Omega 83 (2019) 26–38

Contents lists available at ScienceDirect

Omega

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

Banking credit worthiness: Evaluating the complex relationships☆

Chunguang Bai[a,*], Baofeng Shi[b,*], Feng Liu[c], Joseph Sarkis[d]

[a] School of Management and Economics, University of Electronic Science and Technology of China, No. 2006, Xiyuan Av, Chengdu, 611731, PR China
[b] College of Economics & Management, Northwest A&F University, No. 3 TaiCheng Road, Yangling, Shaanxi 712100, PR China
[c] School of Management Science and Engineering, Dongbei University of Finance & Economics, Jianshan Street 217, Dalian 116025, PR China
[d] School of Business, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609-2280, USA

ARTICLE INFO

Article history:
Received 12 August 2017
Accepted 2 February 2018
Available online 9 February 2018

Keywords:
Or in banking
Credit risk
Fuzzy rough-set
Fuzzy C-means
Farmers
China

ABSTRACT

In developing economies agriculture and farming play crucial roles for economic sustainable development. Farmer credit risk evaluation is an important issue when determining financial support to farmers, improving agricultural supply chain performance, and ensuring profitability of financial institutions. Credit risk evaluation, or creditworthiness, is not a trivial exercise due to various complexities. Honoring complexity is necessary to effectively evaluate and predict farmer creditworthiness. A methodology using fuzzy rough-set theory and fuzzy C-means clustering is used to evaluate and investigate the complex relationships between farmer characteristics, competitive environmental factors, and farmer credit level. The methodology is detailed using actual bank data from 2044 farmers within China. This empirical methodology generates decision rules that provide insight to more complex relationships than can be found through standard econometric multivariate approaches. A rule-based methodological outcome can be used to predict the creditworthiness of farmers and to aid in agricultural loan decision making. Prediction accuracy of the rule-base was 81.16%. A central finding is that education and skills related characteristics are important for determining farmer credit-worthiness. Other implications are presented along with study limitations and future research directions.

1. Introduction

China feeds 21% of the world's population using 7% of the world's arable farmlands [55]. The Chinese central government, in publically available documents from 2004 to 2015, has reinforced the importance of agriculture, with the countryside and farmers receiving much of the spotlight in terms of criticality to China's development [24]. Even though 64.7% of the population in China is rural and agricultural they possess less than 20% of the entire wealth in China [23]. In order to realize a sustainable and equitable society as set by the Chinese government, increasing the wealth of farmers and modernizing villages is paramount. An important and prominent problem is the disconnect and difficulty in managing the relationship between agricultural and financial lending policies.

Farmer credit is necessary for access to working capital and credit loans offered by financial institutions [75]. Finances are important for basic stocks, but are also needed for farming modernization through introduction of advanced agricultural technologies, to build inherent flexibilities and enhance the ability to cope with risks. For these reasons farmer credit is a matter for economic sustainable development and poverty reduction in many developing countries [75]. Bank and financial credit in emerging nations, especially at the "bottom of the pyramid" has started to gain in importance, as exemplified by the Grameen Bank, which has focused on providing microfinance for poor Bangladeshis and has successfully helped millions of poor Bangladeshis overcome poverty [56]. Farmer credit practices can be informal which may result in corruptive, predatory and exploitative practices [39]. A vast majority of Chinese farmers are arguably members of this bottom of the pyramid.

Given the importance of bank credit in emerging economies, especially those that have significant rural and agricultural populations, credit evaluation remains a complex activity [35]. This complex activity requires large amounts data and a relatively complex evaluation process. This complexity is magnified by farmers' characteristics and contextual environmental factors, with interactions and relationships amongst factors needing to be considered. As complexity increases, the ability to use information to evaluate and predict credit level becomes more difficult. Having ways to reduce this complexity and providing clarity to both farmers and lenders may greatly enhance the process, lessening risk and improving fair practices. Understanding the factors that most contribute to successful loans can help financial institutions identify

☆ This manuscript was processed by Associate Editor Prof. B. Lev.

* Corresponding authors.
E-mail addresses: chunguang.bai@gmail.com (C. Bai), shibaofeng@nwsuaf.edu.cn (B. Shi), drliufeng@mail.dlut.edu.cn (F. Liu), jsarkis@WPI.edu (J. Sarkis).

https://doi.org/10.1016/j.omega.2018.02.001
0305-0483/

creditworthy applicants. It will also help farmers focus on those characteristics that allow them to be successful in paying off loans, which implies economic success of their farms. It will also benefit governmental agencies and policy makers by supporting the factors that can lead to economically successful farmers, allowing them to pursue policies to buttress farmers from financial risks.

A comprehensive evaluation of farmer credit policies and practices can provide insights to manage social sustainability in agrarian regions of emerging economies. Using financial information from bank lenders to farmers, this study seeks to determine which farmer characteristics and what contextual environmental factors will lead to good or bad credit for farmers. A formal methodology can help this process by helping to address and identify complexity and provide a general evaluation system. This paper introduces a methodology to identify various relationships, rules, among farmers' characteristics and environmental factors and creditworthiness of farmers.

In this paper fuzzy C-means (FCM) is used to discretize continuous data, making the data amenable to rough set theory [41,42,57]. Dependency degrees, which rely on fuzzy rough set theory (FRST), are used to identify a usable subset of farmer characteristics and environmental factors for creditworthiness with minimal information loss. Using actual loan data from a large bank, rough set theory (RST) is used with Boolean algebra to arrive at rules of various relationships between the conditional attributes (farmer characteristics and environmental factors) and creditworthiness. Rules that will aid farmers, financial institutions, and governmental agencies support improved access and quality of credit decisions.

The contribution of this paper includes the development of a new integrative methodology that combines FRST and FCM. Another contribution is developing insights and relevance of complex relationships between farmer characteristics, contextual environmental factors and creditworthiness in China. More broadly, this study provides initial evidence and relationships for generalized evaluation mechanisms (rules) for financial institutions to predict creditworthiness of farmers. This study also addresses some of the methodological issues facing previous techniques applied to credit evaluation, especially correlative econometric models. For example, the proposed methodology addresses complex relationships of indicators, such as interactions amongst factors; easing parametric assumptions; allowance of equifinality and non-linearity of relationships, all which limit the application and predictive power of previous techniques.

The remainder of this paper begins with a literature review to help set the study's theoretical and practical foundation. Section 3 introduces the joint FRST and FCM methodology. The combined methodology is applied to data from one of China's largest banks and evaluates farmer creditworthiness from a large data perspective in Section 4. Managerial, policy and research implications of the empirical study and methodology are discussed in Section 5. In the final section, a conclusion with limitations and future research directions, is introduced.

2. Background

2.1. Credit evaluation and its complexity

Credit evaluation by financial institutions is used to measure the ability of a borrower to repay their proposed obligation, namely the credit and the interest earned on the credit [73]. Credit risk and credit worthiness rating is a multidimensional, and oftentimes, complex decision-making problem [22,44]. During years immediately preceding the financial crisis of 2008, the rating agencies credit rating models (such as Moody's and Standard and Poor's (S&P)) were biased toward granting higher ratings than merited in order to compete for revenues from debtors, who pay to be rated; these inflated ratings resulted in enormous risks and losses [19,54]. During this period banks provided easy credit even for those applicants who were deemed non-creditworthy. After the crisis, banks and loan institutions tightened their requirements to new levels of credit austerity especially for smaller enterprises and individuals.

Some commonalities and differences exist amongst the various existing crediting systems. The three major rating agencies S&P, Moody, and Fitch, established corporate customer credit rating systems [53,68,34]. Fair Isaac Company's (FICO) credit rating systems evaluate customer credit status from at least five aspects, such as length of customer's building credit time and historical records of customers' paying credit [32]. Small and medium sized customer credit ratings may also be evaluated using the "5C principle": Character, Capital, Capacity, Collateral and Condition of Business [70]. In China the microfinance credit rating systems for customers of the Industrial and Commercial Bank of China includes asset-liability ratios and other indicators [43]. The credit rating system for farmers of the Postal Savings Bank of China considers four factors, i.e. repayment willingness, solvency, basic situation and operation capacity [60]. The Agricultural Bank of China established a credit rating system for farmers that include factors such as living status, the main business income debt ratio and other indicators [3]. Credit rating systems also exist for other commercial banks, for example the Rural Credit Cooperatives of Sichuan Province and Pudong Development Bank credit systems [28,30,62].

Countries and banks typically use varying credit evaluation systems; the variety of these credit evaluation systems is still expanding. Banks need to invest significant labor and resources for collecting these large volumes and high varieties of data. Determining data attributes that are useful or useless (redundant), is still a concern. Credit evaluation is complex making it difficult to effectively assess borrower creditworthiness. However, significant aspects of this complexity are omitted in standard discriminatory and predictive approaches such as multiple regression techniques [33].

Complexity in credit evaluation is exemplified by large attribute and data quantities, non-linear relationships among attributes or objects, and credit result unpredictability. Complexity includes the characteristic of the evaluation process as an open system, which is prone to the impact of various contextual environmental factors. Second, complexity is non-linear in that small changes in one characteristic or environmental factor of the credit system may not necessarily lead to correspondingly linear changes in other factors. Third, complexity includes interdependencies that may not be easily captured by regression and optimization models; making the relationships even less direct and less linear. Fourth, complexity can result from uncertainty (unpredictability). Credit prediction uncertainty sources include insufficient information, redundant information, information uncertainty, and dynamic decision environments. Given this credit system complexity, the difficulty in predicting the creditworthiness level of loan recipients increases.

2.2. Credit evaluation methods

Credit risk evaluation should be able to classify applicants as those with 'good credit' who repay on time and those with 'bad credit' who default. In a review of 214 studies on credit evaluation, Abdou and Pointon [1] determined that there is no best approach for establishing credit scoring models. They affirmed that no best model for all circumstances exists.

