Neural Computing and Applications (2020) 32:9809–9826
https://doi.org/10.1007/s00521-019-04537-7

REVIEW ARTICLE

Machine learning and decision support system on credit scoring

Gernmanno Teles[1] · Joel J. P. C. Rodrigues[1,2,3,4] · Kashif Saleem[5] · Sergei Kozlov[4] · Ricardo A. L. Rabêlo[2]

Received: 8 February 2019 / Accepted: 5 October 2019 / Published online: 24 October 2019

Abstract

Among the numerous alternatives used in the world of risk balance, it highlights the provision of guarantees in the formalization of credit agreements. The objective of this paper is to compare the achievement of fuzzy sets with that of artificial neural network-based decision trees on credit scoring to predict the recovered value using a sample of 1890 borrowers. Comparing with fuzzy logic, the decision analytic approach can more easily present the outcomes of the analysis. On the other hand, fuzzy logic makes some implicit assumptions that may make it even harder for credit-grantors to follow the logical decision-making process. This paper leads an initial study of collateral as a variable in the calculation of the credit scoring. The study concludes that the two models make modelling of uncertainty in the credit scoring process possible. Although more difficult to implement, fuzzy logic is more accurate for modelling the uncertainty. However, the decision tree model is more favourable to the presentation of the problem.

Keywords Machine learning · Decision trees · Fuzzy logic · Credit scoring · Performance evaluation

1 Introduction

Decision-making is an important factor in achieving success in selecting borrowers using large amounts of information and knowledge. Identifying less risky borrowers is crucial in successful credit scoring. The financial sector is not only volatile but also very competitive. The probability of failure to deliver the desired outcomes is high for individual borrowers, and the lenders must take appropriate steps to confront the situation and mitigate the associated risks to achieve favourable debt repayment outcomes [1, 2]. The selection process for the borrowers should lead to the identification of safer borrowers who have the ability and willingness to pay back debt within the set repayment period. Some credit scoring methods currently in existence are criticized for being inadequate in addressing the borrower's ability to repay debt with minimal risk within the given time [3].

Probability models are widely used in the quantification and assessment of risk. These models have become integral in the informed decision-making processes related to uncertainty in many application areas [4–12]. However, such frameworks based on the conventional theory may not be able to describe some uncertainties in the most appropriate manner because imprecise data, insufficient experience datasets, and the complex cause-effect relationships make it difficult to assess the different risk types using only the conventional probability models [13–15]. Of particular concern is the misunderstanding of the cause of the risk and its characteristics [7, 16].

Credit scoring is a collection of many different tasks, processes, and requirements that need to be considered in their entirety [2, 17–23]. Consequently, it is difficult and arduous to make credit lending decisions in such environments [24–33]. Therefore, mechanisms that help to characterize such complex scenarios are needed. The literature suggests that the application of multiple criteria for decision-making can facilitate the resolution of these issues [9]. Other models, such as fuzzy logic, artificial neural networks, and decision tree models explicitly consider the

✉ Gernmanno Teles
germanno.teles@ubi.pt

1 Instituto de Telecomunicações, Universidade da Beira Interior, 6201-001 Covilhã, Portugal

2 Federal University of Piauí, Teresina-PI, Brazil

3 College of Computer and Information Sciences (CCIS), King Saud University (KSU), Riyadh 12372, Saudi Arabia

4 ITMO University, St. Petersburg 197101, Russia

5 Center of Excellence in Information Assurance (CoEIA), King Saud University, Riyadh 11653, Saudi Arabia

underlying relationships and recognize the uncertainties (such as operational risks). In addition to the multi-criteria methods, additional complementary tools such as fuzzy sets or numerical simulations are increasingly being used in the credit scoring process. These tools have been applied in lender decision-making as far as dealing with the aspects of financial risks are concerned. The capability of these tools to deal with uncertainty enables the credit grantor to deal with the issues in a manner that conventional methods would not [6, 34]. The advantage of our approach is that it uses features that are independent to use collateral as a parameter for calculating risk.

The concept of Fuzzy sets also known as the possibility theory was first introduced by Lukasiewicz (cited in Rescher [35]. Later, Zadeh [36] extended the work on this theory to identify Fuzzy sets as a mathematical model for the characterization and quantification of uncertainty [37]. Since then, the fuzzy set theory has been used as a method for modelling of risk in many applications, including credit scoring in the financial sector [32]. The usefulness of this model is such that it can be used to characterize and quantify uncertainty using fuzzy sets in the absence of sufficient data to estimate uncertainty through the conventional statistical estimation of frequencies [38]. The basis of the fuzzy set theory is on a group of elements or data that share some common characteristics within their memberships [39, 40]. As sufficient data on borrowers may not be readily available to the lender, the application of the fuzzy set theory can play an essential role in the quantification of uncertainty in borrower selection decisions.

Then, this work aims to investigate study of collateral as a variable in the calculation of the credit scoring applied to systems that are using credit operations. The main contribution of this work includes a compare the performance of fuzzy sets with that of artificial neural network-based decision trees on credit scoring to predict the recovered value using a sample of 1890 borrowers.

The paper is organized as follows. Section 2 presents theoretical background, introducing the related works and discussing used fuzzy methods and decision trees procedures on classifying credit scoring approach. Section 3 presents results analysis and shows which model is superior in terms of accuracy and which is superior in terms of fitting the data correctly. Section 4 performs the discussion about the issues of fuzzy and decision tree models bring forward a comparison of those two models. Finally, Section 5 provides the conclusion and suggestions for further works.

2 Theoretical background

In the decision trees technique, a set of rules presented as a tree are used to make decisions [41]. For example, one can build a classification tree for credit risk based on a person's income, age, among other parameters. The benefit of this model is that, unlike most of the credibility models, the tree expresses the reasoning process behind the framework. The usual algorithm for decision tree building includes classifying sets under a root node by assigning all the training data to the best splitting attribute (see Fig. 1).

The primary objective of a decision tree is to divide a group of data into fewer portions. On a qualitative data used to build a person's income, for instance, the root of the tree asks if Income < 92.5. If the answer is "yes", it move to the next child, and it pass to the left child of the node; if "no", it moves to the right. Proceeding in this way, it finally succeeds at a final node.

The data comprehend a piece of information on the credit history of customers of a lender. The population consists of existing customer accounts of the lender. A sample of customers is chosen randomly from the data available on their performance. The main characteristics of the customer information are the recovered and whether the value rates are overdue. The data used in the analysis are obtained from the lender's database.

The underlying analysis techniques include fuzzy logic- and artificial neural network-based decision trees to predict credit recovery using a real dataset of 1890 records from a bank in Brazil.

2.1 Credit scoring using fuzzy set theory and fuzzy logic

Definition 1 Suppose that (X, Y, Z) is the target system.

X is a limited set object;

$$X = (x_1 + x_2 \ldots, x_n) \quad (1)$$

Y is the object attribute set

$$Y = (y_1 + y_2 \ldots y_n) \quad (2)$$

Z is the target set

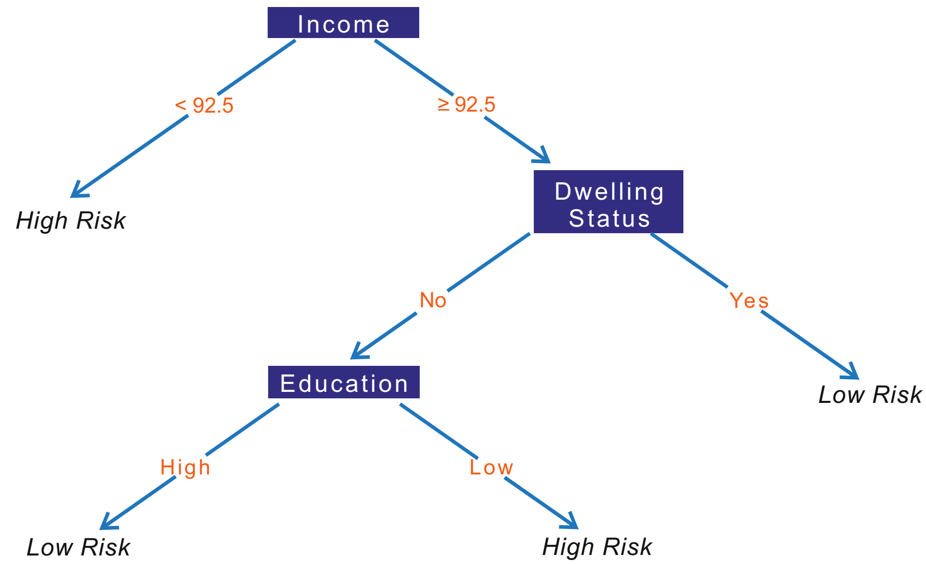
$$Z = (z_1 + z_2 \ldots z_n) \quad (3)$$

and the fuzzy target set is Z_i.

Fuzzy logic is applied because y_j and the degree of effect of y_j on target Z_i in the target system are uncertain.

Definition 2 "Generated real objected states target set is defined as target state set z where one or more than one objected attributes are inputted to the system. And the objected states from the math model of the system are defined as expected states" z^0 [10, 31, 42–44].

Fig. 1 Example of a decision tree model



The fuzzy relation is

$$R : Y \times Z \rightarrow [0, 1] \tag{4}$$

If there is $(y, z) \in Y \times Z, R,\ (y,\ z)$ is the degree of relation from attribute y to state z. R is the binary relation from set Y to set Z.

