RESEARCH ARTICLE

AN EXPERT SYSTEM BASED ON FUZZY LOGIC FOR AUTOMATED DECISION MAKING FOR LOAN APPROVAL

Banty Kumar, *Indranil Bawane, Aishwarya Shirsathe and Prashant Pardeshi

Department of Computer Engineering, MIT College of Engineering, Pune, India

ARTICLE INFO	ABSTRACT

At present, bad debt has become a stubborn enemy of financial industry in every country. So there arises a need of a technique which will identify the risk in loan approval. Fuzzy logic is a computing model which deals with uncertainty rather than the traditional True (1) and False (0) values and has caught some eyes in recent years. Fuzzy logic is a valid and novel means of problem solving with intelligent systems which deals with imprecision, vagueness and uncertainty in the financial applications like loan processing. In this paper, an expert system based on fuzzy logic has been proposed for approval of loans. Process starts with a set of observed input variables that are taken into the system. These inputs are assigned a degree of membership to one, or more, fuzzy sets known as membership functions. These membership functions are then tied to output membership functions by linguistic fuzzy patches known as rule sets.

INTRODUCTION

Fuzzy logic is a computational model in which the input variables are classified in terms of linguistic variables which helps in more human-like reasoning. Traditional programming model uses binary logic where things are either zero or one i.e., True or False. The variables can't take any values out of this binary set but in Fuzzy logic a variable has a membership i.e. a variable does not belong to one particular set but is a member of more than one set. This model helps to understand the real life problems more closely because nothing in this world is one thing or another but they are the combination of many things in continuum. Since the introduction of Fuzzy logic by Zadeh in 1965, it has been successfully implemented in many fields like Medical, Robotics, Decision making in Artificial Systems, operational research and many more applications in almost every field of computing. In this paper we try to discuss the use of fuzzy logic in loan risk identification and credit scoring. Fuzzy logic is especially useful in this field because there are many vague and imprecise variables involved in the assessment of loan risk and the process of classification of borrower into profitable and non-profitable categories is ambiguous in itself because of lack of a clear boundary. Keeping the above point in mind it would be fruitful to apply fuzzy logic rules in this context. As credit risk is increasing gradually and more number of banks and financial institutions are facing huge losses due to inability of recovering loan interest in time, there is a need to improve upon the traditional system of loan analysis than can handle the uncertainties and vagueness more precisely and help the banks to take an intelligent decision.

*Corresponding author: Indranil Bawane,
Department of Computer Engineering, MIT College of Engineering,
Pune, India.

This paper discusses a model in which the fuzzy set and fuzzy rules are applied in calculating the credit score of a borrower and calculating his/her credit-worthiness for sanctioning the loan. The paper is organized as follows: Section II is about literature survey Section III discusses the general implementation of fuzzy logic. In section IV sample processing of an application is given. Section V gives a comparison of fuzzy system with other related technologies. Section VI discuss the future scope of and concludes the paper.

Literature Survey

Various methods have been proposed for the issue of appropriate decision making in loan approval. Some of the primary techniques are use of Artificial Neural Networks, Decision Trees, SVM and the logistic regression. This section highlights some of the advantages and disadvantages of those techniques. To begin with, the model based on artificial neural networks used all the training samples and considered the misclassification costs in order to give a final decision. This technique gives good prediction, however is slower due to the back propagation process. It requires validation and testing (MaRuowei, Yang Deyong *et al.,* 2010). The decision tree model uses the C4.5 algorithm and works by using a training dataset and deciding the partition attribute using the gain ratio.

This method gives a precision of up to 78%. However, the accuracy depends on the quality of the training data. Also large trees might be generated which further requires pruning to simplify them (Rafik Amin *et al*., 2015). Another technique uses the logistic regression and RVM in combination. It gives good performance when used with the RVM, but is useful to classify data in two categories only (Qinrong Meng, 2011).

The application of SVM works better than the regression and neural networks. Here, the type of classification needs to be identified and accordingly linear or non-linear SVM is applied. The error rate depends on the choice of parameters. It tries to reduce the error rate but at the cost of iterative processing (YinxiaoMa and Hong Liu, 2010). All these techniques, in general show that the data is classified in two groups. Hence there becomes a necessity to ignore the strength of certain parameters and classify them into either of the groups. Fuzzy logic approach, on the other hand, considers the membership of all parameters to each of the output classes and gives decision accordingly. Such type of decision making is very closely related to the human thinking.

