Decision Science Letters 12 (2023) 267–278

Contents lists available at GrowingScience

Decision Science Letters

homepage: www.GrowingScience.com/dsl

Integration of factor analysis and Tsukamoto's fuzzy logic method for quality control of credit provisions in rural banks

Yuyun Hidayat[a], Sukono[b], Predy Hartanto[c], Titi Purwandari[b], Riza Andrian Ibrahim[d], Moch Panji Agung Saputra[d] and Jumadil Saputra[e*]

[a]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia
[b]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia
[c]Undergraduate of Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia
[d]Doctoral of Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia
[e]Department of Economics, Faculty of Business, Economics and Social Development, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

C H R O N I C L E

Article history:
Received: November 26, 2022
Received in revised format: December 20, 2022
Accepted: January 30, 2023
Available online: January 30, 2023

Keywords:
Credit risk
Credit risk rate
Factor analysis
Tsukamoto's fuzzy logic method

A B S T R A C T

Giving credit to debtors can pose a default risk. This risk arises because of an error in analyzing the credit risk rate of the debtor. Therefore, this study aims to design a framework for analyzing the credit risk rate of debtors so that the default risk can be reduced. This framework is created using the integration of factor analysis and Tsukamoto's fuzzy logic method. This integration method can group many credit assessment variables into several decisive factors. In addition, the integration method can estimate credit risk rate firmly based on the α-predicate of each basic rule. This analytical framework is simulated on credit application data at a Rural Bank, in Indonesia. The simulation results show that there are three factors and one variable to measure the credit risk rate, namely: factor 1 represents repayment capacity, business length, working capital, and liquidity value; factor 2 represents the age and the difference between the granted and the proposed loan amount; factor 3 represents the stay length, character, and credit history; and one variable represents a dependent number. This research is expected to help credit institutions measure the credit risk rate in making credit decisions for prospective debtors.

© 2023 by the authors; licensee Growing Science, Canada.

1. Introduction

Credit institutions such as banks are needed by the business world from various segments, including oversized, medium, retail, and micro corporations (Mkhaiber & Werner, 2021). Credit institutions must pay special attention to credit rate analysis in providing credit funds, so there are a few default risks. Credit institutions must formulate credit standards, assessment variables, and other additional data required to apply (Mhlanga, 2021). Then, the application submission is processed and analyzed until a credit granting decision is obtained based on the guidelines owned by the institution. Debtors are required to fill out a registration form when applying for credit (Xiao & Wu, 2008). Then, the analyst collects the forms used as data for processing. Analysts also conduct interviews and field surveys with debtors. The form, interview, and survey data were checked for significance. After the inspection is complete, this study measures the credit risk rate of the debtor. At this measurement stage, inconsistency of each analysis may occur because they may only use personal understanding in measuring it (Brown et al., 2018). Therefore, credit risk rate analysis requires a consistent analytical framework. This analytical framework must be easy to perform, fast, efficient, and accurate to minimize the default risk.

* Corresponding author.
E-mail address: jumadil.saputra@umt.edu.my (J. Saputra)

© 2023 by the authors; licensee Growing Science, Canada.
doi: 10.5267/dsl.2023.1.008

Several studies have examined the credit risk analysis framework. Yu et al. (2009) designed a credit risk analysis framework with multi-criteria using fuzzy group decision-making (FGDM). Wu et al. (2014) used a two-stage analysis through the integration of supervised machine learning and preprocessing data. The integration method accuracy is high at 82.96 percent. Then, Yi et al. (2015) analyzed the credit risk level in the port area using an external-internal information fusion model. The prediction results obtained have a lower mean square error than the non-information-fusion model. Bao et al. (2019) proposed an integration strategy of unsupervised learning with supervised learning for credit risk assessment. Pan et al. (2020) designed a credit risk analysis framework for debtors using the integration of genetic algorithm methods and a hybrid kernel support vector machine (SVM). Yangyudongnanxin (2021) introduced a credit risk analysis strategy using a weighted random forest algorithm. Yangyudongnanxin (2021) also compared this strategy with several other methods, and the result is that the accuracy of its strategy is the best. Wu et al. (2021) introduced the constrained logistic regression method to measure the default risk of debtors in making credit decisions.

There are gaps in previous studies. Previous studies have generally grouped the credit assessment variables without any logical reasoning. It can lead to an unbiased credit risk rate because there may be significant and non-significant credit assessment variables in one group of variables. Then, the methods used are generally less based on the concepts of natural logic. It makes the concept difficult to understand. In addition, many explanations of these methods are presented with complex mathematical equations. It is practically also tricky for the general public to understand. Finally, the methods used in previous studies generally do not use standard rules in decision-making. It causes the inferences of decision-making to be fewer firms.

Based on the gaps described previously, this study aims to design a simple framework of credit risk rate analysis that can facilitate practitioners to measure the credit risk rate in making credit decisions for prospective debtors. In designing the framework, we introduce the integration of factor analysis and Tsukamoto's fuzzy logic method. Factor analysis can group credit assessment variables into several significant factors. This grouping is based on the enormous loading value of each variable in the factors involved. This grouping will also not eliminate the role of each credit assessment variable in measuring the credit risk rate (Yang & Islam, 2020). Then, Tsukamoto's fuzzy logic method is straightforward to use. The reason is that this method is explained by simple mathematical equations, which can be easily understood using logic. Tsukamoto's fuzzy logic method can also overcome the lack of inference from the decisions taken (Sudiyatno & Wibowo, 2018). Finally, a simulation using this integration method is also carried out on credit application data at a Rural Bank, in Indonesia. It is expected to provide a more practical guide for users. This research is expected to help credit institutions such as banks to measure the credit risk rate in making credit decisions for prospective debtors.