Fuzzy set on Y and fuzzy set on Z are denoted by A and B, respectively. Then R can be expressed as implication relations;

$$R = A \rightarrow B \tag{5}$$

Then,

$$R(y, z) = (A \rightarrow B)(y, z) \tag{6}$$

And the algorithm regulated,

$$(A \rightarrow B)(y, z) = (1 - A(y)) \curlyvee B(z) \tag{7}$$

Definition 3 Suppose $\lambda \in \left[\frac{1}{2}, 1\right]$

(i) if $\exists(y, z)$ makes $R(y, z) \geq \lambda$, the effect relation R to $(y,\ z)$ is regarded as fuzzy λ true, or as fuzzy λ false.

(ii) if $\forall(y, z)$ makes $R(y, z) \geq \lambda$, the effect relation R is regarded as fuzzy λ true, or as fuzzy λ false.

To calculate the effect characteristic of intersection values y,

$$y = \bigcap_{i=1}^{n} U_i \tag{8}$$

Since Eq. (6) and $\forall U \in E$, membership of fuzzy set A on Y is,

$$Y(y_j) = \begin{cases} 1, & y_j \in U \\ 0, & y_j x \notin U \end{cases} \quad (j = 1, 2\ldots, m) \tag{9}$$

According to formula (7),

$$\begin{aligned} &(A_i \cap A_j \rightarrow B_i \cup B_j)(y, z) = [1 - (A_i \cap A_j)(y)] \curlyvee [(B_i \cup B_j)(y)] \\ &= [(1 - A_i(y)) \curlyvee B_i(z)] \curlyvee [(1 - A_j(y)) \curlyvee B_j(z)] \\ &= (A_i \rightarrow B_i)(y, z) \curlyvee (A_j \rightarrow B_j)(y, z) \end{aligned} \tag{10}$$

Therefore, $\forall \lambda \in \left[\frac{1}{2}, 1\right]$,

$$\begin{aligned} &(A_i \rightarrow B_i)(u, p) \geq \lambda (A_j \rightarrow B_j)(u, p) \geq \lambda \\ &\quad \Leftrightarrow (A_i \cap A_j \rightarrow B_i \cup B_j)(u, p) \geq \lambda \end{aligned}$$

2.2 Fuzzy data analysis procedure

Case–control matching is a traditional procedure used to join records in the "case" representation with related records in a typically much larger "control" example based on a set of essential variables. So to explains the fuzzy extension command for Statistical Package for the Social Sciences (SPSS) that implements this method and some recent improvements to it, which allows the input of a custom function, is used. First, the data is reduced down to only the variables used [38, 42, 45–48]. It was used on this dataset to reduce the select bias and improve the internal validity.

The Polytomous Universal Model (PLUM), an extension of the general linear model to ordinal categorical data, was used to fit the logistic model predicting the probability of the treatment. It uses contract value, collateral value, main value delay, the balance value, tax rate value, tax interest value, client size, seniority level, per cent used, duration in years, duration in days, and delay in days as predictors of the recovered value.

```
*Fitting logit model via PLUM.
PLUM HalfwayHouse WITH NonViol SidewalkCafe TypeC_D
 /CRITERIA=CIN(95) DELTA(0) LCONVERGE(0) MXITER(100) MXSTEP(5) PCONVERGE(1.0E-6)
SINGULAR(1.0E-8)
 /LINK=LOGIT
 /PRINT=FIT PARAMETER SUMMARY
 /SAVE=ESTPROB.
```

The model is not good but as can be seen, the balance value, tax rate value, tax interest value, client size, seniority level, per cent used, and duration are not associated with credit repayment. Now a custom function with which to restrict matches based on the probability of the treatment and period in months is created. In this case, another file is made in python and named PerFun.py in which the following functions are placed:

```
#These functions are for SPSS's fuzzy case control matching
import math
#period under 12, and caliper within 0.02
def PerFun(demander,supplier):
  dy = math.pow(demander[1] - supplier[1],2)
  dz = math.pow(demander [2] - supplier [2],2)
  period = math.sqrt(dy + dz)
  p = abs(d[0] - s[0]) #diference in month
  if per < 12 and p < 0.02:
    t = 1 #for credit default
  else:
    t = 0
  return t
#period over 12, but under 24
def PerBuf(demander,supplier):
  dx = math.pow(demander[1] - supplier[1],2)
  dy = math.pow(demander[2] - supplier[2],2)
  period = math.sqrt(dx + dy)
  p = abs(d[0] - s[0]) #diference in month
  if period  12 and p < 0.02:
    t = 1 # for credit default
  else:
    t = 0
  return t
```

The fuzzy logic algorithm above returns either the value of 1 for credit default or 0 otherwise [2, 9, 32, 39, 46, 49, 50]. Also, the algorithm takes only a fixed set of vectors from the dataset. The first two elements of the first function PerFun account for the period of repayment, and the last element is the probability of treatment. Next, the function returns the euclidean distance [12, 38]. If the period is under 12 months, it returns the value of 0. The second function sets the boundaries of the period under consideration with values not too far away from 12 to 24 months of repayment default.

The program code is then run in Python as follows.

```
import PerFun

#test case
x = [0,0,0.02]
y = [0,23,0.02]
z = [0,24,0.02]
print PerFun.PerFun(x,y)
print PerFun.PerFun(x,z)
END PROGRAM.
```

To this end, a custom function in order to use the fuzzy command is made. The custom function helps one to do more complicated functions such as the buffer function, which takes the probability of the treatment along with the two spatial points of the period variable. The custom function returns the values for all the predictor variables. To conduct the credit score analysis involves a little more data involvement. The cases and controls of the second data of the just matched credit defaulters are reshaped in the long format before merging with the original.

```
*custom function

FUZZY BY=EST2_1 XMonths YMonths ID=PerID IDVARS=Match1 Match2 Match3 GROUP=Period
CUSTOMFUZZ = "PerFun.PerFun"
  EXACTPRIORITY=FALSE
MATCHGROUPVAR=PGroup
/OPTIONS REPLACEMENT=FALSE SHUFFLE=TRUE MINIMIZEMEMORY=TRUE SEED=10.
```

```
*Reshaping and merging the data back as well as the outcome analysis.

DATASET COPY PerMatch.
DATASET ACTIVATE PerMatch.
SELECT IF Period = 1.
CASES /MAKE PerID FROM PerID Match1 Match2 Match3
        /INDEX Type
        /KEEP PGroup.
```

```
* merging with the original.
SORT CASES BY PerID.
MATCH FILES FILE = *
 /TABLE = 'Period_Data'
 /BY PerID.

* merging with the original
SORT CASES BY PerID.
MATCH FILES FILE = *
 /TABLE = 'Period_Data'
 /BY PerID.
*Analysis
T-TEST GROUPS=Period(0 1)
 /MISSING=ANALYSIS
 /VARIABLES=Viol
 /CRITERIA=CI(.95).
```

2.3 Decision trees procedure

The customer data of a lending institution are imported into RapidMiner Software to produce a decision tree for addressing the current problem. RapidMiner is a data science software program built by the organization of the same name that provides an integrated environment for data preparation, machine learning, deep learning, text mining, and predictive analytics [51]. The steps for applying decision trees for credit solutions include;

1. Importing the data from a spreadsheet,
2. Splitting data into training and testing samples,
3. Training the decision tree, and
4. Applying the model to the testing samples and evaluating the performance.

Importing the data is easy using RapidMiner's built-in import operator to automatically load the data from a spreadsheet into the software interface (Fig. 2).

As with all data modelling techniques, the dataset must be split into training and testing sets to ensure that the model works well on both sets [11, 13, 50]. The standard practice is to split the available data into a training set and a validation set [46, 52, 53]. "Typically 70–80% of the original data of the dataset is split into the training set and the remainder is set aside for validation" [26, 30, 54]. So, in this step, stratified sampling is chosen with a split ratio of 0.7 (70% training) to ensure that all variables have similar distributions of class values. Finally, the XL operator output is connected to the Split Validation operator input (the process causes some errors to appear because the process is still incomplete) (Fig. 3).

As discussed earlier, five simple steps are used to create decision trees by developing information contained in the reduced data. The reduced data reduces the uncertainties contained in the data before splitting thereby increasing information through classification. So, the greatest increase in information is achieved by reducing data as much as possible. The three main options available in RapidMiner Studio for decision tree splitting are.

1. *Information gain* is calculated as the difference between the information before and after splitting the data. The problem with this parameter arises when the variables are very few with a large number of classes. In this case, the variables have the tendency of becoming root nodes and it becomes a challenge when the case is extreme.
2. *Gain ratio* This is a better parameter than the information gain whereby it overcomes the problem with information gain because it considers the probable number of branches before splitting the data.
3. *Gini index* provides supplementary information to the gain ratio.

However, there are other parameters such as the "minimal gain" value whose default value is 0.1 but it can be anything from 0 upwards.

The size of the dataset also determines the other decision tree parameters such as the "minimal size for a split", "maximal depth", and "minimal leaf size" indices.

Finally, the training ports and the model ports are joined together (Fig. 4).

The last step involves checking the model validity using performance operators such as accuracy, precision, recall, receiver operating characteristic (ROC) and area under curve (AUC) charts. The common methods of evaluating the performance of classification models include confusion matrices and Gain and Lift charts. A confusion matrix is a table that is able to compare the model predicted and actual classes from the labelled data within the validation set. The main evaluation criteria in the confusion matrix are accuracy, sensitivity, and specificity. Whereas accuracy is the indicator of overall model effectiveness, sensitivity measures the rate of true positives and specificity the rate of false positives (Fig. 5).