Proposed Methodology

The process consists of the following basic steps:

- Extract the necessary parameters
- Assign linguistic labels
- Define membership functions
- Inference using rules
- Combine outputs of all the rules
- Defuzzify the output
- Generate the advice

Extract the necessary parameters

Loan is processed based on some parameters. The bank collects the client's information and requires it to be verified. For this purpose, it consults third party entities too. Parameters like name, address, income, credit history, etc. are verified. If all the parameters are found to be authentic, then the loan application is sent for further processing.

In this paper, parameters we have considered are –

- Income
- Expenditure/Income
- Savings
- Employment Experience
- CIBIL Score
- Loan Purpose
- Duration with bank
- Investment
- Age

Assign Linguistic Labels

Define the range and assign linguistic labels for the classes of each parameter.

Input variables

Income = {low, average, high}
Expenditure/Income = {low, average, high}
Savings = {low, average, high}
Employment Experience = {no, low, average, high}
CIBIL Score = {low, average, high}
Loan Purpose = {Car(new), Car(old), Bike, Vacation, other}
Duration with bank = {short, medium, long}
Investment = {low, average, high}
Age = {young, adult, senior}

Output variable

Advice = {risky, less risky, advisable}

Define membership functions

Considering the problem scenario and parameters involved, following membership functions are found to be appropriate

a. Triangular Functions
b. Trapezoidal Functions

Taking an example of the parameter age
Age is assigned the labels as: young, adult and senior. Each of these labels can be assigned a suitable membership function.
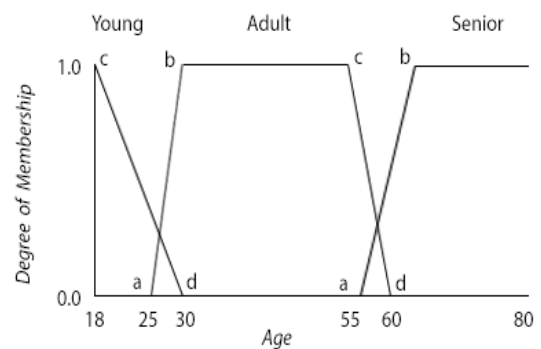

Fig. 1. Trapezoidal MF

Fig. 2. Triangular MF
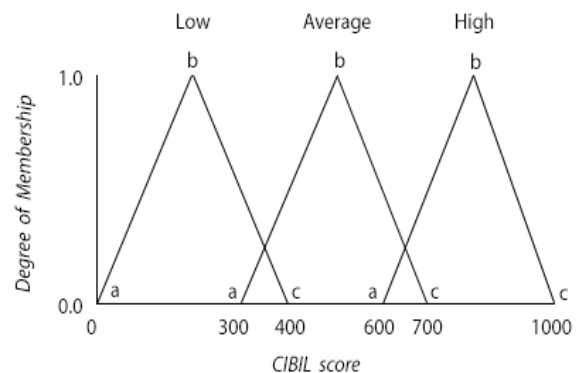
Degree of membership for the inputs can be calculated as

Triangular MF -
Triangular (x: a,b,c) =
$$
\begin{cases}
0 & x < a \\
(x-a)/(b-a) & a <= x <= b \\
(c-x)/(c-b) & b <= x <= c \\
0 & x > c
\end{cases}
$$

Trapezoidal MF-
Trapezoidal(x: a,b,c,d) =
$$
\begin{cases}
0 & x < a \\
(x-a)/(b-a) & a <= x < b \\
1 & b <= x < c \\
(d-x)/(d-c) & c <= d <= d \\
0 & x >= d
\end{cases}
$$

Inference using rules

Rules used for the inference process have the following structure:

- IF age is YOUNG, income is HIGH, savings is LOW THEN advice is ADVISABLE.
- However here, if all the parameters are to be considered, the rule may become too complex to process. Hence, we can perform the processinglevel by level.
- In first level, four parameters - income, expenditure/income, savings and investment are considered to form the rules. The output produced, circumstance, is considered as input for the next level.

- In the next level, age, employment experience, duration with bank and circumstance are considered to form the rules. The output produced, consistency, is considered as input for the third level.
- Finally the last level will have CIBIL score, loan purpose and consistency as inputs and will generate degree of membership for output labels of advice. The rule base can be formed by consulting experts. Number of experts can be consulted so that accuracy of rules can be improved.

