Fuzzy Credit Risk Scoring Rules using FRvarPSO

Patricia Jimbo Santana[1], Laura Lanzarini[2], and Aurelio F. Bariviera [3]

[1]Ecuador Central University - Faculty of Administrative Sciences, Career Accounting and Auditing, Quito, Ecuador
[2]Institute for Research in Computer LIDI, Faculty of Computer Science, National University of La Plata, La Plata, Buenos Aires, Argentina
[3]Universitat Rovira i Virgili, Department of Business, Reus, Spain

Abstract

There is consensus that the best way for reducing insolvency situations in financial institutions is through good risk management, which involves a good client selection process. In the market, there are methodologies for credit scoring, each analyzing a large number of microeconomic and/or macroeconomic variables selected mostly depending on the type of credit to be granted. Since these variables are heterogeneous, the review process carried out by credit analysts takes time. The objective of this article is to propose a solution for this problem by applying fuzzy logic to the creation of classification rules for credit granting. To achieve this, linguistic variables were used to help the analyst interpret the information available from the credit officer. The method proposed here combines the use of fuzzy logic with a neural network and a variable population optimization technique to obtain fuzzy classification rules. It was tested with three databases from financial entities in Ecuador – one credit and savings cooperative and two banks that grant various types of credits. To measure its performance, three benchmarks were used: accuracy, number of classification rules generated, and antecedent length. The results obtained indicate that the hybrid model that is proposed performs better than its previous versions due to the addition of fuzzy logic. At the end of the article, our conclusions are discussed and future research lines are suggested.

Keywords: Fuzzy Rules; varPSO (Variable Particle Swarm Optimization); Credit Risk.

1 Introduction

Credit risk is defined as the probability of loss due to non-compliance by the borrower with the required payments in relation to any type of debt. Credit risk assessment is the process carried out by financial institutions to establish the ability of borrowers to meet their financial obligations. "Credit scoring" models are mathematical or statistical econometric methods used to measure the risk and/or the probability of non-compliance of the borrower [1]; therefore, they are considered as a decision-making model where credit applications from good and bad clients are accepted and the ultimate goal is to reject, through the use of statistical procedures, or machine learning algorithms, those credit requests that involve a higher risk.

In general, credit risk assessment is linked to several characteristics of the borrower, such as reputation, leverage, revenue volatility, intended use of the credit, guarantees, and so forth. Traditional credit scoring techniques allow assessing borrowers based on characteristics such as ability to pay, backup available capital, collateral as loan guarantee, general economy conditions. All of these characteristics can be combined in the microeconomic and macroeconomic variable analysis. It should be noted that granted credit data contain interesting relations between the characteristics of credit holders and their delinquency rates. These characteristics also depend on the general

aurelio.fernandez@urv.cat

economy of the country where they live and, as a result, the macroeconomic context of the credit holder should also be considered in the analysis. However, financial institutions oftentimes worry about recovering their money, since credit holders usually behave in a confusing and unpredictable manner. It is difficult for financial institutions to assess the credit score of clients with just a few input characteristics, since in that case, the assessment would become almost intuitive and the end result could be lacking in accuracy. To attract more clients, the granting process should be quick. As a result, banks need to improve accuracy when approving loans to avoid non-compliance risks and at the same time they need to reduce the time required to grant such loans to achieve a competitive advantage. Credit risk analysis research typically has considered two issues; namely, unbalanced data and deciding whether to grant the credit or not. The first issue, unbalanced data, stems from the fact that the minority group is always that of credits that should not have been granted or clients who are bad payers, which, in the best of cases, represents 20% of the total body. The second issue, deciding whether to grant the credit or not, relates to the selection of client characteristics to be analyzed. These characteristics can be various in nature and different from one financial institution to another. This article focuses on the second issue. Data mining presents a set of models that have been applied in the credit scoring field to characterize borrowers and help decide whether the credit should be granted or not. In the literature, there are many success cases where these models are an alternative that is better than the subjective assessment usually applied to decide credit granting [2].

There are several models that link delinquency and borrower characteristics, such as the ones described by Reinke [4] and Zeller [5]. However, these models have several issues, including a lack of robustness due to their being based on small-sized samples and using variables that are not always available. Some of these aspects have been solved by using machine learning techniques. Such is the case of Qiao [6], who used a SOM competitive neural network to group clients based on their most representative characteristics. In that article, clients were split into several groups and the characteristics of the clients in each group were then analyzed. Interesting conclusions were obtained, such as observing that the group with the highest risk was that corresponding to the clients that had the highest average revenue. One of the reasons why clients in this group have easy access to credits is precisely due to their higher revenues, but the analysis of microeconomic variables is not adequate and macroeconomic variables are not considered.

The model proposed in this article corresponds to the predictive model category, since it can return a response to a new credit request and it uses a set of fuzzy classification rules. A classification rule is an expression with the following format:

IF condition THEN Result = Possible answer

where "condition" is a combination of expressions with the format (variable = value). The variables in the antecedent can be qualitative or quantitative. Since these are classification rules, the consequent always shows the same attribute – the one used to refer to the expected response [3]. In this paper, classification rules operate on numeric information in a fuzzy manner, thus avoiding the use of hard thresholds when establishing the conditions that are included in the antecedent of the rule. The rules are generated through a model that combines a competitive neural network with an optimization technique based on a particle swarms, or PSO (Particle Swarm Optimization). The neural network is used to identify the starting point for the optimization technique. It should be noted that this technique, unlike the traditional methods for building rules, is not based on successively splitting available information, but rather on simultaneously selecting those conditions whose combination creates the antecedent of each rule. The rule set obtained is characterized for its low cardinality and higher accuracy than the method proposed in [23, 24], where fuzzy logic was not used. Also, unlike [23, 24], information about the economic situation of the country was included as part of the information about the borrower. The characteristics of the borrower also included aspects relating to delinquency.