The main credit evaluation methods can be divided into three categories. The first category represents credit rating approaches relying on parametric methods associated with regression and econometrics techniques, such as statistical discrimination techniques [5], linear discriminant analysis [31,48,66], multiple discriminant analysis [20], logistic regression analysis [45], and multinomial regression models [46]. These credit rating models use stochastic probability approaches to describe debtor's probability of credit rating change [38], or to measure default probability [29]. These variants of econometrically based models, fall into many of the problems associated with correlative-based regression models such as the concerns of non-linearity, lack of equifinality, parametric underlying assumptions, and multiple interactions.

The second category of credit rating models is based on artificial intelligence methods. Although there is a wealth of research proposing credit evaluation based on parametric methods, researchers suggest that artificial intelligence approaches have better performance in credit rating [2,21]. For example, Malhotra and Malhotra [49] compared the performance of neural networks and multiple discriminant analysis (MDA) in identifying good loan applicants. They found that the neural network credit evaluation model performed better than the MDA model.

Moula et al. [52] measured customers' credit default risk by using six different methods. The empirical study showed that the proposed support vector machine (SVM) model was superior to classification and regression tree (CART) with discriminant analysis (DA), and more robust than the other approaches, i.e. logistic regression (LR), multilayer perceptron (MLP), and radial basis function (RBF). Xia et al. [71] built a credit rating model by combining boosted decision tree and Bayesian hyper-parameter optimization approaches. Bayesian models [18], similar to other stochastic approaches, have difficulty in identifying data parameters and complex interactions. Tools such as fuzzy neuro-nets [4], evolutionary computing [50] and learning algorithms [74] continue to be investigated in credit risk assessment. These credit evaluation approaches do not take into account indicator multicollinearity, which leaves larger data sets and redundant indicators. Larger data or redundant indicators increase the time complexity for calculation and even have an impact on the final credit rating results.

Third, nonparametric models are another classifier for credit risk evaluation. Compared with conventional methods such as multiple discriminant analysis (MDA), logistic regression analysis (LRA), and neural networks (NN) for applicant default recognition, which require extra a priori information; data envelopment analysis (DEA) as a credit evaluation model can calculate applicant's credit score by using ex-post information [51]. A number of other tools, including multi-attribute and fuzzy non-parametric techniques have also been applied in this area. Xu and Zhang [72] proposed a credit evaluation method by combining the analytic hierarchy process (AHP) and the set pair analysis (SPA). Fuzzy approach and expert systems are also important for credit rating [25,37,64,40]. Chi and Zhang [27] developed a nonparametric credit risk evaluation model based on the rank sum test and entropy weight. The analysis results showed that the proposed model could avoid the assumption of normality (and other distributional requirements) associated with other credit rating methods. These credit rating methods do not consider the complexity and relationships associated with indicator development, such as interactions amongst factors and non-configurational methods. Therefore, the applicability of these techniques is limited.

The drawbacks of existing approaches and studies lie along four dimensions. First, many of these credit rating system techniques do not consider microfinance issues faced by farmers, especially in emerging economies such as China; this is more of a limitation of previous study focus rather than the methodological techniques. Second, most works did not identify the key indicators that significantly influence the credit default based on farmers' real credit data. Many of the rating systems fail to analyze the complex relationships amongst factors including farmer basic environmental factors, operations capacity, guarantee and joint guarantees, and macro environmental contextual issues. Third, existing credit evaluation approaches do not take into account the reduction of indicators and reduction of data magnitude. Larger data or redundant data sets consume more time and resources; much of this data does not impact the final results. Fourth, the methodological issues related to complexity and relationships to indicator development, such as interactions amongst factors, parametric assumptions, consideration of equifinality, and non-linearity of relationships, or non-configurational approaches, limit the application of these techniques. We address many of these concerns in this study and introduce techniques that can address these methodological and modeling issues in this new study setting.

3. Fuzzy rough set (FRS) and fuzzy C-means (FCM) applicability

In light of above gaps, the purpose of this paper is to build a credit evaluation methodology using fuzzy rough-set theory and fuzzy C-means clustering to evaluate and investigate the complex relationships between farmer characteristics, competitive environmental factors and farmer credit level.

RST is capable of addressing a variety of complexities either as a standalone tool or as a complement to other techniques [10]. First, attribute reduction is a major RST methodological application that can be used to address large attribute and data quantity complexity characteristics. Second, rough set techniques can determine relationships related to creditworthiness evaluation. Relationships amongst variables include complex relationships and interactions resulting in non-linear relationships. The non-parametric aspects of this technique do not require a priori distribution characteristics for determining these relationships, providing expanded applicability. Third, rules generation from RST can be used to address credit systems unpredictability complexities. For example, RST decision rules can be generated to determine and identify which attributes relate to performance outcomes [7]. RST has been utilized in various fields, including credit risk for financial information, sustainable and green supply chain management, and other operations management concerns [11,14,17,59,63,69]. Fuzzy Rough Set (FRS) is an extension of RST based on measures of inclusion and has greater flexibility in what variables can be included in an analysis when compared to basic rough set [12]. Thus, FRS is adopted to evaluate and mitigate various credit system complexities.

Even though a variety of other techniques could have been selected for this relationship evaluation, such as regression or fuzzy systems [9], the main advantage of rough set theory, as stipulated above, is no requirement for parametric assumptions or additional information about data such as possibility values used in fuzzy set theory [7]. However, rough set cannot effectively integrate continuous numeric data. Fuzzy C-means can be used to effectively discretize continuous data, making the data amenable to rough set theory [8,42]. The FCM algorithm can also classify the data objects by grouping similar objects into the specified quantity clusters. Traditional hard clustering requires well-defined boundaries between clusters, which is not the characteristic of most real world applications. Appendix A provides additional the FCM methodology detail.

The strict classification of rough set is not suitable for direct application of large data, thus there is a need to improve the classification (lower and upper set) of rough sets to fuzzy rough set based on an inclusion threshold value. This classification allows for flexibility and ambiguity (fuzziness). This joint FRS/FCM methodology is used to evaluate farmer creditworthiness within China.

4. Identifying relationships to farmer creditworthiness

A multistep FRS/FCM methodology to investigate the complex relationships between various demographic farmer characteristic factors, environmental contextual factors, and creditworthiness

Table 1
Original information table for evaluation of farmers creditworthiness level.

Farmers	Basic situation			Operation capacity		Willingness of repayment		Guarantee and joint guarantee		Macro environment		Creditworthiness level
	Age	Education	...	Skills Status	...	Living Status	...	Whether or Not to Have Guarantee	...	Net Per Capita Income	...	
Farmer 01	57	5	...	3	...	1	...	0	...	3502.9	...	1
Farmer 02	38	3	...	4	...	1	...	1	...	4121.21	...	2
Farmer 03	40	6	...	3	...	1	...	0	...	3502.9	...	1
Farmer 04	40	4	...	4	...	1	...	1	...	4121.21	...	1
Farmer 05	51	4	...	3	...	1	...	0	...	4795.46	...	1
Farmer 06	33	3	...	2	...	1	...	1	...	9257.93	...	2
Farmer 07	42	2	...	3	...	1	...	0	...	3502.9	...	2
...	...	...	...	...	...	...	...	...	...	...	...	...
Farmer 2044	24	4	...	4	...	1	...	1	...	7356.47	...	4

levels is now presented. There are 10 steps in this methodology with commensurate study results presented at each stage.

Step 1: Collect and Screen Data

For this study data are collected and screened from a Chinese state-owned commercial bank. The bank has over 39,000 business outlets, with more than 70% of its business outlets located in rural regions representing the broadest geographic dispersion of any bank within China. The case study bank saw the largest loan growth amongst Chinese banking institutions. Loans of 198.6 billion RMB occurred in 2012, increasing at a 25% rate, and benefiting 2.2 million agricultural households [61].

The study sample included bank supplied data for all 2044 farmer loans that were due on October 9, 2009 [26]. For this data set, the maximum loan amount is 200,000 Yuan (approximately USD31,000 or 6.5 Yuan for each U.S. Dollar), and the minimum loan amount is 10,000 Yuan. Data quality verification was completed by comparing data sets to the bank's credit information system. The 2044 farmer samples included 28 provincial administrative regions in China, summarized in Appendix B. This sample covers almost all the administrative provinces in the eastern, central and western regions of China to providing geographic data diversity. This broad nation-level dispersion allows for identifying regional macro-economic (environmental contextual) factors on farmers' credit risk. This final information system has no missing data.

The 2044 farmers have 43 conditional attributes (characteristics) $A=\{a_j,\ j=1,\ 2,\ 3,\ \ldots,\ 43\}$ each, as shown in Appendix C. The conditional attributes include five categories: basic situation, operation capacity, willingness of repayment, guarantee and joint guarantee, and macro environment. Using a number of banking operations experts currently working as loan officers in banks[1] an outcome value D is determined. D is the *creditworthiness valuation outcome,* a proxy for creditworthiness, and is assigned one of four levels (1-Very Low Creditworthiness, 2-Low Creditworthiness, 3-High Creditworthiness, and 4-Very High Creditworthiness). Due to limited space, Table 1 summarizes the data (information system).

Step 2: Clustering Objects using the Fuzzy C-Means (FCM) Algorithm

The data includes continuous variables and variables that have a large range of values (i.e. greater than five values). To use the FRS approach more efficiently, the data needs to be discretized. FCM is used to discretize the continuous data and data with large ranges, such as age. Discretization occurred for each of the 2044 farmers and each numerical conditional attribute. Initial values for the clusters (c) are set to 5 (Very High, High, Medium, Low, Very Low) for each numerical conditional attribute; the order of fuzziness (m) is set to 2.

