credit characteristics of impoverished farmers who face difficulties in obtaining microfinance from formal financial institutions in China.

Furthermore, this paper utilizes a sample of 1298 microfinance farmers from this institution (called farmer dataset), consisting of 1180 non-default samples and 118 default samples. It highlights the imbalance in microfinance dataset, with a significantly higher number of farmers having non-default microfinance compared to those with default microfinance. This dataset contains real farmer loan dataset from March 11, 2017 to June 30, 2018, with an average loan term of 11.39 months, a maximum loan amount of 300,000 yuan, and a minimum loan amount of 500 yuan. Referring to the classic credit evaluation features of credit rating agencies such as S&P and Fitch (Standard&Poor's (SP)., 2011; Fitch Ratings (FR), 2013), as well as financial institutions such as Bank of China and China Construction Bank (China Construction Bank (CCB), 2007; Bank of China (BC), 2010), and sorting out the classic features of credit risk in classic literature (Bai et al., 2019; Medina-Olivares et al., 2022; Dong et al., 2023). We further removed samples with a missing rate of over 30 % based on the observability of the data and processed non-stationary samples in the data according to the principle of adding or subtracting three times the standard deviation of the mean (Chai et al., 2024). Under the premise of ensuring that high-frequency classical features are not missed, 33 features were obtained. And we conducted robustness analysis of the model using two publicly available datasets (German dataset and Australian dataset). The total number of samples in the German dataset is 1000, while the Australian dataset consists of 690 samples. German dataset is available at https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data. Australian dataset is available at https://archive.ics.uci.edu/dataset/143/statlog+australian+credit+approval. The imbalance ratio of default to non-default samples in these two datasets is relatively lower compared to that of the farmer dataset. The ratio of default to non-default samples and number of features for all three datasets is presented in Table 2. Existing credit risk assessment models primarily use these two publicly available loan datasets for empirical validation of their performance, with significantly less focus on real loan datasets from vulnerable farmers, particularly in developing countries. To provide a more comprehensive discussion of the performance and generalizability of the model designed in this paper, we further analyzed its performance based on these two datasets.

3.2. Farmers' microfinance credit risk feature set

This paper constructed a farmers' microfinance credit risk feature set including 33 features such as "Per capita GDP, Per capita disposable income of rural residents, Per capita disposable income of rural residents, Labor share". And the descriptive statistics of the sample are shown in Table 3. In addition, these features provide feedback on the basic situation, family characteristics, financial information, and other aspects of loan farmers. The original data and standardized data details are shown in Table 4. Table 5 describes the results of quantifying qualitative features.

4. Methodology

4.1. The bias variance decomposition

Excellent ensemble methods should have both low variance and low bias features. This is due to the significant correlation between variance and bias and the generalization error of the model. The generalization error is as follows,

$$E_D\left[(y-\widehat{f}_D(x))^2\right] = Bias[\widehat{f}_D]^2 + Var[\widehat{f}_D] + \varepsilon^2 \tag{1}$$

where, y represents default status of farmers; $f_D(x)$represents predicted values of farmers' default status; ε represents noise; $E_D\left[(y-\widehat{f}_D(x))^2\right]$ represents the generalization error of the model, or can be referred to as bias variance decomposition; The smaller

Table 2
Dataset details.

(1) Dataset	(2) Number of default samples	(3) Number of non-default samples	(4) Number of features	(5) Sample imbalance ratio	(6) Train dataset/test dataset
Farmer	118	1180	33	1:10	9:1
German	300	700	24	3:7	9:1
Australian	383	307	14	55.5:44.5	9:1

Note: this table describes the number of default samples, non-default samples, features, sample imbalance ratio, and training test set ratio included in farmer, German and Australian datasets, which list in columns (1)–(6).

Table 3
Descriptive statistics.

(1) Feature	(2) Minimum	(3) Maximum	(4) Median	(5) Mean	(6) SD	(7) Skewness	(8) Kurtosis
Age	20.5	63	38.5	39.276	8.909	0.291	−0.684
Labor share	0.3	1	0.75	0.75	0.17	0.247	−1.033
Total assets	44.2	51,300,000	500,000	742,496.43	1,595,231.656	24.993	779.766
Total liability	0	28,058,422	50,000	155,386.864	823,259.124	30.299	1020.044
Debt Asset ratio	0	250	0.102	0.357	6.937	35.985	1295.937
Monthly disposable income	1000	109,525	8492.5	11,298.885	9962.951	4.168	26.256
Per capita GDP	0	62,681	31,585	30,797.496	22,523.097	−0.205	−1.365
Per capita disposable income of rural residents	0	18,340	9720	8945.18	6110.85	−0.455	−1.155
…	…	…	…	…	…	…	…
Deposit to loan ratio	0	0.999	0.425	0.429	0.048	9.527	113.192
Consumer Price Index	0	102.9	101.3	73.727	45.343	−1.013	−0.975
Regional agricultural gross output value	0	532.82	179.780	166.301	145.789	0.4	−0.677
Engel coefficient of rural households	0	532.82	179.78	166.301	145.789	0.4	−0.677

Note: this table reports descriptive statistics for all the variables. Our sample covers China-listed. The columns from one to six list the feature names, the minimum, the maximum, the median, the mean, the standard deviation (SD), the skewness, and the kurtosis, respectively.

$E_D\left[\left(y-\widehat{f}_D(x)\right)^2\right)\Big]$ is, the better the prediction effect of the model on farmers' credit risk is. The model bias $Bias[\widehat{f}_D]$ refers to the average deviation between the model's predictions $f_D(x)$ on different training datasets and the true values y, measuring the model's fit or accuracy to the training data. A high bias model underfits the training data and has a weak ability to capture the features of farmer microfinance and default status in the training data. The model variance $Var[\widehat{f}_D]$ is used to describe the sensitivity or instability of the model to the training data. A high variance model is prone to overfitting the training data but performs poorly on new datasets.

4.2. ADASYN-LCE-Shapash hybrid ensemble model

Existing research has also shown that Bagging has the advantage of reducing model variance, Boosting has the advantage of reducing model bias, and Local Cascade has the advantage of dividing and conquer (Simsek et al., 2021). Divide and conquer refers to parallel computing, scalability, accuracy, and flexibility, thereby making the model computationally efficient and accurate prediction. Therefore, the paper combines the advantages of these three methods to construct a Local Cascade Ensemble credit evaluation model to improve the generalization performance of the farmer credit risk evaluation model. In addition, to alleviate the impact of sample imbalance on model performance, ADASYN is used to augment the data. Therefore, this paper combines ADASYN and LCE to build a hybrid ensemble model to evaluate microfinance credit risk of farmers, which is called the ADASYN-LCE model.

4.2.1. Adaptive synthetic sampling (ADASYN) model

ADASYN overcomes the problem of insufficient identification of defaulting farmers in imbalanced farmers microfinance data by augmenting the minority class samples, i.e., defaulting farmers. Increasing the number of default samples can ensure the formation of balanced sample data. A balanced sample refers to having an equal number of defaulting and non-defaulting samples in the training set derived from historical loan data. Both defaulting and non-defaulting samples pertain to applicants who have been granted credit by the bank. Specifically, non-defaulting applicants are those who have repaid their loans on time or in full, while defaulting applicants are those who have not.

We focus on applicants who were accepted by the bank because these individuals have clear outcomes regarding default and non-default status. In contrast, applicants who were not accepted represent the bank's loan decisions based on existing credit evaluation systems. However, it is impossible to determine whether these rejected applicants truly carry default risk. On the other hand, historical samples of granted loans, which include outcomes indicating timely repayment or default, provide a more effective basis for training the model.

In practice, default samples are relatively scarce, which can lead to inadequate learning about default features, ultimately resulting in poorer performance in identifying default risk (Liu et al., 2022b; Chen et al., 2024). Therefore, balanced samples become critically important. The details of ADASYN are shown as follows:

Let U_{train} represent the training set of the credit risk assessment model, where the non-default sample set is U_{large}, the default sample set is U_{small}, and x_i is the i-th default sample in U_{small}.

***Step* 1**. Calculate the number of newly generated default samples N_{new}.

If the number of non-default samples in U_{large} is N_{large} and the number of default samples in U_{small} is N_{small}, then a new number of default samples needs to be generated,

$$N_{new} = N_{large} - N_{small} \tag{2}$$

***Step* 2**. Calculate the proportion of non-defaulting samples r_i in the K-nearest neighbors of each defaulting sample x_i.

The number of non-defaulting samples in the K-nearest neighbors of x_i is Δ_i, and the proportion of non-defaulting samples in the K-

Table 4
Raw data and standardized data of credit evaluation features for farmers.

No.	Criterion	Feature	Original Data						Standard Data					
			Non-defaulting farmer			Defaulting farmer			Non-defaulting farmer			Defaulting farmer		
			(1)	…	(1180)	(1181)	…	(1298)	(1)	…	(1180)	(1181)	…	(1298)
1	Basic information of the farmer	Education level	College degree or above	…	Middle school	High school	…	High school	1.000	…	0.400	0.600	…	0.600
…		…	…	…	…	…	…	…	…	…	…	…	…	…
14	Characteristics of the farmer's family	Number of students	1	…	1	1	…	1	0.500	…	0.500	0.500	…	0.500
…		…	…	…	…	…	…	…	…	…	…	…	…	…
19		Labor force proportion	0.382	…	0.800	0.525	…	0.525	0.382	…	0.639	0.525	…	0.525
20	Financial information of the farmer	Total assets	–	…	1,250,000	380,000	…	800,000	0.101	…	0.256	0.078	…	0.164
…		…	…	…	…	…	…	…	…	…	…	…	…	…
24	External macro conditions	Per capita GDP	38,083.84	…	35,442	31,585	…	23,018	0.510	…	0.315	0.264	…	0.150
…		…	…	…	…	…	…	…	…	…	…	…	…	…
33		Engel coefficient of rural households	166.60	…	27.20	179.78	…	278.78	0.000	…	0.007	0.000	…	0.007
34	–	Default status	0	…	0	1	…	1	0	…	0	1	…	1

Note: this table reports the raw data and standardized data corresponding to 33 features in non-defaulting farmers (column (1) to column (1180)) and defaulting farmers (column (1181) to column (1298)), as well as the 4 criterion layers to which each feature belongs.