3 Results analysis

The two models explored in this study show that both can be applied effectively to credit scoring. Decision trees can be used for the classification where it has discrete data (continuous or nominal values). In the case where both the input attributes and the output decision are continuous, decision trees can be used to generate fuzzy rules. For

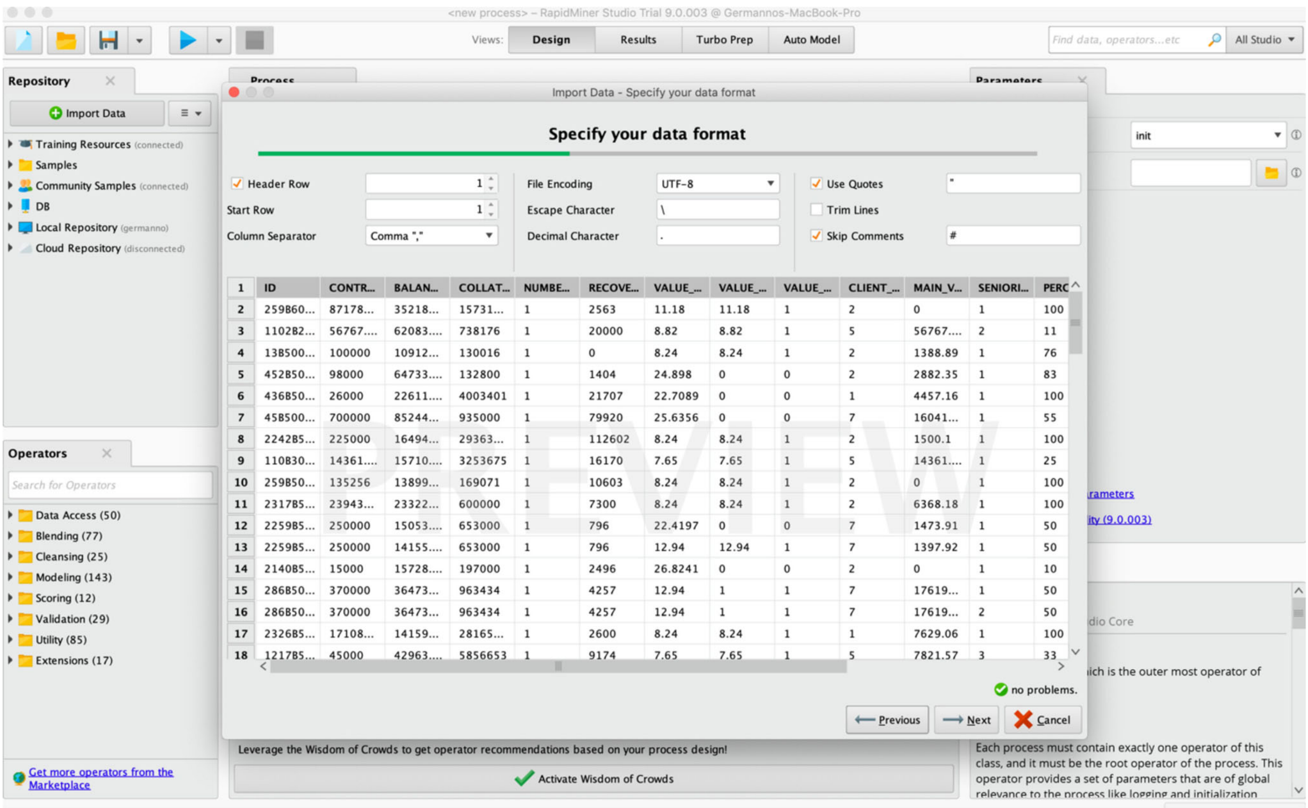


Fig. 2 Step 1: showing how to import external data

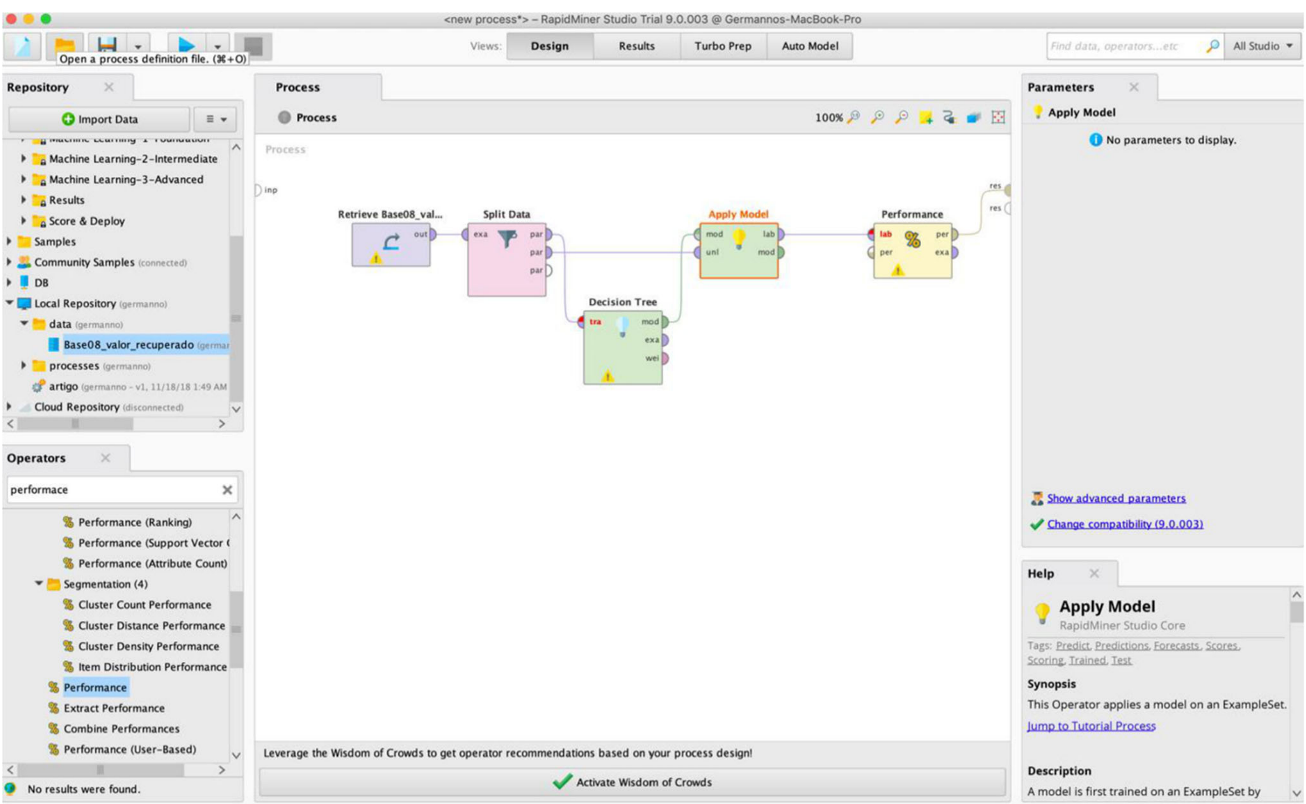


Fig. 3 Step 2: splitting and validating data

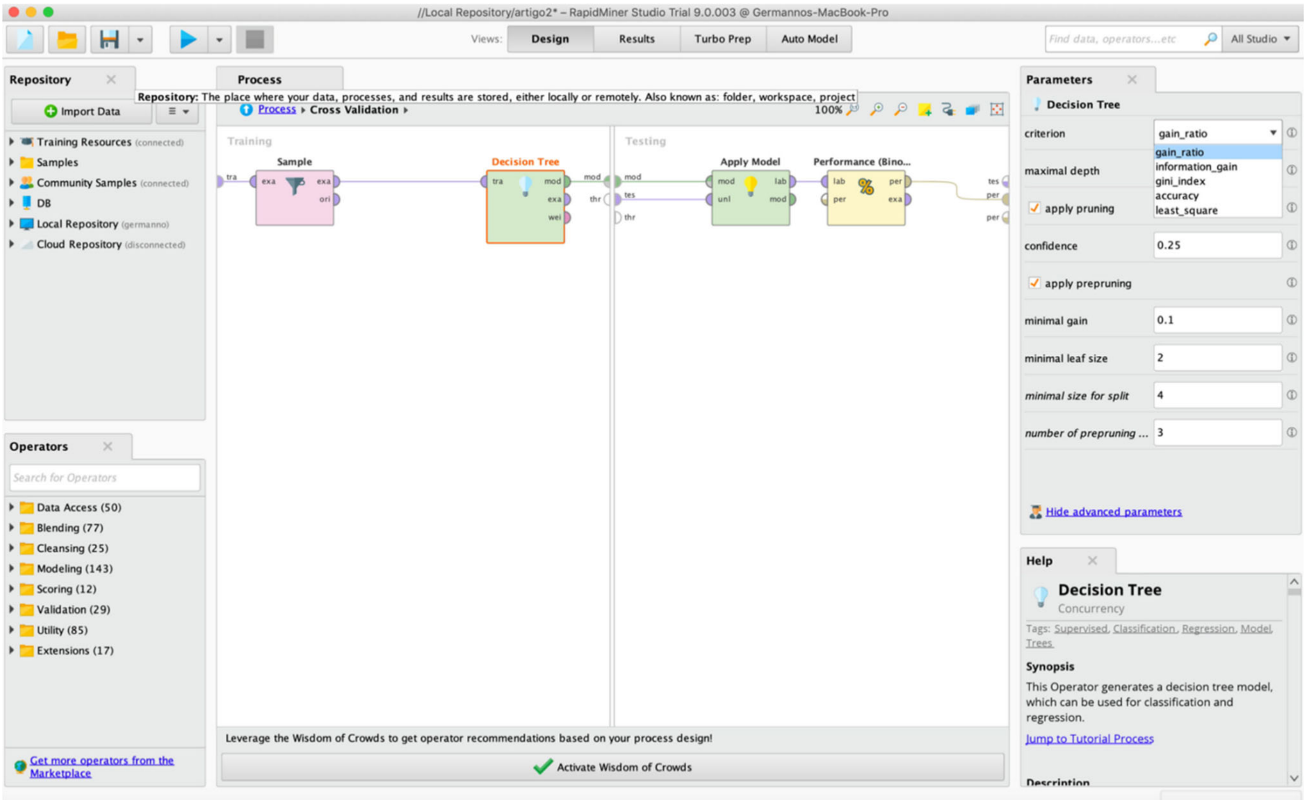


Fig. 4 Step 3: setting up the decision tree parameters

example, if the sets provide output values that provide at most 2 sets, each training example will belong to 1 or 2 output classes or fuzzy sets.

