Detection of Loan Risk Using Fuzzy Association Rules

Tohida Rehman

Jadavpur University, Kolkata, India

ABSTRACT

Nowadays, loans put a lender at risk by granting credit to someone with a poor credit score. Furthermore, the number of transactions in the banking sector is continually increasing, and vast data volumes are available that represent consumer behaviour and loan risks. The vast volume of data and information that must be processed and integrated necessitates the use of specific approaches and technologies. As a result, a financial organization's or bank's selection of a new loan application is a critical issue. To predict credit risk patterns, banks have traditionally employed static models incorporating demographic or static characteristics. There are numerous aspects that influence the decision to grant a loan to an application. But employing the fuzzy association rules, Banks and other private financers can be selecting the best applicant with the least amount of risk during the loan process. We invoked the expert knowledge who are specialists in the finance domain to design all fuzzy rules. There are no hard and fast rules to calculate the accurate loan risk. Hence, Fuzzy association rules can provide the best tools to invoke human Knowledge, human perception, vagueness, and imprecision to predict the exact selection result.

Keywords : Membership Function, Fuzzy Logic, Fuzzy controller, Expert Knowledge, Loan, Risk Assessment.

I. INTRODUCTION

A commercial bank, sometimes known as a bank, is a type of financial organization that accepts deposits, makes business loans, and offers basic investment products. Because these are economic factors, banks must invariably assume risks when providing loans to customers. When it comes to borrowing money, there are many different sorts of loans to choose from, and it's vital to be aware of your alternatives, like profitability or high risk. The practice of evaluating loan collections and assigning loans to groups or grades based on perceived danger and other related loan features is known as loan classification. Sometimes, approved loan to a customer gets in a wrong way to recovery the loan amount. In practice, all banking institutions may use different intuitions to select their customer for granting a loan but some factors are common for all banks. We use the MATLAB fuzzy logic tool to assess a new loan applicant's age, total assets, employment stability, no of years working experience, total income, loan amount, loan period, among other things(Like credit score).

II. RELATED WORK

The definition of fuzzy in the Longman Dictionary of Contemporary English is "inaccurate & ambiguous". Aristotle's law has influenced our philosophical sense of what is true and what is wrong for almost two thousand years. Professor Lotfi A. Zade, proposed fuzzy theory in 1960 to explain confusing and fuzzy real-world occurrences [1]. Unlike Boolean logic, which is based on zeros and ones or True and False, fuzzy logic is based on an element's degree of

membership in a fuzzy set specified by a membership function [8]. A crisp set's characteristic function assigns a value of 0 or 1 to each individual in the universe set, allowing members and non-members of the crisp set to be distinguished [1, 2]. MATLAB fuzzy logic tools can be used to implement a variety of concepts. The fuzzy logic tool assists in meeting new technological advancements [3]. Fuzzy set theory is a mathematical tool for modelling the inaccuracy and ambiguity of the real world and human thought. The risk to granting loan to the customer can be assessed using fuzzy set theory. Many risk analysts have been drawn to fuzzy logic because of its popularity. The advantage of the fuzzy method is that it allows for the processing of variables with ambiguous definitions and variables whose relationships are not mathematically defined. Fuzzy logic is a way to make machines more intelligent to deal with uncertain, imprecise or qualitative decision-making problems like humans [5]. Fuzzy logic has been employed in handling inexact and vague information because of its ability to utilize natural language in terms of linguistic variables in different applications including banking and finance [4]. Banks and financial institutions confront credit risks to conduct their business [6]. Fuzzy Expert System is designed based on the selected rules from association rules to specify the credit degree of banks' customers [7].

III. PROPOSED WORK

The most significant aspect in the process of developing a loan risk prediction tool is to cover the characteristics that influence a financial company's choice on new loan applications. As a result, we have picked eight separate parameters that have the most impact on loan decisions. With discussion of experts, we have considered these important parameters such as loan applicant's age, total assets, employment stability, no of years working experience, total income, loan amount, loan period, among other things (credit score). In the Matlab toolbox named

Fuzzy Logic Toolbox, there are mainly "Mamdani" and "Sugeno" types of mostly used fuzzy inference systems. So, in this study, we employed the "Mamdani" kind of FIS(fuzzy inference systems) to design a fuzzy logic controller to assess the likelihood of customer loan applications. Apply fuzzification on input values, applications of fuzzy operators on the antecedent part of fuzzy rules, evaluation of fuzzy rules, aggregation of fuzzy sets across rules, i.e. combining the evaluation result and the defuzzification of the aggregated fuzzy set to generate crisp output values are the core five processes in designing any fuzzy logic controller. Here, we have used eight input parameters and one output parameter. Here, I assumed that the range of all membership functions and the input and output variables are [0, 100]. Each input and output variable has five membership functions. Here I have used the "Triangular membership function" to fuzzify the input values. Input-wise membership functions are given in table 1.