2. A Brief Explanation of Credit Risk

Credit risk is the risk of loss due to the debtor failing to fulfil its obligations to pay the loan and interest (Spuchľáková et al., 2015). These failures are like late payments and inability to pay. There are two common reasons for default from debtors, namely external and internal reasons (Llorca, 2017). External reasons include the debtor's unwillingness to pay due to the debtor's character, the institution's weakness in identifying the debtor's feasibility, and the debtor's declining business condition due to mismanagement. Then, internal reasons include weak institutional control and control systems, ineffective management processes, and the existence of bad faith from institutional administrators. Bartholomew (1985) classified the credit risk rates as follow:

(1) The lower credit risk rate. The lower credit risk rate can practically be seen from the excellent business prospects and can be controlled by sound management and integrity.
(2) The moderate credit risk rate. A moderate credit risk rate can practically be viewed from a moderate business prospect and supported by sufficient collateral.
(3) The high credit risk rate. A high credit risk rate can practically be viewed from business activities with doubtful prospects.

3. Factor Analysis and Tsukamoto Fuzzy-Logic Method

3.1. Mathematical Notation

The mathematical notations used are as follows:

(1) M represents the number of credit assessment variables involved.
(2) X_m represents the m-th credit assessment variable, where $m = 1,2, \dots, M$.
(3) Y_m represents the standardized credit assessment variable X_m.
(4) N represents the number of factors that group the assessment variables.
(5) F_n represents the n-th credit assessment factor, where $n = 1,2, \dots, N$.
(6) K_n represents the number of sets that represent the factor F_n.

(7) $S_{n,k}$ is the k-th set representing of F_n, where $k = 1, 2, \dots, K_n$. It is also known as the k-th fuzzy set, which represents F_n.
(8) $\mu_{S_{n,k}}(\cdot)$ is a membership function that maps $S_{n,k}$ into a real number interval [0, 1].
(9) J represents many fuzzy rule bases.
(10) α_j represents the fire strength value of the j-th fuzzy rule basis, where $j = 1, 2, \dots, J$. It is also known as the α-predicate of the j-th fuzzy rule base.
(11) r_j represents the credit risk rate of the j-th fuzzy rule basis.
(12) R represents the final credit risk rate.

3.2. Factor Analysis

There are many credit assessment variables involved in measuring the credit risk rate. It is so that the measurement of the credit risk rate is not biased (Bartholomew, 1985). However, many credit assessment variables can complicate the analysis process. Factor analysis is one of the analytical frameworks that can be used to overcome these difficulties. Factor analysis can make the measurement of the credit risk rate of the debtor simpler. This analysis is carried out by grouping the many credit assessment variables involved in several factors without eliminating their role (Ludvigson & Ng, 2007). The unit of credit assessment variable X_m involved in factor analysis must be uniform. Suppose the units of the credit assessment variable X_m vary. In that case, each credit assessment variable X_m is standardized first, resulting in a standardized credit assessment variable Y_m. If standardization is not carried out, this will cause unbiasedness in measuring the credit risk rate (Bartholomew, 1985). After X_m is standardized to Y_m, Y_m must meet the assumption of variable suitability in factor analysis. This compatibility check can be done using the Kaiser-Meyer-Olkin (KMO) test. The standardized credit assessment variable Y_m is expressed as a linear combination of the factors involved. Mathematically, Velicer & Jackson (1990) written as follows:

$$Y_m = B_{m,1}F_1 + B_{m,2}F_2 + \cdots + B_{m,N}F_N;\ m = 1, 2, \dots, M, \tag{1}$$

where $B_{m,n}$ states the regression coefficient of the m-th standardized credit assessment variable to the factor F_n. Some assumptions must be met. The assumption is that $\{Y_m, m = 1,2, \dots, M\}$ is not independent (Saul & Rahim, 2000). In other words, Y_{m_1} and Y_{m_2} with $m_1 \neq m_2$ has a non-zero correlation. In Eq. (1), the determination of the value of $F_n, n = 1, 2, \dots, N$ can be determined by making it a linear combination of standardized credit assessment variables. Mathematically, it can be expressed as follows (McDade & Adair, 2001):

$$F_n = W_{n,1}Y_1 + W_{n,2}Y_2 + \cdots + W_{n,M}Y_M;\ n = 1, 2, \dots, M, \tag{2}$$

where $W_{n,m}$ represents the regression coefficient of the n-th factor on the standardized credit assessment variable Y_m. The number of significant factors can be sequentially reviewed. The first order starts from one factor first. Then, the following sequence continues to involve two factors and so on. The number of factors chosen is M, so the variance value of the model in Eq. (1) is more than 10 for the last time (Corrigan et al., 2004). Another alternative is to review the eigenvalues of the correlation matrix (Miwakeichi et al., 2004). The number of factors chosen is M, so the eigenvalue of the correlation matrix Eq. (1) is more than 1 for the last time. After many factors have been determined, the next step is to determine the members of each factor. It can be conducted using the principal component analysis (PCA) method with or without rotational operations (Lenka et al., 2021). Each Y_m variable can be classified based on the highest factor loading value in each row in the factor matrix. The factor matrix contains elements containing the loading factor between the Y_m variable and the F_n factor. Ensuring that the many factors used are representative of the population is essential. For that, validation is the last thing to do. This validation is conducted by dividing the sample data into several parts of the same size. Make sure that the sample pieces are at least 50 in size. It is so that the sample pieces are representative. After the sample is cut, perform a factor analysis similarly on each sample piece. If the final result of the factor analysis of each sample piece is the same, then the result of the factor analysis of the sample as a whole is valid.