The remaining sections of this article are organized as follows: Section 2 briefly mentions some related works, Section 3 describes the method proposed, Section 4 discusses the results obtained applied to real databases from financial institutions, and Section 5 summarizes our conclusions and proposes some future lines of work.

2 Related works

In relation to credit assessment, there are several computer-based techniques that can be mentioned, including the statistical approaches that use techniques such as linear discriminant [7, 8] and logistic regression [9, 10, 11]. These approaches work with ideal assumptions, since variables are considered to have a multivariate normal distribution, which limits the performance of the models.

Machine learning offers other solutions, such as the C4.5 method [12], which consists in generating a classification tree, or the PART method [13], which allows obtaining a rule set. In either case, it is essential that the rule set obtained covers the examples with a preset error level, regardless of the number of rules generated.

In recent years, researchers have proposed new approaches to improve the performance of credit scoring models. Li et al. [14] propose a semi-supervised SVM to improve credit model performance; the performance of the model proposed is higher than reference measurements. Zhao et al. [15] also propose an improvement for a model that uses a multi-layer Perceptron neural network based on back propagation to improve credit scoring, which results in increased model performance. However, this approach has issues when working with unbalanced classes. Studies have shown that that the combination of models yields better results than using a single technique to rank credit risk [16]. Wang and Ma [17] propose a new technique based on SVM and two set classification strategies, and the performance of this model was better than that of other classification models. Oreski and Oreski [25] use a hybrid algorithm based on genetic algorithms and neural networks to predict credit granting classification accuracy. Zieba et al. [26] use Extreme Gradient Boosting to train a set of decision trees, used to predict financial loss.

There are other approaches, such as fuzzy sets theory, stemming from Zadeh's work [35], which is very useful in cases such as credit scoring, where limits are not very well-defined, which requires the knowledge of the credit expert. The authors in [33] discuss the importance of using fuzzy logic in machine learning, considering it a key element in the development of artificial intelligence. Uncertainty representation and management is one of the topics in which fuzzy logic can better contribute by reasonably supplementing the probability theory, since not all types of uncertainty are probabilistic in machine learning. For this reason, the role of fuzzy sets in data learning must be accurately assessed. The authors in [28, 29, 30] highlight the importance of using fuzzy logic for understanding the models. This is also true in the method proposed in this article.

Chia [28] proposes a method based on PSO to build a fuzzy rules system. In this method, each individual in the population represents an entire fuzzy system. This is an alternative approach to the method proposed in this article that obtains rules individually as examples are covered. Thus, there is no need to define beforehand the maximum number of rules to be generated. As a side note, it should be noted that the authors of [28] state that the options used demands quite a lot of time for training.

Jang [31] used an adaptive partition in the input space to design a fuzzy system based on ANFIS (adaptive-network-based fuzzy inference system). In this case, the network's adaptive partitioning scheme has the disadvantage that the number of fuzzy sets for each input variable is preset. Wong and Wong [18] applied a binary-coded genetic algorithm (GA) to determine an appropriate number of fuzzy sets for each input variable and membership functions. However, to obtain accurate central positions for the membership functions, lengthy coding is required, which results in an increase in training time. Additionally, the fuzzy partition used for the input space has drawbacks, since the number of fuzzy rules increases considerably as the number of dimensions increases, generating fictional rules due to the lack of training data in fuzzy regions. Each fuzzy region maps the premise of a fuzzy rule that is associated to several membership functions. Several grouping techniques have been used to determine the premises of a fuzzy system, such as the fuzzy algorithm c-mean (FCM) [39, 40] and the method based on ART [20]. The latter proposes a method to generate the regions that will determine the premise of a fuzzy system. Basically, it consists in grouping input data into clusters, the number of fuzzy rules being the same as the number of clusters.

Sugeno and Yasukawa [5] apply the FCM algorithm to determine the premise of the fuzzy system. However, the disadvantage of FCM is that the number of clusters must be preset. Additionally, the results of the FCM grouping algorithm are affected by the selected initial centers and the distance.

Fuzzy rules are used in [32] to control class imbalance by improving the performance of the fuzzy rule-based scoring system. To do this, a starting linguistic rule base is built and a genetic algorithm is then applied to it to select the best rules. One of the issues here is that the method used to generate the initial rules produces a rule set that is too large, which hinders understability.

In previous research works, we have combined a neural network and an optimization technique to obtain this type of model [21, 22, 23, 24]. The solution presented then did not use fuzzy logic. The results obtained when applying it to the database used in the Ecuadorian financial market yielded an accuracy that was 1% lower than conventional methods; however, its simplicity made using it worth it. No macroeconomic information was used.

The variation presented in this article introduces fuzzy logic as a mechanism to guide how solutions are improved through the optimization technique. This helps improve the accuracy of the method. Finally, rule antecedent will consist of nominal and/or linguistic variables, which will assist credit officers when interpreting the results.

3 Methodology

This article presents a hybrid methodology that uses fuzzy variables, combining a competitive neural network with a variable population optimization technique to obtain a set of fuzzy classification rules. Therefore, before describing the methodology itself in Section 3.4, a brief description of each of these concepts will be provided.