Farmers are assigned to each numerical conditional attribute by using the fuzzy memberships of the FCM algorithm. This step is completed by utilizing expressions (A1) to (A6). After FCM clustering, each farmer will be associated with a membership value u_{ik} for each cluster. These degrees of membership have values in the range [0, 1]. Higher values indicate greater strength of the association between that object and a particular cluster.

Second, a farmer is assigned to a cluster for which it has the highest membership value. In the empirical case, for example the membership value of Farmer 01 for the five *Age*_clusters are 93.46%, 3.64%, 1.54%, 0.85%, and 0.50% respectively; thus Farmer 01 is best assigned to *Age*_Very High (5). Table 2 shows cluster results for all farmers and attributes. This information from across all 2044 farmers will be used to reduce attributes using FRS that will be used to generate the core attributes set for given farmers' characteristics and environmental factors.

Step 3: Compute Farmers' Conditional Relations for Each Conditional Attribute

The conditional relations between each of the farmers on each of the conditional attributes is now determined. Farmers conditional relations are used to group sets of farmers based on similarity or indistinguishability relationships. This initial calculation will result in forty-three 2044×2044 conditional relational matrices, because each attribute will be evaluated separately, and there are forty-three attributes. The conditional relational matrix for conditional attribute a is defined as:

$$M_a(N) = (r_{ij})_{n\times n},\ \text{where } r_{ij} = \begin{cases} 1, & x_i = x_j, \\ 0, & otherwise. \end{cases} \tag{1}$$

For the conditional attribute *Age* we know that $r_{12} = r_{21} = 0$ because $x_1 \neq x_2$. An abridged version $M_{Age}(2044) = (r_{ij})_{2044*2044}$ is shown in Table 3.

Step 4: Compute Farmers Decision Relation for Each Outcome Attribute

This step calculates the decision relational matrix for the farmers who have the same creditworthiness level score, $M_a^D(N)$, using expression (2) to populate the matrix.

$$M_a^D(N) = (r_{ij})_{n\times n};\ \text{where } r_{ij} = \begin{cases} 1, & x_i = x_j, \quad D_i = D_j \\ 0, & otherwise. \end{cases} \tag{2}$$

For our example from the last column of Table 2, *Farmer* 03 and *Farmer* 04 both have *creditworthiness valuation* $D=1$. Thus $r_{34} = r_{43} = 1$. An abridged relational matrix $M_{Age}^D(2044)$ is shown in Table 4.

Step 5: Determine the Inclusion and Lower Rough Sets for Each Conditional Attribute.

There are four sub-steps in this step to derive the inclusion and lower rough set for each conditional attribute.

[1] Input from managers and economists at the Postal Savings Bank of China (PSBC), the China Banking Regulatory Commission (CBRC), and the Dalian Branch of China Banking Regulatory Commission (DBCBRC) were utilized to determine rankings. This group included three executives from the risk management departments. In total, values were aggregated from 17 experts from the three banking institutions.

Table 2
Clustered information table for evaluation of farmers creditworthiness level.

Farmers	Basic situation			Operation capacity		Willingness of repayment		Guarantee and joint guarantee		Macro environment		Credit level
	Age	Education	…	Skills status	…	Living status	…	Whether or not to have guarantee	…	Per capita income	…	
Farmer 01	5	2	…	3	…	5	…	2	…	1	…	1
Farmer 02	2	4	…	2	…	5	…	1	…	1	…	2
Farmer 03	2	1	…	3	…	5	…	2	…	1	…	1
Farmer 04	2	3	…	2	…	5	…	1	…	1	…	1
Farmer 05	4	3	…	3	…	5	…	2	…	2	…	1
Farmer 06	1	4	…	4	…	5	…	1	…	5	…	2
Farmer 07	3	5	…	3	…	5	…	2	…	1	…	2
…	…	…	…	…	…	…	…	…	…	…	…	…
Farmer 2044	1	3	…	2	…	5	…	1	…	5	…	4

Table 3
The conditional relation matrix for the *age* characteristic attribute between farmers' clustered scores.

Farmers	Farmer 01	Farmer 02	Farmer 03	Farmer 04	Farmer 05	Farmer 06	Farmer 07	Farmer 08	Farmer 09	Farmer 10	…	Farmer 2044	$\delta_{Age}(x_i)$
Farmer 01	1	0	0	0	0	0	0	0	0	0	…	0	133
Farmer 02	0	1	1	1	0	0	0	0	0	0	…	0	445
Farmer 03	0	1	1	1	0	0	0	0	0	0	…	0	445
Farmer 04	0	1	1	1	0	0	0	0	0	0	…	0	445
Farmer 05	0	0	0	0	1	0	0	0	0	0	…	0	365
Farmer 06	0	0	0	0	0	1	0	0	1	0	…	1	561
Farmer 07	0	0	0	0	0	0	1	1	0	1	…	0	540
Farmer 08	0	0	0	0	0	0	1	1	0	1	…	0	540
Farmer 09	0	0	0	0	0	1	0	0	1	0	…	1	561
Farmer 10	0	0	0	0	0	0	1	1	0	1	…	0	540
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Farmer 2044	0	0	0	0	0	1	0	0	1	0	…	1	561

Table 4
The decisional relation matrix and information granules for each farmer for the *age* characteristic attribute and $D_{farmeri}$ creditworthiness level.

Farmers	Farmer 01	Farmer 02	Farmer 03	Farmer 04	Farmer 05	Farmer 06	Farmer 07	Farmer 08	Farmer 09	Farmer10	…	Farmer 2044	$\delta^D_{Age}(x_i)$	$I(\delta(x_i), X)$
Farmer 01	1	0	0	0	0	0	0	0	0	0	…	0	95	71.4%
Farmer 02	0	1	0	0	0	0	0	0	0	0	…	0	133	29.9%
Farmer 03	0	0	1	1	0	0	0	0	0	0	…	0	278	62.5%
Farmer 04	0	0	1	1	0	0	0	0	0	0	…	0	278	62.5%
Farmer 05	0	0	0	0	1	0	0	0	0	0	…	0	111	30.4%
Farmer 06	0	0	0	0	0	1	0	0	1	0	…	0	166	29.6%
Farmer 07	0	0	0	0	0	0	1	0	0	1	…	0	168	31.1%
Farmer 08	0	0	0	0	0	0	0	1	0	0	…	0	290	53.7%
Farmer 09	0	0	0	0	0	1	0	0	1	0	…	0	166	29.6%
Farmer 10	0	0	0	0	0	0	1	0	0	1	…	0	168	31.1%
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Farmer 2044	0	0	0	0	0	0	0	0	0	0	…	1	348	62.0%

Sub-step 1: Compute the information granule ($[x_i]^a_R$) for each farmer for the given conditional attribute a. For example we can compute $[Farmer01]^{Age}_R$ from Table 3 by simply summing the values in the *Farmer* 01 row $[Farmer01]^{Age}_R = 133$. The information granule of *Farmer* 01 represents the number of farmers who have the same conditional attribute value.

Sub-step 2: Determine the information granule for a given Farmer i, a conditional attribute a and creditworthiness valuation D. Illustratively, we seek the information granule for the *Age* characteristic and the *creditworthiness level* of $D_{Farmeri}$. In other words, this step is completed by summing the rows of $M^D_{Age}(2044)$. Thus $[Farmer01]^{Age\&D=1}_R = 95$.

Sub-step 3: Determine the measure of inclusion, $I([x_i]^a_R, X)$, for each Farmer I using expression (3).

$$I(A, B) = \frac{|A \cap B|}{|A|},\ where\ A \neq \emptyset. \tag{3}$$

where $|*|$ is the cardinality of a set $*$.

For *Farmer* 01 the numerator will be $\delta^{D=1}_{Age}(Farmer01)$ while the denominator is $[Farmer01]^{Age}_R$. That is, for *Farmer* 01 there are 94 other farmers with creditworthiness valuation $D = 1$ and 39 additional farmers whose creditworthiness valuation is $D \neq 1$. Thus, the characteristic of inclusion, $I([Farmer01]^{Age}_R, X_{D=1}) = 95/133 = 71.4\%$. The last column of Table 4 shows the degree of inclusion for each farmer.

Sub-step 4: Determine the lower approximations for a given degree of inclusion, each *Farmer*, and the selected inclusion threshold value k using expression (4).

$$\underline{F^k}X = \{x_i | I([x_i]^a_R, X) \geq k, x_i \in U\}, \tag{4}$$

where $1 \geq k \geq 0.5$.

For this case the initial setting is $k = 0.70$. That is, when the degree of inclusion for a given *Farmer* has $I(\delta_a(x_i),\ X) \geq 0.70$, this given *Farmer* will be included in the lower approximations. Table 4 shows that *Farmer* 01 with an inclusion value of 71.4% means $Farmer01 \in POS_{Age}(D)$. For this study, the number of farmers who meet the threshold value is $|POS_{Age}(D)| = 95$.

Step 6: Compute the Dependency Degree.

For each $a_j \in A$, expression (5) is used to calculate the dependency degree for each creditworthiness valuate score (D) on a conditional attribute (a_j). This dependency degree is used to eliminate unimportant attributes and determine the information significance for each remaining conditional attribute in the Step 7.

$$\gamma_a(D) = \frac{|POS_a(D)|}{|U|} \tag{5}$$

Table 5
The core attributes reduction process.