Table 1 includes the Chi-Square goodness of fit test of 392.741 on 1 df (degrees of freedom), significant beyond .001. Therefore, the null hypothesis that adding the variables to the model has not significantly increased our ability to predict defaulters is rejected. Table 2 is the model summary that includes the Pseudo-R^2. It sees that the -2 log likelihood statistic is .000. The -2 log likelihood measures how poorly the model predicts the decisions whereby the smaller the statistic, the better the model. Adding the variables reduced the -2 log likelihood statistic by $392.741 - .000 = 392.741$. As such, the result shows that the model is superior in terms of accuracy great. It can also see that Nagelkerke's R^2 is 1.00, which indicates that the model is superior in terms of fits the data perfectly. Cox and Snell's R^2 is .26, which can be explained as 26% probability of the prediction is explained by the logistic model.

Table 3 illustrates "The cut value is .500" whereby the likelihood of a case being classified into the "default" category is greater than .500; otherwise, the case is classified as in the "no default" category. The classification tables show the percentage accuracy in classification (PAC) of cases correctly classified as "no default" with the independent variables added. It also indicates the sensitivity, which is the percentage of cases that had the observed characteristic that was correctly predicted by the model. It also includes the specificity, which is the percentage of cases that did not have the marked characteristic and were also accurately predicted as not having the observed feature. Table 3 shows that 1278/1278 = 100% of the subjects in default were observed. In other words, the sensitivity of the prediction was 100%. It also sees that 45/45 = 100% were correctly classified into the "no default category." Thus, the specificity of prediction of non-occurrences correctly predicted was 100%.

Table 4 shows the contribution of each predictor variable to the model and its statistical significance. The results show that all the independent variables did not add significantly to the model ($p > .05$). Although insignificant, the largest effect size was due to the number of collateral ($b = -24.53$, $p = 1.00$), followed by duration in months ($b = -13.56$, $p = 1.00$), value tax interest rate ($b = 6.56$, $p = 1.00$), and duration in years ($b = 2.71$, $p = 1.00$). The results show that contract value, balance value, recovered value, main value delay, and collateral values did not have any effect at all, while DELAY in days had a very minimal effect.

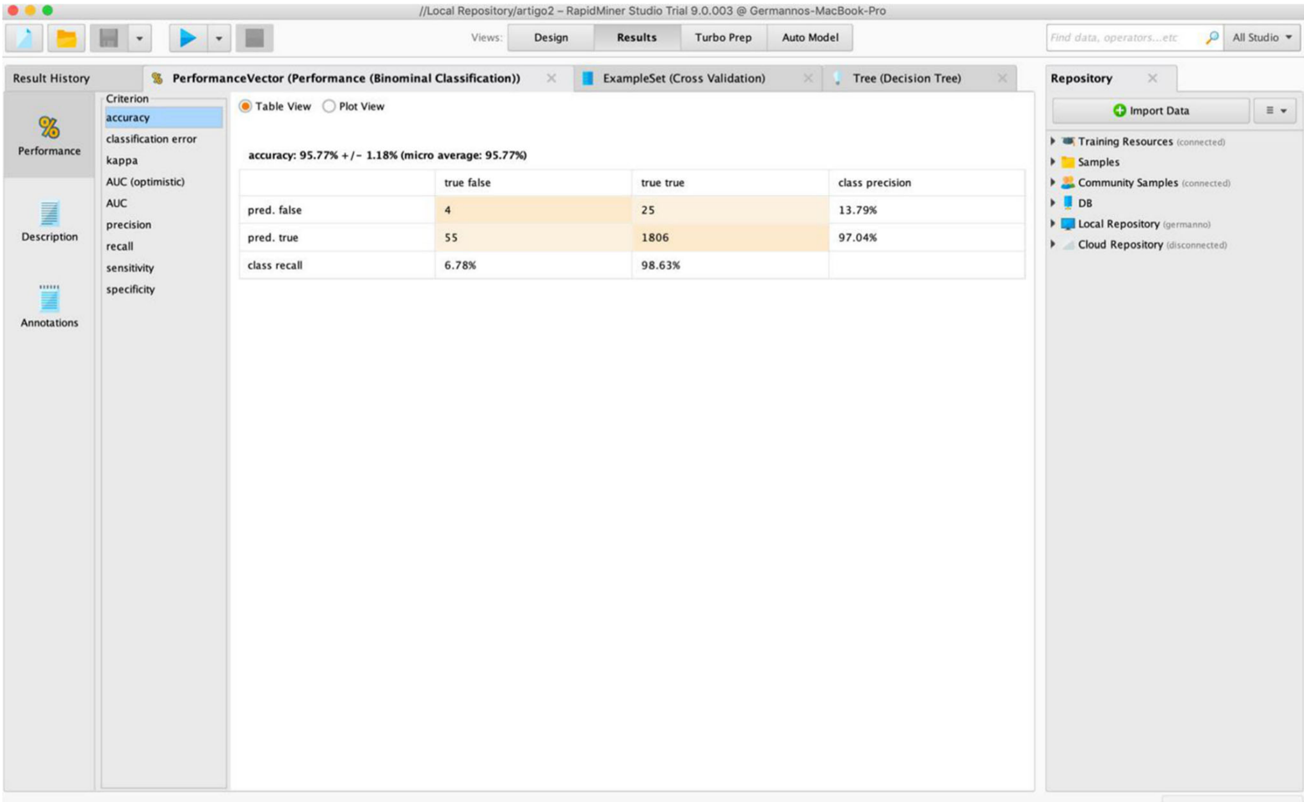


Fig. 5 Step 4: evaluating the performance

Table 1 Omnibus tests of model coefficients

	Chi square	*df*	Sig.
Step 1			
Step	392.741	14	.000
Block	392.741	14	.000
Model	392.741	14	.000

Table 2 Model summary statistic functions

Step	− 2 log likelihood	Cox and Snell R^2	Nagelkerke R^2
1	.000[a]	.257	1.000

[a]Estimation terminated at iteration number 20 because maximum iterations have been reached. The final solution cannot be found

Table 3 Classification table of percentage accuracy in classification (PAC)[a]

Observed	Predicted		
	OVERDUE		Percentage correct
	No default	In default	
Step 1			
OVERDUE			
No default	45	0	100.0
In default	0	1278	100.0
Overall percentage			100.0

[a]The cut value is .500

Table 5 shows bootstrapping results that descend robust approximations of regular errors and confidence periods for the regression coefficients. It reveals that value tax rate and value tax interest rate significantly added to the model ($p < .05$). In other words, number of collateral ($b = -24.53$, $p = .05$), value tax rate ($b = -2.34$, $p = .02$) and value tax interest rate ($b = 6.56$, $p = .02$) play an influential role in determining whether the clients will default on their repayments. Again, contract value, balance value, recovered value, main value delay, and collateral values, did not have any effect at all, while delay in days had a very minimal effect.

Tables 6, 7, 8, and 9 present results of the expert dataset results, which show the results are more or less the same as those of the training dataset. Table 6 shows that 1837/1837 = 100% of the subjects "in default" were observed. The overall percentage of the correct subjects that were identified by the model in the expert dataset or the sensitivity of the prediction was 97.2%. Comparatively,

Table 4 Variables statistical significance contribution

	B	SE	Wald	df	Sig.	Exp (B)	95% CI for EXP (B)	
							Lower	Upper
Step 1[a,b]								
CONTRACT_VALUE1	.000	.003	.000	1	.988	1.000	.993	1.007
BALANCE_VALUE1	.000	.004	.000	1	.986	1.000	.993	1.007
COLLATERAL_VALUE1	.000	.000	.000	1	.996	1.000	1.000	1.000
NUMBER_OF_COLLATERAL1	− 24.529	25,929.218	.000	1	.999	.000	.000	–
RECOVERED_VALUE1	.000	.002	.000	1	.995	1.000	.996	1.004
VALUE_TX_RATE1	− 2.337	253.262	.000	1	.993	.097	.000	3.649E + 214
VALUE_TX_INTEREST_RATE1	6.557	587.232	.000	1	.991	704.368	.000	–
VALUE_RATE_OVERDUE1	2.644	359.890	.000	1	.994	14.069	.000	3.068E + 307
CLIENT_SIZE1	.930	965.552	.000	1	.999	2.534	.000	–
MAIN_VALUE_DELAY1	.000	.011	.000	1	.985	1.000	.979	1.021
SENIORITY_LEVEL1	− .716	1316.866	.000	1	1.000	.489	.000	–
PERCENT_USED1	− .116	27.920	.000	1	.997	.891	.000	5.188E + 308
DURATION_IN_YEARS1	2.712	4453.449	.000	1	1.000	15.054	.000	–
DURATION_IN_MONTHS1	− 13.557	3412.583	.000	1	.997	.000	.000	–
DURATION_IN_DAYS1	.451	119.137	.000	1	.997	1.569	.000	4.031E + 101
DELAY_IN_DAYS1	− .019	2.233	.000	1	.993	.981	.012	78.131
Constant	23.385	27,376.297	.000	1	.999	14,325,648,374.858		

[a]Variable(s) entered on step 1: CONTRACT VALUE, BALANCE VALUE, COLLATERAL VALUE, NUMBER OF COLLATERAL, RECOVERED VALUE, VALUE TX RATE, VALUE TX INTEREST RATE, VALUE RATE OVERDUE, CLIENT SIZE, MAIN VALUE DELAY, SENIORITY LEVEL, PERCENT USED, DURATION IN YEARS, DURATION IN MONTHS, DURATION IN DAYS, DELAY IN DAYS