IV. DESIGN METHODOLOGY

There are numerous other factors that influence whether or not to select an applicant for a loan. As a result, selecting the most risk-free candidate to give the loan is a difficult task. Using some expert knowledge in financial domain, all important factors are group together into eight input variables and one output variable. Output variable is considered as a percentage of risk to give a loan to a customer. The impact of risk can be Very Low, Low, Medium, High, and Very High. After describing the inputs and outputs of fuzzy sets in membership functions, rules are used.

We have added a total of 100 rules to predict the risk to grant a loan. In table 2, we used the following abbreviations as, AA-applicant's age, TA-total assets, ES-employment stability, EX-no of years working experience, TI-total income, LA-loan amount, LP-

loan period, CS-among other things(credit score), CR-
risk to grant a loan. In table 2, we have shown some

parts of the rules.

Table 1 : Input parameters for applicant loan risk

Parameters	Impact details	Explanation of impact
applicant's age	Low, Medium, High	Salary candidates between the ages of 23 and 58 are preferred by lenders.
total assets	Low, Medium, High	A borrower's assets can only be attached when the proper procedure has been followed.
employment stability	Low, Medium, High	As proof of income and stability, previous employment history is also considered.
no of years working experience	Low, Medium, High	If a person has a history of changing jobs frequently, a bank may be hesitant to grant their loan.
total income	Low, Medium, High	When applying for any loan, you should think about how much money you have leftover each month and your ability to make EMI payments.
loan amount	Low, Medium, High	Before accepting your loan amount, banks will look at your entire financial history, repayment capabilities, and other factors.
loan period	Low, Medium, High	Applicants that choose a shorter payback time are frequently preferred by banks.
Among other things(like credit score)	Low, Medium, High	Your credit score is based on your financial history and capacity to repay a loan.