Characteristic attributes	Dependency degree	Atr (Dependency)	Atr (Redundancy)	Information significance	Atr (Significance)
Age	4.7%	Keep	Keep	1.77%	Keep
Education	78.2%	Keep	Keep	29.38%	Keep
Marital Status	0.0%	Delete			
Gender	0.0%	Delete			
Family Size	0.0%	Delete			
Labor Forces	0.0%	Delete			
Family Size/Labor Forces	16.1%	Keep	Keep	6.05%	Keep
Dependent Population	16.2%	Keep	Delete		
Loan Purpose	0.0%	Delete			
Own House Value	1.9%	Keep	Delete		
House Value	2.5%	Keep	Keep	0.94%	Delete
Skills Status	60.2%	Keep	Keep	22.61%	Keep
Net Business Income	2.2%	Keep	Delete		
Net Income Per Year Per Capita GDP[a]	0.3%	Keep	Keep	0.11%	Delete
Net Income	4.1%	Keep	Keep	1.54%	Keep
Daily Living Expenses	0.0%	Delete			
Total Expenses	5.5%	Keep	Keep	2.07%	Keep
Expenses/Income	0.0%	Delete			
Total Property	3.9%	Keep	Keep	1.47%	Keep
Agricultural Net Income	0.7%	Keep	Keep	0.26%	Delete
Agricultural Production Income	1.5%	Keep	Keep	0.56%	Delete
Agricultural Production Expenses	2.4%	Keep	Delete		
The Proportion of Non-agricultural Income	25.5%	Keep	Keep	9.58%	Keep
Children's Education Spending	26.3%	Keep	Keep	9.88%	Keep
Living Status	1.1%	Keep	Delete		
Living Stability	0.2%	Keep	Keep	0.08%	Delete
Amounts of Unpaid Loans	0.0%	Delete			
Whether or Not to Have Unpaid Loan	0.0%	Delete			
Deposit in Bank	0.1%	Keep	Delete		
Whether or Not to Have Private Loan	0.0%	Delete			
Default or Not	0.5%	Keep	Delete		
Number of Historical Apply Loan	0.1%	Keep	Delete		
Loaning Records	0.5%	Keep	Delete		
Whether or Not to Have Guarantee	0.0%	Delete			
Power of Guarantor	0.5%	Keep	Keep	0.19%	Delete
Whether or Not to Have Joint Guarantee	0.0%	Delete			
The Relationship of Joint Guarantee Member	0.0%	Delete			
Net Per Capita Income	7.9%	Keep	Keep	2.91%	Keep
Per Capita Agricultural Output Value	4.1%	Keep	Keep	1.54%	Keep
Regional GDP Growth Rate	0.1%	Keep	Keep	0.04%	Delete
Consumer Price Index	0.0%	Delete			
Savings Deposit Balance	0.0%	Delete			
Engel Coefficient[b]	23.9%	Keep	Keep	8.98%	Keep

[a] GDP (Gross domestic product) is a monetary measure of the market value of all final goods and services produced in a period (quarterly or yearly) of time.
[b] The Engel Coefficient is defined as the proportion of total food expenditure to total personal consumption expenditure. This indicator shows that as income rise the proportion of income spent on food falls, even if absolute expenditure on food rises.

where $POS_a(D)$ is the lower approximation for the outcome attributes D based on the attribute a.

From the results in step 5, we can see the numerator, $|POS_{Age}(D)|$, has a value of 95 farmers who meet the threshold value of $k = 0.70$. For *Age*, the dependency degree of D on the *Age* characteristic attribute is: $\gamma_{Age}(D1) = |POS_{Age}(D)|/|U| = 95/2044 = 0.047 = 4.7\%$

The remaining dependency degrees for each of the characteristic, conditional, attributes in this study data set are shown in Table 5.

Step 7: Determine the Similarity and Information Significance between Pairs of Conditional Attributes

In this step, information redundancy (correlation and overlap) between every pair of conditional attributes is determined. That is, the similarity of two conditional attributes a_j and a_l is determined in three sub-steps.

First, expression (6) is used to determine the information content, or information entropy, across an attribute (a_j) [47,65,67].

$$IC(a_j) = 1 - \frac{1}{|U|^2}\sum_{i=1}^{|U|}|X_i^{a_j}| \tag{6}$$

where $IC(a_j)$ is the information content for an attribute (a_j). $|U|$ is the cardinality of the universe of farmers. $|X_i^{a_j}|$ is the information granule of farmers with similar attributes levels across the attribute (a_j) for Farmer i. It is also defined as the number of members within the attribute (a_j) for Farmer i.

Second, expression (7) is used to determine the joint information content between an attribute (a_j) and another attribute (a_l).

$$IC(a_j\&a_l) = 1 - \frac{1}{|U|^2}\sum_{i=1}^{|U|}|X_i^{a_j\&a_l}| \tag{7}$$

where $IC(a_j\&a_l)$ is the information content for two attributes (a_j) and (a_l). $|X_i^{a_j\&a_l}|$ is the information granule of farmers with similar attributes levels across two attributes (a_j) and (a_l) for Farmer i.

Third, expression (8) is used to determine the incremental information content of attribute (a_j) when compared to an attribute (a_l).

$$Sig(a_j, a_l) = IC(a_j\&a_l) - IC(a_l) \tag{8}$$

The greater the increment information content, the greater the difference between attributes (a_j) and (a_l), which denotes that attribute (a_l) can provide different information than attribute (a_j).

Smaller incremental information content, means greater similarity between attributes (a_j) and (a_l), which denotes that the attribute (a_l) has greater overlap with attribute (a_j). Highly redundant, very similar, pairs of attributes mean that one of them can be deleted with little, if any, information loss.

For example, the information content across the *Age* characteristic attribute is $IC(Age) = 0.771$, and the information content between the *Age* characteristic attribute and *Education* characteristic attribute is $IC(Age\&Education) = 0.918$. The overlap between the *Age* characteristic attribute and the *Education* characteristic attribute is $sig_{01,02} = 0.918 - 0.771 = 0.147$. For every two conditional attributes a similarity matrix is completed.

Step 8: Reduce the Conditional Attribute Set

Four sub-steps are used in this step to reduce the conditional attribute set using the information significance of a conditional attribute.

The first sub-step in this process is to exclude those attributes not directly related to a specific creditworthiness value outcome attribute. Using the dependency degree value calculated in step 6, a conditional attribute a_k that satisfies (9) is selected:

$$Atr = \{a_k | \gamma_{a_k}(D) > 0\} \tag{9}$$

For example, in this study the dependency degrees for the (*Family Size)/(Family members eligible for labor force (or "Labor force"))* attribute and *Dependent Population* attributes are 16.1% and 16.2%, respectively. Thus, both these attributes will initially be included in the reduced attribute set *Atr*. Whereas, the *Marital Status* attribute, will not be kept (deleted) because its dependency degree is equal to 0. Table 5 shows which attributes will be kept after the first sub-step for further consideration in the second sub-step.

The second sub-step is to exclude those attributes that do not provide a minimal level of information. Using the overlap or similarity value calculated in step 7, we keep attributes a_l that satisfy (10):

$$Atr = \{a_l | sig(a_l, a_j) \le 0.07 \ \& \ sig(a_l, a_j) < sig(a_j, a_l)\} \tag{10}$$

where $sig(a_l, a_j) \le 0.07$ means that if the similarity significance $sig(a_l, a_j)$ is smaller than 0.07, then a_j is assumed to be similar with a_l. In other words, a_j only provides a small amount of information when compared to attribute a_l. If $sig(a_l, a_j) < sig(a_j, a_l)$, it means that attribute a_j provides a smaller amount of information when compared with a_l. Thus, a_l is retained in the *Atr* set.

As an example, we found that the incremental information content of *Family Size/Labor Force* is 0.042 when adding it as a pairing with the *Dependent Population* attribute. The incremental information is 0.114 when adding the *Dependent Population attribute* as a pairing with the *Family Size/Labor Force* attribute. That means the *Family Size/Labor Force* attribute will be kept in the reduced attribute set *Atr* and the *Dependent Population* attribute is no longer needed in the *Atr* set.

The third sub-step is to determine the information significance of an attribute. The dependency degree for reducing characteristic attributes is used to determine the information significance of a characteristic attribute a_l that satisfies (11):

$$w(a_l) = \frac{\gamma_{a_l}(D)}{\sum_{a_j \in Atr} \gamma_{a_j}(D)} \tag{11}$$

As an example, the information significance of *Age* attribute is:

$$w(Age) = \frac{\gamma_{Age}(D)}{\sum_{a_j \in Atr} \gamma_{a_j}(D)} = 1.77\%.$$

The forth sub-step is to exclude those attributes which have small information significance values. For this study, we set the information significance for adding an attribute a_l that satisfies (12):

$$Atr = \{a_l | w(a_l) \ge 1\%\} \tag{12}$$

The final reduced attribute set is $Atr = \{$*Age, Education, Family Size/Labor Force, Skills Status, Net Income, Total Expenses, Total Property, The Proportion of Non-agricultural Income, Children's Education Spending, Net Per Capita Income, Per Capita Agricultural Output Value, Engel Coefficient*$\}$. Table 5 summarizes the process and final core, reduced attribute set; also called reducts. The creditworthiness rules are produced using this reduct set.

Step 9: Identifying Relationship Rules between Attributes and Outcome

In this step we take the reduct set determined in the previous steps and relate them to the outcome through a series of rules developed to discern the intrinsic relationships. To complete this step, expressions (13) to (14) are used.

The first sub-step is to form a matrix titled the discernibility matrix whose elements are defined as follows:

$$\alpha_G(x, y) = \begin{Bmatrix} g \in G, \ where \ V_g(x) \ne V_g(y), \ x, y \in U \\ \emptyset, \ where \ D(x) = D(y), \ x, y \in U \end{Bmatrix} \tag{13}$$

where an element of $\alpha_G(x, y)$ is a set of reducted attributes between Farmer *x* and Farmer *y*. $V_g(x)$ is the value of reduct attribute *g*. $D(x)$ is the value of outcome attribute *D* for Farmer *x*. In other words, an element of the discernibility matrix is the set of all reduced attributes that discern Farmer *x* and Farmer *y* and at the same time do not have the same outcome value of *D*. If the outcome value of *D* is the same for two farmers then the corresponding matrix element is a null set as defined by expression (13).