[b]Variable(s) entered on step 1: CONTRACT VALUE, BALANCE VALUE, COLLATERAL VALUE, NUMBER OF COLLATERAL, RECOVERED VALUE, VALUE TX RATE, VALUE TX INTEREST RATE, VALUE RATE OVERDUE, CLIENT SIZE, MAIN VALUE DELAY, SENIORITY LEVEL, PERCENT USED, DURATION IN YEARS, DURATION IN MONTHS, DURATION IN DAYS, DELAY IN DAYS

this model predicted 97.2% of the expert dataset correctly as compared with 100% in the training dataset (see Table 3). Table 7 shows the contribution of each predictor variable to the model and its statistical significance in the expert dataset. Like in the training dataset (i.e., Table 3), the results show that all the independent variables did not add significantly to the model ($p > .05$). Although insignificant, the largest effect size was due to the number of collateral ($b = -34.79$, $p = 1.00$), followed by duration in years ($b = -14.97$, $p = 1.00$), per cent used ($b = 12.68$, $p = 1.00$), and value tax interest rate ($b = 11.65$, $p = 1.00$). The results show that contract value, balance value, recovered value, main value delay, client size, and collateral values did not have any effect at whatsoever while the rest had negligible effects. Tables 8 and 9 show the correlation of the various variables in the training set and the expert dataset with the recovered value. In Table 8, contract value, value tax interest rate, client size, duration in years, and per cent used were all negatively correlated with recovered value. All the other variables had a positive correlation with the recovered value. Similarly, in Table 9, contract value, value tax interest rate, client size, duration in years, and per cent used were all negatively correlated with recovered value. All the other variables had a positive correlation with the recovered value.

Table 10 shows the Kolmogorov–Smirnov test results for the various independent-samples tests. The nonparametric tests show that the contract value is not the same across the categories of decisions ($p > .05$). Therefore, the null hypothesis is accepted. The same is observed for the number of collateral ($p > .05$) and collateral value ($p > .05$). There are significant differences in the means of all the other variables, in which case the null hypotheses are rejected and the alternative hypotheses accepted.

Table 11 shows the results of the decision approach. The decision tree results in Table 11 show that the model correctly predicted 97.4% of defaulters. This is slightly higher than what was predicted by the previous model that correctly predicted 97.2% of defaulters (see Table 6).

Table 5 Bootstrap for variables with standard errors and confidence intervals for the regression coefficients[a]

	B	Bootstrap				
		Bias	SE	Sig. (2-tailed)	95% confidence interval	
					Lower	Upper
CONTRACT VALUE	.000	.000[b]	.000[b]	.071[b]	.000[b]	.000[b]
BALANCE VALUE	.000	.000[b]	.000[b]	.059[b]	.000[b]	.000[b]
COLLATERAL VALUE	.000	.000[b]	.000[b]	.099[b]	.000[b]	.000[b]
NUMBER_OF COLLATERAL	− 24.529	5.933[b]	11.456[b]	.053[b]	− 41.839[b]	− 1.944[b]
RECOVERED VALUE	.000	.000[b]	.000[b]	.109[b]	.000[b]	.000[b]
VALUE TX RATE	− 2.337	− .159[b]	1.135[b]	.016[b]	− 5.631[b]	− 1.090[b]
VALUE TX INTEREST RATE	6.557	− 1.083[b]	.953[b]	.016[b]	3.764[b]	7.455[b]
VALUE RATE OVERDUE	2.644	− 1.145[b]	.674[b]	.245[b]	.315[b]	2.764[b]
CLIENT SIZE	.930	.369[b]	1.930[b]	.703[b]	− 3.101[b]	4.680[b]
MAIN VALUE DELAY	.000	.000[b]	.000[b]	.059[b]	.000[b]	.000[b]
SENIORITY LEVEL	− .716	− .948[b]	1.598[b]	.712[b]	− 5.223[b]	1.266[b]
PERCENT USED	− .116	.032[b]	.079[b]	.156[b]	− .243[b]	.069[b]
DURATION IN YEARS	2.712	− .029[b]	6.348[b]	.687[b]	− 10.105[b]	17.086[b]
DURATION IN MONTHS	− 13.557	7.428[b]	7.107[b]	.400[b]	− 22.353[b]	6.407[b]
DURATION IN DAYS	.451	− .250[b]	.230[b]	.447[b]	− .187[b]	.734[b]
DELAY IN DAYS	− .019	.009[b]	.006[b]	.393[b]	− .021[b]	.000[b]
Constant	23.385	− 1.731[b]	15.498[b]	.128[b]	− 5.105[b]	52.794[b]

[a]Unless otherwise noted, bootstrap results are based on 1223 bootstrap samples

[b]Based on 1035 samples

Table 6 Classification table of fuzzy model[a]

Observed	Predicted		
	OVERDUE		Percentage correct
	No default	In default	
Step 1			
OVERDUE			
No default	0	53	.0
In default	0	1837	100.0
Overall percentage			97.2

[a]The cut value is .500

4 Discussion

4.1 Issues with decision trees in credit scoring

Decision trees clearly have many advantages. The experience with decision tree modelling in this study shows that it is an excellent tool for representing graphics decision alternatives efficiently, i.e., visual representation of complex alternatives can be expressed quickly and clearly. The visual appearance is particularly important in understanding subsequent decisions and outcome dependencies. Therefore, the model can be used to compare the relationships between the changing input values and the various decision alternatives. Furthermore, it appears that decision tree models can play a complementary role to other scoring tools. For example, decision can be used to easily explain complex alternatives to non-professional users in ways that fuzzy logic cannot. Not only do decision trees are capable of representing any discrete-value classifier but also are capable of handling datasets with errors and missing values. However, three disadvantages were encountered notably; the algorithm requires that the target attribute has only discrete values, poor performance with the presence of more complex interactions, and over-sensitivity to the training set, irrelevant attributes, and to noise. Poor performance is mostly associated with the need to redraw the tree every time the model is updated, i.e., data classified by an already-trained Tree is added to the Tree as a training data point. Unlike in most other supervised learning algorithms, the addition of training instances is not incremental. Put in other words, training of decision trees cannot occur online, but rather only in batch mode. This limitation became obvious when the classifier was updated. This is significant because for other supervised learning algorithms, for example, they begin classifying data once they

Table 7 Contribution of each predictor variable

	B	SE	Wald	*df*	Sig.	Exp (*B*)	95% CI for EXP (*B*)	
							Lower	Upper
Step 1[a]								
CONTRACT VALUE	.000	.003	.001	1	.978	1.000	.994	1.006
BALANCE VALUE	.000	.003	.001	1	.977	1.000	.994	1.006
COLLATERAL VALUE	.000	.000	.000	1	.998	1.000	1.000	1.000
NUMBER_OF COLLATERAL	− 34.789	15,751.026	.000	1	.998	.000	.000	–
RECOVERED VALUE	.000	.002	.000	1	.991	1.000	.997	1.003
VALUE TX RATE	− 3.641	137.979	.001	1	.979	.026	.000	7.357E + 115
VALUE TX INTEREST RATE	11.645	362.399	.001	1	.974	114,119.379	.000	–
VALUE RATE OVERDUE	.070	323.855	.000	1	1.000	1.072	.000	4.974E + 275
CLIENT SIZE	.000	.008	.001	1	.972	1.000	.986	1.015
MAIN VALUE DELAY	− .001	605.169	.000	1	1.000	.999	.000	–
SENIORITY LEVEL	− .023	14.433	.000	1	.999	.977	.000	1.885E +12
PERCENT USED	12.683	1642.600	.000	1	.994	322,122.914	.000	–
DURATION IN YEARS	− 14.970	1813.071	.000	1	.993	.000	.000	–
DURATION IN MONTHS	.468	55.962	.000	1	.993	1.597	.000	6.892E + 47
DURATION IN DAYS	− .012	1.110	.000	1	.992	.988	.112	8.708
Constant	36.288	15,941.435	.000	1	.998	5.752E + 15		

[a]Variable(s) entered on step 1: CONTRACT VALUE, BALANCE VALUE, COLLATERAL VALUE, NUMBER OF COLLATERAL, RECOVERED VALUE, VALUE TX RATE, VALUE TX INTEREST RATE, VALUE RATE OVERDUE, CLIENT SIZE, MAIN VALUE DELAY, SENIORITY LEVEL, PERCENT USED, DURATION IN YEARS, DURATION IN MONTHS, DURATION IN DAYS, DELAY IN DAYS

have been trained. The data can also be used to prepare the already-trained classifier. However, the entire dataset needs to be trained with decision trees, i.e., the original dataset plus new instances need to be retrained.

As the data is classified stepwise one node a time until the terminal node, only two possibilities are possible (left–right) with decision trees. Therefore, there are other relationships between some variables that the decision trees just are not able to learn. This is significant in the sense that in the logistic regression, for example, it is able to see the individual impacts of the variables in the model which is impossible with the decision trees.

A practical limitation that was encountered with the use of decision tree modelling technique is that it is not possible to use it in regression mode most probably because it is predominantly used for prediction of discrete outcomes.

4.2 Issues with fuzzy logic in credit scoring

This study finds fuzzy logic as the most convenient method for risk analyses. However, it was important to gain a comprehensive understanding of the system processes in order to conduct credit scoring using fuzzy logic. This included identifying the sources of defaulting correctly and consistently as well as identifying the input data for credit scoring. The most important feature of the credit scoring based on fuzzy logic principles is that the entire process leads to the creation of mechanisms that can effectively reduce the risk of default. Because of the analysis's exact output and mechanisms to ameliorate the risks, it can repeat the scoring process on a regular basis with a valuable output.