Combining outputs of all the rules

The output degrees of all of the fuzzy rules are combined to obtain one fuzzy output distribution. This is done by using the ORing the output from each rule.
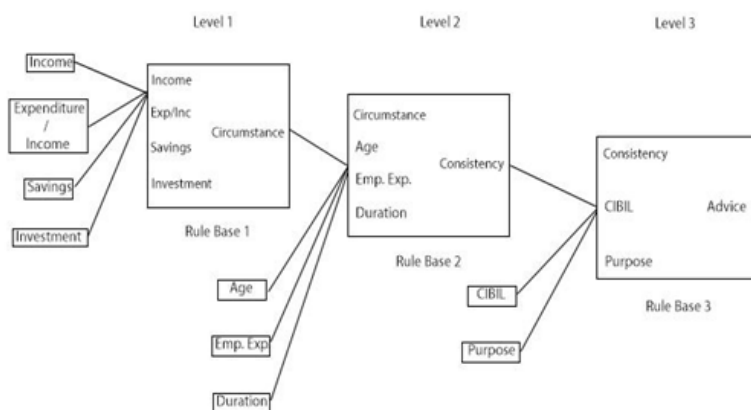

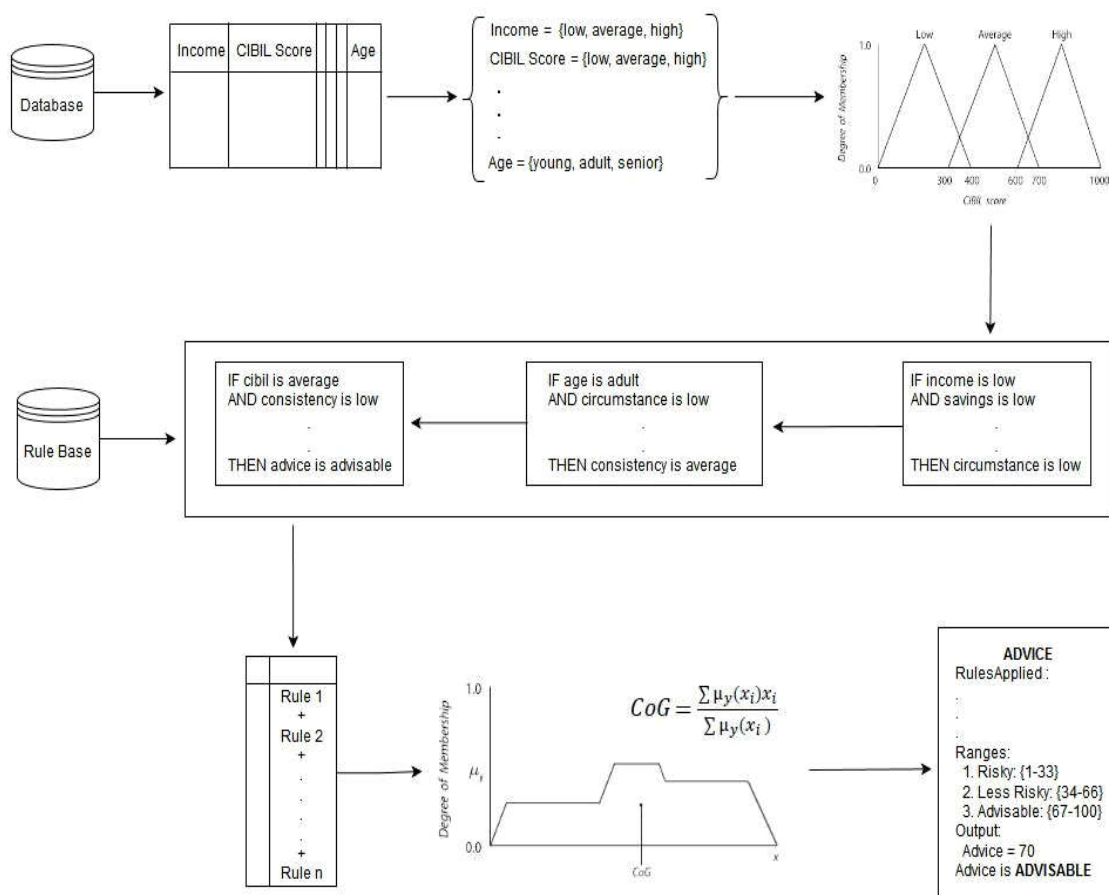**Fig. 3. Level wise Architecture of System**

$$CoG = \frac{\sum \mu_y(x_i)x_i}{\sum \mu_y(x_i)}$$

Fig. 4. Architecture of System

Defuzzify the output

A crisp number is obtained in this process. The Centroid Defuzzification method can be used to find the crisp output which provides the decision to the loan officer. The centroid is represented mathematically as

CoG = (S(m.x))/(Sm)

where,

S= Summation
m= membership value in the membership function
x = center of membership function.

Generate the advice

Output classes - less risky, advisable, and risky - can have predefined ranges. The advice is that class to which crisp number belongs.

Sample Processing of an Application

Consider an application where the parameters have following given values along with their membership in the classes of (low, average, high) respectively. (In case of duration, (short, medium, long)).