3.3. Tsukamoto's Fuzzy Logic Method in Credit Risk Rate Analysis

Tsukamoto's fuzzy logic method can provide an overview of the debtor's credit risk rate based on predetermined basic rules. Credit risk rate analysis using Tsukamoto's fuzzy logic method consists of three stages: fuzzification, inference, and defuzzification (Suharjito et al., 2017). Fuzzification is the transformation stage of the F_n factor into fuzzy form and the F_n categorization stage. Next is the inference stage. This stage contains mapping each value of the F_n factor per base rule using certain membership functions so that the α-predicate in the base rule is obtained. Finally, defuzzification is the stage of processing each base rule's α-predicate values so that the credit risk rate is obtained. A more detailed description of the three stages is presented in sections 3.3.1 to 3.3.2.

3.3.1. Fuzzification

Each value of the factor F_n is determined first. It is done using a weighting rule for each standardized credit assessment variable included in the factor (Setyono & Aeni, 2018). The fuzzification process can practically be done using IBM

Statistics SPSS 23 software. Furthermore, the F_n factor is divided into several categories. For example, the factor F_n is divided into K_n categories. Each category is assumed to be $S_{n,k}$ with $k = 1, 2, \dots, K_n$.

3.3.2. Inference

After the categories of $S_{n,k}$ are obtained, the next step is determining the membership function that maps $S_{n,k}$. The membership function is a function that mapped $S_{n,k}$ into a closed interval between 0 to 1. The mapping results are later used to determine the α-predicate in the inference stage. There are several types of membership functions. The most popular are the ascending linear membership function, the descending linear membership function, and the triangular membership function (Adriyendi, 2018). The linear ascending membership function is expressed as follows:

$$\mu_{S_{n,k}}(a) = \begin{cases} 0 & ; a \le a_{n,k}^{min} \\ \dfrac{a - a_{n,k}^{min}}{a_{n,k}^{max} - a_{n,k}^{min}} & ; a_{n,k}^{min} < a < a_{n,k}^{max}, \\ 1 & ; a \ge a_{n,k}^{max} \end{cases} \tag{3}$$

where $a \in S_{n,k}$, $a_{n,k}^{min}$ represents the minimum value of $S_{n,k}$, dan $a_{n,k}^{max}$ represents the maximum value of $S_{n,k}$. Then, the descending linear membership function is expressed as follows:

$$\mu_{S_{n,k}}(a) = \begin{cases} 1 & ; a \le a_{n,k}^{min} \\ \dfrac{a_{n,k}^{max} - a}{a_{n,k}^{max} - a_{n,k}^{min}} & ; a_{n,k}^{min} < a < a_{n,k}^{max}, \\ 0 & ; a \ge a_{n,k}^{max} \end{cases} \tag{4}$$

Finally, the triangular membership function is expressed as follows:

$$\mu_{S_{n,k}}(a) = \begin{cases} 1 & ; a = a_{n,k}^{med} \\ \dfrac{a - a_{n,k}^{min}}{a_{n,k}^{med} - a_{n,k}^{min}} & ; a_{n,k}^{min} < a < a_{n,k}^{med} \\ \dfrac{a_{n,k}^{max} - a}{a_{n,k}^{max} - a_{n,k}^{med}} & ; a_{n,k}^{med} < a < a_{n,k}^{max} \\ 0 & ; a \le a_{n,k}^{min} \vee a \ge a_{n,k}^{max} \end{cases}, \tag{5}$$

where $a_{n,k}^{med}$ represents the median of $S_{n,k}$.

Next is the determination of the value of the α-predicate for each base rule. First, the basic rules must be defined. Each of these base rules contains the implications of the categories in each factor and their inferences. The number of these base rules is $J = K_1 \times K_2 \times \dots \times K_N$ (Bandyopadhyay et al., 2013). Let the α-predicate of the 1st base rule (α_1) be the implication of the category $S_{n,1}$. α_1 is mathematically determined using the following equation:

$$\alpha_1 = \bigcap_{n=1}^{N} \mu_{S_{n,1}}(a) = \min_{n=1,2,\dots,N} \mu_{S_{n,1}}(a) \tag{6}$$

In this study, the α-predicate for the j-th basis rule is denoted by α_j, where $j = 1, 2, \dots, J$.

3.3.3. Defuzzification

Defuzzification is the final stage in credit risk rate analysis. The credit risk rate of each basic rule is processed so that the average credit risk rate is obtained. The average credit risk rate is the final credit risk rate obtained (Ardika et al., 2017).

α_j of a j-th basis rule is used to map the credit risk rate function. This credit risk rate function is based on the rules used in financial institutions. For example, suppose that the credit risk level is divided into five: watchlist, marginal, average, moderate, and high. Watchlist, marginal, average, moderate, and high credit risk levels respectively at intervals $[0, 20]\%$,