The discernibility rules (functions) $\Delta^D(x)$ can now be determined. The discernibility rule for an object (farmer) *x* for an outcome value of *D* exists if and only if

$$\Delta^D(x) = \wedge\{\vee a_G(x, y) : y \in \{z \in U, D(z) \notin \partial_G(x)\}, a_G(x, y) \ne \phi\} \tag{14}$$

where $\partial_G(x) = \{f^w(y) | y \in S_G(x)\}$ is the generalized outcome in *D* (the outcome matrix or table) and determines to which decision classes the Farmer *x* may be classified based on the available information on that farmer.

This step is completed by utilizing the Rough Set Exploration System (RSES) software [15]. Initially the 2044 farmers are separated into two groups. The first group, which contains 1800 farmers, is used to generate the rules set, this group will be called the generator group. The second group, called the predictive data set, contains 244 farmers and is used to examine the predictive performance of those rules. For rough set approaches, any outcome table may be regarded as a set of generalized decision rules of the form:

$$\wedge(g, v) \to \vee(D, w), \ \ where \ g \in G, v \in V_g, w \in W_D, D \notin G \tag{15}$$

where $\wedge(g, v)$ is called the condition or premise; which in our illustrative example is a characteristic (versus environmental) attribute. The outcome or conclusion is represented by $\vee(D, w)$, which in our illustrative example is the outcome value *D*. The $\wedge$, $\vee$are Boolean notations for "and" and "or" respectively. Indices *v*, *w* represent the values (belonging to the sets *V* and *W*) for reduct attribute *g* and outcome variable *D*, respectively.

Table 6 summarizes some of the rules generated. As an example, consider the following rule: $(Age = 5) \wedge (Education = 3) \wedge (Skills\ Status = 2) \wedge (Total\ Expenses = 5) \wedge (Children's\ Education\ Spending = 5) \to (Creditworthiness\ \ Valuation = 1\ \ [95])$. This rule indicates that if a farmer belongs to $Age = 5$, $Education = 3$, $Skills\ Status = 2$, $Total\ Expenses = 5$, and $Children's\ Education\ Spending = 5$,

Table 6
Rough set decision rules for some parts of creditworthiness level $(D) = 1$.

Rules	Number of occurrences
(Age = 5)&(Education = 3)&(Skills Status = 2)&(Total Expenses = 5)&(Children's Education Spending = 5) = > (Creditworthiness level = 1 [95])	95
(Education = 3)&(Skills Status = 2)&(Total Property = 1)&(Children's Education Spending = 5)&(Per Capita Agricultural Output Value = 5) = > (Creditworthiness level = 1[76])	76
(Age = 5)&(Education = 3)&(Net Income = 1)&(Children's Education Spending = 5) = > (Creditworthiness level = 1[73])	73
(Education = 3)&(Skills Status = 2)&(Total Expenses = 5)&(Children's Education Spending = 5)&(Per Capita Agricultural Output Value = 5) = > (Creditworthiness level = 1[72])	72
(Education = 3)&(Skills Status = 2)&(Children's Education Spending = 5)&(Engel Coefficient = 3) = > (Creditworthiness level = 1[66])	66
(Education = 3)&(attr2 = 4)&(Skills Status = 2)&(Children's Education Spending = 5) = > (Creditworthiness level = 1[65])	65
(Education = 3)&(Net Income = 1)&(The Proportion of Non-agricultural Income = 1)&(Children's Education Spending = 5) = > (Creditworthiness level = 1[62])	62
(Education = 3)&(Skills Status = 2)&(The Proportion of Non-agricultural Income = 1)&(Children's Education Spending = 5)&(Per Capita Agricultural Output Value = 5) = > (Creditworthiness level = [62])	62
(Age = 5)&(Education = 3)&(Total Property = 1)&(Children's Education Spending = 5)&(Per Capita Agricultural Output Value = 5) = > (Creditworthiness level = 1[61])	61
(Education = 3)&(Children's Education Spending = 5)&(Net Per Capita Income = 2)&(Per Capita Agricultural Output Value = 5) = > (Creditworthiness level = 1[61])	61

then this farmer's creditworthiness is low. This is a relatively strong rule with 95 observations generating this rule. A total of 13,774 rules were generated.

Step 10: Using Rules to Predict an Outcome

The creditworthiness level of 244 farmers in the predictive set is completed through the rules developed in step 9. It is observed that many times a farmer's characteristics may match multiple rules. Multiple rules may result in a number of conflicting outcomes. In that case, a farmer's characteristics are assigned to the rule with the largest quantity of occurrences.

As an example, Farmer 1840 can match the following rules:

$$\begin{aligned}&(Age = 5) \wedge (Education = 4) \wedge (NetIncome = 2)\\ &\quad \wedge (Total\ Property = 4) \wedge (Children's\ Education\ Spending = 4)\\ &\quad \wedge (Per\ Capita\ Agricultural\ Output\ Value = 1)\\ &\quad \rightarrow (Creditworthiness\ Valuation = 3).\end{aligned}$$

and

$$\begin{aligned}&(Age = 5) \wedge (Education = 4) \wedge (\text{Skills Status} = 3)\\ &\quad \wedge (Net\ Income = 2) \rightarrow (Creditworthiness\ Valuation = 2).\end{aligned}$$

The first of these two rules is generated 8 times from the generator data set. The second rule is generated 2 times. In that case, creditworthiness outcome for Farmer 1840 will be 3 based on the first rule since it is generated more frequently.

The outcomes for each of the other farmers in the 244 farmer predictive set (farmers 1801–2044), were also determined using this process. The predicted result showed that 81.16% of the time the predicted outcome was the same as the original outcome.

5. Results and discussion

5.1. Relationships between creditworthiness level, farmer characteristics and contextual factors

The generated rules can provide insights to relationships of creditworthiness and various contextual and farmer characteristics. These rules show a number of possible evaluation strategies that can identify creditworthiness level. It would be helpful to decision makers to delineate the relationship between farmers' characteristics, environmental factors, and creditworthiness. There were 13,774 rules generated. This large set of rules can help develop a rule base which be used to predict the creditworthiness level of farmers. The rule base can be updated over time as new farmer samples are introduced.

The results showed that had a 93.7% predictive accuracy level, using the predictive data set, for a farmer creditworthiness of level 1 (the lowest level of creditworthiness). The most inconsistent predictive results were for farmers with a creditworthiness level of 4 (high creditworthiness level). One reason for this occurrence may be the number of rules generated for each level, with a greater number of rules generated representing a more likely relationship. For example, 7182 rules exist for *Very Low Creditworthiness Level, 4365* rules for *Low Creditworthiness Level, 1867* for *High Creditworthiness Level* and *only 360* rules related to *Very High Creditworthiness Level.* A small relative set of rules for the Very High Creditworthiness Level means that additional data needs to be collected for these farmer samples. These results also imply that the rule set can be used to eliminate poor performers, but cannot necessarily identify good performers.

Additional testing is completed to further support the supposition that additional data can increase the number of rules generated and predictive accuracy of the methodology. Three scenarios are considered. The first scenario is to arbitrarily delete 6 farmers for each creditworthiness level from the generator group. The second scenario is to make an even smaller data set arbitrarily deleting another 6 farmers for each creditworthiness level from the generator group. In this way we can show the accuracy results as the number of rules reduces. The results of all three scenarios (including the initial scenario case illustration) are shown in Table 7.

From Table 7, it can be observed that more objects generate more rules. More rules result in a higher predictive accuracy rate. As a result, for greater accuracy, a substantive data set is required or the accuracy of this methodology could be called into question.

5.2. Identification of the key characteristics

With an emphasis on sustainable development and supply focusing on impoverished and low income regions, financial institutions and agricultural supply chain managers have become concerned with the creditworthiness level of farmers. But, creditworthiness evaluation is complex. Therefore, in order to maintain effective creditworthiness valuation, financial institutions and policy makers identify what farmer characteristics and environmental (contextual) factors play a significant role in prediction and determination of creditworthiness.

From the original set of 43 attributes only 12 core attributes were found to help monitor and evaluate farmer creditworthiness level. This smaller set of core attributes helps to reduce the complexity of the evaluation process, but allows for a concentrated set of rules. With fewer characteristics and environmental factors information process savings occur. The rules from this study's

Table 7
The predicted result for various scenarios.

Creditworthiness level	The number of farmers in prediction (second) group	Initial scenario		Scenario 1		Scenario 2	
		Rules	Accuracy	Rules	Accuracy	Rules	Accuracy
Very low creditworthiness level	143	7182	93.70%	7012	92.30%	6976	90.90%
Low creditworthiness level	74	4365	64.90%	4187	62.20%	4083	61.50%
High creditworthiness level	21	1867	66.70%	1858	58.10%	1726	57.70%
Very high creditworthiness level	6	360	33.30%	303	16.70%	231	13.70%
Sum/average	244	13,774	81.16%	13,360	78.37%	13,016	77.23%

methodology show more complex relationships than if standard regression and statistical averaging (correlative) approaches are used.

5.3. The significance of core attributes

Five proposed farmer characteristic groupings including basic situation, operations capacity, willingness of repayment, guarantee and joint guarantee, and macro environment were used in this study. Using fuzzy rough set the six most important farmer attributes for farmer creditworthiness determination, based on information significance are: *Education* (29.38%), *Skills Status* (22.61%), *Children's Education Spending* (9.88%), *The Proportion of Non-agricultural Income* (9.58%), *Engel Coefficient* (8.98%) and *Family Size/Labor Force* (6.05%). These six attributes represent 86.48% of the total information significance. *Education* and *Skills Status* are the two most significant attributes at 51.99% of information significance. Conventional education and skill education are essential for identifying farmer creditworthiness level.