Because of the application of quantitative data in the fuzzy logic algorithms and methods, subjectivity is reduced to acceptable levels. Therefore, it can better control the process of creating relationships and dependencies between input data and credit scoring. However, this should not be misconstrued to imply that subjectivity is eliminated entirely from the process of risk analysis.

Risk analysis basically implies an assessment of risk, which is an activity associated with the measurement of the strength of the overall system. Consequently, it provides the information required for planned improvement of the company's operation based on the information obtained from credit scoring. One of the benefits of using fuzzy logic in credit scoring is that the whole system is very flexible [8]. Although every situation that can be solved by fuzzy logic can be solved using other methods, fuzzy logic is the most efficient of all. Unlike in decision trees, the modification of the fuzzy logic system requires only adding some

Table 8 Results of negative correlation with recovered value

	Constant	CONTRACT VALUE	BALANCE VALUE	COLLATERAL VALUE	NUMBER OF COLLATERAL	RECOVERED VALUE	VALUE TX RATE	VALUE TX INTEREST RATE
Constant	1	0.127	− 0.134	− 0.148	− 0.988	− 0.044	− 0.051	0.131
CONTRACT VALUE	0.127	1	− 0.989	− 0.446	− 0.113	− 0.195	− 0.066	0.745
BALANCE VALUE	− 0.134	− 0.989	1	0.465	0.12	0.158	0.073	− 0.791
COLLATERAL VALUE	− 0.148	− 0.446	0.465	1	0.158	0.099	− 0.07	− 0.692
NUMBER_OF COLLATERAL	− 0.988	− 0.113	0.12	0.158	1	0.024	0.019	− 0.154
RECOVERED VALUE	− 0.044	− 0.195	0.158	0.099	0.024	1	0.094	− 0.121
VALUE TX RATE	− 0.051	− 0.066	0.073	− 0.07	0.019	0.094	1	− 0.038
VALUE TX INTEREST RATE	0.131	0.745	− 0.791	− 0.692	− 0.154	− 0.121	− 0.038	1
VALUE RATE OVERDUE	− 0.13	− 0.304	0.306	0.261	0.025	0.16	− 0.301	− 0.163
CLIENT SIZE	0.125	0.474	− 0.499	− 0.127	− 0.053	− 0.271	− 0.45	0.305
MAIN VALUE DELAY	− 0.128	− 0.002	0.005	− 0.19	0.005	0.055	0.128	0.168
SENIORITY LEVEL	− 0.212	− 0.414	0.429	0.632	0.132	0.169	0.103	− 0.399
PERCENT USED	0.02	− 0.067	0.071	0.549	0.067	− 0.02	− 0.41	− 0.334
DURATION IN YEARS	0.088	0.216	− 0.238	− 0.735	− 0.136	− 0.022	0.048	0.456
DURATION IN MONTHS	− 0.092	− 0.214	0.237	0.719	0.136	0.023	− 0.021	− 0.443
DURATION IN DAYS	− 0.062	− 0.386	0.409	0.674	0.113	0.035	0.328	− 0.687

	VALUE RATE OVERDUE	CLIENT SIZE	MAIN VALUE DELAY	SENIORITY LEVEL	PERCENT USED	DURATION IN YEARS	DURATION IN MONTHS	DURATION IN DAYS
Constant	− 0.13	0.125	− 0.128	− 0.212	0.02	0.088	− 0.092	− 0.062
CONTRACT VALUE	− 0.304	0.474	− 0.002	− 0.414	− 0.067	0.216	− 0.214	− 0.386
BALANCE VALUE	0.306	− 0.499	0.005	0.429	0.071	− 0.238	0.237	0.409
COLLATERAL VALUE	0.261	− 0.127	− 0.19	0.632	0.549	− 0.735	0.719	0.674
NUMBER_OF COLLATERAL	0.025	− 0.053	0.005	0.132	0.067	− 0.136	0.136	0.113
RECOVERED VALUE	0.16	− 0.271	0.055	0.169	− 0.02	− 0.022	0.023	0.035
VALUE TX RATE	− 0.301	− 0.45	0.128	0.103	− 0.41	0.048	− 0.021	0.328
VALUE TX INTEREST RATE	− 0.163	0.305	0.168	− 0.399	− 0.334	0.456	− 0.443	− 0.687
VALUE RATE OVERDUE	1	− 0.437	0.299	0.481	− 0.004	0.131	− 0.14	− 0.232
CLIENT SIZE	− 0.437	1	− 0.241	− 0.405	0.255	− 0.048	0.035	− 0.057

Table 8 (continued)

	VALUE RATE OVERDUE	CLIENT SIZE	MAIN VALUE DELAY	SENIORITY LEVEL	PERCENT USED	DURATION IN YEARS	DURATION IN MONTHS	DURATION IN DAYS
MAIN VALUE DELAY	0.299	− 0.241	1	0.368	− 0.653	0.193	− 0.148	− 0.422
SENIORITY LEVEL	0.481	− 0.405	0.368	1	0.043	− 0.421	0.432	0.227
PERCENT USED	− 0.004	0.255	− 0.653	0.043	1	− 0.634	0.584	0.37
DURATION IN YEARS	0.131	− 0.048	0.193	− 0.421	− 0.634	1	− 0.998	− 0.531
DURATION IN MONTHS	− 0.14	0.035	− 0.148	0.432	0.584	− 0.998	1	0.519
DURATION IN DAYS	− 0.232	− 0.057	− 0.422	0.227	0.37	− 0.531	0.519	1

other variables rather than changing all or most of what has already been done [9, 16].

The experience associated with the use of fuzzy logic had its own drawbacks as well. These limitations are most likely associated with the principles of the method, i.e., the rules of combining membership functions into the min–max rule for conjunctive (AND) and disjunctive (OR) reasoning. The major drawback with these rules is that they are not robust enough. This view is augmented by the findings of other studies that proposed different rules of combining the AND and OR clauses. Some studies have proposed that instead of applying the minimum or the maximum of the membership functions, the geometric mean should be considered. There are many possibilities of rules, but all-in-all, they are just arbitrary. One of the ways that these limitations were minimized was by using enough training data to train the fuzzy logic system to choose the best rule for the classification. Lack of adequate data can have deleterious consequences on the robustness of the output. As such, fuzzy logic seems to be best suited for big data classification and predictive analytics.

Perhaps the biggest challenge encountered with the use of fuzzy logic is associated with the fact that the rules assign similar importance to all factors that need to be aggregated. For instance, it is possible that the role of the value tax rate is not of the same importance to financial defaulting as the role of seniority level. Therefore, this problem is solved by not insisting on all membership functions in the fuzzy logic model taking values between 0 and 1.

Consequently, fuzzy logic models can be integrated with neural networks leading to superior prediction accuracy. In other words, the incorporation of qualitative data into the system increases the scope of the model. Furthermore, instead of using standalone fuzzy logic models, it can be used as a module in credit rating whereby the retail customers can be rated and investigated further using artificial intelligence techniques to make the model self-learning. Integration of fuzzy logic models with other neural networks has been used in credit scoring to isolate risky borrowers from those with high creditworthiness successfully.

4.3 Comparison of the two models

The study suggests that fuzzy logic is preferable over the decision trees method. However, care should be taken in the interpretation of the effectiveness of this model. First, the identification of the defaulters was done using dynamic modelling, which was assumed to be perfectly discriminatory. This is probably not true in real life.

All the assumptions were stated a priori in the decision tree modelling. However, assigning equal weights, both to

Table 9 Results of positive correlation with the recovered value

	Constant	CONTRACT VALUE	BALANCE VALUE	COLLATERAL VALUE	COLLATERAL	RECOVERED VALUE	TX RATE	TX INTEREST RATE
Constant	1	0.127	− 0.134	− 0.148	− 0.988	− 0.044	− 0.051	0.131
CONTRACT VALUE	0.127	1	− 0.989	− 0.446	− 0.113	− 0.195	− 0.066	0.745
BALANCE VALUE	− 0.134	− 0.989	1	0.465	0.12	0.158	0.073	− 0.791
COLLATERAL VALUE	− 0.148	− 0.446	0.465	1	0.158	0.099	− 0.07	− 0.692
NUMBER_OF COLLATERAL	− 0.988	− 0.113	0.12	0.158	1	0.024	0.019	− 0.154
RECOVERED VALUE	− 0.044	− 0.195	0.158	0.099	0.024	1	0.094	− 0.121
VALUE TX RATE	− 0.051	− 0.066	0.073	− 0.07	0.019	0.094	1	− 0.038
VALUE TX INTEREST RATE	0.131	0.745	− 0.791	− 0.692	− 0.154	− 0.121	− 0.038	1
VALUE RATE OVERDUE	− 0.13	− 0.304	0.306	0.261	0.025	0.16	− 0.301	− 0.163
CLIENT SIZE	0.125	0.474	− 0.499	− 0.127	− 0.053	− 0.271	− 0.45	0.305
MAIN VALUE DELAY	− 0.128	− 0.002	0.005	− 0.19	0.005	0.055	0.128	0.168
SENIORITY LEVEL	− 0.212	− 0.414	0.429	0.632	0.132	0.169	0.103	− 0.399
PERCENT USED	0.02	− 0.067	0.071	0.549	0.067	− 0.02	− 0.41	− 0.334
DURATION IN YEARS	0.088	0.216	− 0.238	− 0.735	− 0.136	− 0.022	0.048	0.456
DURATION IN MONTHS	− 0.092	− 0.214	0.237	0.719	0.136	0.023	− 0.021	− 0.443
DURATION IN DAYS	− 0.062	− 0.386	0.409	0.674	0.113	0.035	0.328	− 0.687