3.1 Neural networks - SOM and LVQ

The competitive neural networks used in this work were SOM (Self Organization Map) and Learning Vector Quantization (LVQ). Considering only their common characteristics, it can be said that both are aimed at grouping input data and that, as a result, they offer a set of centroids. Each centroid represents a set of input data based on similarity or distance metrics previously defined. As regards the differences between these networks, they are mainly related to their learning methods, since SOM uses non-supervised training and LVQ is the opposite. SOM is aimed at providing additional information about how groups are organized. This is better than any other grouping technique of the "winner-take-all" type, such as k-means. On the other hand, since LVQ uses the information from the expected response during training, its centroids are better adjusted. In this article, both architectures were used, LVQ being the one that yielded a better performance. For this reason, its training procedure is briefly described below. LVQ is based on a competitive neural network formed by three layers. The first layer receives input information, the second layer is where the competition takes place, and finally, the last layer is responsible for the classification process and returning the output values. Each neuron in the competitive layer is associated to a number vector whose dimension is the same as that of the input examples, and a label that indicates the class that it is going to represent. Once the adaptive process finishes, these vectors contain the information related to the classification centroids or prototypes. In LVQ, when the algorithm is started, the number K of centroids to be used must be indicated. This defines network architecture, since this value matches the number of competitive neurons that will form the intermediate layer. These centroids are initialized taking K random examples. Then, the examples are entered one at a time and centroids are adapted as described below. For each example, the closest centroid is calculated. Then, if it belongs to the same class, it is “moved closer” by strengthening its representation; otherwise, it is “moved away”. These movements are done by means of a factor or adaptation speed. The process is repeated until the modifications to be done are below a preset threshold, or until the examples are identified with the same centroids in two consecutive iterations, whichever happens first. For more information about both neural networks, see [34].

3.2 Fuzzy logic

Fuzzy logic is multivalued logic that allows mathematically representing uncertainty and vagueness. It was proposed by Zadeh in 1965 [35]. Unlike classic logic, fuzzy logic uses the concept of degree of membership as a tool to express the degree of truth for a given event to occur. This led to the concept of fuzzy sets, adding a membership function to the conventional definition of a set. This function matches each input value with a real value within the interval $[0, 1]$, thus establishing the degree of membership of each element to the corresponding fuzzy set.

We refer to [36] for a deeper analysis on the development of fuzzy logic, and a detailed survey of established and prospective applications of it in various areas of human affairs.

Risk analysis is one of the areas where the usefulness of fuzzy logic was first acknowledged, and it was used for the first time in Clements [27]. To achieve this, a method based on linguistic variables was developed, where the subjective assessments of risk experts could be properly used together with any other (objective) data available. In this article, the membership functions associated with each fuzzy set were defined by credit risk experts.

Nowadays, neuro-fuzzy systems have attracted increasing interest from researchers in various scientific and engineering fields [37] due to their effective learning and reasoning abilities. Neuro-fuzzy systems are highly significant, since they combine the learning power of artificial neural networks and the explicit knowledge representation features of fuzzy inference systems. There are three essential characteristics that models have to analyze: computer speed, accuracy and complexity. Using fuzzy logic can directly improve reasoning and machine learning inference. Qualitative knowledge, even though it lacks accuracy, can be modeled to allow automated learning symbolic expression using fuzzy logic. The use of neural networks complements learning with capacity, robustness and mass parallelism in the system. For this reason, the knowledge representation and automated learning ability in a neuro-fuzzy system make this type of systems into a powerful framework to solve machine learning problems. Among neuro-fuzzy systems based on learning algorithms, there are systems that are trained in gradients, hybrid systems, systems that use populational techniques and systems that use techniques based on Support Vector Machine (SVM). The same as back-propagation and other supervised learning algorithms that use the gradient-based technique, neuro-fuzzy systems can also run into the issue of local minima and premature convergence.

Neuro-fuzzy techniques based on particle swarm optimization (PSO) are used due to their implementation simplicity, reduced use of memory, and reduced time required to converge to the optimal solution. This neuro–fuzzy technique has been used in financial predictions with the extraction of fuzzy variables based on PSO [38]. After studying several neuro-fuzzy techniques in [37], it was concluded that those methods that use PSO are better than those that use gradient-based techniques. ELM or SVM are among the best, since their convergence speed is low, algorithm complexity is medium, classification ability is also medium range, their accuracy is good, and they do not get trapped in local minima, all of which is a safeguard for the model developed here.

3.3 Particle Swarm Optimization (PSO)

Particle Swarm Optimization or PSO is a population-based metaheuristic algorithm proposed by Kennedy and Eberhart in [41] where each individual in the population, called particle, represents a possible solution to the problem and changes following three factors: its current knowledge (its ability to solve the problem), its historical knowledge or previous experiences (its memory), and the historical knowledge or previous experiences of neighboring individuals (its social knowledge).

In this paper, PSO was used to obtain fuzzy classification rules. For simplicity, each particle represents the antecedent of a single rule. The consequent is determined by a value of the attribute corresponding to the class chosen beforehand, and it is the same for all particles. Therefore, when the adaptive process ends, the best individual in the population will contain the best fuzzy classification rule for the examples and the class indicated at the beginning.

The antecedent consists of two parts – one that is binary and another one that is continuous. The binary part is used to choose the pairs (qualitative attribute = value) or (linguistic variable = value) using the binary PSO version defined in [42, 43, 44]. The real part is only applied to fuzzy variables, and its purpose is logging the degree of membership to each fuzzy set. This determines the value of the fuzzy variable that will be used if the corresponding position of the binary vector calls for it. As regards the movement of the continuous part, the version defined in [22] is modified by using, in addition to inertia and particle cognitive and social values, a "voting" factor to update speed. This factor is obtained by calculating the average degree of membership for the examples that meet the conditions of the rule. Thus, individuals prioritize movement in the direction of the highest value, which is the one that received the highest "vote".

The adaptive process is repeated, updating the positions of the particles until it detects that the best solution cannot be improved for a certain number of attempts. At this point, the optimal value is considered to have been found and the best particle or rule found is used as the solution.

This process is repeated a certain number of times reducing the set of input examples and selecting the appropriate class each time. This is described in greater detail in the following section.