$(20, 40]$%, $(40, 60]$%, $(60, 80]$%, and $(80,100]$%. Mathematically, the credit risk rate function of each level is written as follows:

$$\mu_{watchlist}(r) = \begin{cases} 1 & ; r = 10 \\ \dfrac{r-0}{10-0} & ; 0 < r < 10 \\ \dfrac{20-r}{20-10} & ; 10 < r \leq 20 \\ 0 & ; r \leq 0 \vee r > 20 \end{cases}, \tag{7}$$

$$\mu_{marginal}(r) = \begin{cases} 1 & ; r = 30 \\ \dfrac{r-20}{30-20} & ; 20 < r < 30 \\ \dfrac{40-r}{40-30} & ; 30 < r \leq 40 \\ 0 & ; r \leq 20 \vee r > 40 \end{cases}, \tag{8}$$

$$\mu_{average}(r) = \begin{cases} 1 & ; r = 50 \\ \dfrac{r-40}{50-40} & ; 40 < r < 50 \\ \dfrac{60-r}{60-50} & ; 50 < r \leq 60 \\ 0 & ; r \leq 40 \vee r > 60 \end{cases}, \tag{9}$$

$$\mu_{moderate}(r) = \begin{cases} 1 & ; r = 70 \\ \dfrac{r-60}{70-60} & ; 60 < r < 70 \\ \dfrac{80-r}{80-70} & ; 70 < r \leq 80 \\ 0 & ; r \leq 60 \vee r > 80 \end{cases}, \tag{10}$$

and

$$\mu_{high}(r) = \begin{cases} 1 & ; r = 90 \\ \dfrac{r-60}{70-60} & ; 80 < r < 90 \\ \dfrac{80-r}{80-70} & ; 90 < r \leq 100 \\ 0 & ; r \leq 80 \vee r > 100 \end{cases}. \tag{11}$$

After the credit risk rate function is designed, now is the calculation of the credit risk rate of each basic rule. First, take α_j as the value of the credit risk rate function. Then, determine the inverse of the credit risk rate function (Fajri et al., 2017). That is what the credit risk rate of the base rule is. In this case, it is denoted by r_j. Finally, to determine the final credit risk rate, calculate the average credit risk rate of each base rule using the following equation (Akhirina et al., 2019):

$$R = \frac{\sum_{j=1}^{J} \alpha_j r_j}{\sum_{j=1}^{J} \alpha_j}, \tag{12}$$

4. Simulation

4.1. Data Description

The data used in this simulation are primary and secondary data types. The primary data is the result of interviews with officers from a Rural Bank, in Indonesia, regarding the financial institution's credit condition. Meanwhile, the secondary data used are as follows: (1) rural business loan form, (2) customer visit result form, (3) collateral appraisal reports, (4) balance reports, (5) profit reports, (6) loss reports, (7) photos of guarantees, (8) business credit decisions, (9) credit disbursement, and (10) customer coaching/supervision form. We have collaborated with the bank and have been permitted to process the data in this research. The size of the data used is 100. For company purposes, we cannot share the data openly.

4.2. Determination of Credit Assessment Variables

The variables used are data from the field survey of prospective debtors by an official of the Rural Bank, in Indonesia. The assessment variables are as follows:

(1) $X_1 \in \mathbb{Z}^+$ represents the age of the prospective debtor in years when applying for credit.
(2) $X_2 \in \mathbb{Z}^+$ represents the length of stay of the prospective debtor in years at the place of residence.
(3) $X_3 \in \mathbb{Z}^+$ represents the number of dependents with units of people.
(4) $X_4 \in \{1, 2, 3, 4\}$ represents the character of a prospective debtor. The numbers "1", "2", "3", and "4," respectively, represent very good, good, moderate, and poor characters.
(5) $X_5 \in \{1, 2, 3, 4\}$ represents the credit history of the prospective debtor. The numbers "1", "2", "3", and "4", respectively, represent instalment payments that are always on time, in arrears and whole, in arrears and not yet paid off, and not paying at all.
(6) $X_6 \in \mathbb{R}^+$ represents the ability to return payments in IDR/month.
(7) $X_7 \in \mathbb{R}^+$ represents the amount of own funds owned by prospective debtors in IDR units.
(8) $X_8 \in \mathbb{Z}^+$ represents the length of time the prospective debtor has been in business, from starting the business until submitting a credit application in years unit.
(9) $X_9 \in \mathbb{R}^+$ represents rounding off the value of collateral liquidity submitted by prospective debtors in IDR units.
(10) $X_{10} \in \mathbb{R}^+$ represents the difference between the loan size received and the one proposed by the debtor in IDR units.

Finally, there are additional assumptions that we use as follows:

(1) There is no difference between the credit risk rates of male and female prospective debtors. This assumption is used to make calculating the credit risk rate easier.
(2) Prospective debtors are required to own a house. This assumption is used so that prospective debtors can focus on paying their credit with a sense of security so that the default risk can be reduced.
(3) The business analyzed by the debtor does not consider the type of business, business status, and business conditions of the prospective debtor. This assumption is used to simplify determining the credit risk rate.

4.3. Checking of Non-Independent Assumption

First, standardizing the values of each credit assessment variable X_m is carried out. This standardization is carried out using IBM SPSS Statistics 23 software. The standardization results in the Y_m variables. After standardization, the next step is to check the non-independent assumption of Y_m. To check it, we used the Bartlett statistical test with a significance value of 0.05. In short, the test statistic value obtained is 0.001. The value of this test statistic is smaller than its significance value. It indicates that Y_m is not independent of each other (Glaser, 1976). Therefore, the non-independent assumption of Y_m is satisfied.

4.4. Variable Suitability Test in Factor Analysis

The suitability test for the Y_m variable in factor analysis is performed using the Kaiser-Meyer-Olkin (KMO) test. In short, the statistical value of the KMO test obtained is 0.453. This value is less than 0.500. It indicates that the Y_m variable is not suitable to be studied by factor analysis (Liu & Wang, 2021). The solution to this problem is to exclude one variable with the smallest anti-image correlation value. In short, the standardized number of dependents Y_3 has the smallest anti-image correlation value among the others, which is 0.368. It indicates that this variable cannot be included in the factor analysis. After Y_3 is ignored, the next step is to repeat the suitability test of the variable Y_m with $m \neq 3$ in factor analysis using the KMO test. In short, the statistical value of the KMO test obtained is 0.602. This value is more significant than 0.500. It indicates that the variable Y_m is suitable to be studied by factor analysis (Liu & Wang, 2021). Thus, the variable Y_m with $m \neq 3$ is grouped as factors in factor analysis, while the variable Y_3 is not. Y_3 becomes the single variable in measuring the credit risk rate in the final stage.