	VALUE RATE OVERDUE	CLIENT SIZE	VALUE DELAY	SENIORITY LEVEL	PERCENT USED	DURATION IN YEARS	DURATION IN MONTHS	DURATION IN DAYS
Constant	− 0.13	0.125	− 0.128	− 0.212	0.02	0.088	− 0.092	− 0.062
CONTRACT VALUE	− 0.304	0.474	− 0.002	− 0.414	− 0.067	0.216	− 0.214	− 0.386
BALANCE VALUE	0.306	− 0.499	0.005	0.429	0.071	− 0.238	0.237	0.409
COLLATERAL VALUE	0.261	− 0.127	− 0.19	0.632	0.549	− 0.735	0.719	0.674
NUMBER_OF COLLATERAL	0.025	− 0.053	0.005	0.132	0.067	− 0.136	0.136	0.113
RECOVERED VALUE	0.16	− 0.271	0.055	0.169	− 0.02	− 0.022	0.023	0.035
VALUE TX RATE	− 0.301	− 0.45	0.128	0.103	− 0.41	0.048	− 0.021	0.328
VALUE TX INTEREST RATE	− 0.163	0.305	0.168	− 0.399	− 0.334	0.456	− 0.443	− 0.687
VALUE RATE OVERDUE	1	− 0.437	0.299	0.481	− 0.004	0.131	− 0.14	− 0.232
CLIENT SIZE	− 0.437	1	− 0.241	− 0.405	0.255	− 0.048	0.035	− 0.057

Table 9 (continued)

	VALUE RATE OVERDUE	CLIENT SIZE	VALUE DELAY	SENIORITY LEVEL	PERCENT USED	DURATION IN YEARS	DURATION IN MONTHS	DURATION IN DAYS
MAIN VALUE DELAY	0.299	− 0.241	1	0.368	− 0.653	0.193	− 0.148	− 0.422
SENIORITY LEVEL	0.481	− 0.405	0.368	1	0.043	− 0.421	0.432	0.227
PERCENT USED	− 0.004	0.255	− 0.653	0.043	1	− 0.634	0.584	0.37
DURATION IN YEARS	0.131	− 0.048	0.193	− 0.421	− 0.634	1	− 0.998	− 0.531
DURATION IN MONTHS	− 0.14	0.035	− 0.148	0.432	0.584	− 0.998	1	0.519
DURATION IN DAYS	− 0.232	− 0.057	− 0.422	0.227	0.37	− 0.531	0.519	1

Table 10 Kolmogorov–Smirnov test results regarding the fuzzy model

Hypothesis test summary

	Null hypothesis	Test	Sig.	Decision
1	The distribution of CONTRACT_VALUE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.910	Retain the null hypothesis
2	The distribution of BALANCE_VALUE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.039	Reject the null hypothesis
3	The distribution of COLLATERAL_VALUE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.163	Retain the null hypothesis
4	The distribution of NUMBER_OF_COLLATERAL1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.810	Retain the null hypothesis
5	The distribution of RECOVERED_VALUE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
6	The distribution of VALUE_TX_RATE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
7	The distribution of VALUE_TX_INTEREST_RATE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
8	The distribution of VALUE_RATE_OVERDUE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
9	The distribution of CLIENT_SIZE1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.002	Reject the null hypothesis
10	The distribution of MAIN_VALUE_DELAY1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
11	The distribution of SENIORITY_LEVEL1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.043	Reject the null hypothesis
12	The distribution of PERCENT_USED1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.042	Reject the null hypothesis
13	The distribution of DURATION_IN_YEARS1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
14	The distribution of DURATION_IN_MONTHS1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
15	The distribution of DURATION_IN_DAYS1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis
16	The distribution of DELAY_IN_DAYS1 is the same across categories of OVERDUE	Independent-samples Mann–Whitney *U* test	.000	Reject the null hypothesis

Asymptotic significances are displayed. The significance level is .05

Table 11 Confusion matrix of a model developed using the decision tree method

	Predicted		
	0	1	$\sum$
Actual			
0	NA	2.6%	35
1	NA	97.4%	1290
$\sum$		1325	1325

defaulters and non-defaulters, may not truly reflect the true values of the research population. Each estimate was assumed to be independent and equally weighted. These assumptions were addressed by conducting the sensitivity analysis to affirm the robustness of the findings.

The fuzzy logic made no assumptions about the probabilities even though it was assumed that the expert estimates were comparable in scale thus limiting its discriminatory effect on dimensions. However, a sensitivity analysis cannot be performed to ascertain the robustness of the fuzzy logic model. As such, if the expert data analysis were not accurate and verifiable, this could result in a large variation in the model's outcome.

As compared with fuzzy logic, it is conceivable that a decision analytic approach can more easily present the outcomes of the analysis in a decision tree or a diagram. Furthermore, the decision analytic approach explicitly states the hypotheses and can be easily verified using sensitivity tests. However, it can easily discourage its use especially for those who are not familiar with probabilistic concepts. Nevertheless, fuzzy logic makes some implicit assumptions that may make it even harder for credit-grantors to follow the logical decision-making process.

5 Conclusion and future work

This paper compares the performance of fuzzy sets with that of artificial neural network based decision trees on a credit scoring to predict the recovered value. The specific objective is to determine the best model for predicting the recovered value. The findings of the study show that artificial neural network-based decision trees are excellent for representing graphics decision alternatives efficiently. The particular strength of artificial neural network-based decision trees is in their tendency to help to comprehend sequential decisions and outcome dependencies. The model can play a complementary role to other scoring tools such as fuzzy assets whereby the classes it creates can be used as fuzzy sets. However, a decision tree algorithm requires that the target attribute has only discrete values. Another disadvantage is that it performs poorly in terms of complex interactions in which the decision trees are redrawn every time new data is added to the model. Also, decision trees are over-sensitive to the training set, irrelevant attributes, and to noise. Coupled with the fact that data is classified stepwise one node a time until the terminal node, it is difficult to add a regression function to the model making it is predominantly a classification model rather than a predictive one.

On the other hand, the findings of the fuzzy logic show it to be the most convenient method for credit scoring. Our ability to add the regression feature on the fuzzy logic models increases its predictive capabilities. Because of this predictive capability, the data miner can follow an entire process that leads to the creation of mechanisms that can effectively reduce the risk of default. Unlike the decision trees, fuzzy logic allows for additive analysis of data using the selected model. The model is also useful in credit scoring because it helps in the better control of the process of creating relationships and dependencies between the datasets. In addition, it leads to the reduction of subjectivity to acceptable levels thanks in part for the application of quantitative data in the fuzzy logic algorithms. Furthermore, risk analysis based on fuzzy logic models provides the information required for planned operational improvement based on the information obtained from credit scoring.

One of the most important concerns about fuzzy logic credit scoring is that the rules of combining membership functions are too simplistic which makes the model not robust at all. This limitation is minimized by using other arithmetic functions such as the mean instead of the minimum or the maximum of the membership functions. However, the classification gains more credibility with increasing the size of the training data to train the fuzzy logic system to choose the best rule for the classification. Lack of adequate data has adverse consequences on the robustness of the output. It finds that the biggest challenge encountered with the use of fuzzy logic is associated with the fact that the rules assign similar importance to all factors that need to be added together. The solution to this problem can be found in using different values for the different membership functions in the fuzzy logic model rather than just taking values between 0 or 1.

It observes that both fuzzy logic- and artificial neural network-based decision trees have benefits as well as drawbacks. However, the most successful model for credit scoring in financial institutions is the one that incorporates various qualitative consumer aspects in addition to the quantitative ones. Indeed, most of these models rely on quantitative data for classification and modelling of creditworthiness.

In conclusion, the two models made it possible to model uncertainty in the credit scoring process. Although more difficult to implement, fuzzy logic was the best for modelling the uncertainty. However, the decision tree model is more favourable to the presentation of the problem. The assumptions were slightly different for both models. It was difficult to determine which model produces the best results. The two models explored in this study show that both can be applied effectively to credit scoring, although decision trees can complement fuzzy logic models by generating fuzzy rules for these models.

For future works, fuzzy logic models can be integrated, for example, with neural networks leading to superior prediction accuracy. Incorporating qualitative data expands the scope of the model thereby increasing its robustness. In this regard, it is not surprising that standalone fuzzy logic models are hardly used in contemporary credit scoring practices, but rather they are integrated with neural networks for machine learning of consumer behaviour. For example, the fuzzy support vector machine [55, 56] may provide more favourable outcomes than the original version of the algorithm. This practice of integrating standalone classification and predictive models with other neural networks has been used in credit scoring to successfully evaluating creditworthy from non-creditworthy borrowers.

Acknowledgements This work was supported by the National Funding from the FCT - Fundação para a Ciência e a Tecnologia through the UID/EEA/50008/2019 Project; by the Government of the Russian Federation, Grant 08-08; by Brazilian National Council for Research and Development (CNPq) via Grant No. 309335/2017-5; by Ciência sem Fronteiras of CNPq, Brazil, process number 200450/2015-8; and by the International Scientific Partnership Program ISPP at King Saud University through ISPP #0129.