Table 5
Quantitative standards for qualitative features of farmers.

No.	(1) Criterion	(2) Feature	(3) Option label	(4) Text options content	(5) Scoring
1	Basic information of the farmer	Gender	1	Male	1
			0	Female	0.8
…		…	…	…	…
…	…	…	…	…	…
			1	≥ 4	1
			2	3	0.75
17	Characteristics of the farmer's family	Number of Labor	3	2	0.5
			4	1	0.25
			5	0	0

Note: column (1) reports the criterion layer (in column (2)) to which the classification features belong. Column (2) reports the feature classification labels for different features (see column 4), and column (5) corresponds to the score of the former.

nearest neighbors of x_i is,

$$r_i = \frac{\Delta_i}{K} \tag{3}$$

***Step* 3.** Normalize r_i to obtain distribution function,

$$\Gamma_i = \frac{r_i}{\sum_{i=1}^{N_{small}} r_i} \tag{4}$$

***Step* 4.** Calculate the newly synthesized sample number g_i for defaulting sample x_i, according to the distribution function Γ_i,

$$g_i = \Gamma_i N_{new} \tag{5}$$

***Step* 5.** Utilize SMOTE (Synthetic Minority Over-sampling Technique) to generate new defaulting samples (See Fig. 1),

Let x_{zi} represent a randomly selected defaulting farmer from the K-nearest neighbors of x_i; λ is a random number in the range [0,1]; The newly generated defaulting farmer sample s_i can be represented as,

$$s_i = x_i + \lambda(x_{zi} - x_i) \tag{6}$$

4.2.2. Local Cascade ensemble (LCE) model

LCE combines the bagging and boosting methods for handling the trade-off between bias and variance (see Eq. (2)), using LC as an implicit method to learn the prediction errors of different parts of the training data for farmers microfinance. The bias-variance tradeoff defines the ability of the learning algorithm to generalize outside the training set and the objective of the bias-variance tradeoff is to minimize both the bias and variance. The bias refers to the average deviation between the model's predictions on different training datasets and the true values, measuring the model's fit or accuracy to the training data. A high bias model underfits the training data and has a weak ability to capture the features of farmer microfinance and default status in the training data, as shown in Fig. 2. The variance is used to describe the sensitivity or instability of the model to the training data. A high variance model is prone to overfitting the training data but performs poorly on new datasets, as shown in Fig. 2.

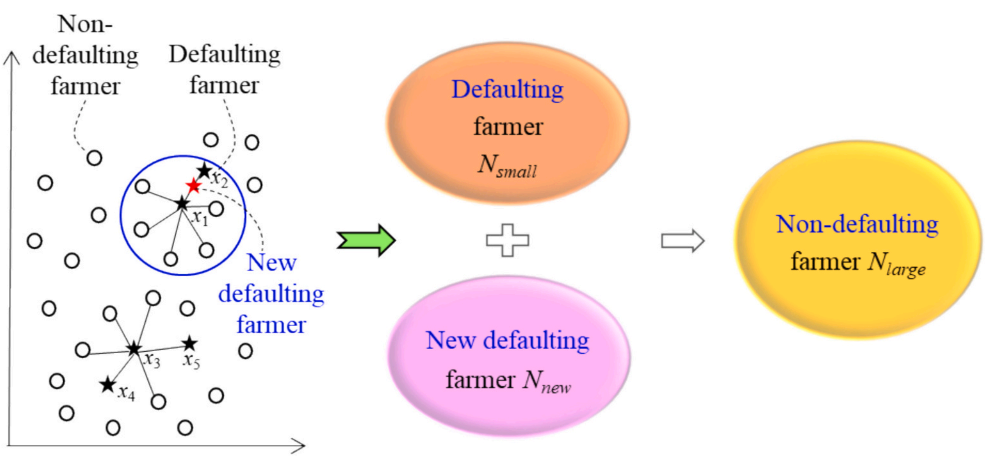


Fig. 1. Expansion process of default samples based on ADASYN.
Note: Fig. 1 presents the expansion process of default samples based on ADASYN. The red star shows the new defaulting farmer based on ASASYN.

Bagging plays a major role in reducing variance. Bagging is to predict the test set by multiple classifiers and aggregate the prediction results of multiple classifiers to get the classification results of whether the loan farmers default or not. The way to reduce the variance is to train the classifier to get the predicted value on the basis of different samples formed by bootstrap, and then get the average predicted value. Boosting plays a major role in reducing bias. The boosting training process is ladder-like, weak classifiers are trained one by one in order, and the training set of weak classifiers is transformed every time according to a certain strategy. Performing linear synthesis on the prediction results of all weak classifiers to generate the final prediction result. The way to reduce the bias is to generate the basic classifier by iteration, that is, the main iteration of the previous generation classifier did not perform well. Local cascading (ie. LC) is a decision tree-based classification algorithm, which has the advantage of dividing and conquer (Carcillo et al., 2018). By adopting a divide-and-conquer method to learn different parts of the training data and capture new relationships that cannot be discovered globally, it helps to simplify given complex problems. The model construction strategy is to use the classification probability of the base classifier in the leaf node as a new feature and propagate it down the decision tree.

Based on all above analysis, LCE model can combine the bagging and boosting methods for handling the trade-off between bias and variance (Fauvel et al., 2022), using LC as an implicit method to learn the prediction errors of different parts of the training data for farmers microfinance. (1) LCE adopts a divide-and-conquer strategy (decision tree) to apply local cascade generalization, sequentially using a set of base classifiers, with each cascade base classifier trained on top of the previous cascade to improve the overall performance of the model. (2) LCE selects boosting as the base classifier and adds the boosting's probability of correctly predicting the default status of farmers as a new feature that is propagated down to the next tree level and used as a weighted scheme on the training data (which focuses more on misclassifying farmers with default status), further reducing the bias in the divide-and-conquer method. (3) LCE uses bagging to create multiple decision trees from randomly sampled different subsets of the original farmer dataset, mitigating overfitting caused by boosting, as shown in Fig. 3.

$$argminE_D\left[\left(y-\widehat{f}_D(x)\right)^2\right]=argmin\left\{Bias\left[\widehat{f}_D\right]^2+Var\left[\widehat{f}_D\right]+\varepsilon^2\right\} \tag{7}$$

where $argminE_D\left[\left(y-\widehat{f}_D(x)\right)^2\right]$ represents the minimum generalization error of the model. As mentioned earlier, LCE achieves the

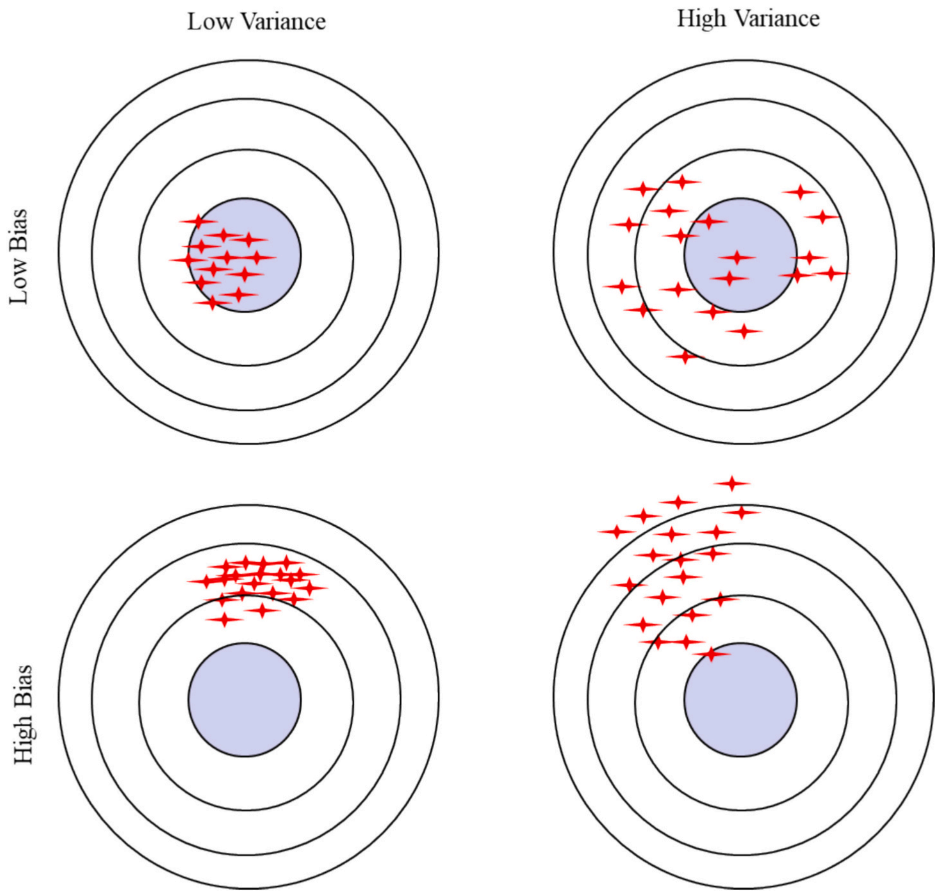


Fig. 2. Bias and variance.
Note: Fig. 2 shows the four different phenomena of variance and bias combinations, including low bias and low variance, low bias and high variance, high bias and low variance, and high bias and high variance.

minimum generalization error of the model by combining bagging and boosting methods, which ensures that LCE can accurately assess the credit risk of farmer microfinance.