According to this result, the most effective factor related to creditworthiness is *Education*. This result shows that the education level of farmers has great relevance but also captures information from other factors in its relationship to creditworthiness. Overall, more highly educated farmers will be strongly related to higher levels of *creditworthiness level*. For example, there are 7 farmers with *creditworthiness level* of 4, and 11 farmers who have a 3 *creditworthiness level* for a total of 20 farmers with a 6 education score level (2 farmers with 6 education level have a 2 *creditworthiness level*). In contrast, there are only 2 farmers with a 4 level of *creditworthiness level* from amongst all 205 farmers with a 0 education score.

The dependency degree value shows a strong relationship between low education of farmers and the low *creditworthiness level* which is more significant than the relationship between high education of farmers and high *creditworthiness level*. For example, 51.0% of farmers with a 5 and 6 education score value, have greater than a 3 *creditworthiness level*. 78.5% farmers with a 1 education and 88.3% of 2 education have a 1 *creditworthiness level*. Similarly, *Children's Education Spending* is also very important and ranked third amongst all characteristic attributes.

The *Skills Status* attribute is the second most strongly related factor. A higher *Skills Status* valuation for a farmer relates to a higher *creditworthiness level*. The 72.7% farmers with a *Skills status* score level of 1, 73.9% farmers with a *Skills status* score level of 2; and 81.8% with a 3 *Skills status* level have a 1 *creditworthiness level*.

5.4. A framework for farmer credit worthiness evaluation

A summary of the above relationships and framework for farmer creditworthiness is shown in Fig. 1. The "Strong correlation characteristics" represent characteristics having strong relationships with the farmer creditworthiness. Education, skills, and children education spending seemed to be the most important characteristics. This result points to education and skills development as important aspects to support creditworthiness of farmers. The "Strong correlation redundancy characteristics" represent characteristics with strong relationships to farmer creditworthiness, but these other strong characteristics were replaced because of similar information content. For example, "Dependent Population" and "Family Size/Labor Force" provide significant overlap in information. These additional characteristics can be used as an important reference for creditworthiness evaluation, if for some reason other data is not available. The "weak correlation characteristics" represent characteristics with limited relationships for credit risk of farmers.

This model can guide bankers to determine which data should or should not be collected for their databases. It may also provide insights to various agencies, governmental or non-governmental social agencies, seeking to aid farmers identify various supportive programs. For example, education and educational support should be encouraged for farmers. Education for farmers has not historically received the largest priority since farming has been characterized as physical labor and investment in education may not be viewed as productive. These results prove otherwise, that education and skills development, whether for the current or future generations can provide significant benefits in terms of successfully meeting financial obligations, and by connection successful farming.

It should also be observed that it is not only one characteristic alone that relates to a low or high creditworthiness outcome. As can be seen the relationships are relatively complex with significant interaction amongst these three characteristics. Equifinality is shown by having at least 10 different complex relationships that can help identify creditworthiness levels. Some are quite specific, and some combination and/or aggregation of the rules can provide more insights.

Theoretical insights can also be gained in the type of resources and capabilities necessary to function in a micro-enterprise and family enterprise situation. In a study exploring farming families' responses to change, conducted across a wide range of economic conditions in different countries, findings show that profit-seeking is not a key driver [6]. Broadly, research has shown that family farms are successfully transferred generation-to-generation, the financial and frequently emotional survival of each generation is linked to the farm's success [36]. These concepts fit within the socio-emotional wealth of family business perspective [16]. Our findings, although not part of the complete set of dimensions and constructs of socio-emotional wealth, do support the idea that some of the more important success characteristics are 'softer' in terms of education and skills development of farmers, especially in China. But, the additional characteristics and factors such as generational concerns, technological capabilities, and other factors, from other studies, can be added to the large set of factors used in this study.

5.5. The advantage of this methodology

The multistep methodology is advantageous for three reasons. (1) Advanced functionality: it can not only effectively evaluate farmer creditworthiness, but it can also form rules to accurately

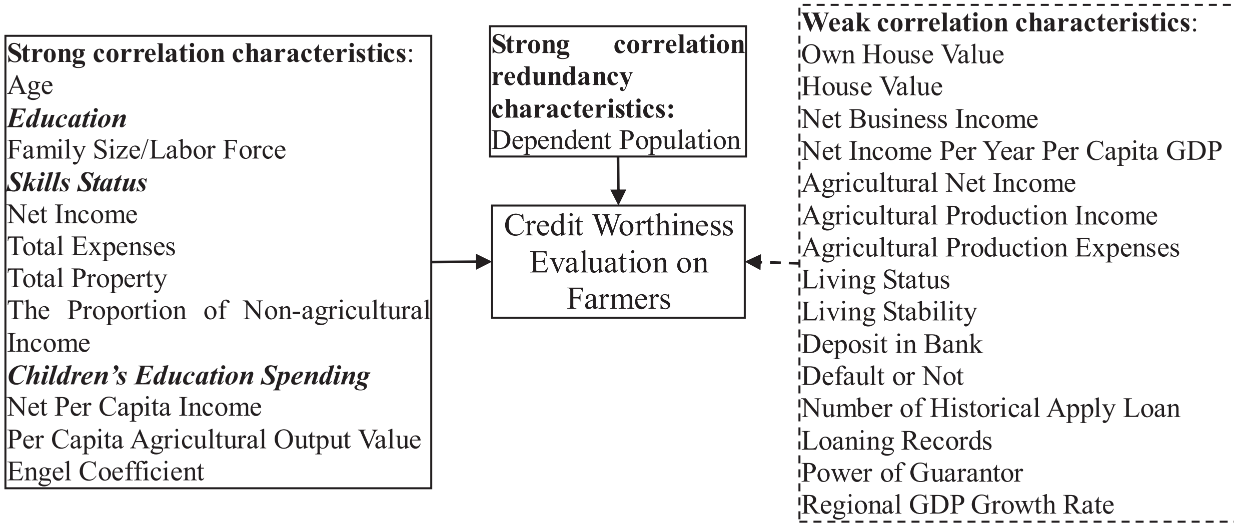

Fig. 1. The framework for credit worthiness evaluation on farmers.

predict farmer creditworthiness. In the evaluation process, the methodology reveals a variety of complex relationships between various factors and farmer creditworthiness. Furthermore, the methodology can prioritize factors to facilitate decision and policy making to improve farmer creditworthiness. (2) Ease application: it allows banks to determine farmer creditworthiness without any additional information beyond what currently exists in their data sets. Financial institutions or agricultural supply chain managers can use this methodology to handle large data. (3) Computational efficiency: it reduces the characteristic attribute sets requiring less extensive acquisition of data by financial institutions or agricultural supply chain managers and lessened computational requirements. We reduced the characteristic attribute sets from the initial 43 attributes to 12 attributes, which could reduce computation requirements by 72.1%. When compared to traditional econometric models, the complexity of relationships (interactions), no parametric requirements, and many potential relationships (equifinality) are all advantages of this methodology.

6. Conclusion and future work

Increased sustainability efforts by Chinese financial institutions and government agencies have increased concerns of farmer creditworthiness evaluation, especially for small farmers. One of the core issues facing the design and development of farmer creditworthiness and risk evaluation is the identification and investigation of relationships amongst various farmer characteristics, contextual environmental factors and farmer creditworthiness levels. Farmer loan by banking organizations is still a relatively new concept in China. This investigation is one of the first to provide insights into farmer creditworthiness in China. This research study uses Chinese micro agribusiness, small farmer, and bank data to evaluate complex relationships.

In this exploratory study the findings can help both financial institutions, agricultural supply chain managers who are dependent on these farmers and government development agencies that may be investing a large portion of their managerial budgets on monitoring and managing farmers. A joint Fuzzy rough set and Fuzzy C-means methodology was introduced in this study to determine a core set of farmer characteristic attributes for creditworthiness. This study represents one of the first to explicitly address these issues for this environment from a formal modeling perspective.

The results obtained from this methodology provide various practical insights facing financial institutions and farmers. Education and skills related characteristic attributes are essential to enhance the farmers' credit worthiness level. Sustainable business theories on building capacities in family-owned and micro-organizations suggest that socio-emotional issues such as education and skills are important factors [36].

Although this exploratory study provides some valuable results and insights, there are limitations and room for further investigation. Greater parametric evaluations and sensitivity of the solution can provide a broader set of potential relationships, although the number of relationships found in this study is quite extensive. Given that this data set is from one bank, although regionally dispersed, that results may not be generalizable enough for all banks in China. When investigating sustainability issues, there are also concerns of more relevant data. We introduced general characteristic attributes into this framework, additional resources and capability dimensions of family farms in China may provide additional results. This methodology can easily incorporate additional characteristic attributes. Thus management/researchers can easily determine impacts of additional attributes not used in this study. We encourage researchers to use this methodology to further investigate practical and theoretical relationships for creditworthiness of banking customers in general.

Acknowledgments

The research was supported by National Natural Science Foundation of China (Nos. 71772032, 71503199, 71502026, 71472031], the Key Projects of National Natural Science Foundation of China (No. 71731003), the Credit Risks Evaluation and Loan Pricing for Petty Loan Funded for the Head Office of Post Savings Bank of China (No. 2009-07), the Youth Talent Cultivation Program of Northwest A&F University (No. Z109021717).