4.5. Application of Factor Analysis

Many factors used are analyzed first. It is conducted as described in section 3.2. The variance value and the eigenvalues of the correlation matrix from Eq. (1) are used to select the number of factors. The variance value and the eigenvalues of the correlation matrix of Eq. (1) are presented in Table 1.

Table 1
The Variance Value and the Eigenvalues of the Correlation Matrix of Eq. (1).

M	The Eigenvalues	The Variance Value
1	2.343	26.029
2	1.561	17.341
3	1.309	14.541
4	0.887	9.852
5	0.788	8.756
6	0.683	7.587
7	0.644	7.155
8	0.524	5.821
9	0.263	2.917

Table 1 shows that at $M = 3$, the last eigenvalue is greater than 1, and the last variance value is greater than 10. It indicates that the number of factors that are best used is 3. Therefore, the many factors we use are 3. Next is the determination of the members of each factor. The determination of the members of each of these factors is carried out using the principal component analysis (PCA) method with rotation, as described in section 3.2. The factor matrix of each variable Y_m obtained using IBM Statistics SPSS 23 software is presented in Table 2.

Table 2
Factor Matrix

m	F_1	F_2	F_3
1	0.129	0.760	0.297
2	0.207	0.134	0.659
4	-0.004	-0.536	0.567
5	0.027	0.038	-0.660
6	0.789	0.127	0.345
7	0.530	0.334	0.163
8	0.828	-0.154	0.078
9	0.674	-0.133	-0.275
10	-0.139	0.711	-0.163

Table 2 shows that the most significant loading factor of the Y_1 variable is the F_2 factor, which is 0.760. It indicates that Y_1 is suitable to be classified into factor F_2. Then, the most significant loading factor of the Y_2 variable is the F_3 factor, which is 0.659. It indicates that Y_2 is suitable to be classified into factor F_3. Similarly, the results of the classification of each variable Y_m, $m \neq 3$ are presented in Table 3.

Table 3
Classification of Variables to Factors.

F_1	F_2	F_3
Y_6	Y_1	Y_2
Y_7	Y_{10}	Y_4
Y_8		Y_5
Y_9		

Table 3 shows that factor F_1 contains variables Y_6, Y_7, Y_8, and Y_9, each of which is a standardization of the ability to pay monthly debts, the amount of money owned, the length of business, and the value of liquidity. Factor F_2 contains variables Y_1 and Y_{10}, each of which is standardization of age and the difference between loans received and applied. Factor F_3 contains variables Y_2, Y_4, and Y_5, each of which is standardization of length of stay, character, and credit history. The last is the validation stage. Since the sample size of the data is 100, we randomly divided it into two subsamples, A and B. The size of the two subsamples is 50. In brief, the results of the final factor analysis of subsamples A and B are the same as the last factor analysis of the whole. It indicates that many factors in it are valid (Nugraha et al., 2019).

4.6 Application of Tsukamoto's Fuzzy Logic Method

On the basis of the results of the previous analysis, the credit risk rates in this simulation are determined based on factors F_1, F_2, F_3, and the variable Y_3. Factor values F_1, F_2, and F_3 are obtained from each standardized credit assessment variable's weighting, while the value of the Y_3 variable did not change. Next is determining the category division of each factor and variable and its membership function. More details about this are summarized in Table 4.

Table 4 shows that the values of the factor F_1 are divided into three sets, namely $S_{1,1}$, $S_{2,1}$, and $S_{3,1}$. $S_{1,1}$, $S_{2,1}$, and $S_{3,1}$, respectively, represent the set of lower, moderate, and high values of factor F_1. Then, $S_{1,1}$, $S_{2,1}$, and $S_{3,1}$ are respectively mapped using a descending linear, triangular, and ascending linear membership function as shown by Eq. (4), Eq. (5), and Eq. (3). The interpretations for the other factors and variables are the same.

The next stage is the determination of the credit risk level. This study divides the credit risk level into five: watchlist, marginal, average, moderate, and high. Watchlist credit risk level is at an interval $[0, 20]$%, marginal credit risk level is at an interval $(20, 40]$%, the average credit risk level is at an interval $(40, 60]$%, the moderate credit risk level is at interval $(60, 80]$%, and high credit risk level is in the interval $(80, 100]$%. Mathematically, each level's credit risk level function is written in Eq. (7) to Eq. (11). The next stage is the determination of the base rules. This base rule amounts to $J = 3 \times 3 \times 3 \times 5 = 135$ rules. It is based on the multiplication of the number of categories in the set of factors and variables. Because 135 base rules are too many to present in this article, we provide access to view them at the following link: https://bit.ly/3LsWMBp. As a quick overview, we present snippets of the 1-st to 3-rd base rules and the 133-th to 135-th base rule snippets in Table 5.