4.2.3. Interpretability of ADASYN-LCE

Shapash is a Python library designed to provide a straightforward and efficient way to interpret and visualize the prediction results of machine learning models. The primary objective of Shapash is to streamline the interpretability of models, consequently bolstering their credibility and aiding decision-makers in comprehending the predictions more effectively. Shapash uses LIME and SHAP as the backend and explores the logic between global and local through the use of web applications to quickly understand the model and understand how various key points work. The flowchart of the Shapash is shown in Fig. 4, and the procedures are as follows.

Step **1**. Standardize credit data and divide the standardized data into training and test data sets.

Step **2**. Train the ADASYN-LCE model using the training data set.

Step **3**. Import SmartExplainer and compile it using the trained model from Step 2 and test data set.

Step **4**. Compute Shapash, get WebApp and charts to understand the model.

The definitions of key concepts related to the ADASYN-LCE-Shapash hybrid ensemble model are presented in Appendix Table 1.

5. Experimental results

5.1. ADASYN-LCE VS classical models based on farmer dataset

Table 6 presents the results of various credit risk evaluation models used in the paper. Those models are XGBoost (eXtremeGradient Boosting), LightGBM (Light Gradient Boosting Machine), AdaBoost, GBDT (Gradient Boosting Decision Tree), and RF (Random Forest), which are considered classical models for credit risk evaluation (in column (1)) (Xie et al., 2023). This paper also incorporates ADASYN, a data augmentation technique with these classical models. The evaluation measures include AUC, recall-1, recall-2 and Gmean in column (2)–(5), and the bigger they are, the better the classification effect of the model for microfinance farmers will be (Fenech et al., 2016; Xie et al., 2023). The AUC of ADASYN-LCE is 0.784, which is the highest in all models, and the larger the value, the better the ADASYN-LCE has the best comprehensive ability to identify defaulting and non-defaulting farmers, and Gmean also showed the same findings. The recall-2 of ADASYN-LCE is 0.62, which is the best among other models, that is, the model has a better identification effect on defaulting farmers. Although the recall-1 of ADASYN-LCE performed poorly compared with other models, according to the results of Average rank, that is, the evaluation results of these four types of evaluation measures, it was found that ADASYN-LCE ranked the highest and was also the model with the best overall performance.

And we perform the non-parametric Friedman test to investigate the statistical significance of the differences in performance (Yuan et al., 2022). The Friedman test can be used to rank the classifiers for each dataset independently. The null hypothesis of Friedman's test is that the difference among the rankings of different models is accidental without any significance. The null hypothesis of Friedman test is rejected at the 0.1 significance level regardless of the evaluation measures, which means that there are differences in the models. This further indicates that ADASYN-LCE has the best comprehensive ability to identify the credit risk of loan farmers.

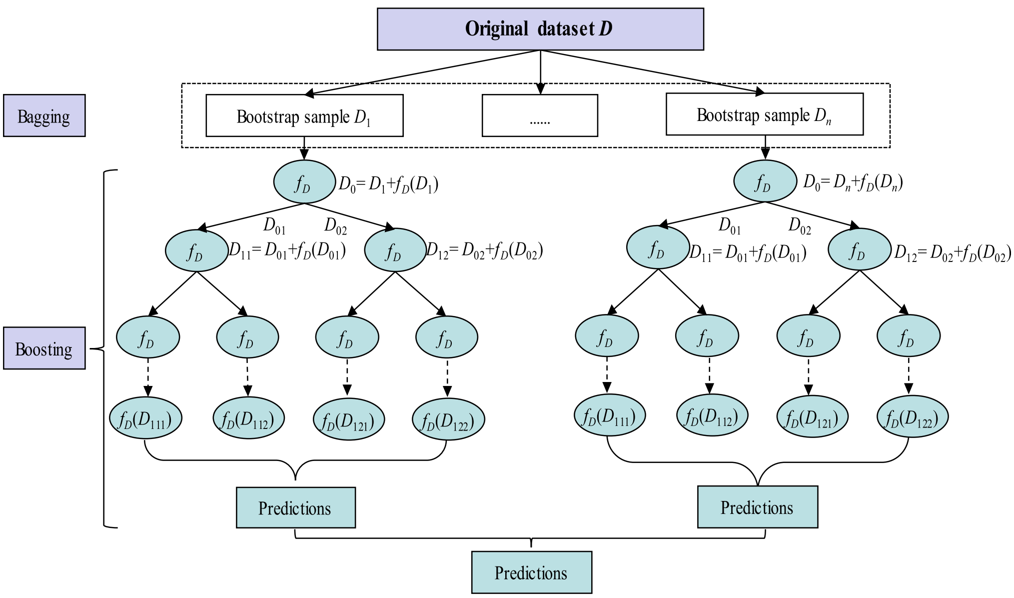


Fig. 3. The framework of Local Cascade Ensemble.
Note: Fig. 3 presents the principle of LCE, which include the bagging and boosting process.

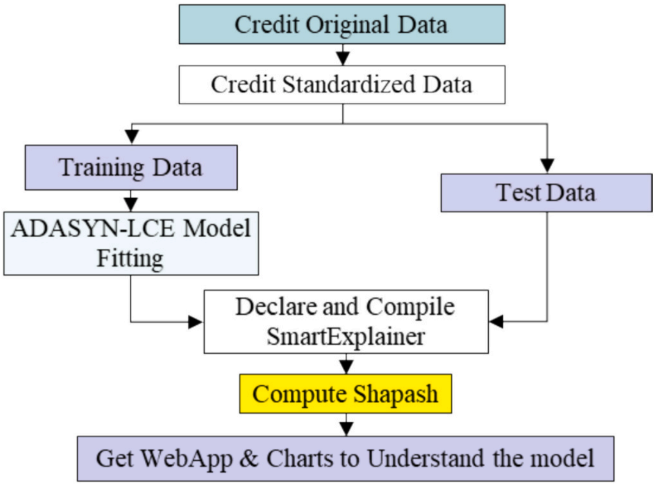


Fig. 4. The flowchart of Shapash.
Note: Fig. 4 presents the flowchart of Shapash, which explains how ADASYN-LCE can be interpreted.

Table 6
Performance of credit risk evaluation models based on farmer dataset.

(1) Models	(2) AUC	(3) recall-1	(4) recall-2	(5) Gmean	(6) Average rank
ADASYN-LCE	**0.784**	0.94	**0.62**	**0.768**	**2.25**
ADASYN- XGBoost	0.755	0.89	0.62	0.744	3.75
ADASYN-LightGBM	0.588	0.93	0.25	0.481	5.75
ADASYN-AdaBoost	0.568	0.89	0.25	0.47	7.5
ADASYN-GBDT	0.613	**0.98**	0.25	0.494	4
ADASYN-RF	0.588	0.93	0.25	0.481	5.75
ADASYN-DT	0.622	0.87	0.38	0.571	5.5
XGBoost	0.5	1	0	0	7.75
LightGBM	0.5	1	0	0	7.75
AdaBoost	0.509	0.89	0.12	0.334	9.25
GBDT	0.48	0.96	0	0	9.5
RF	0.5	1	0	0	7.75
DT	0.592	0.93	0.25	0.483	5.25
Friedman test	–	–	–	–	19.174* (*p*-value $<$ 0.1)

Note: columns (2) to (6) demonstrate the recognition performance of different models (in column (1)) on credit risk of farmers. * is significant at the confidence level of 10 %.

Table 7
Performance of credit risk evaluation models based on German and Australian dataset.

Dataset	(1) Models	(2) AUC	(3) recall-1	(4) recall-2	(5) Gmean	(6) Average rank
German	**ADASYN-LCE**	**0.828**	0.86	**0.8**	**0.828**	**1.75**
	XGBoost	0.726	0.89	0.57	0.708	2.75
	LightGBM	0.736	0.87	0.6	0.723	2.25
	AdaBoost	0.69	0.81	0.57	0.679	4.25
	GBDT	0.676	0.79	0.57	0.667	5
	RF	0.705	**0.94**	0.47	0.663	4.5
	DT	0.64	0.71	0.57	0.636	6
	Friedman test	–	–	–	–	15.308** (*p*-value<0.05)
Australian	**ADASYN-LCE**	**0.915**	**0.94**	**0.89**	**0.915**	**1**
	XGBoost	0.87	0.87	0.87	0.87	2.5
	LightGBM	0.86	0.9	0.82	0.858	3
	AdaBoost	0.791	0.87	0.71	0.787	6.25
	GBDT	0.857	0.87	0.84	0.856	3.75
	RF	0.846	0.9	0.79	0.844	4.25
	DT	0.83	0.87	0.79	0.829	5.25
	Friedman test	–	–	–	–	19.245*** (*p*-value<0.01)

Note: columns (2) to (6) demonstrate the performance of different models (in column (1)) based on German and Australian dataset to identify customers' credit risk. **, *** are significant at the confidence level of 1 % and 5 %.

Furthermore, the performance of models combining ADASYN and classical method (e.g., ADASYN-XGBoost) is better than that of single classical method (e.g., XGBoost). This suggests that ADASYN is an effective technique for handling imbalanced data, which is a common issue in credit risk evaluation for farmers. In conclusion, the results indicate that ADASYN-LCE performs better than other models in credit risk evaluation for farmers. Additionally, combining ADASYN with classical methods improves performance compared to using classical methods alone. This highlights the effectiveness of ADASYN as a method for addressing the challenge of imbalanced data in evaluating credit risk for farmers.