Appendix A. Fuzzy C-means background

Fuzzy C-Means (FCM) uses clustering to generate groups of subsets of data so that objects in each cluster are more similar to each other than objects from different clusters [7,58]. It partitions a set of n objects $X = \{x_1, x_2, \ldots, x_n\} \subset R_P$, a P-dimensional, real-number space into c $(1 < c < C)$ fuzzy clusters with $H = \{h_1, h_2, \ldots, h_c\}$ cluster centers or centroids. The FCM algorithm assigns objects to each category using fuzzy memberships, which are calculated as follows:

$$u_{ik} \in [0, 1] \quad \forall k = 1, 2, \ldots, n; \quad \forall i = 1, 2, \ldots, c; \tag{A1}$$

$$\sum_{i=1}^{c} u_{ik} = 1, \quad \forall k = 1, 2, \ldots, n; \tag{A2}$$

$$0 \le \sum_{k=1}^{n} u_{ik} \le n \quad \forall i = 1, 2, \ldots, c; \tag{A3}$$

where u_{ik} indicates the degree of association or membership function of the kth object with the ith cluster. The fuzzy clustering of objects is described by a fuzzy matrix $U = \{u_{ik}\} i = 1, \ldots, c; k = 1, \ldots, n$ with n rows and c columns in which n is the number of data objects and c is the number of clusters.

The objective of the FCM algorithm is defined by expression (A4)

$$\min J(U, H) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^{e} (\|x_k - h_i\|)_A, \tag{A4}$$

where e $(e > 1)$ is a scalar term for the weighting exponent and controls the fuzziness of the resulting clusters; $\| * \|_A A = I$ is the Euclidian distance from object x_k to the cluster center h_i.

Alternatively, the solution of the constrained optimization problem in expression (A4) and the cluster centers h_i and related membership functions u_{ik} are given in expressions (A5) and (A6) for every iteration number t, respectively.

$$h_{i,t} = \frac{\sum_{k=1}^{n} (u_{ik})^e x_k}{\sum_{k=1}^{n} (u_{ik})^e} \le n, i = 1, 2, \ldots, c \tag{A5}$$

$$u_{ik,t} = \left[\sum_{j=1}^{c} \left(\frac{\|x_k - h_{i,t-1}\|_A}{\|x_k - h_{j,t-1}\|_A} \right)^{2/(e-1)} \right]^{-1}, i \ne j \tag{A6}$$

This solution of clustering data objects is obtained by iteratively minimizing an objective function that is a function of the sum of Euclidian distances between the objects and the center of all the clusters based on weighting the degree of association or membership function of a particular with a particular cluster. The degree of membership is quantified by a value in the interval [0, 1], which indicates the closeness of the data objects to the cluster centers. The highest degree of membership level represents the most association between that object and a particular cluster centroid, and indicates that this object is the most possibly belong to this particular cluster. FCM clustering has been applied in fields such as geology, medical imaging, target recognition, image segmentation, and performance evaluation [7,13,58].

Appendix B. The information of 28 provincial administrative regions in China

No.	The information of provincial administrative regions
1	Tianjin Municipality, Shanxi Province, Hebei Province and Inner Mongolia Autonomous region of Northern China
2	Liaoning Province, Jilin Province, and Heilongjiang Province of Northeastern China
3	Shanghai Municipality, Shandong Province, Jiangsu Province, Jiangxi Province, Zhejiang Province, and Anhui Province of Eastern China
4	Fujian Province, Hainan Province, and Guangdong Province of Southern China
5	Shanxi Province, Gansu Province, Qinghai Province, Ningxia Hui Autonomous region, Xinjiang Uygur Autonomous region of Northwestern China
6	Henan Province, Hunan Province, and Hunan Province of Central China
7	Henan Province, Hunan Province, and Hunan Province of Central China

Appendix C. The characteristic attributes of farmer credit evaluation

Categories	Characteristic Attributes	Explanation
Basic Situation	Age	Age when the farmer obtains loan
	Education	Education background when the farmer obtains loan
	Marital Status	Marital status when the farmer obtains loan
	Gender	Male or female
	Family Size	Total number of family members
	Labor Forces	The number of labor forces in the family
	Family Size/Labor Forces (%)	The members of farm family eligible to be in labor force
	Dependent Population	The number of people supported by the family
	Loan Purpose	The purpose of the farmer to apply for a loan
	Own House Value (Yuan)	The discount amount of self-owned house bank loan
	House Value (Yuan)	The discount amount of self-owned house and storage room in apply for a loan
Operations Capacity	Skills Status	The agricultural technologies (plantation, cultivation, etc.) grasped by the loan applicant and family members
	Net Business Income (Yuan)	The net yearly income from farmer's business activities
	Net Income Per Year Per Capita GDP (%)	The percentage of farmer's net yearly income in the local capita GDP
	Net Income (Yuan)	The net yearly income including farmer's operational and non-operational activities
	Daily Living Expenses (Yuan)	The last year daily expenditure of the family when the farmer applies loan
	Total Expenses (Yuan)	The last year total expenditure of the family when the farmer applies loan
	Expenses/Income (%)	The percentage of family total expenditure in the income when the farmer applies loan
	Total Property (Yuan)	The total offset value of property when the farmer applies loan
	Agricultural Net Income (Yuan)	The last year total income from agricultural operational activities when the farmer applies loan
	Agricultural Production Income (Yuan)	The last year total income from agricultural production activities when the farmer applies loan
	Agricultural Production Expenses (Yuan)	The last year total expenditure of agricultural production activities when the farmer applies loan
	The Proportion of Non-agricultural Income (%)	The last year percentage of non-agricultural income in total income when the farmer applies loan
	Children's Education Spending (Yuan)	The last year expenditure for children's education when the farmer applies loan
Willingness of Repayment	Living Status	The living house is self-owned, mortgage-purchased, collective-owned or borrowed when the farmer applies loan
	Living Stability (Year)	The local residential time when the farmer applies loan
	Amounts of Unpaid Loans (Yuan)	The unpaid bank loan when the farmer applies loan
	Whether or Not to Have Unpaid Loan	Whether there exists unpaid bank loan when the farmer applies loan (yes or no)
	Deposit in Bank (Yuan)	The deposit amount when the farmer applies loan
Willingness of Repayment	Whether or Not to Have Private Loan	Whether there exists private loan when the farmer applies loan
	Default or Not	Whether there exists default in historical loan (yes or no)
	Number of Historical Apply Loan	The number of previous loan applications when the farmer applies loan
	Loaning Records	The number of bank loan lending and default situations when the farmer applies loan
Guarantee and Joint Guarantee	Whether or Not to Have Guarantee	Whether guaranteed when the farmer applies loan
	Power of Guarantor (Yuan)	The monthly average income of the guarantor if it is guaranteed
	Whether or Not to Have Joint Guarantee	Whether it is five members joint guarantee
	The Relationship of Joint Guarantee Member	The relationship between five members in joint guarantee (neighbor, business partner, relatives, or others)
Macro Environment	Net Per Capita Income (Yuan)	The capita net income of homeplace when the farmer applies loan
	Per Capita Agricultural Output Value (%)	The capita total agricultural output of homeplace when the farmer applies loan (total agricultural output divided by total population)
	Regional GDP Growth Rate (%)	The GDP growth rate of homeplace when the farmer applies loan
	Consumer Price Index	The consumer price index of homeplace when the farmer applies loan
	Capita Savings Deposit Balance (Yuan)	The capita deposit balance of homeplace when the farmer applies loan
	Engel Coefficient	The Engel coefficient of homeplace when the farmer applies loan

References

[1] Abdou HA, Pointon J. Credit scoring, statistical techniques and evaluation criteria: a review of the literature. Intell Syst Account Finance Manage 2011;18(2-3,):59–88.

[2] Abellán J, Mantas CJ. Improving experimental studies about ensembles of classifiers for bankruptcy prediction and credit scoring. Expert Syst Appl 2014;41(8):3825–30.

[3] Agricultural Bank of China (ABC). Management of farmer's credit rating of agricultural bank of China. Beijing: Agricultural Bank of China; 2008.

[4] Akkoc S. 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 Operat Res 2012;222(1):168–78.

[5] Altman EI, Haldeman RG, Narayanan P. ZETATM analysis: a new model to identify bankruptcy risk of corporations. J Banking Finance 1977;1(1):29–54.

[6] Austin EJ, Deary IJ, Gibson GJ, McGreggor MJ, Dent JB. Attitudes and values of Scottish farmers: 'Yeoman' and 'Entrepreneur' as factors, not distinct types. Rural Sociol 1996;61(3):464–74.

[7] Bai C, Dhavale D, Sarkis J. Complex investment decisions using rough set and fuzzy c-means: an example of investment in green supply chains. Eur J Operat Res 2016;248(2):507–21.

[8] Bai C, Rezaei J, Sarkis J. Multicriteria green supplier segmentation. IEEE Trans Eng Manage 2017;64(4):515–28.

[9] Bai C, Sarkis J. A grey-based DEMATEL model for evaluating business process management critical success factors. Int J Prod Econ 2013;146(1):281–92.

[10] Bai C, Sarkis J. Improving green flexibility through advanced manufacturing technology investment: modeling the decision process. Int J Prod Econ 2017;188:86–104.

[11] Bai C, Sarkis J, Wei X, Koh L. Evaluating ecological sustainable performance measures for supply chain management. Supply Chain Management 2012;17(1):78–92.

[12] Bai C, Sarkis J. Determining and applying sustainable supplier key performance indicators. Supply Chain Management 2014;19(3):275–91.

[13] Bai C, Dhavale D, Sarkis J. Integrating Fuzzy C-Means and TOPSIS for performance evaluation: an application and comparative analysis. Expert Syst Appl 2014;41(9):4186–96.

[14] Barbati M, et al. Optimization of multiple satisfaction levels in portfolio decision analysis. Omega 2017:1–13. doi:10.1016/j.omega.2017.06.013.