References

1. Akkoç S (2012) An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: the case of Turkish credit card data. Eur J Oper Res 222:168–178. https://doi.org/10.1016/j.ejor.2012.04.009
2. Lahsasna A, Ainon RN, Wah TY (2010) Credit scoring models using soft computing methods: a survey. Int Arab J Inf Technol 7:115–123
3. Chen W, Xiang G, Liu Y, Wang K (2012) Credit risk evaluation by hybrid data mining technique. Syst Eng Procedia 3:194–200. https://doi.org/10.1016/j.sepro.2011.10.029
4. Ibn UF, Panford JK, Ben H-J (2014) Fuzzy logic approach to credit scoring for micro finances in Ghana (A case study of KWIQPLUS money lending). Int J Comput Appl 94(8):8887. https://doi.org/10.5120/16362-5772
5. Zamula A, Kavun S (2017) Complex systems modeling with intelligent control elements. Int J Model Simul Sci Comput 08:1750009. https://doi.org/10.1142/S179396231750009X
6. Abiyev RH (2014) Credit rating using type-2 fuzzy neural networks. Math Probl Eng. https://doi.org/10.1155/2014/460916
7. Buikstra E, Fallon AB, Eley R (2007) A bi-level neural-based fuzzy classification approach for credit scoring problems. Rural Remote Health 7:543. https://doi.org/10.1002/cplx
8. Darwish NR, Abdelghany AS (2016) A fuzzy logic model for credit risk rating of Egyptian commercial banks. Int J Comput Sci Inf Secur 14:11–19
9. Louzada F, Ara A, Fernandes GB (2016) Classification methods applied to credit scoring: a systematic review and overall comparison. Surv Oper Res Manag Sci 21:117–134. https://doi.org/10.1016/j.sorms.2016.10.001
10. Mammadli S (2016) Fuzzy logic based loan evaluation system. Procedia Comput Sci 102:495–499. https://doi.org/10.1016/j.procs.2016.09.433
11. Rao TVN, Reddy K (2017) Application of fuzzy logic in financial markets for decision making. Int J Adv Res Comput Sci 8(3). https://doi.org/10.26483/ijarcs.v8i3.3020
12. Ye J (2017) Aggregation operators of trapezoidal intuitionistic fuzzy sets to multicriteria decision making. Int J Intell Inf Technol 13:1–22. https://doi.org/10.4018/IJIIT.2017100101
13. Ren S (2017) Multicriteria decision-making method under a single valued neutrosophic environment. Int J Intell Inf Technol 13:23–37. https://doi.org/10.4018/IJIIT.2017100102
14. Mohmad Hassim YM, Ghazali R (2013) Functional link neural network—artificial bee colony for time series temperature prediction. Springer, Berlin, pp 427–437
15. Li EY (1994) Artificial neural networks and their business applications. Inf Manag 27:303–313. https://doi.org/10.1016/0378-7206(94)90024-8
16. Liao SH, Chu PH, Hsiao PY (2012) Data mining techniques and applications—a decade review from 2000 to 2011. Expert Syst Appl 39:11303–11311
17. Goonatilake S, Treleaven PC, Philip C (1995) Intelligent systems for finance and business. Wiley, Hoboken
18. Bentley PJ, Kim J, Jung G-H, Choi J-U (2000) Fuzzy darwinian detection of credit card fraud. In: 14th Annual fall symposium of the Korean information processing society, vol 14
19. Hoffmann F, Baesens B, Martens J et al (2002) Comparing a genetic fuzzy and a neurofuzzy classifier for credit scoring. Int J Intell Syst 17:1067–1083. https://doi.org/10.1002/int.10052
20. Bojadziev G, Bojadziev M (2007) Fuzzy logic for business, finance, and management. World Scientific, Singapore
21. Laha A (2007) Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring. Adv Eng Inform 21:281–291. https://doi.org/10.1016/j.aei.2006.12.004
22. Hájek P, Olej V (2015) Intuitionistic fuzzy neural network: the case of credit scoring using text information. Springer, Cham, pp 337–346
23. Khashei M, Mirahmadi A (2015) A soft intelligent risk evaluation model for credit scoring classification. Int J Financ Stud 3:1–12
24. Malhotra R, Malhotra DK (2003) Evaluating consumer loans using neural networks. Omega 31:83–96. https://doi.org/10.1016/S0305-0483(03)00016-1
25. Kogut B, MacDuffie JP, Ragin C (2004) Prototypes and strategy: assigning causal credit using fuzzy sets. Eur Manag Rev 1:114–131. https://doi.org/10.1057/palgrave.emr.1500020
26. Ong C, Huang J, Tzeng G (2005) Building credit scoring models using genetic programming. Expert Syst Appl 29:41–47. https://doi.org/10.1016/j.eswa.2005.01.003
27. Dahal K, Hussain Z, Hossain MA (2005) Loan Risk analyzer based on fuzzy logic. In: 2005 IEEE international conference on e-technology, e-commerce and e-service. IEEE, pp 363–366
28. Jiao Y, Syau Y-R, Lee ES (2007) Modelling credit rating by fuzzy adaptive network. Math Comput Model 45:717–731. https://doi.org/10.1016/J.MCM.2005.11.016
29. Lahsasna A (2009) Evaluation of credit risk using evolutionary-fuzzy logic scheme. University of Malaya, Kuala Lumpur
30. Wang S (2010) A comprehensive survey of data mining-based accounting-fraud detection research. In: 2010 international

conference on intelligent computation technology and automation. IEEE, pp 50–53
31. Grace AM, Williams SO (2016) Comparative analysis of neural network and fuzzy logic techniques in credit risk evaluation. Int J Intell Inf Technol. https://doi.org/10.4018/IJIIT.2016010103
32. Sohn SY, Kim DH, Yoon JH (2016) Technology credit scoring model with fuzzy logistic regression. Appl Soft Comput 43:150–158. https://doi.org/10.1016/J.ASOC.2016.02.025
33. Paul U, Biswas A (2017) Consumer credit limit assignment using bayesian decision theory and fuzzy logic—a practical approach. J Manag 4:11–18
34. Zurada J (2010) Could decision trees improve the classification accuracy and interpretability of loan granting decisions? In: 2010 43rd Hawaii international conference on system sciences. IEEE, pp 1–9
35. Rescher N (1969) Many-valued logic. McGraw-Hill, New York
36. Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353. https://doi.org/10.1016/S0019-9958(65)90241-X
37. Baghban A, Jalali A, Shafiee M et al (2019) Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids. Eng Appl Comput Fluid Mech 13:26–39. https://doi.org/10.1080/19942060.2018.1542345
38. Ohno-Machado L, Lacson R, Massad E (2000) Decision trees and fuzzy logic: a comparison of models for the selection of measles vaccination strategies in Brazil. In: Proceedings AMIA symposium, pp 625–629
39. Martens D, Baesens B, Van Gestel T, Vanthienen J (2007) Comprehensible credit scoring models using rule extraction from support vector machines. Eur J Oper Res 183:1466–1476. https://doi.org/10.1016/j.ejor.2006.04.051
40. Yaseen ZM, Sulaiman SO, Deo RC, Chau KW (2019) An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction. J Hydrol 569:387–408. https://doi.org/10.1016/j.jhydrol.2018.11.069
41. Wu DD, Olson DL (2014) A decision support approach for accounts receivable risk management. IEEE Trans Syst Man Cybern Syst 44:1624–1632. https://doi.org/10.1109/TSMC.2014.2318020
42. Massad E, Burattini MN, Ortega NR (1999) Fuzzy logic and measles vaccination: designing a control strategy. Int J Epidemiol 28:550–557
43. Zhang H, He H, Zhang W (2018) Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring. Neurocomputing 316:210–221. https://doi.org/10.1016/J.NEUCOM.2018.07.070
44. Yu L, Wang S, Lai KK (2009) An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: the case of credit scoring. Eur J Oper Res 195:942–959. https://doi.org/10.1016/j.ejor.2007.11.025
45. Freeling ANS (1980) Fuzzy sets and decision analysis. IEEE Trans Syst Man Cybern 10:341–354. https://doi.org/10.1109/TSMC.1980.4308515
46. Malhotra R, Malhotra DK (2002) Differentiating between good credits and bad credits using neuro-fuzzy systems. Eur J Oper Res 136:190–211. https://doi.org/10.1016/S0377-2217(01)00052-2
47. Hoffmann F, Baesens B, Mues C et al (2007) Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms. Eur J Oper Res 177:540–555. https://doi.org/10.1016/j.ejor.2005.09.044
48. Capotorti A, Barbanera E (2012) Credit scoring analysis using a fuzzy probabilistic rough set model. Comput Stat Data Anal 56:981–994. https://doi.org/10.1016/j.csda.2011.06.036
49. Muslim MA, Nurzahputra A, Prasetiyo B (2018) Improving accuracy of C4.5 algorithm using split feature reduction model and bagging ensemble for credit card risk prediction. In: 2018 International conference on information and communications technology (ICOIACT). IEEE
50. Tang T-C, Chi L-C (2005) Predicting multilateral trade credit risks: comparisons of logit and fuzzy logic models using ROC curve analysis. Expert Syst Appl 28:547–556. https://doi.org/10.1016/J.ESWA.2004.12.016
51. Mierswa I, Wurst M, Klinkenberg R et al (2006) Yale: rapid prototyping for complex data mining tasks. In: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining—KDD'06. ACM Press, New York, p 935
52. Huang JJ, Tzeng GH, Ong CS (2006) Two-stage genetic programming (2SGP) for the credit scoring model. Appl Math Comput 174:1039–1053. https://doi.org/10.1016/j.amc.2005.05.027
53. Chen W, Ma C, Ma L (2009) Mining the customer credit using hybrid support vector machine technique. Expert Syst Appl 36:7611–7616. https://doi.org/10.1016/j.eswa.2008.09.054
54. Rahman N (2018) Data mining techniques and applications. Int J Strateg Inf Technol Appl 9:78–97. https://doi.org/10.4018/IJSITA.2018010104
55. Wang S, Li Y, Shao Y et al (2016) Detection of dendritic spines using wavelet packet entropy and fuzzy support vector machine. CNS Neurol Disord: Drug Targets 16:116–121. https://doi.org/10.2174/1871527315666161111123638
56. Zhang YD, Yang ZJ, Lu HM et al (2016) Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. IEEE Access 4:8375–8385. https://doi.org/10.1109/ACCESS.2016.2628407