Table 4
Division Category Set of F_1, F_2, F_3, and Y_3

Factor/ Variable	Set Category	Membership Function Type	Membership Function Parameter
F_1	$S_{1,1}$: The lower-value set of F_1	Linear Descending	$a_{1,1}^{min} = 1.493, a_{1,1}^{max} = 3.529$
	$S_{1,2}$: The moderate-value set of F_1	Triangular	$a_{1,2}^{min} = -1.493, a_{1,2}^{max} = 3.529, a_{1,2}^{med} = 1.018$
	$S_{1,3}$: The high-value set of F_1	Linear Ascending	$a_{1,3}^{min} = -1.493, a_{1,3}^{max} = 3.529$
F_2	$S_{2,1}$: The lower-value set of F_2	Linear Descending	$a_{2,1}^{min} = -1.918, a_{2,1}^{max} = 2.738$
	$S_{2,2}$: The moderate-value set of F_2	Triangular	$a_{2,2}^{min} = -1.918, a_{2,2}^{max} = 2.738, a_{2,2}^{med} = 0.410$
	$S_{2,3}$: The high-value set of F_2	Linear Ascending	$a_{2,3}^{min} = -1.918, a_{2,3}^{max} = 2.738$
F_3	$S_{3,1}$: The lower-value set of F_3	Linear Descending	$a_{3,1}^{min} = -2.052, a_{3,1}^{max} = 2.683$
	$S_{3,2}$: The moderate-value set of F_3	Triangular	$a_{3,2}^{min} = -2.052, a_{3,2}^{max} = 2.683, a_{3,2}^{med} = 0.316$
	$S_{3,3}$: The high-value set of F_3	Linear Ascending	$a_{3,3}^{min} = -2.052, a_{3,3}^{max} = 2.683$
Y_3	$S_{4,1}$: The zero dependent set of Y_3	Linear Descending	$a_{4,1}^{min} = 0, a_{4,1}^{max} = 5$
	$S_{4,2}$: The little dependent set of Y_3	Triangular	$a_{4,2}^{min} = 0, a_{4,2}^{max} = 5, a_{4,2}^{med} = 2$
	$S_{4,3}$: The moderate dependent set of Y_3	Triangular	$a_{4,3}^{min} = 0, a_{4,3}^{max} = 5, a_{4,3}^{med} = 3$
	$S_{4,4}$: The many dependent set of Y_3	Triangular	$a_{4,4}^{min} = 0, a_{4,4}^{max} = 5, a_{4,4}^{med} = 4$
	$S_{4,5}$: The very many dependent set of Y_3	Linear Ascending	$a_{4,5}^{min} = 0, a_{4,5}^{max} = 5$

Table 5
Snippets of Base Rules

j		Value Category of F_1	Value Category of F_2	Value Category of F_3	Value Category of Y_3		Credit Risk Level
1		Lower	Lower	Lower	Zero		Average
2		Lower	Lower	Lower	Little		Average
3	IF	Lower	Lower	Lower	Moderate	ELSE	Marginal
⋮		⋮	⋮	⋮	⋮		⋮
133		High	High	High	Moderate		Moderate
134		High	High	High	Many		Average
135		High	High	High	Very Many		Average

Table 5 shows that the 1st basis rule with a lower value of F_1, a lower value of F_2, a lower value of F_3, and a zero value of Y_3 has an average credit risk level. The interpretation of the other base rules is like that. Now a case study is carried out to determine the credit risk level of one of the debtors. The debtor is referred to as Mr. X. Credit assessment variable data from Mr. X is presented in Table 6.

Table 6
Credit Assessment Variable Data from Mr. X

Credit Assessment Variable	Value	Credit Assessment Variable	Value
X_1	60 years	X_6	4,574,625.00 IDR/Month
X_2	15 years	X_7	233,299,978.00 IDR
X_3	1 person	X_8	18 years
X_4	3 for "moderate"	X_9	76,300,00.00 IDR
X_5	1 for "always on time"	X_{10}	50,000,000.00 IDR

How to determine the credit risk rate of Mr. X is transform the data in Table 6 so that the values of F_1, F_2, F_3, and Y_3 are obtained. The values of F_1, F_2, F_3, and Y_3 are 0.092, 1.447, -0.382, and 1, respectively. Also, map the values of F_1, F_2, F_3, and Y_3 using the membership function for each category. The results of mapping the values of F_1, F_2, F_3, and Y_3 are presented in Table 7.

Table 7
The Results of Mapping the Values of F_1, F_2, F_3, and Y_3

Factor/ Variable	Value Category	Membership Function Map
F_1	Lower	0.684
	Moderate	0.632
	High	0.316
F_2	Lower	0.277
	Moderate	0.555
	High	0.723
F_3	Lower	0.647
	Moderate	0.705
	High	0.352
Y_3	Zero	0.800
	Little	0.800
	Moderate	0
	Many	0
	Very Many	0.200

(1) Determine α_j for each $j = 1, 2, \ldots, 135$ using equation (6). The results of the determination of α_j can be seen at the following link: https://bit.ly/3LsWMBp. We present the snippets α_j for j from 1 to 3 and for j from 132 to 135 in Table 8.

Table 8. The Snippets α_j

j	α_j
1	0.277
2	0.277
3	0
⋮	⋮
133	0
134	0
135	0.200

(2) Determine r_j for each $j = 1,2, \ldots ,135$ using the inverse of the credit risk rate function in equations (7) to (11). The results of determining r_j can be seen at the following link: https://bit.ly/3LsWMBp. We present the snippets of r_j for j from 1 to 3 and for j from 132 to 135 in Table 9.

Table 9
The Snippets r_j

j	r_j
1	13.865
2	13.865
3	0
⋮	⋮
133	0
134	0
135	10

(5) Determine the credit risk level of Mr. X using equation (12). Briefly, the credit risk rate obtained for Mr. X is 19.215%. The credit risk rate of Mr. X is low. It can be a consideration for financial institutions to provide loans to Mr. X.