3.4 Proposed method: FRvarPSO

The goal of the method proposed, FRvarPSO, is obtaining a set of low-cardinality classification rules that has an adequate accuracy level and is easy to interpret. To achieve this, two important aspects were considered – the first of these aspects in relation to the method's ability to operate with linguistic variables, and the second aspect in relation to the insertion of information based on degrees of membership, both for fitness evaluation as well as for the search process using the optimization technique.

The method starts by establishing the fuzzy sets and the membership functions for each fuzzy variable associated with each of the sets. These functions are triangular and were defined by a credit risk expert. The sets used are defined so that the value of a fuzzy variable can only simultaneously belong to two of them at most. Then, training examples are further grouped using a competitive neural network. Network size is an input parameter whose value depends on whether PSO is used with fixed- or variable-sized population. The similarity measurement used is the Euclidean distance. Once the network is trained, the classification rules are obtained one by one by repeatedly running the chosen PSO version. In each opportunity, first the class for which the fuzzy rule will be built is chosen based on the number of uncovered examples (highest first). Then, the particle population is initialized using centroid information from the competitive neural network mentioned above. Thus, instead of starting at random positions, the search process starts at promising regions, requiring only a few steps to find an adequate antecedent.

To evaluate the performance of each particle, the following fitness function is used:

*Fitness Function = (support * confidence* factor1) – (factor2 * NumAttribsAntecedent / MaxAttribs)*

Where *support* and *confidence* are the values corresponding to the rule that represents the particle; *factor1* is a penalization value in case the support is not found within the ranges set in the algorithm; the second term in the fitness function reflects the significance given to the number of attributes included in the antecedent, and *factor2* is a constant.

Once the first rule has been obtained, the degree of membership for each of the examples that meet the conditions of the rule is also obtained G_{p_i} , where G_{p_i} is the degree of membership of example i that meets the conditions of the rule, which is given by a t-standard that uses the minimum operator between the degrees of membership of the fuzzy attributes involved in the antecedent of the rule. Then, the Vote Criterion CV is calculated, which is given by the average of the degrees of membership of the examples that meet the conditions of the rule.

$$CV = \frac{\sum_{i=1}^{n} G_{p_i}}{n} \tag{1}$$

This vote criterion is used in the movement for the next particle, added to the speed, and a new position is obtained for the particle, which is updated as follows:

$$V(t+1) = WV(t) + \phi_1 rand_1(pBest - p) + \phi_2 rand_2(localBest - p) + CV \tag{2}$$

where

W is the inertia factor, which increases as algorithm iterations are performed.

$V(t)$ is the speed value of particle at time t.

ϕ_1, ϕ_2 are constant values that represent the cognitive and social factors.

$rand_1, rand_2$ are values within interval $[0, 1]$, which are set randomly.

$pBest$ stores the best solution found by the particle.

p is the position of the particle.

$localBest$ gives the position of the closest neighbor with a fitness function value higher than this one.

The vote criterion, used when selecting fuzzy attributes, is taken into account for the movement of the individual, considering that individuals will prioritize towards the highest value, i.e., the value that obtained the highest "vote", where fuzzy classification plays a direct role. If this rule is met with the appropriate confidence and support, it is added to the rule set. Then, those examples that are correctly covered by that rule are removed from the input dataset and the class selection process starts over to determine the consequent for the next rule. The process keeps iterating until the maximum number of iterations is reached or until a sufficient number of examples has been covered, whichever happens first. The steps required in the design of the fuzzy system are shown in the pseudocode presented in Figure 1.

4 Data and results

The method proposed was tested with three credit risk databases. The first one is a database from a credit and savings cooperative in Ecuador in Segment 2 of the Superintendency of Popular and Solidarity Economy, with assets in the range between 20,000,000.00 and 80,000,000.00. The other two databases come from banking institutions – one of them in the microcredit business, specialized in mass credit placement, and the other working in the consumer credit and productive or business credit market The information from all three financial institutions corresponding to credit operations between January 2012 and December 2016 was analyzed.

The following fuzzy microeconomic variables were considered: granted credit value, interest rate, equity, household value, income, expenditures, and delinquency. The fuzzy macroeconomic variables used were Consumer Price Index (CPI), Unified Base Salary (UBS) and value of favorite

Algorithm 1 Pseudocode of the proposed method

```
Determine fuzzy variables
Represent fuzzy variable knowledge (degree of membership) of all training examples by means of appropriate fuzzy set definitions
Train the competitive neural network using the training examples
Determine the minimum support for each class
while termination criterion is not reached do
    Choose the class with the higherst number of examples that have not been covered
    Build a population considering neural network centroids
    Evaluate the fitness value of each particle as in [4]
    while the particle population does not reach a stable status do
        Identify the best solution found so far
        for each particle do
            Calculate voting criterion (average degree of membership of the examples that meet the rule indicated by the particle)
            Calculate the speed and add it to the vote criterion mentioned above.
            Obtain the new position of the particle by adding the speed mentioned above and limit as appropriate.
        end for
        If using elitism, recover the best solution from the previous iteration.
    end while
    Obtain the best rule for the population
    if the rule meets support and confidence requirements then
        Add the rule to the ruleset
        Remove from the input set those examples that are correctly covered
        Recalculate the minimum support for the class that has been considered
    end if
end while
```

stocks. Additionally, the variables credit type, province, and credit granting year and month were also considered.

A study of credit behavior after granting the credit was carried out. Among the data analyzed, client delinquency was reviewed, and whether delinquency exceeded sixty days, or cases where credit recovery was minimal or almost non-existent, which is known as written-off portfolio. If these conditions were present, it meant that credit holder behavior was not appropriate, meaning that the credit should have not been granted. The resulting class should have been "Denied".