5.2. ADASYN-LCE VS classical models based on German and Australian dataset

Based on the German data, Table 7 presents that the average rank of ADASYN-LCE is performance than other models in credit risk evaluation of farmers. The null hypothesis of Friedman test is rejected at the 0.05 and 0.01 significance level regardless of the evaluation measures, which means that there are differences in those models. It can be seen that the credit risk evaluation model ADASYN-LCE designed in this paper performs well and robustly in the credit risk evaluation of other farmers. Meanwhile, in the Australian dataset, the ADASYN-LCE model itself also outperforms other models. Specifically, similar to the farmer dataset, the AUC of ADASYN-LCE of the two datasets were the highest among all models, which were 0.828 and 0.915, respectively, indicating that the model has the best comprehensive strength for discriminating between defaulting and non-defaulting farmers, and the G-mean also showed the same findings. From the perspective of single-class defaulting farmers, ADASYN-LCE performs best on the 2-type dataset. Although ADASYN-LCE performed slightly worse on the German dataset for non-defaulting farmers, the average rank of ADASYN-LCE on both datasets was the highest and passed the significance test.

And a performance comparison between the proposed hybrid ensemble model and benchmark hybrid ensemble models proposed by Marqués et al., 2012, Jadhav et al., 2018, Xiao et al., 2020, Shen et al., 2021, Xiao et al., 2021, Yang et al., 2024 are presented in Table 8. The main finding is that the proposed model performs better than the benchmark models based on German and Australian dataset. Especially in the Australian dataset, the recall-1, recall-2 and Gmean of proposed model perform better than the basic models. Table 8 is provided to further illustrate the performance of the proposed model compared to similar hybrid ensemble models, particularly to evaluate whether the integration of the advantages of the LC, Boosting, and ensemble models presents a more advantageous approach.

5.3. Sensitive analysis of ADASYN-LCE model

The ADASYN-LCE model is a hybrid ensemble model based on trees, with relatively robust results and few parameters to optimize. The key parameters are the number of classical models and maximum depth, as defined and described in Table 9. Fig. 5 illustrates how changes in these parameters impact model performance, while Table 10 outlines the optimal values for different datasets. We reached the following conclusions. (1) The optimal number of classical models varies across datasets, with values of 20 for the farmer dataset, 15 and 20 for the Australian dataset, and 60 for the German dataset. (2) The optimal maximum depth remains consistent across datasets at 2, balancing between underfitting and overfitting risks. (3) Changes in the number of classical models and maximum depth have a limited impact on the discrimination power of non-default samples. The study primarily focuses on evaluating AUC, recall-2, and G-mean. (4) In addition to the number of classical models and maximum depth, the model also considers parameters like Learning rate and gamma, but their influence on model performance is relatively minor compared to the former parameters. (5) The recall-2 and G-mean metrics are more sensitive to variations in the number of classical models and maximum depth, especially with larger values potentially leading to overfitting. This insight can assist financial institutions in setting parameter ranges for identifying credit risks among farmers effectively.

Table 8
Performance comparison between the proposed model and benchmark models.

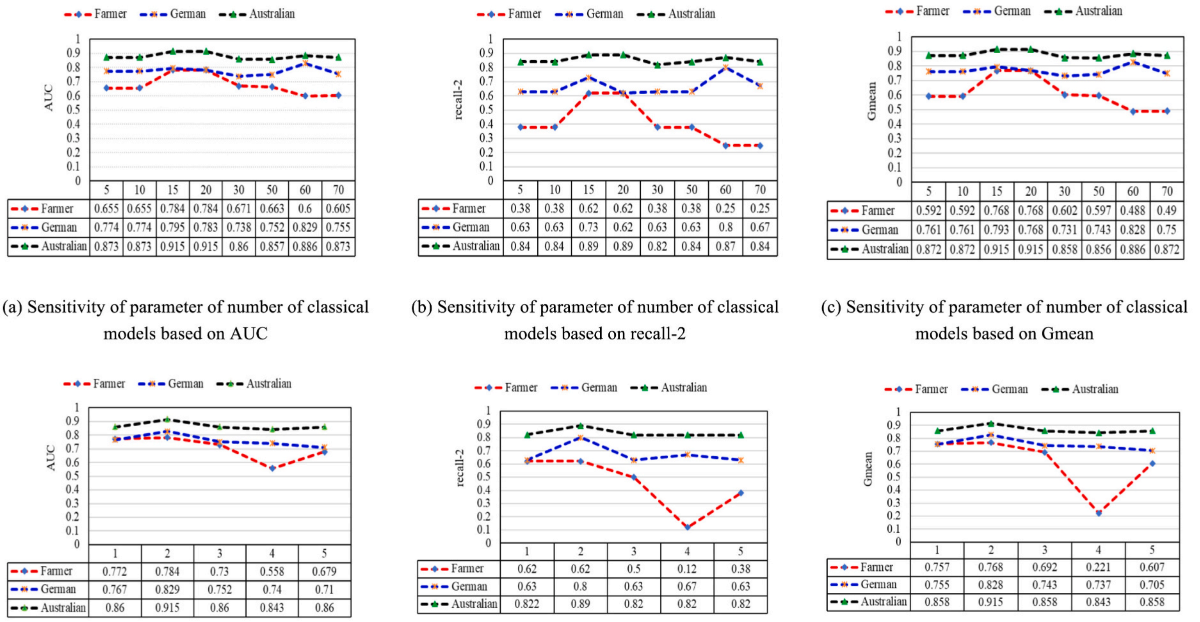
(1) Dataset	(2) Papers	(3) Models	(4) AUC	(5) recall-1	(6) recall-2	(7) Gmean
German	The proposed model	ADASYN-LCE model	**0.828**	**0.86**	0.8	**0.821**
	Marqués et al., 2012	Two-level classifier ensembles	0.79	0.43	0.89	0.619
	Jadhav et al., 2018	IGDFS model	0.767	/	/	/
	Xiao et al., 2020	GCSSE model	0.756	/	/	/
	Shen et al., 2021	Smote-LSTM-Adaboost model	0.796	/	/	/
	Xiao et al., 2021	RUS model	/	/	/	0.717
	Yang et al., 2024	FE-soft voting weight optimization model	0.795	/	/	/
Australian	The proposed model	ADASYN-LCE model	0.915	**0.94**	**0.89**	**0.915**
	Marqués et al., 2012	Two-level classifier ensembles	0.93	0.86	0.86	0.86
	Jadhav et al., 2018	IGDFS model	0.913	/	/	/
	Xiao et al., 2020	GCSSE model	0.912	/	/	/
	Xiao et al., 2021	RUS model	/	/	/	0.874
	Yang et al., 2024	FE-soft voting weight optimization model	0.917	/	/	/

Notes: column (1) shows two datasets; Column (3) displays models designed in different papers (in column (2)); Column (4) to (7) present the performance of identification of customer default risk using different models based on German and Australian dataset. “/” indicates that the paper does not present the performance of its designed model in a certain aspect.

Table 9
Recommended optimal parameter values for ADASYN-LCE model on different datasets.

(1) Parameter	(2) Definition	(3) Function
Number of classical models	Indicates the size of the number of classical models	The larger the number, the stronger the model's learning ability
Maximum depth	The maximum depth of the tree	Avoid overfitting of the model
Learning rate	Controlled the step size of each update step	Avoid overfitting of the model
Gamma	The minimum loss function descent required for node splitting	Adjusting the loss function

Note: columns (2) and (3) show the definition and function of the parameters (see column (1)), respectively.



(a) Sensitivity of parameter of number of classical models based on AUC

(b) Sensitivity of parameter of number of classical models based on recall-2

(c) Sensitivity of parameter of number of classical models based on Gmean

(d) Sensitivity of parameter of maximum depth based on AUC

(e) Sensitivity of parameter of maximum depth based on recall-2

(f) Sensitivity of parameter of maximum depth based on Gmean

Fig. 5. Sensitivity map of vital parameters of ADASYN-LCE based on the different datasets.
Note: Fig. 5 shows the performance of ADASYN-LCE under different data and parameters.

Table 10
Recommended optimal parameter values for ADASYN-LCE model on different datasets.

(1) Dataset	(2) Number of classical models	(3) Maximum depth	(4) Learning rate	(5) gamma
Farmer	20	2	0.1	0.1
German	60	2	0.2	0.01
Australian	15; 20	2	0.1	0.1

Note: columns (2) to (5) show the optimal values of different parameters for the ADASYN-LCE model in three datasets.

5.4. Interpretation of ADASYN-LCE model

Improving the interpretability of machine learning based credit scoring models is one of the important paths to solving the credit crisis caused by machine learning black boxes. This study discusses the interpretability of the ADASYN-LCE model from a global/local perspective in Section 4.2.3 above. Table 11 describes the definition of feature. Fig. 6 analyzes the global interpretability of the model, with it showcasing the top 15 features contributing to the prediction of default status among loan farmers. Notably, X_2, X_{32}, X_{25}, X_{12}, and X_{19} are the top five key factors that have a significant impact on predicting the default status of loan farmers. Among them, X_2 represents marital status, X_{32} represents regional agricultural gross output value, X_{25} represents per capita disposable income of rural residents, X_{12} represents customer gender, and X_{19} represents total liabilities.

Fig. 7 provides insights into the nuanced effects of a single feature, specifically X_2 Marital Status, on the model's prediction regarding farmers' default status. The *x*-axis delineates distinct categories corresponding to X2, where 0 signifies 'others,' 0.25 denotes 'widowhood,' 0.5 represents 'unmarried,' 0.75 indicates 'divorce,' and 1 signifies 'married and remarried.' On the other hand, the *y*-axis portrays the extent to which each category influences the model's predictions. A point represents a sample, and the color of the

Table 11
The definition of feature.