5. Discussion

Factor F_1 consists of variables Y_6, Y_7, Y_8, and Y_9, each of which is a standardization of the ability to pay monthly debts, the amount of money owned, the length of business, and the value of liquidity. If the correlation rate of the four variables is examined, all values are positive. It means that these four variables contribute to reducing the credit risk rate. The greater the value of the four variables, the lower the credit risk rate of the debtor, and vice versa. This interpretation is logical for variables Y_6, Y_7, and Y_8, but it is odd for Y_9. Supposedly, the higher the Y_9 liquidity value, the higher the credit risk rate, and vice versa.

Factor F_2 consists of variables Y_1 and Y_{10}, each of which is a standardization of age and the difference between loans received and applied. If the correlation rate of the two variables is examined, all values are also positive. It means that these two variables contribute to reducing the credit risk rate. The greater the value of the two variables, the lower the credit risk rate of the debtor, and vice versa. This interpretation is odd. Supposedly, the higher the age and the difference between the loan received and applied, the higher the credit risk rate of the debtor, and vice versa.

Finally, factor F_3 consists of variables Y_2, Y_4, and Y_5, which are standardization of length of stay, character, and credit history, respectively. If the correlation rate of the three variables is examined, all values are also positive. These three variables contribute to reducing the credit risk rate. The greater the value of the three variables, the lower the credit risk rate of the debtor, and vice versa. This interpretation is logical for the variable Y_2 but odd for the variables Y_4 and Y_5. Supposedly, the higher the character Y_4 and credit history Y_5, the higher the credit risk rate, and vice versa.

6. Conclusions

This study provides a new framework for measuring the credit risk rate from debtors through the integration of factor analysis and Tsukamoto's fuzzy logic method. Factor analysis can group credit assessment variables into several significant factors. This grouping is based on the enormous loading value of each variable in the factors involved. This grouping will also not eliminate the role of each credit assessment variable in measuring the credit risk rate. Then, Tsukamoto's fuzzy logic method has a simple system to use. It can be seen from the simplicity of the mathematical equations used. In addition, Tsukamoto's fuzzy logic method can also overcome the lack of inference from the decisions taken.

The analytical framework is simulated on credit application data at a Rural Bank, in Indonesia. The simulation obtains three factors and one variable to measure the credit risk rate. Factor 1 represents the repayment capacity, length of business, working capital, and liquidity value, factor 2 represents the age and the difference between the loan amount granted and proposed, factor 3 represents the variable length of stay, character, and credit history, and one variable represents the number of dependents. The estimation of the credit risk rate generated in the case study also appears reasonable. The analysis framework of the credit risk rate designed in this study is expected to facilitate risk analysts in measuring the credit risk rate of debtors. This analytical framework allows analysts to use credit assessment variables without reducing them. In addition, analysts can obtain an estimate of the credit risk rate from prospective debtors more strictly based on the standard rules they have designed.

Funding

This research was funded by Universitas Padjadjaran through Riset Data Pustaka dan Daring. Number: 2203/UN6.3.1/PT.00/2022.

Acknowledgments

The author thanks the Referees for the useful suggestions and comments. The paper was improved greatly because of them. All remaining errors are mine.

References

Adriyendi. (2018). Fuzzy Logic using Tsukamoto Model and Sugeno Model on Prediction Cost. *International Journal of Intelligent Systems and Applications*, *10*(6), 13–21. https://doi.org/10.5815/ijisa.2018.06.02

Akhirina, T. Y., Rusmardiana, A., Yulistyanti, D., & Pauziah, U. (2019). The Comparison of K-Nearest Neighbor (K-NN) Algorithm and Fuzzy Tsukamoto Logic in the Determination of SMA Students Majors in Banten. *Journal of Physics: Conference Series*, *1175*(1), 012068. https://doi.org/10.1088/1742-6596/1175/1/012068

Ardika, B. S., Setianingrum, A. H., & Hakiem, N. (2017). Funding eligibility decision support system using fuzzy logic Tsukamoto: (Case: BMT XYZ). *2017 Second International Conference on Informatics and Computing (ICIC)*, 1–7. https://doi.org/10.1109/IAC.2017.8280622

Bandyopadhyay, S., Mistri, H., Chattopadhyay, P., & Maji, B. (2013). Antenna array synthesis by implementing non-uniform amplitude using Tsukamoto fuzzy logic controller. *2013 International Conference on Advanced Electronic Systems (ICAES)*, 19–23. https://doi.org/10.1109/ICAES.2013.6659353

Bao, W., Lianju, N., & Yue, K. (2019). Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. *Expert Systems with Applications*, *128*, 301–315. https://doi.org/10.1016/j.eswa.2019.02.033

Bartholomew, D. J. (1985). Foundations of factor analysis: Some practical implications. *British Journal of Mathematical and Statistical Psychology*, *38*(1), 1–10. https://doi.org/10.1111/j.2044-8317.1985.tb00811.x

Brown, A. W., Kaiser, K. A., & Allison, D. B. (2018). Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Sciences*, *115*(11), 2563–2570. https://doi.org/10.1073/pnas.1708279115

Corrigan, P. W., Salzer, M., Ralph, R. O., Sangster, Y., & Keck, L. (2004). Examining the Factor Structure of the Recovery Assessment Scale. *Schizophrenia Bulletin*, *30*(4), 1035–1041. https://doi.org/10.1093/oxfordjournals.schbul.a007118

Fajri, D. M. N., Mahmudy, W. F., & Anggodo, Y. P. (2017). Optimization of FIS Tsukamoto using particle swarm optimization for dental disease identification. *2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 261–268. https://doi.org/10.1109/ICACSIS.2017.8355044