If client delinquency was between zero and sixty days, it was considered that the credit involved some risk. The advantage of this analysis of data is that credit holder behavior was reviewed for a period of time, including client past behavior, also considering the financial situation of the country through the macroeconomic variables, which allowed predicting credit holder behavior.

An expert in credit risk was consulted to define the fuzzy sets for each of the variables, based on the economy of Ecuador. All fuzzy variables were assigned to three fuzzy sets: LOW, MEDIUM and HIGH. The membership functions associated to each of these fuzzy sets were triangular.

To determine if the method proposed is effective, the performance of several methods was tested first. These methods combined two types of PSO, one with fixed population and one with variable population, initialized with two different competitive neural networks: LVQ and SOM.

Twenty-five separate runs were done for each of the methods. The method proposed and the control methods were compared as regards how they find the classification rules they use, and the proposed method, unlike the others, uses in its algorithm random values that result in particle movement being not excessively deterministic. The most important characteristic of the results obtained is the improvement in accuracy versus the results reported in [4], which is obtained by generating the set of fuzzy classification rules from micro- and macroeconomic variables.

Below, six tables are presented, each of them with a confusion matrix indicating the number of classes predicted. This matrix contains the information for correct classifications, i.e., those credits that were correctly classified in granted or denied credits. Those credits that were denied when they should have been accepted based on the analysis of the credit holder, which are the credits that meet the condition of the antecedent of the classification rule but not its consequent, are classified as a Type I error, also called False Positive. Type II errors, or False Negatives, are those credits that should have been denied for not meeting the conditions of the antecedent of the classification rule but were granted instead. These are the credits that have a record of delinquency, and are the ones that represent the highest risk for the financial institution. For each of these values, the corresponding standard deviation is indicated.

Then, accuracy is also included, which is calculated as the ratio between correct predictions and the total number of predictions (correct and incorrect). The average number of rules obtained can also be seen, as well as the average number of variables used to form the antecedent of the rule.

For the database of the credit and savings cooperative, the LVQ+varPSO method has an accuracy of 78.94%; for the microcredit bank, the LVQ+varPSO method has an accuracy of 97.78%; and finally, for the financial institution granting consumption and productive or business credits, the LVQ+varPSO method has an accuracy of 98.80

Using the fuzzy models with the credit and savings cooperative database, the FRvarPSO method yielded an accuracy of 79.88% with 1% fewer rules than the non-fuzzy method. For the microcredit bank, specialized in mass credit placement, the FRvarPSO method yielded an accuracy of 98.80% with 13.4% fewer rules than the non-fuzzy method. Finally, for the financial institution granting consumption and productive or business credits, the FRvarPSO method yielded an accuracy of 85.01% with 4.62% fewer rules.

As it can be seen in Table 7, for all three financial institutions, the accuracy of the classification obtained using fuzzy FRvarPSO is better, compared with the other methods that do not use fuzzy logic. Thus, the goal of allowing credit officers to respond quickly and with a high level of accuracy, using microeconomic and macroeconomic variables, to clients applying for credit is achieved.

We consider that this method that uses fuzzy logic is an excellent alternative for use in financial institutions, since it allows obtaining the profile of a good client, which results in credit risk reduction through reduced delinquency rates. In the future, it is going to be able to sort clients into “good clients”, “risky clients”, and “bad clients”.

Table 1: Results obtained with rules using macroeconomic and microeconomic variables with a database from a credit and savings cooperative in Ecuador in Segment 2 of the Superintendency of Popular and Solidarity Economy, with assets in the range between 20,000,000.00 and 80,000,000.00

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + PSO	Denied	**0.6023** ±0.0029	0.1358 ±0.0017	0.7718	6.2966	4.3952
	Accepted	0.0923 ±0.0019	**0.1695** ±0.0022	±0.0019	±0.2990	±0.1439
SOM + varPSO	Denied	**0.607** ±0.0042	0.1439 ±0.0047	0.7701	5.6899	4.8471
	Accepted	0.0859 ±0.0037	**0.1631** ±0.0049	±0.0023	±0.2801	±0.17714
LVQ + PSO	Denied	**0.6113** ±0.0038	0.1147 ±0.0041	0.7871	5.4772	5.2956
	Accepted	0.0982 ±0.0029	**0.1758** ±0.0045	±0.0031	±0.1837	±0.1834
LVQ + varPSO	Denied	**0.6136** ±0.0046	0.0996 ±0.0044	0.7894	5.3535	4.9001
	Accepted	0.1109 ±0.0043	**0.1758** ±0.0045	±0.0023	±0.2650	±0.3441

5 Conclusions

In this article, a new method for obtaining classification rules that have fuzzy variables in their antecedent has been presented. These rules were applied to credit risk analysis using microeconomic and macroeconomic variables.

The method introduces fuzzy logic, and it is based on combining a competitive neural network with an optimization technique based on variable population particle swarms. To build antecedents that are easier to understand and improve model accuracy significantly, fuzzy (microeconomic and macroeconomic) variables were used at the beginning of the process, and a fuzzy vote criterion was defined that uses the degree of membership of fuzzy variables, directly affecting the speed vector that controls particle movement.

The FRvarPSO method, the same as its previous version, LVQ+varPSO, still offers a set of low cardinality rules that are easy to interpret thanks to the reduced length of their antecedents. The improvement obtained with this new version lies in an increased rule set accuracy. The variables available for credit scoring are fuzzy in nature, which is why the use of fuzzy logic in the method allows obtaining better results, as shown in Table 7.

The results obtained allow stating that fuzzy logic, used to guide particle movement, which is reflected on the fitness function, significantly improved rule set accuracy, in addition to making rules easier to understand since, by using fuzzy variables, they are not limited by hard values.