(1) Name	(2) Feature	(3) Definition
X_1	Educational level	The education level of the borrowing farmer: Elementary school, middle school, high school, etc
X_2	Marital status	The marital status of the borrowing farmer: Married, remarried, divorced, etc
X_3	Types of poverty	Whether they are poor farmers or not?
X_4	Nation	Han or other ethnic groups
X_5	Occupation	What is the occupation of a loan farmer? Such as technical personnel, waiter
X_6	Profession	What industry does the loan farmer belong to? Such as education, construction industry and so on
X_7	Local or not	Whether the loan farmer is local or not?
X_8	Number of students	How many students are attending school in the household of the loan farmer?
X_9	Number of preschool children	How many children are there in the family of the loan farmer?
X_{10}	Number of people who are sick or disabled	How many sick or disabled people are there in the family of the loan farmer?
X_{11}	Customer's age at the time of the loan	The age of the farmer at the time of the loan
X_{12}	Gender	Whether the loan farmer is male or female
X_{13}	Real estate ownership certificate	Does the loan farmer have a real estate ownership certificate?
X_{14}	Family size	How many people are there in the family of the debtor farmer?
X_{15}	Labor force population	How many people are able to work?
X_{16}	Labor share	The proportion of the number of workers in the household to the total number of households
X_{17}	Type of household registration	Whether the lender is a household or a non-household account
X_{18}	Total assets	The total amount of assets owned by the loaned households
X_{19}	Total liabilities	The amount of debt in the loan farmers
X_{20}	Debt Asset ratio	The ratio of liabilities of loaned households to their total assets
X_{21}	Monthly disposable income	The amount of monthly disposable income of the farmers who borrowed money
X_{22}	Length of local residence	The number of years that a household has lived in its locality
X_{23}	Purpose of loan	What is the use of loans for farmers? Such as building farmhouses, spending money
X_{24}	Per capita GDP	Reflects the level of wealth and economic development of a region
X_{25}	Per capita disposable income of rural residents	Reflects the disposable income level of farmers in the region
X_{26}	RMB savings deposits per capita in urban and rural areas	Reflects the deposit level of urban and rural residents in the region
X_{27}	Per capita deposit balance of financial institutions	Reflects the level of deposits accepted by financial institutions
X_{28}	Per capita loan balance of financial institutions	Reflects the per capita loan level of financial institutions
X_{29}	The registered urban unemployment rate	Feedback the unemployment situation of the town where the loan farmer is located
X_{30}	Deposit to loan ratio	Feedback the share of loans to deposits in the region where the farmer resides
X_{31}	Consumer Price Index	The trend and degree of change in the price level of goods and services purchased by residents
X_{32}	Regional agricultural gross output value	Feedback the total agricultural output value of the area where the loan farmer resides
X_{33}	Engel coefficient of rural households	Feedback the consumption structure of farmers in the region where the loan farmers reside

Note: column (3) shows the definition of features (see columns (1) and (2)).

point represents the size of the predicted value. The more the point color leans towards orange, the greater the probability that the customer belongs to non-default. A point located above the x-axis indicates that the category will have a positive impact on predicting features as non-default, while a point located below the x-axis indicates that the category will have a negative impact on predicting farmers as non-default. Orange (or gray) peaks represent the distribution of non-defaulting (or defaulting) samples. From Fig. 7, it can be seen that categories 0, 0.25, 0.5, and 0.75 have a negative impact on the model prediction of features as non-default, while category 1 has a positive impact on the prediction of farmers as non-default, which is consistent with the reality. For category 0.5, its mean contribution to the model prediction of farmers as non-default (in the middle of the orange peak) is significantly higher than its mean contribution to the model prediction of farmers as default. For category 1, these two distributions seem very close.

Figs. 8 and 9 delve into the local interpretability of the model, showcasing features that positively and negatively impact predictions of non-defaulting and defaulting farmers, respectively (taking the top 8 features respectively). In Fig. 8, the top 8 features that positively influence the model's prediction of non-defaulting farmers are X_{12}, X_{25}, X_2, X_{26}, X_5, X_{21}, X_1 and X_6, with their impact decreasing sequentially. Conversely, features X_{16}, X_{15}, X_{18}, X_{23}, X_{14}, X_{19}, X_{32} and X_{20} have a negative impact on the model prediction that the farmer belongs to non-default and the degree of impact increases sequentially. On the other hand, Fig. 9 shows the feature that positively and negatively impact on the model prediction of the farmer being in default (taking the top 8 features respectively). Features X_2, X_{24}, X_{32}, X_{12}, X_{19}, X_{26}, X_{23} and X_{33} positively impact the prediction of defaulting farmers, with their impact decreasing sequentially. Conversely, features X_6, X_{16}, X_{11}, X_{14}, X_{15} and X_{25} exhibit a negative impact on the model prediction that the farmer belongs to default.

Furthermore, the article includes a visualization in Fig. 10 that showcases the results of utilizing ADASYN-LCE to forecast a test dataset of farmers. The x-axis represents the sample type (0 for non-defaulting and 1 for defaulting), while the y-axis displays the predicted values for each sample. In the visualization, black dots represent samples with incorrect predictions, while red dots denote samples with correct predictions. The analysis in Fig. 10 indicates those 7 non-defaulting samples and 3 defaulting samples were inaccurately predicted, shedding light on the model's performance and areas for refinement.

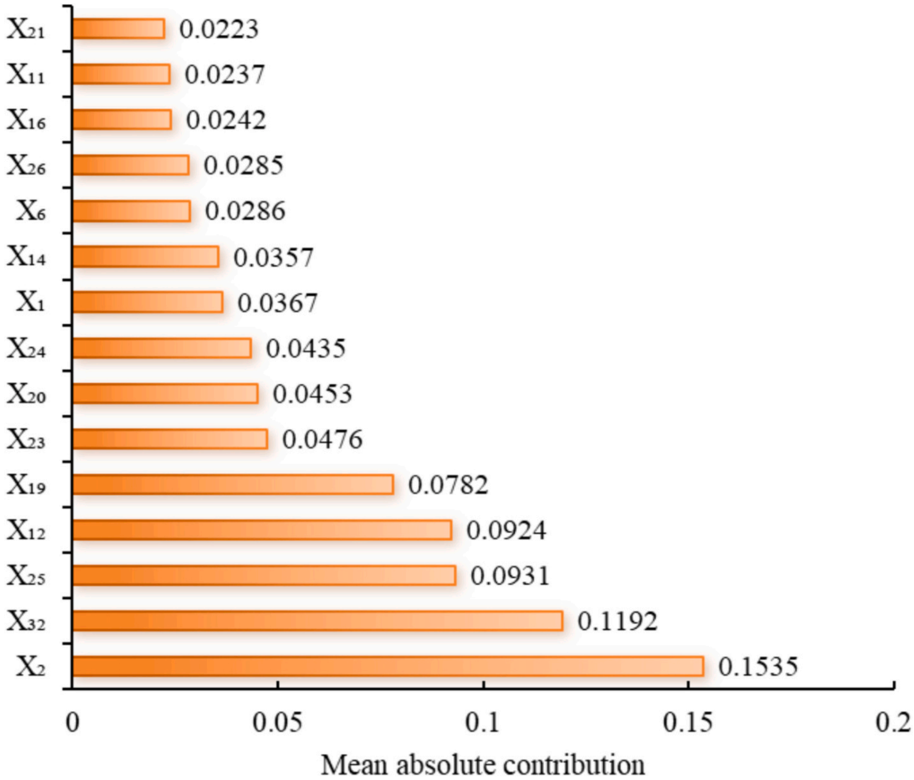


Fig. 6. The average absolute impact of features on default status.
Note: Fig. 6 illustrates the relative importance of features for ADASYN-LCE in identifying the default status of farmers.

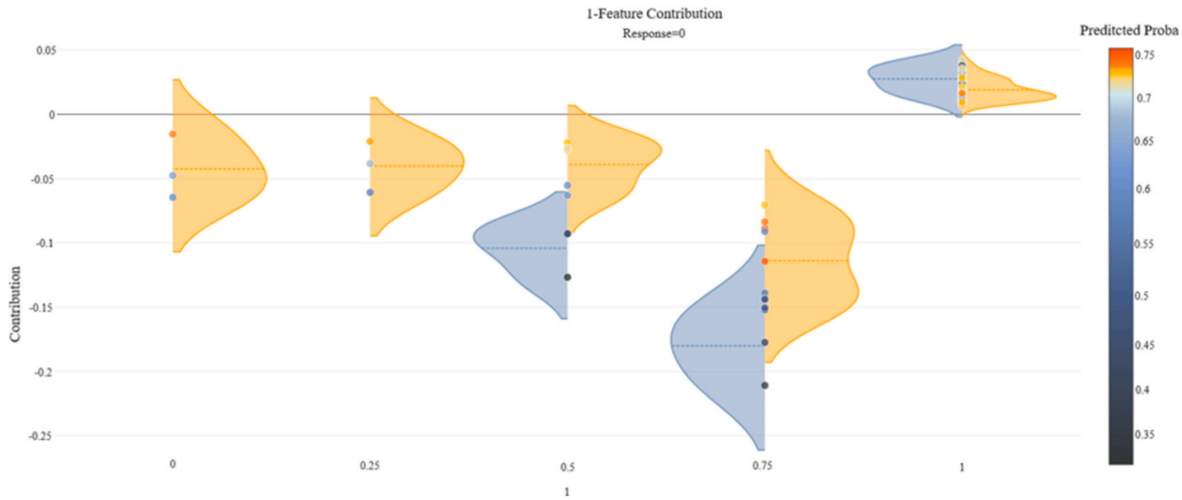


Fig. 7. The impact of X_2 on the model's prediction of farmers being non-defaulting farmers.
Note: Fig. 7 shows the contribution of different classification attributes of classification feature X_2 to predicting farmers as non-defaulting farmers.