[15] Bazan JG, Szczuka M. The rough set exploration system. Trans Rough Sets III 2005:37–56.

[16] Berrone P, Cruz C, Gómez-Mejia LR. Socioemotional wealth in family firms: theoretical dimensions, assessment approaches, and agenda for future research. Family Bus Rev 2012;25:258–79.

[17] Beynon MJ, Peel MJ. Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega 2001;29(6):561–76.

[18] Bijak K, Thomas LC. Modelling LGD for unsecured retail loans using Bayesian methods. J Oper Res Soc 2015;66(2):342–52.
[19] Burns J. Former ratings-firm officials blame conflicts for rosy views. Wall Street J 2008:A4.
[20] Calantone RJ, Di Benedetto CA, Errunza VR. The use of discrete variable selections for credit evaluations. Omega 1988;16(5):469–80.
[21] Carmona R, Fouque JP, Vestal D. Interacting particle systems for the computation of rare credit portfolio losses. Finance Stochastics 2009;13(4):613–33.
[22] Castellani B, Hafferty F. Sociology and complexity science. London: Springer; 2009.
[23] Central People's Government Website of the People's Republic of China (CPGW). 2012 Chinese non-agricultural registered permanent residence population account for 35.29%. (March 6, 2013) http://www.gov.cn/zhuanti/2013-03/06/content_2593537.htm.
[24] Central People's Government Website of the People's Republic of China (CPGW). Central Document No.1 of the Central People's Government and State Council of the People's Republic of China from 2004 to 2015. (June 1, 2017) http://new.sousuo.gov.cn/s.htm?t=gov&q=%E4%B8%80%E5%8F%B7%E6%96%87%E4%BB%B6.
[25] Chen LH, Chiou TW. A fuzzy credit-rating approach for commercial loans: a Taiwan case. Omega 1999;27(4):407–19.
[26] Chi Guotai Research Group of Dalian University of Technology. Credit risks decision and evaluation system of peasant petty loan for the postal savings bank of China. Dalian: Dalian University of Technology; 2011.
[27] Chi GT, Zhang Z. Multi criteria credit rating model for small enterprise using a nonparametric method. Sustainability 2017(9):1–22.
[28] China Construction Bank (CCB). Small business customers evaluation approaches of China construction bank. Beijing: China Construction Bank; 2007. p. 1–8.
[29] Crosbie P, Bohn J. Modeling default risk. Moody's KMV Corporation; 2003. p. 5–30.
[30] Emei Credit Cooperatives (ECC). The operating rules and regulations of microcredit loans to farmers for rural credit cooperatives of Sichuan Province. (March 7, 2012) http://www.dyaudit.gov.cn/newsshowdetails.aspx?newsid=539 .
[31] Falbo P. Credit-scoring by enlarged discriminant models. Omega 1991;19(4):275–89.
[32] Fair Isaac Corporation (FICO). Free credit scores estimator: get your estimated FICO scores range. (January 4, 2016) http://www.myfico.com/fico-credit-score-range-estimator/ .
[33] Fiss PC. Building better causal theories: a fuzzy set approach to typologies in organizational research. Acad Manage Rev 2011;54(2):393–420.
[34] Ratings Fitch. Fitch ratings global corporate finance 2014 transition and default study. Credit Market Res – Fitch Ratings 2014:1–19.
[35] Fuglseth AM, Grønhaug K. IT-enabled redesign of complex and dynamic business processes: the case of bank credit evaluation. Omega 1997;25(1):93–106.
[36] Glover JL, Reay T. Sustaining the family business with minimal financial rewards: how do family farms continue? Family Bus Rev 2013;28(2).
[37] Griffiths B, Beynon MJ. Expositing stages of VPRS analysis in an expert system: application with bank credit ratings. Expert Syst Appl 2005;29(4):879–88.
[38] Gupton GM, Finger CC, Bhatia M. Creditmetrics: technical document. New York: J. P. Morgan & Co. Incorporated; 1997. p. 5–22. 43.
[39] Gupta MR, Chaudhuri S. Formal credit, corruption and the informal credit market in agriculture: a theoretical analysis. Economica 1997;64(254):331–43.
[40] Hajek P. Predicting corporate investment/non-investment grade by using interval-valued fuzzy rule-based systems — a cross-region analysis. Appl Soft Comput 2018;62:73–85.
[41] He Y, Liao N, Zhou Y. Analysis on provincial industrial energy efficiency and its influencing factors in China based on DEA-RS-FANN. Energy 2018;142(1):79–89.
[42] Hu QH, Xie ZX, Yu DR. Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation. Pattern Recognit 2007;40:3509–21.
[43] Industrial and Commercial Bank of China (ICBC). The notice about printing the evaluation method of the small business corporate clients credit rating of China industrial and commercial bank, Beijing: Industrial and Commercial Bank of China; 2005. No. [2005]78.
[44] Jardin P. A two-stage classification technique for bankruptcy prediction. Eur J Operat Res 2016;254(1):236–52.
[45] Ju YH, Sohn SY. Updating a credit-scoring model based on new attributes without realization of actual data. Eur J Operat Res 2014;234(1):119–26.
[46] Kim Y, Sohn SY. Random effects model for credit rating transitions. Eur J Operat Res 2008;184(2):561–73.
[47] Liang J, Shi Z, Li D, Wierman MJ. Information entropy, rough entropy and knowledge granulation in incomplete information systems. Int J Gen Syst 2006;35(6):641–54.
[48] Malhotra R, Malhotra DK. Differentiating between good credits and bad credits using neuro-fuzzy systems. Eur J Operat Res 2002;136(1):190–211.
[49] Malhotra R, Malhotra DK. Evaluating consumer loans using neural networks. Omega 2003;31(2):83–96.
[50] Marques AI, García V, Sanchez JS. A literature review on the application of evolutionary computing to credit scoring. J Oper Res Soc 2012;64(9):1384–99.
[51] Min JH, Lee YC. A practical approach to credit scoring. Expert Syst Appl 2008;35(4):1762–70.
[52] Moula FE, Chi G, Abedin MZ. Credit default prediction modeling: an application of support vector machine. Risk Manage 2017;19(2):158–87.
[53] Moody's Investors Service (2009), Global credit research, Moody's Investors Service, pp. 136–147.
[54] Murphy A. An analysis of the financial crisis of 2008: causes and solutions 2-28, Electronic copy available at. (January 8, 2018) http://ssrn.com/abstract=1295344 .
[55] National Bureau of Statistics of the People's Republic of China (NBS). International statistical yearbook – 2013. Beijing: China Statistics Press; 2013.
[56] Oksan B. Muhammad Yunus, Grameen bank and the nobel peace prize: what political science can contribute to and learn from the study of microcredit. Int Studies Rev 2008;10(3):525–47.
[57] Ozer M. User segmentation of online music services using fuzzy clustering. Omega 2001;29(2):193–206.
[58] Ozer M. Fuzzy c-means clustering and Internet portals: a case study. Eur J Operat Res 2005;164(3):696–714.
[59] Pawlak Z. Rough sets. Int J Comput Inf Sci 1982;11(5):341–56.
[60] Postal Savings Bank of China (PSBC). Merchant credit rating table of postal savings bank of china. Beijing: Postal Savings Bank of China; 2009.
[61] Postal Savings Bank of China (PSBC). Social responsibility report and "Three rural" financial service report for postal savings bank of China in 2012. Postal Savings Bank of China Co., LTD; 2013.
[62] Pudong Development Bank (PDB). Credit rating and credit application form of Pudong development bank. Shanghai: Pudong Development Bank; 2007.
[63] Rowshannahad M, Dauzère-Pérès S, Cassini B. Capacitated qualification management in semiconductor manufacturing. Omega 2015;54:50–9.
[64] Shi B, Chen N, Wang J. A credit rating model of microfinance based on fuzzy cluster analysis and fuzzy pattern recognition: empirical evidence from Chinese 2,157 small private businesses. J Intell Fuzzy Syst 2016;31(6):3095–102.
[65] Shi B, Meng B, Yang H, Wang J, Shi W. A novel approach for reducing attributes and its application to small enterprise financing ability evaluation. Complexity 2018;2018:1032643.
[66] Shi B, Wang J, et al. A novel imbalanced data classification approach based on Logistic regression and Fisher discriminant. Math Problems Eng 2015;2015:1–12.
[67] Shi B, Yang H, Wang J, Zhao J. City green economy evaluation: empirical evidence from 15 sub-provincial cities in China. Sustainability 2016;8(6):1–39.
[68] Standard & Poor's Ratings Services (S&P's). S&P's study of China's top corporates highlights their significant financial risks. Standard & Poor's Ratings Services; 2012. p. 175–99.
[69] Tseng TLB, Huang CC. Rough set-based approach to feature selection in customer relationship management. Omega 2007;35(4):365–83.
[70] Wang TC, Chen YH. Applying rough sets theory to corporate credit ratings. In: IEEE international conference: service operations and logistics, and informatics; 2006. p. 132–6.
[71] Xia Y, Liu C, Li Y, Liu N. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. Expert Syst Appl 2017;78:225–41.
[72] Xu Y, Zhang Y. A online credit evaluation method based on AHP and SPA. Commun Nonlinear Sci Numer Simul 2009;14(7):3031–6.
[73] Yurdakul M, Ic YT. AHP approach in the credit evaluation of the manufacturing firms in Turkey. Int J Prod Econ 2004;88(3):269–89.
[74] Zhong H, Miao C, Shen Z, Feng Y. Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings. Neurocomputing 2014;128(3):285–95.
[75] Zhu X, Li ZN. Heterogeneous impact of farmer credit: an empirical investigation based on IVQR model. Syst Eng – Theory Pract 2007;27(2):68–75.