To check method performance, two bank databases and one database from a credit and savings cooperative from the Ecuadorian financial system were used, applying the fuzzy logic approach, dynamically analyzing clients, and incorporating expert input to decrease risk when granting credits.

Table 2: Results obtained with fuzzy rules using macroeconomic and microeconomic variables with a database from a credit and savings cooperative in Ecuador in Segment 2 of the Superintendency of Popular and Solidarity Economy, with assets in the range between 20,000,000.00 and 80,000,000.00

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + fuzzy PSO	Denied	**0.6341** ±0.0048	0.1167 ±0.0045	0.7857	5.9874	7.733
	Accepted	0.0976 ±0.0036	**0.1516** ±0.0045	±0.0026	±0.2429	±0.3362
SOM + fuzzy varPSO	Denied	**0.6235** ±0.0034	0.087 ±0.0030	0.7891	5.692	7.9917
	Accepted	0.1238 ±0.0033	**0.1656** ±0.0032	±0.0021	±0.2245	±0.2953
LVQ + fuzzy PSO	Denied	**0.6571** ±0.0031	0.1122 ±0.0029	0.7922	5.4958	7.624
	Accepted	0.0955 ±0.0036	**0.1351** ±0.0027	±0.0029	±0.2128	±0.2888
LVQ + fuzzy varPSO	Denied	**0.6499** ±0.0021	0.1145 ±0.0017	0.7988	5.299	8.3969
	Accepted	0.0867 ±0.0016	**0.1489** ±0.0028	±0.0029	±0.1907	±0.3913

Table 3: Results obtained with rules using macroeconomic and microeconomic variables with the database of a bank in Ecuador that grants microcredits and is specialized in mass credit placement.

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + PSO	Denied	**0.7973** ±0.0019	0.0167 ±0.0019	0.9688	7.4939	5.8892
	Accepted	0.0044 ±0.0025	**0.1815** ±0.0020	±0.0033	±0.2928	±0.3901
SOM + varPSO	Denied	**0.8005** ±0.0028	0.0114 ±0.0031	0.97	7.0936	5.9966
	Accepted	0.0085 ±0.0029	**0.1795** ±0.0027	±0.0026	±0.2642	±0.3005
LVQ + PSO	Denied	**0.802** ±0.0035	0.0105 ±0.0031	0.9737	7.6993	5.6234
	Accepted	0.0058 ±0.0031	**0.1817** ±0.0032	±0.0037	±0.2307	±0.3866
LVQ + varPSO	Denied	**0.8032** ±0.0024	0.0076 ±0.0021	0.9778	7.5894	5.1209
	Accepted	0.0046 ±0.0018	**0.1846** ±0.0015	±0.0018	±0.1880	±0.3032

Table 4: Results obtained with fuzzy rules using macroeconomic and microeconomic variables with the database of a bank in Ecuador that grants microcredits and is specialized in mass credit placement.

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + fuzzy PSO	Denied	**0.8028** ±0.0036	0.0086 ±0.0049	0.9819	6.993	5.7707
	Accepted	0.0065 ±0.0032	**0.1821** ±0.0030	±0.0037	±0.2121	±0.3480
SOM + fuzzy varPSO	Denied	**0.8038** ±0.0034	0.0071 ±0.0032	0.984	6.1888	5.6533
	Accepted	0.0089 ±0.0029	**0.1802** ±0.0035	±0.0032	±0.2828	±0.2547
LVQ + fuzzy PSO	Denied	**0.8024** ±0.0016	0.0076 ±0.0011	0.9869	6.7588	5.9593
	Accepted	0.0034 ±0.0019	**0.1865** ±0.0023	±0.0024	±0.2642	±0.2997
LVQ + fuzzy varPSO	Denied	**0.9037** ±0.0023	0.0089 ±0.0024	0.988	6.3972	5.1549
	Accepted	0.0031 ±0.0027	**0.0843** ±0.0029	±0.0026	±0.1915	±0.2048

Table 5: Results obtained with rules using macroeconomic and microeconomic variables with the database of a bank in Ecuador that grants consumption and productive or business credits.

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + PSO	Denied	**0.6634** ±0.0019	0.1091 ±0.0019	0.8133	8.4994	6.7538
	Accepted	0.0775 ±0.0015	**0.1499** ±0.0017	±0.0024	±0.1907	±0.2547
SOM + varPSO	Denied	**0.6675** ±0.0038	0.1057 ±0.0035	0.8264	8.2951	6.5513
	Accepted	0.0679 ±0.0031	**0.1589** ±0.0034	±0.0041	±0.2121	±0.3928
LVQ + PSO	Denied	**0.6523** ±0.0015	0.1221 ±0.0021	0.8356	6.4926	6.2304
	Accepted	0.0522 ±0.0015	**0.1733** ±0.0032	±0.0028	±0.3145	±0.3854
LVQ + varPSO	Denied	**0.667** ±0.0034	0.1118 ±0.0022	0. 8415	6.2801	5.9569
	Accepted	0.0566 ±0.0027	**0.1645** ±0.0028	±0.0035	±0.3536	±0.2601

Table 6: . Results obtained with fuzzy rules using macroeconomic and microeconomic variables with the database of a bank in Ecuador that grants consumption and productive or business credits.