6. Discussion

Based on the empirical results from Sections 5.1, 5.2, 5.3, 5.4, we present the following discussions:

(1) A discussion on credit risk and default, as well as the differences between microfinance and traditional banking operations, along with an analysis of the necessary conditions for applying machine learning to microfinance prediction. Credit risk refers to the potential economic loss faced by lenders if borrowers fail to meet their debt obligations (Basel Committee on Banking

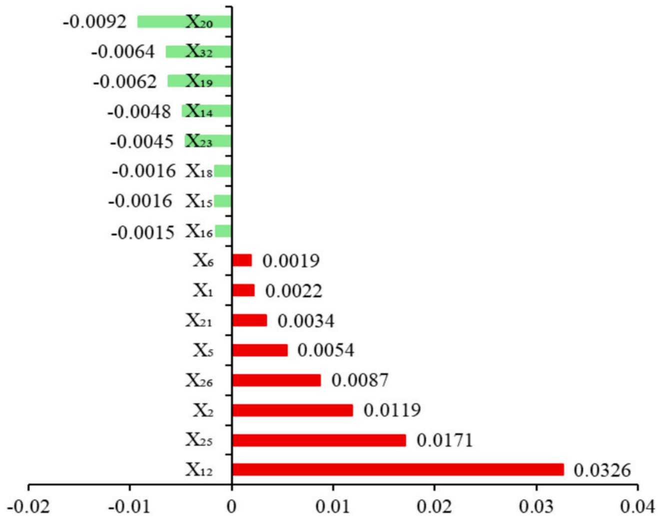


Fig. 8. Features that have positive or negative impacts on the model prediction of a non-defaulting farmer as a non-defaulting farmer.
Note: Fig. 8 shows the contribution of different classification attributes of classification feature X_2 to predicting farmers as non-defaulting farmers.

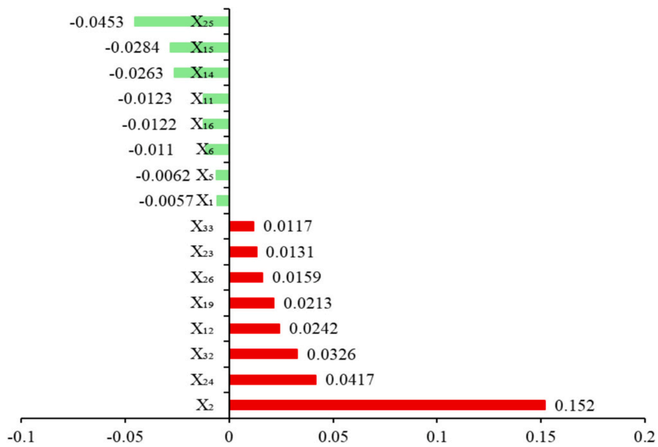


Fig. 9. Features that have positive or negative impacts on the model's prediction of a defaulting farmer as a defaulting farmer.
Note: Fig. 9 shows the different impacts on the model's prediction of a defaulting farmer as a defaulting farmer.

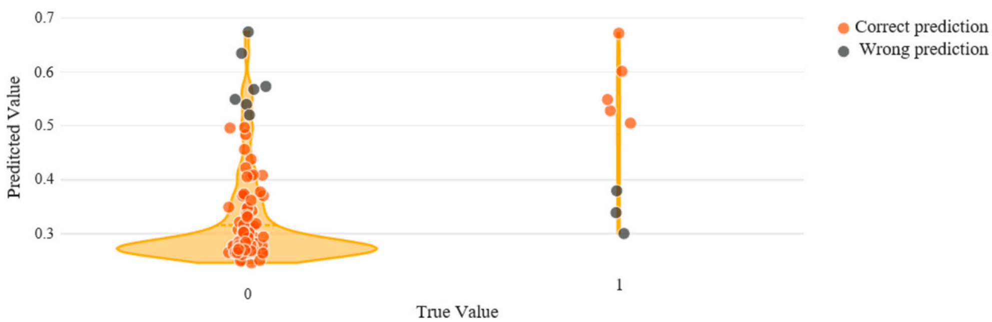


Fig. 10. The visualization of prediction results for non-default and default farmers from test data.
Note: Fig. 10 shows the prediction performance for non-default and default farmers from test data.

Supervision (BCBS), 2010). At the banking level, credit risk primarily involves the potential for loan applicants to default on their loans (Xu et al., 2024). If borrowers fail to make timely repayments or declare bankruptcy, the bank must absorb the losses associated with the outstanding loans. This situation necessitates banks implementing stringent credit review and risk control mechanisms to assess the creditworthiness of borrowers and mitigate potential losses.

Microfinance represents a new credit product developed by banks targeting vulnerable borrowing clients, such as farmers, who historically been excluded from credit markets. These groups typically consist of low-net-worth individuals (Wang et al., 2021). In contrast to traditional loan operations, identifying credit risk in microfinance relies more heavily on non-financial information. This is intrinsically linked to the objective reality that farmers often lack financial information, which is one of the underlying reasons contributing to their financing constraints (Medina-Olivares et al., 2022). This perspective aligns with the conventional understanding that a borrower's financial information is a key determinant in assessing loan default risk (Altman and Sabato, 2007). Moreover, microfinancing is characterized by short lending periods, small loan amounts, and simple approval processes, which necessitate a more precise and stringent identification of credit risk.

Chai et al. (2024) thoroughly discuss the applicability of machine learning versus simpler econometric methods for assessing credit risk in farmer microcredit, concluding that machine learning offers greater potential advantages in this domain. The reason for this advantage is that machine learning is well-suited for capturing the complex nonlinear relationships between non-financial information and the credit risk of borrowing farmers (Dong et al., 2023). It exhibits high accuracy in predicting risk and can lower information acquisition costs, thereby improving the accessibility of microloans (Wang et al., 2024a). However, the machine learning models designed need to strike an appropriate balance between accuracy and interpretability. The ADASYN-LCE model developed in this study achieves this trade-off effectively.

(2) The model designed in this paper performs well in terms of AUC, recall-2, and G-mean compared to other classical models, but shows poor performance in terms of recall-1. However, the better performance of recall-2 in the model designed in this paper aligns more closely with the requirements of financial institutions focusing on identifying default samples. Of course, recall-1 should not differ significantly from other typical models to prevent financial institutions from losing potential customers and incurring substantial losses. Existing research on the trade-off between recall-1 and recall-2 aims to ensure that recall-2 performance is relatively good while maintaining recall-1 at a similar level (Marqués et al., 2012; Liu et al., 2022b). This study's conclusions are similar to existing research and meet the requirements of financial institutions for credit risk evaluation models.

(3) The model designed in this paper enhances credit risk identification for microfinance targeting farmers, but its performance is not better than that in data of ordinary customers such as loan customers in Germany and Australia dataset. This may be due to the fact that dataset for microfinance targeting farmers includes more indirect feedback on their credit characteristics and lacks many direct features related to production, sales, costs, market, profit, and risk compared to data for ordinary customers. This part of the feature, also known as soft information, is generally relatively hard information (direct features, such as financial information), and the value of soft information is relatively high when the information of loan customers is not transparent (Cornée, 2019; Brighi et al., 2019). However, this study performs well in credit risk identification for microfinance targeting farmers compared to existing research. For example, in this paper, the characteristics of macro-environment in the area where the loan farmers are located account for nearly 1/3 of the total characteristics. The model designed by Medina-Olivares et al. (2022) had an AUC value of 0.7414, while the model designed in this paper achieved an AUC value of 0.784. It is worth noting that the former used data on microfinance targeting farmers from the same data source as this paper. Of course, the datasets of these three countries vary greatly from country to country, especially in terms of regulatory frameworks and standards, market information transparency, and credit culture. However, the fundamental reason for the supplementary use of Germany and Australian datasets in this paper is that these two sets of data belong to public datasets and are the fundamental basis for the existing literature to conduct experiments and discriminate model performance and their applicability, for example, the literature in Table 8 above uses these two sets of public data for design model verification.

(4) Discussion on the optimal parameters of the model and the performance evaluation of the model under different training and testing set ratios. The integration of LC, bagging, and boosting in the LCE model introduced in this article results in relatively high model complexity. To balance model performance with computational cost, we employ a step-by-step parameter adjustment process to identify the optimal parameters (Xiao et al., 2017). Referring to Gunnarsson et al. (2021) and Zedda (2024), as well as considering the specific characteristics of our empirical data, we propose the following ranges for hyperparameters: Learning rate: {0.1, 0.2, 0.3}; Maximum depth: {2, 3, 4.5}; Number of classical models: {5, 10, 15, 20, 60}; and gamma: {0.01, 0.1, 0.5}. Among these parameters, the influence of the Learning rate and Maximum depth on model performance is relatively stable. Therefore, we first fix the other parameters and incrementally adjust the values of the Learning rate and Maximum depth from their recommended ranges to determine their optimal values in the training set. Subsequently, we fix the Learning rate and Maximum depth to find the optimal values for the remaining two parameters. The optimal parameter values are summarized in Table 10. Due to the complexity of the parameter optimization process, it cannot be fully encapsulated in a table, and thus it remains somewhat of a black box operation. This is a common limitation associated with many machine learning and artificial intelligence methods (Zedda, 2024). Furthermore, this paper discusses the performance of the proposed model under different training and testing set ratios, as shown in Table 12. We found that the proposed model consistently demonstrates the best performance in identifying the credit risk of loan clients across all datasets, with the average rank consistently positioned at the top (see column (8) in Table 12). Additionally, at a 9:1 ratio, the ADASYN-LCE model performs the best compared to other ratios, particularly in identifying defaulting clients.