Glaser, R. E. (1976). Exact Critical Values for Bartlett's Test for Homogeneity of Variances. *Journal of the American Statistical Association*, *71*(354), 488–490. https://doi.org/10.1080/01621459.1976.10480374

Lenka, S. R., Bisoy, S. K., Priyadarshini, R., Hota, J., & Barik, R. K. (2021). An Effective Credit Scoring Model Implementation by Optimal Feature Selection Scheme. *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 106–109. https://doi.org/10.1109/ESCI50559.2021.9396911

Liu, H., & Wang, N. (2021). Research on the Present Situation of Professional Identity of Young University Teachers Based on the KMO Sample Suitability Test and Bartlett Spherical Test. *2021 2nd International Conference on Big Data and Informatization Education (ICBDIE)*, 366–369. https://doi.org/10.1109/ICBDIE52740.2021.00089

Llorca, M. (2017). *External debt sustainability and vulnerabilities: evidence from a panel of 24 Asian countries and prospective analysis*. ADBI Working Paper.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach☆. *Journal of Financial Economics*, *83*(1), 171–222. https://doi.org/10.1016/j.jfineco.2005.12.002

McDade, T. W., & Adair, L. S. (2001). Defining the "urban" in urbanization and health: a factor analysis approach. *Social Science & Medicine*, *53*(1), 55–70. https://doi.org/10.1016/S0277-9536(00)00313-0

Mhlanga, D. (2021). Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment. *International Journal of Financial Studies*, *9*(3), 39. https://doi.org/10.3390/ijfs9030039

Miwakeichi, F., Martínez-Montes, E., Valdés-Sosa, P. A., Nishiyama, N., Mizuhara, H., & Yamaguchi, Y. (2004).

Decomposing EEG data into space–time–frequency components using Parallel Factor Analysis. *NeuroImage*, *22*(3), 1035–1045. https://doi.org/10.1016/j.neuroimage.2004.03.039

Mkhaiber, A., & Werner, R. A. (2021). The relationship between bank size and the propensity to lend to small firms: New empirical evidence from a large sample. *Journal of International Money and Finance*, *110*(5), 102281. https://doi.org/10.1016/j.jimonfin.2020.102281

Nugraha, E., Wibawa, A. P., Hakim, M. L., Kholifah, U., Dini, R. H., & Irwanto, M. R. (2019). Implementation of fuzzy tsukamoto method in decision support system of journal acceptance. *Journal of Physics: Conference Series*, *1280*(2), 022031. https://doi.org/10.1088/1742-6596/1280/2/022031

Pan, S., Wei, J., & Pan, H. (2020). Study on Evaluation Model of Chinese P2P Online Lending Platform Based on Hybrid Kernel Support Vector Machine. *Scientific Programming*, *2020*, 1–7. https://doi.org/10.1155/2020/4561834

Saul, L. K., & Rahim, M. G. (2000). Maximum likelihood and minimum classification error factor analysis for automatic speech recognition. *IEEE Transactions on Speech and Audio Processing*, *8*(2), 115–125. https://doi.org/10.1109/89.824696

Setyono, A., & Aeni, S. N. (2018). Development of Decision Support System for Ordering Goods using Fuzzy Tsukamoto. *International Journal of Electrical and Computer Engineering (IJECE)*, *8*(2), 1182. https://doi.org/10.11591/ijece.v8i2.pp1182-1193

Spuchľáková, E., Valašková, K., & Adamko, P. (2015). The Credit Risk and its Measurement, Hedging and Monitoring. *Procedia Economics and Finance*, *24*, 675–681. https://doi.org/10.1016/S2212-5671(15)00671-1

Sudiyatno, S. I., & Wibowo, F. W. (2018). *Fuzzy VRIO and THES based model of university competitive advantage*. unpublished.

Suharjito, Diana, Yulyanto, & Nugroho, A. (2017). Mobile Expert System Using Fuzzy Tsukamoto for Diagnosing Cattle Disease. *Procedia Computer Science*, *116*, 27–36. https://doi.org/10.1016/j.procs.2017.10.005

Velicer, W. F., & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research*, *25*(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Wu, H.-C., Hu, Y.-H., & Huang, Y.-H. (2014). Two-stage credit rating prediction using machine learning techniques. *Kybernetes*, *43*(7), 1098–1113. https://doi.org/10.1108/K-10-2013-0218

Wu, M., Cheng, G., & Gao, J. (2021). Research on the Measurement of Subject Credit Risk of Chinese Port Enterprises by Constrained Logistic Regression. *Asia-Pacific Journal of Operational Research*, *38*(03), 2040016. https://doi.org/10.1142/S0217595920400163

Xiao, J. J., & Wu, G. (2008). Completing debt management plans in credit counseling: An application of the theory of planned behavior. *Journal of Financial Counseling and Planning*, *19*(2).

Yang, S., & Islam, M. T. (2020). Principal component analysis and factor analysis for feature selection in credit rating. *ArXiv Preprint ArXiv:2011.09137*, 09137.

Yangyudongnanxin, G. (2021). Financial Credit Risk Control Strategy Based on Weighted Random Forest Algorithm. *Scientific Programming*, *2021*, 1–9. https://doi.org/10.1155/2021/6276155

Yi, G., Lei, H., & Ziqiang, L. (2015). Port Customer Credit Risk Prediction Based on Internal and External Information Fusion. *The Open Cybernetics & Systemics Journal*, *9*(1), 1323–1328. https://doi.org/10.2174/1874110X01509011323

Yu, L., Wang, S., & Lai, K. K. (2009). An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring. *European Journal of Operational Research*, *195*(3), 942–959. https://doi.org/10.1016/j.ejor.2007.11.025

© 2023 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).