Method	Prediction type	Denied	Accepted	Precision	#rules	Antecedent length
SOM + fuzzy PSO	Denied	**0.6723** ±0.0019	0.1023 ±0.0012	0.8316	6.4415	6.5538
	Accepted	0.0721 ±0.0015	**0.1533** ±0.0027	±0.0022	±0.1814	±0.3469
SOM + fuzzy varPSO	Denied	**0.667** ±0.0038	0.0999 ±0.0022	0.8385	6.399	6.2832
	Accepted	0.0016 ±0.0025	**0.1615** ±0.0034	±0.0033	±0.2099	±0.3496
LVQ + fuzzy PSO	Denied	**0.6599** ±0.0025	0.0744 ±0.0019	0.8455	6.6904	6.1818
	Accepted	0.0945 ±0.0021	**0.1711** ±0.0032	±0.0029	±0.3021	±0.3176
LVQ + fuzzy varPSO	Denied	**0.6627** ±0.0034	0.0903 ±0.0028	0. 8501	5.9901	5.9623
	Accepted	0.0669 ±0.0030	**0.1801** ±0.0022	±0.0033	±0.3112	±0.2854

Table 7: Comparison of accuracy between the hybrid models and those that use fuzzy logic.

	Method	Accuracy
Credit and savings cooperative	SOM + PSO	0.7718
	SOM + varPSO	0.7701
	LVQ + PSO	0.7871
	LVQ + varPSO	0.7894
	Fuzzy SOM + PSO	0.7857
	Fuzzy SOM + varPSO	0.7891
	Fuzzy LVQ + PSO	0.7922
	FRvarPSO	**0.7988**
Bank in Ecuador that grants microcredits and is specialized in mass credit placement	SOM + PSO	0.9688
	SOM + varPSO	0.97
	LVQ + PSO	0.9737
	LVQ + varPSO	0.9778
	Fuzzy SOM + PSO	0.9819
	Fuzzy SOM + varPSO	0.984
	Fuzzy LVQ + PSO	0.9869
	FRvarPSO	**0.988**
Bank in Ecuador that grants consumption and productive or business credits	SOM + PSO	0.8133
	SOM + varPSO	0.8264
	LVQ + PSO	0.8256
	LVQ + varPSO	0.8315
	Fuzzy SOM + PSO	0.8316
	Fuzzy SOM + varPSO	0.8385
	Fuzzy LVQ + PSO	0.8455
	FRvarPSO	**0.8501**

The analysis carried out in this article is highly relevant for credit officials, since in addition to the use of microeconomic variables to assess credit scoring, which affect the borrower macroeconomic variables are also used, which take into account the conditions affecting the general environment where the credit is granted, which directly affect delinquency rates.

In future research lines, the use of the degree of membership corresponding to credit holder analysis as a condition for meeting a rule should be considered as an addition to the model, as a way of assessing current credit risk in granted credits. We will also explore how sudden changes in macroeconomic variables (e.g. a large financial crisis) could affect the model's accuracy.

References

References

[1] X. Wang, M. Xu and O.T. Pusatli, A Survey of Applying Machine Learning Techniques for Credit Rating: Existing Models and Open Issues, in: *Neural Information Processing 22nd International Conference ICONIP*, eds. S. Arik, T. Huang, W.K. Lai and Q. Liu (Springer Publishing Company, Cham, 2015), pp. 122-132.

[2] L. C. Thomas, A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers, *Int. J. Forecasting* **16**(2) (2000) 149–172.

[3] C. Aggarwal, *Data Mining: The Textbook* (Springer Publishing Company, Heidelberg, 2015).

[4] J. Reinke, How to Lend like Mad and make a profit: A Paradigm Micro-credit versus the Start-Up Fund in South Africa, *J. Dev. Stud.* **34** (1998) 44–61.

[5] M. Zeller, Determinants of Repayment Performance in Credit Groups: The Role of Program Design, Intra-Group Risk Pooling, and Social Cohesion, *Econ. Dev. Cult. Change* **46** (1998) 599–620.

[6] Q-L. Chen and J-B. Lin, Integrating Business Intelligence and CRM of in Banks: An Empirical Study of SOM Applied in Customer Personal Loans in Taiwan, in *2015 International Conference on Fuzzy Theory and Its Applications, iFUZZY* (Yilan, 2015), pp. 68–73.

[7] H. A. Abdou, . and J. Pointon, Credit scoring, statistical techniques and evaluation criteria: A review of the literature, *Intell. Syst. in Acc. Fin. Management* **18**(2011) 59–88.

[8] N. Mahmoudi and E. Duman, Detecting credit card fraud by Modified Fisher Discriminant Analysis,*Expert Syst. Appl.* **42** (2015) 2510–2516.

[9] . G. Daqi, D. Ahmed , G. Lili , W. Zejian and W. Zhe, Pseudo-inverse linear discriminants for the improvement of overall classification accuracies, *Neural Networks* **8** (2016) 59–71.

[10] G. Donga, K. Keung and J. Yenc, Credit scorecard based on logistic regression with random coefficients, *International Conference on Computational Science, ICCS* (2010), pp. 2463–2468

[11] G. Nie, W. Rowe, L. Zhang, Y. Tian and Y. Shi., Credit card churn forecasting by logistic regression and decision tree, *Expert Syst Appl.* **38** (2011) 15273–15285.

[12] J.R. Quinlan *C4.5: programs for machine learning* (Morgan Kaufmann Publishers, San Francisco, 1993).

[13] E. Frank and I.H Witten, Generating accurate rule sets without overall optimization,*Proc. 15th International Conference on Machine Learning, ICML '98* (1998), pp. 144–151.

[14] Z. Li, Y. Tian, K. Li, F. Zhou and W. Yang, Reject inference in credit scoring using Semi-supervised Support Vector Machines. *Expert Syst. Appl.* **74** (2017) 105–114.

[15] Z. Zhao, S. Xu, B. H. Kang, M. M. J. Kabir, Y. Liu and R. Wasinger, Investigation and improvement of multi-layer perceptron neural networks for credit scoring, *Expert Syst. Appl.* **42** (2015) 3508–3516.