Finally, in the characteristics of microfinance for farmers, the features representing the basic personal information of farmers play a crucial role in identifying their own credit risks, while the importance of the characteristics of the farmers' family is not high. As shown in Fig. 6 above, the characteristics of farmers' personal information such as education level, marital status, profession, age, and gender contribute the most to the identification of credit risks for farmers' microfinance, with a value of 0.3349 (=0.0367 + 0.1535 + 0.0286 + 0.0237 + 0.0924). This is consistent with the conclusion of Bai et al. (2019) regarding credit feature analysis related to the creditworthiness of farmers, namely, that personal related features of farmers such as education level and skill status are crucial in determining the credit level of farmers. The characteristics of farmers' families such as family size and labor share contribute the least to the identification of credit risks for farmers microfinance, with an impact value of 0.0599 (=0.0357 + 0.0242).

7. Conclusion

Empowering microfinance with digital technology unleashes its unlimited potential in serving weak-quality clients such as impoverished farmers, making it a crucial tool for various financial institutions to implement inclusive finance strategies. The most important technology in this regard is credit risk evaluation based on machine learning, which provides an essential decision-making basis for further microfinance disbursement. However, developing a credit risk evaluation model for microfinance targeting farmers that possesses features like superior performance, robustness, and high interpretability remains a critical issue that needs to be addressed urgently. To tackle this challenge, this study introduces a new hybrid ensemble ADASYN-LCE model with data balance, bias-variance minimization, and enhanced interpretability.

And the paper conducts an empirical analysis using data from 1298 farmers' microfinance in Chinese poverty alleviation institutions, which called as CHONGHO BRIDGE. The purpose is to verify the performance of the proposed ADASYN-LCE model for assessing credit risks of farmers. The empirical results show that the ADASYN-LCE model has an average rank of 2.1, indicating better performance in ranking compared to ensemble methods like ADASYN-XGBoost. Furthermore, the null hypothesis of the Friedman test is rejected at a significant level of 0.01 for all evaluation measures, indicating differences among the models. Additionally, the robustness of the ADASYN-LCE model is tested using German data and Australian data, and the results demonstrate that the ADASYN-LCE model performs the best in assessing credit risks of farmers' microfinance.

Finally, from the perspective of global interpretability analysis, the paper explored that characteristic such as Marital status, Regional academic gross output value, per capital disposable income of rural residents, Gender, and Total liabilities are urgently important for identifying credit risks in microfinance for farmers. Especially, the characteristics of feedback on the personal information of loan farmers play the greatest role in identifying credit risks in microfinance. Based on local interpretability, among the most influential characteristics of marital status, Married or reminded farmers are more likely to be non-defaulting farmers.

Furthermore, this study has potential policy implications. In terms of establishing credit risk evaluation models with low bias, low variance, and interpretability, our research suggests that policymakers need to focus on collecting individual features data of farmers when establishing farmer credit data, which may help microfinance institutions mitigate the situation where they are unwilling to provide microfinance due to information asymmetry with loan farmers. And policymakers also need to pay attention to the interpretability of credit risk evaluation models to avoid issues such as unfair lending. In addition, the model prosed in this study can also provide new ideas for microfinance institutions to establish their own credit risk evaluation systems. In the future, we will further explore how to help farmers improve their reputation and assist them in obtaining microfinance more effectively. In addition, although the data sets used in this study are from developing countries and developed countries, the essential differences in the credit risk features of different groups in these two types of countries due to different regulatory systems and mechanisms are not deeply explored. We will conduct further research on this issue on the basis of more transparency in the credit risk characteristics of different datasets.

Ethics approval and consent to participate

Not applicable.

Table 12
ADASYN-LCE's performance based on different proportions of training and testing sets.

(1) Dataset	(2) Train dataset/Test dataset	(3) AUC	(4) recall-1	(5) recall-2	(6) Gmean	(7) Average rank	(8) Is it ranked first (Yes or No)
Farmer	9:1	0.784	0.94	0.62	0.768	2.25	Yes
	8:2	0.706	0.93	0.48	0.669	2.5	Yes
	7:3	0.598	0.95	0.25	0.486	2.5	Yes
	9:1	0.828	0.86	0.8	0.828	1.75	Yes
German	8:2	0.761	0.92	0.6	0.744	1.25	Yes
	7:3	0.727	0.9	0.56	0.708	1.25	Yes
	9:1	0.915	0.94	0.89	0.915	1	Yes
Australian	8:2	0.907	0.89	0.92	0.907	1.25	Yes
	7:3	0.882	0.9	0.86	0.881	1.25	Yes

Note: columns (3) to (8) show the ADASYN-LCE performance based on different proportions (in column (2)) of the training and test sets under different datasets (in column (1)). 'Yes' represents that the proposed model ADASYN-LCE performs the best compared to other models such as ADASYN XGBoost, ADASYN LightGBM, and ADASYN AdaBoost. On the contrary, 'No' indicates that it is not the optimal model.

Consent for publication

All authors are very positive to publish this manuscript on this journal.

Funding

This paper was supported by the Major Program of the National Social Science Foundation of China (Grant No. 23&ZD175), the National Natural Science Foundation of China (Grant Nos. 72173096, 72303139, 71873103), the Postdoctoral Fellowship Program (Grade B) of China Postdoctoral Science Foundation (Grant No. GZB20240546), the Shandong Provincial Natural Science Foundation (Grants No. ZR2024QG002).

CRediT authorship contribution statement

Nana Chai: Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mohammad Zoynul Abedin:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lian Yang:** Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Baofeng Shi:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Conceptualization.

Declaration of competing interest

There is no competing interest among the authors.

Acknowledgments

We are very grateful to the editors and anonymous reviewers of this journal, whose valuable feedback has played an important guiding role in this manuscript.

Appendix A. Definition of some key concepts

Appendix Table 1

Name	Definition
Defaulting farmer	Farmers who have obtained loans have not repaid the loans on time or in full.
Non-defaulting farmer	Farmers who have obtained loans have repaid the loans on time and in full.
SMOTE	A technique used to solve the problem of class imbalance in classification tasks, aimed primarily at increasing the number of minority class samples by generating synthetic examples, thereby balancing the proportions of various classes and improving the learning effects and predictive performance of the model.
K-nearest neighbors	A commonly used non-parametric classification and regression algorithm, whose fundamental idea is to determine the class of a new sample or to predict its value by measuring the distance between different data points.
Divide-and-conquer method	An algorithm design strategy that involves breaking down a complex problem into multiple smaller sub-problems, solving these sub-problems recursively, and then combining their solutions to obtain a solution to the original problem.
LIME	A model-agnostic interpretability tool designed to help understand the decision-making process of black box machine learning models. The core idea of LIME is to explain the local behavior of complex models by building a simple interpretable model around a specific prediction.
SHAP	A framework for interpreting machine learning model predictions based on Shapley values from game theory. The core idea of SHAP is to allocate the model's output (prediction) to the input features, quantifying each feature's contribution to the final prediction, thus providing clear and consistent explanations.
SmartExplainer	A model interpretability tool designed to provide intuitive and easy-to-understand explanations for black box machine learning models. It integrates multiple interpretability techniques, helping users understand the decision-making process of the model by analyzing the relationships between the model's outputs and the input features.
WebApp	In Shapash, it refers to the visual user interface provided by the tool, which allows users to interact more conveniently with the outputs and explanations of machine learning models. Shapash's WebApp aims to combine model interpretability with user experience, allowing users to display model predictions, feature importance analysis, SHAP values, and other information through a visual interface.
Charts	In Shapash, Charts refer to the graphs and figures used to visualize the explanations and output results of machine learning models. Shapash helps users understand the predictions, feature importance, and other relevant data of the model more intuitively through the Charts module. These charts are typically generated based on methods like SHAP values and are designed to provide users with clear and easy-to-understand visual representations

Data availability

Our data will be available on request.

References

Akkoc, S., 2012. An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: the case of Turkish credit card data. Eur. J. Oper. Res. 222 (1), 168–178.

Altman, E.I., Sabato, G., 2007. Modelling credit risk for smes: evidence from the US market. Abacus 43 (3), 332–357.

Altman, E.I., Balzano, M., Giannozzi, A., Srhoj, S., 2023. Revisiting SME default predictors: the omega score. J. Small Bus. Manag. 61 (6), 2383–2417.

Bai, C.G., Shi, B.F., Liu, F., Sarkis, J., 2019. Banking credit worthiness: evaluating complex relationships. Omega Int. J. Manag. Sci. 83, 26–38.

Bank of China (BC), 2010. Credit Rating Index System and the Definition of Scoring Standard for Custom in Bank of China. Bank of China Ltd.

Basel Committee on Banking Supervision (BCBS), 2010. Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems. Bank for International Settlements, Basel, Switzerland.

Benami, E., Carter, M.R., 2021. Can digital technologies reshape rural microfinance? Implications for savings, credit, & insurance. Appl. Econ. Perspect. Policy 43 (4), 1196–1220.

Brighi, P., Lucarelli, C., Venturelli, V., 2019. Predictive strength of lending Technologies in Funding SMEs. J. Small Bus. Manag. 57 (4), 1350–1377.

Buchen, T., Wohlrabe, K., 2011. Forecasting with many predictors: is boosting a viable alternative? Econ. Lett. 113 (1), 16–18.

Buera, F.J., Kaboski, J.P., Shin, Y., 2021. The macroeconomics of microfinance. Rev. Econ. Stud. 88 (1), 126–161.

Caigny, A.D., Bock, K.W.D., Verboven, S., 2024. Hybrid black-box classification for customer churn prediction with segmented interpretability analysis. Decis. Support. Syst. 181, 114217.

Carcillo, F., Dal, P.A., Le, B.Y., Caelen, O., Mazzer, Y., Bontempi, G., 2018. SCARFF: a scalable framework for streaming credit card fraud detection with spark. Inf. Fusion 14, 182–194.