Income =   3 lpa   (0.7, 0.2, 0)
Expenditure/Income = 0.7   (0, 0.6, 0.1)
Savings =   30,000pa   (0.7,0.1,0)
Employment
Experience =   2 years   (0.8,0.2,0)
CIBIL Score =   800   (0,0.1,0.8)
Purpose =   Car(new)   (1, 0, 0, 0, 0)
Duration with bank = 6 years   (0,1,0)
Investment =   60,000pa   (0,0.7,0.3)
Age =   26 years   (0.9,0.1,0)

1. Considering we have the following rule in our rule base for level 1:

IF income is low AND exp/inc ratio is average AND savings is low AND investment is average THEN circumstance is average. Now, we consider the label with the maximum membership for each of the parameter (Fuzzy OR operation). Hence this rule can be applied. The membership found for circumstance is say, (0.2,0.5,0). Similarly, if we apply another rule for level 2 as: IF age is young AND emp_exp is low AND duration is medium AND circumstance is average THEN consistency is average. Say the membership for consistency is found as (0.1,0.6,0). And finally the third rule which we have in our rule base could be applied as: IF cibil is high AND purpose is car (new) AND consistency is average THEN advice is advisable. Thus, the applicant's request for loan would be approved.

Comparison of Fuzzy System with Traditional System For Loan Processing

The Traditional Approach

- The customer information is collected from the customer himself in the form of an application.

- The provided information is verified using the intra-bank and inter-bank softwares as well as consulting various third party firms.
- The bank has some rules laid down in granting of loan approval. If any of the criteria is not met, the loan is not approved.
- The rules are strictly defined rules. For example, the income ranges are considered from Rs. x to Rs. y in granting the loan for a specific purpose. That is, strict ranges are considered.
- This entire evaluation of analyzing the application is done manually by the loan officer. Thus time taken for the analysis is long and varies from person to person.
- There is a possibility that a wrong decision is given and a customer is rejected for loan inspite of being capable to repay.

The Fuzzy Approach

- The initial two steps followed in traditional processing are followed in this approach too. Since the verification involves a lot of activities where humans can't be replaced (For example: visiting the customer's house or his shop or business place or the construction site in case of home loans etc.), these steps are recommended not to be automated.
- If any of the condition is not met, the loan is not rejected directly. It checks the repayment ability considering the magnitude of effect of all the parameters. (By using the membership functions)
- There are no strict ranges for classifying the parameters. It considers the belongingness of an input parameter to each of the range, thus simulating real life human thinking. (By using the membership functions)
- This analysis is done by the system not involving any human interference. It would certainly save our time.
- Rules are collected consulting a number of experts. Hence the error rate is reduced greatly.

Conclusion and Future Scope

The proposed system tries to make the loan approval system more efficient by reducing the time involved in processing as well as the risk associated with erroneous decisions. It thus helps in improving the customer's and the bank's satisfaction.

Future Scope

This system is developed for determining the worthiness of an individual for personal loan approval only. By considering more parameters, the system can be extended to make an approval system for all types of loans combined. However, with an increase in number of parameters there is a possibility of affecting the efficiency of system. Though a level wise approach is used, the number of level may increase, which needs to be taken care of.

REFERENCES

Lotfi A. Zadeh, 1996. IEEE transactions on fuzzy systems, VOL. 4, NO. 2, MAY 103, Fuzzy Logic = Computing with Words, Life Fellow, IEEE
MaRuowei and YangDeyong, 2010. Building Default Predicting Model on Firm's Short-term Loan Data with

Artificial Neural Network - Considering Qualitative Indexes and Misclassification Costs, *International Conference on E-Business and E-Government.*

Rafik Amin, Indwiarti and Yuliant Sibaroni, 2015. Implementation of Decision Tree Using C4.5 Algorithm in Decision Making of Loan Application by Debtor (Case Study: Bank Pasar of Yogyakarta Special Region), 3rd *International Conference on Information and Communication Technology* (ICoICT)

Qinrong Meng, 2011. Evaluation of Residential Loan by Combining RVM and Logistic Regression, IEEE

YinxiaoMa and Hong Liu, 2010. Research of SVM Applying in the Risk of Bank's Loan to Enterprises, IEEE

Keshav Dahal, Zabeer Hussain, M. Alamgir Hossain, 2005. "Loan Risk Analyzer based on Fuzzy Logic", IEEE International Conference,

Umar Farouk Ibn Abdulrahman, Joseph Kobina Panford, James Ben and Hayfron Acquah, 2014. "Fuzzy Logic Approach to Credit Scoring for Micro Finances in Ghana (A Case Study of KWIQPLUS Money Lending*)", International Journal of Computer Applications*.

****