[16] S. Lessmann, B. Baesens, H. V. Seow and L.C. Thomas, Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research, *Eur. J. Oper. Res.* **247** (2015) 124–136.

[17] G. Wang and J. Ma, A hybrid ensemble approach for enterprise credit risk assessment based on support vector machine, *Expert Syst. Appl.* **39** (2012) 5325–5331.

[18] C. C. Wong and C. C. Chen, A GA-based method for constructing fuzzy systems directly from numerical data, *IEEE Trans. on Systems, Man and Cybernetics—Part B: Cybernetics*, **30**(6) (2000) 904–911.

[19] M. Sugeno and T. Yasukawa, A fuzzy-logic-based approach to qualitative modeling, *IEEE Trans. on Fuzzy Systems*, **1** (1993) 7–31.

[20] P. K Simpson, Fuzzy min-max neural networks—Part 1: Classification, *IEEE T. Neural Networ.*,**3** (1992) 776–786.

[21] L. Lanzarini, A. Villa Monte, A. Aquino and G. De Giusti, Obtaining classification rules using lvqPSO, in: *Advances in Swarm and Computational Intelligence*, eds. Y. Tan, Y. Shi, F. Buarque, A. Gelbukh, S. Das, A. Engelbrecht (Springer-Verlag, Berlin, 2015), pp. 183–193.

[22] L. Lanzarini L, A. Villa-Monte, A.F. Bariviera and P. Jimbo Santana, Obtaining Classification Rules Using LVQ+PSO: An Application to Credit Risk, in: *Scientific Methods for the Treatment of Uncertainty in Social Sciences*, eds. J. Gil-Aluja, A. Terceno-Gómez, J. Ferrer-Comalat, J. Merigó-Lindahl, S. Linares-Mustarós (Springer, Cham, 2015), pp. 383–391.

[23] L. Lanzarini, A. Villa-Monte, A.F. Bariviera and P. Jimbo Santana, Simplifying rules using credit scoring LVQ+PSO, *Kybernetes* **46** (2017) 7–16.

[24] P. Jimbo Santana, L. Lanzarini, A. F. Bariviera, Extraction of Knowledge with Population-Based Metaheuristics Fuzzy Rules Applied to Credit Risk, in: *Advances in Swarm Intelligence. ICSI 2018*, eds. Y. Tan, Y. Shi and Q. Tang (Springer, Cham, 2018), pp. 153-163.

[25] S. Oreski and G. Oreski, Genetic algorithm-based heuristic for feature selection in credit risk assessment, *Expert Syst. Appl.* **41** (2014) 2052–2064.

[26] M. Zieba, S. K. Tomczak and J. M. Tomczak, Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, *Expert Syst. Appl.* **58** (2016) 93–101.

[27] D. Clements, "Fuzzy models for computer security system metrics", Ph.D. thesis (Univ of California at Berkeley, 1977).

[28] Ch-Ch. Chen, A PSO-based method for extracting fuzzy rules directly from numerical data, *Cybernet. Syst.* **37**(2006) 707–723.

[29] R. R. Yager, L. A. Zadeh (editors). *An Introduction to Fuzzy Logic Applications in Intelligent Systems* (Springer, New York, 2012).

[30] I. Witten, E. Frank, M. Hall and C. Pal. *Data Mining - Practical Machine Learning Tools and Techniques, 4th Edition* (Morgan Kaufmann Publishers, San Francisco, CA, 2016).

[31] J.-R. Jang, ANFIS: adaptive-network-based fuzzy inference system, *IEEE T. Sys. Man Cyb.* **23** (1993) 665–685.

[32] A. Fernandez, M.J. del Jesús and F. Herrera, Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, *Int. J. Approx. Reason.* **50** (2009) 561–577.

[33] E. Hüllermeier, Does machine learning need fuzzy logic?, *Fuzzy Set. Syst.* **281** (2015) 292–299.

[34] T. Kohonen *Self-Organizing Maps* (Springer, Heidelberg, 2012).

[35] L. A. Zadeh, Fuzzy sets, *Inform. Control*, **8** (1965) 338–353.

[36] R. Bělohlávek, J. W. Dauben and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective* (Oxford University Press, Oxford, 2017).

[37] K. V. Shihabudheen and G. N. Pillai, Recent advances in neuro-fuzzy system: A survey, *Knowl-Based Syst.*, **152** (2018) 136–162.

[38] A. Bagheri, H. Mohammadi and M. Akbari, Financial forecasting using ANFIS networks with quantum-behaved particle swarm optimization, *Expert Syst. Appl.*, **41** (2014) 6235–6250.

[39] J. C. Dunn, A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, *J. Cybernetics*, **3** (1973) 32–57.

[40] J. C. Bezdek *Pattern Recognition with Fuzzy Objective Function Algoritms* (Plenum Press, New York, 1981).

[41] J. Kennedy and R. Eberhart, Particle swarm optimization, *IEEE Int. C. Neural Networ.*, **4** (1995) 1942–1948.

[42] J. Sadri and Ch. Y. Suen, A genetic binary particle swarm optimization model, *2006 IEEE C. Evol. Comput.*, (Vancouver, BC,2006), pp. 656–663.

[43] A.P.Engelbrecht, G. Pampara and N. Franken, Combining particle swarm optimization with angle modulation to solve binary problems, *2005 IEEE C. Evol. Computat.* (Edinburgh, Scotland, 2005), pp. 89–96.

[44] A. Marandi, F. Afshinmanesh, M. Shahabadi and F. Bahrami, Boolean Particle Swarm Optimization and Its Application to the Design of a Dual-Band Dual-Polarized Planar Antenna, *2006 IEEE C. Evol. Computat.* (Vancouver, BC, 2006), pp. 3212–3218.