Chai, N.N., Shi, B.F., Hua, Y.T., 2023. Loss given default or default status: which is better to determine farmers' credit ratings? Financ. Res. Lett. 53, 103674.

Chai, N.N., Abedin, M.Z., Wang, X.L., Shi, B.F., 2024. Growth potential of machine learning in credit risk predicting of farmers in the industry 4.0 era. Int. J. Financ. Econ. https://doi.org/10.1002/ijfe.3010.

Chen, Y.J., Calabrese, R., Martin-Barragan, B., 2024. Interpretable machine learning for imbalanced credit scoring datasets. Eur. J. Oper. Res. 312 (1), 357–372.

China Construction Bank (CCB), 2007. Enterprise Customer's Credit Rating Rules of CCB. China Construction Bank, Beijing.

Chongho Bridge (CB), 2023. Every Hardworking Individual has the Right to Pursue a Better Life. https://www.chongho.net/client.html.

Cornée, S., 2019. The relevance of soft information for predicting small business credit default: evidence from a social bank. J. Small Bus. Manag. 57 (3), 699–719.

Dong, Y., Gou, Q., Qiu, H., 2023. Big tech credit score and default risk——evidence from loan-level data of a representative microfinance company in China. China Econ. Rev. 81, 102010.

Dumitrescu, E., Hué, S., Hurlin, C., Tokpavi, S., 2022. Machine learning for credit scoring: improving logistic regression with non-linear decision-tree effects. Eur. J. Oper. Res. 297 (3), 1178–1192.

Ezgi, Z., Selma, A., 2016. A hybrid approach of differential evolution and artificial bee colony for feature selection. Expert Syst. Appl. 62, 91–103.

Fauvel, K., Fromont, É., Masson, V., Faverdin, P., Termier, A., 2022. XEM: An explainable-by-design ensemble method for multivariate time series classification. Data. Min. Knowl. Discov. 36 (3), 917–957.

Fenech, J.P., Yap, K.Y.K., Shafik, S., 2016. Modelling the recovery outcomes for defaulted loans: a survival analysis approach. Econ. Lett. 145, 79–82.

Fitch Ratings (FR), 2013. Fitch Ratings Global Corporate Finance 2012 Transition and Default Study. Credit Market Research-Fitch Ratings, New York.

Gao, W., Liu, Y., Yin, H., Zhang, Y., 2022. Social capital, phone call activities and borrower default in mobile micro-lending. Decis. Support. Syst. 159, 113802.

Gunnarsson, B.R., Vanden Broucke, S., Baesens, B., Oskarsdottir, M., Lemahieu, W., 2021. Deep learning for credit scoring: do or don't? Eur. J. Oper. Res. 295, 292–305.

Hayashi, Y., 2016. Application of a rule extraction algorithm family based on the re-RX algorithm to financial credit risk assessment from a Pareto optimal perspective. Oper. Res. Persp. 3, 32–42.

Jadhav, S., He, H., Jenkins, K., 2018. Information gain directed genetic algorithm wrapper feature selection for credit rating. Appl. Soft Comput. 69, 541–553.

Karlan, D., Zinman, J., 2011. Microcredit in theory and practice: using randomized credit scoring for impact evaluation. Science 332 (6035), 1278–1284.

Koutanaei, F., Sajedi, H., Khanbabaei, M., 2015. A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit. J. Retail. Consum. Serv. 27, 11–23.

Lee, T., Chiu, C.C., Lu, C.C., Chen, I.F., 2002. Credit scoring using the hybrid neural discriminant technique. Expert Syst. Appl. 23 (3), 245–254.

Leif, A.B., Roy, M., Trond, R., 2014. The association between microfinance rating scores and corporate governance: a global survey. Int. Rev. Financ. Anal. 35, 268–280.

Li, K., Niskanen, J., Kolehmainen, M., Niskanen, M., 2016. Financial innovation: credit default hybrid model for SME lending. Expert Syst. Appl. 61, 342–355.

Liu, Y., Yang, M.L., Li, A.Z., 2022a. Applying machine learning algorithms to predict default probability in the online credit market: evidence from China. Int. Rev. Financ. Anal. 79, 101971.

Liu, W., Fan, H., Xia, M., Xia, M., 2022b. A focal-aware cost-sensitive boosted tree for imbalanced credit scoring. Expert Syst. Appl. 208, 118158.

Liu, W.A., Fan, H., Xia, M., 2023. Tree-based heterogeneous cascade ensemble model for credit scoring. Int. J. Forecast. 39 (4), 1593–1614.

Machado, M.R., Karray, S., 2022. Assessing credit risk of commercial customers using hybrid machine learning algorithms. Expert Syst. Appl. 200, 116889.

Marqués, A.I., Garcia, V., Sanchez, J.S., 2012. Two-level classifier ensembles for credit risk assessment. Expert Syst. Appl. 39 (12), 10916–10922.

Medina-Olivares, V., Calabrese, R., Dong, Y.Z., Shi, B.F., 2022. Spatial dependence in microfinance credit default. Int. J. Forecast. 38 (3), 1071–1085.

Papouskova, M., Hajek, P., 2019. Two-stage consumer credit risk modelling using heterogeneous ensemble learning. Decis. Support. Syst. 118, 33–45.

Shen, F., Zhao, X.C., Kou, G., Alsaadi, F., 2021. A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique. Appl. Soft Comput. 98, 106852.

Sigrist, F., Hirnschall, C., 2019. Grabit: gradient tree-boosted Tobit models for default prediction. J. Bank. Financ. 102, 177–192.

Simsek, S., Dag, A., Tiahrt, T., Oztekin, A., 2021. A Bayesian belief network-based probabilistic mechanism to determine patient no-show risk categories. Omega Int. J. Manag. Sci. 100, 102296.

Simumba, N., Okami, S., Kodaka, A., Kohtake, N., 2022. Multiple objective metaheuristics for feature selection based on stakeholder requirements in credit scoring. Decis. Support. Syst. 155, 113714.

Standard&Poor's (SP)., 2011. Standard&Poor's Guide to Credit Rating Essentials. Standard&Poor's, New York.

Sun, S.L., Liang, H., 2021. Globalization and affordability of microfinance. J. Bus. Ventur. 36 (1), 106065.

Teng, H.W., Kang, M.H., Lee, I.H., Bai, L.C., 2024. Bridging accuracy and interpretability: a rescaled cluster-then-predict approach for enhanced credit scoring. Int. Rev. Financ. Anal. 91, 103005.

The Central People's Government of the People's Republic of China (CPGPRC), 2021. Opinions of the Central Committee of the Communist Party of China and the State Council on Comprehensively Promoting Rural Revitalization and Accelerating Agricultural and Rural Modernization. https://www.gov.cn/zhengce/2021-02/21/content_5588098.htm.

Trivedi, S.K., 2020. A study on credit scoring modeling with different feature selection and machine learning approaches. Technol. Soc. 63, 101413.

Wang, B., Yu, Y.Y., Yang, Z., Zhang, X., 2021. Microfinance institutions and peer-to-peer lending: what does microfinance competition bring? Pac. Basin Financ. J. 67, 101557.

Wang, J.J., Zhuang, Z.Z., Gao, D.M., 2023. An enhanced hybrid model based on multiple influencing factors and divide-conquer strategy for carbon price prediction. Omega Int. J. Manag. Sci. 120, 102922.

Wang, Z., Xiao, J., Wang, L., Yao, J., 2024. A novel federated learning approach with knowledge transfer for credit scoring. Decis. Support. Syst. 177, 114084.

Wang, H.J., Du, S.A., Chen, G., Wu, W.T., 2024a. Does digital credit alleviate household income vulnerability? Pac. Basin Financ. J. 88, 102542.

Xiao, Y., Liu, C., Li, Y.Y., Liu, N., 2017. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. Expert Syst. Appl. 78, 225–241.

Xiao, J., Zhou, X., Zhong, Y., Xie, L., Gu, X., Liu, D.H., 2020. Cost-sensitive semi-supervised selective ensemble model for customer credit scoring. Knowl.-Based Syst. 189, 105118.

Xiao, J., Wang, Y.D., Chen, J., Xie, L., Huang, J., 2021. Impact of resampling methods and classification models on the imbalanced credit scoring problems. Inf. Sci. 569, 508–526.

Xie, X.F., Zhang, J., Luo, Y.X., Gu, J., Li, Y., 2023. Enterprise credit risk portrait and evaluation from the perspective of the supply chain. Int. Trans. Oper. Res. 1–31.

Xu, X., Zhou, C., Wang, Z., 2009. Credit scoring algorithm based on link analysis ranking with support vector machine. Expert Syst. Appl. 36 (2), 2625–2632.

Xu, Y., Kou, G., Peng, Y., Ding, K., Ergu, D., Alotaibi, F., 2024. Profit- and risk-driven credit scoring under parameter uncertainty: a multiobjective approach. Omega 125, 103004.

Yang, D.Q., Xiao, B.Q., Cao, M., Shen, H., 2024. A new hybrid credit scoring ensemble model with feature enhancement and soft voting weight optimization. Expert Syst. Appl. 238, 122101.

Yi, Z.L., Liang, Z.M., Li, F., 2023. Financial risk prediction in supply chain finance based on buyer transaction behavior. Decis. Support. Syst. 170, 113964.

Yuan, K.P., Chi, K.P., Zhou, Y., Yin, H., 2022. A novel two-stage hybrid default prediction model with k-means clustering and support vector domain description. Res. Int. Bus. Financ. 59, 101536.

Zedda, S., 2024. Credit scoring: does XGboost outperform logistic regression? A test on Italian SMEs. Res. Int. Bus. Financ. 70, 102397.

Zhang, X.M., Yu, L., 2024. Consumer credit risk assessment: a review from the state-of-the-art classification algorithms, data traits, and learning methods. Expert Syst. Appl. 237, 121484.