Table 2 : Some rules to predict the risk to grant loan

| #	AA	TA	ES	EX	TI	LA	LP	CS	CR
1	Low	Low	Low	Low	Low	Low	Low	Low	Very High
2	Low	Low	Low	Low	Low	Low	Low	Medium	High
3	Low	Low	Low	Low	Low	Low	Medium	Medium	High
4	Low	Low	Low	Low	Low	Medium	Medium	Low	High
5	Low	Low	Low	Low	Medium	Medium	Medium	Low	Medium
6	Low	Low	Low	Medium	Medium	Medium	Medium	Low	Medium
7	Low	Low	Low	Medium	High	Medium	Medium	Low	Low
8	Low	Low	Low	Medium	High	Low	Medium	Low	Very Low
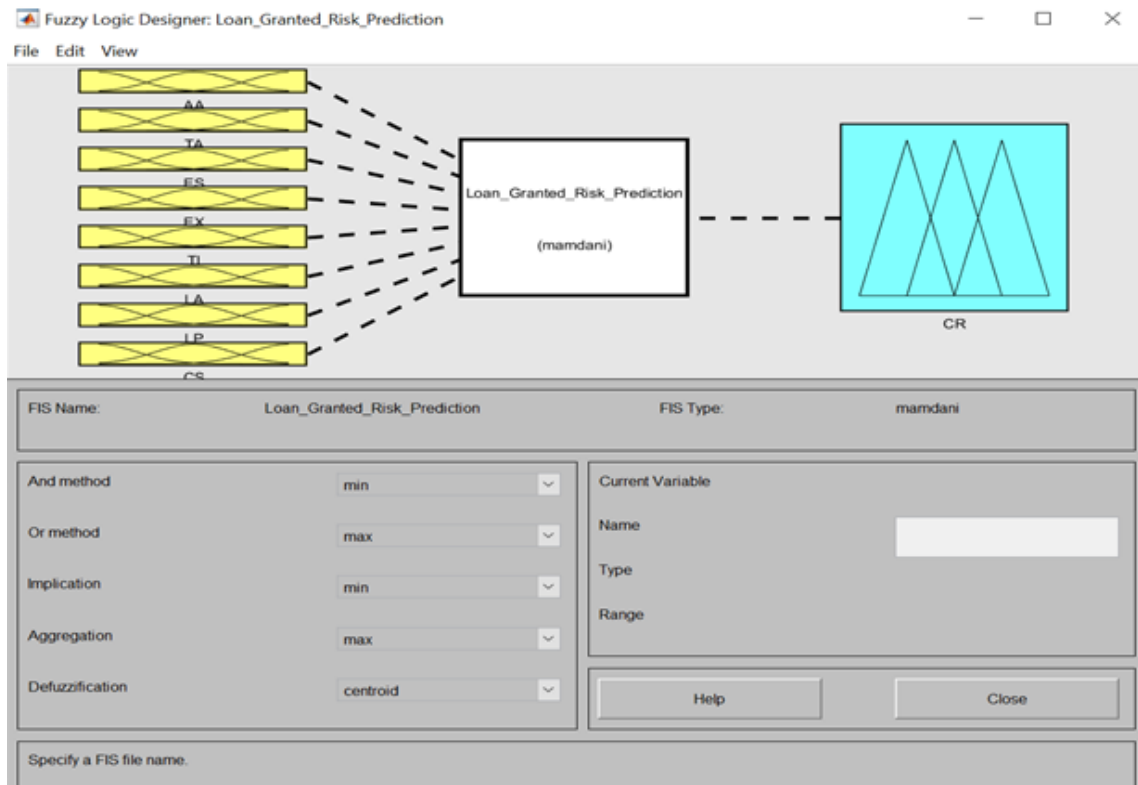
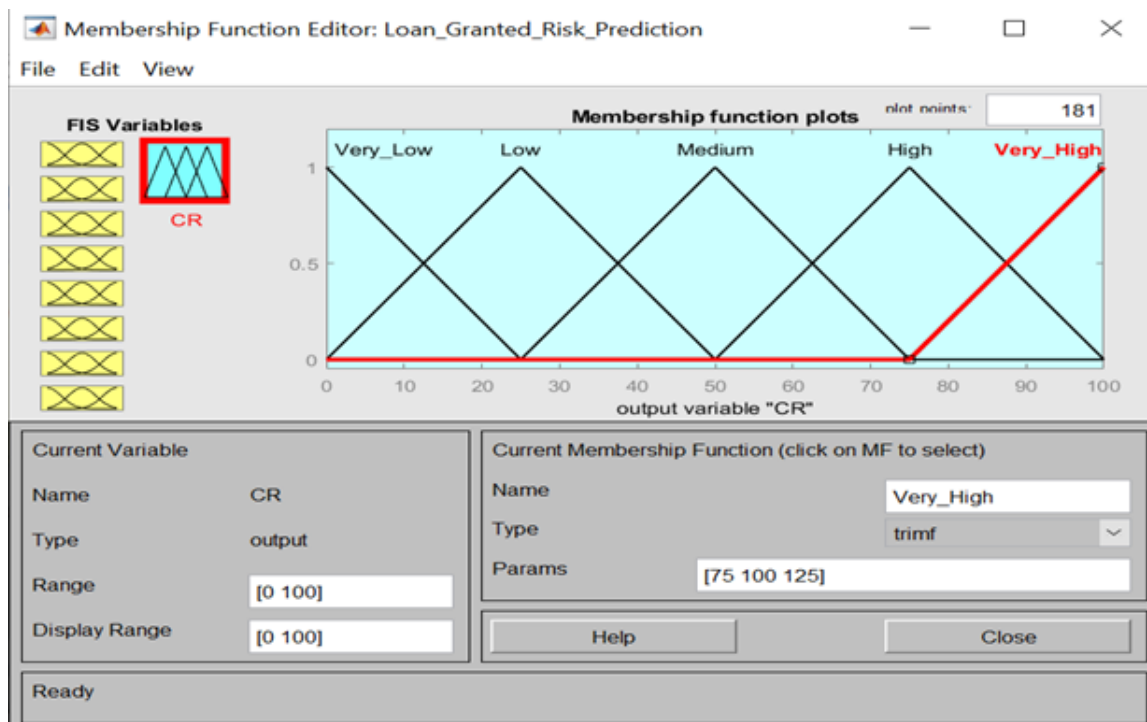Figure 1: Input, Output variables in details

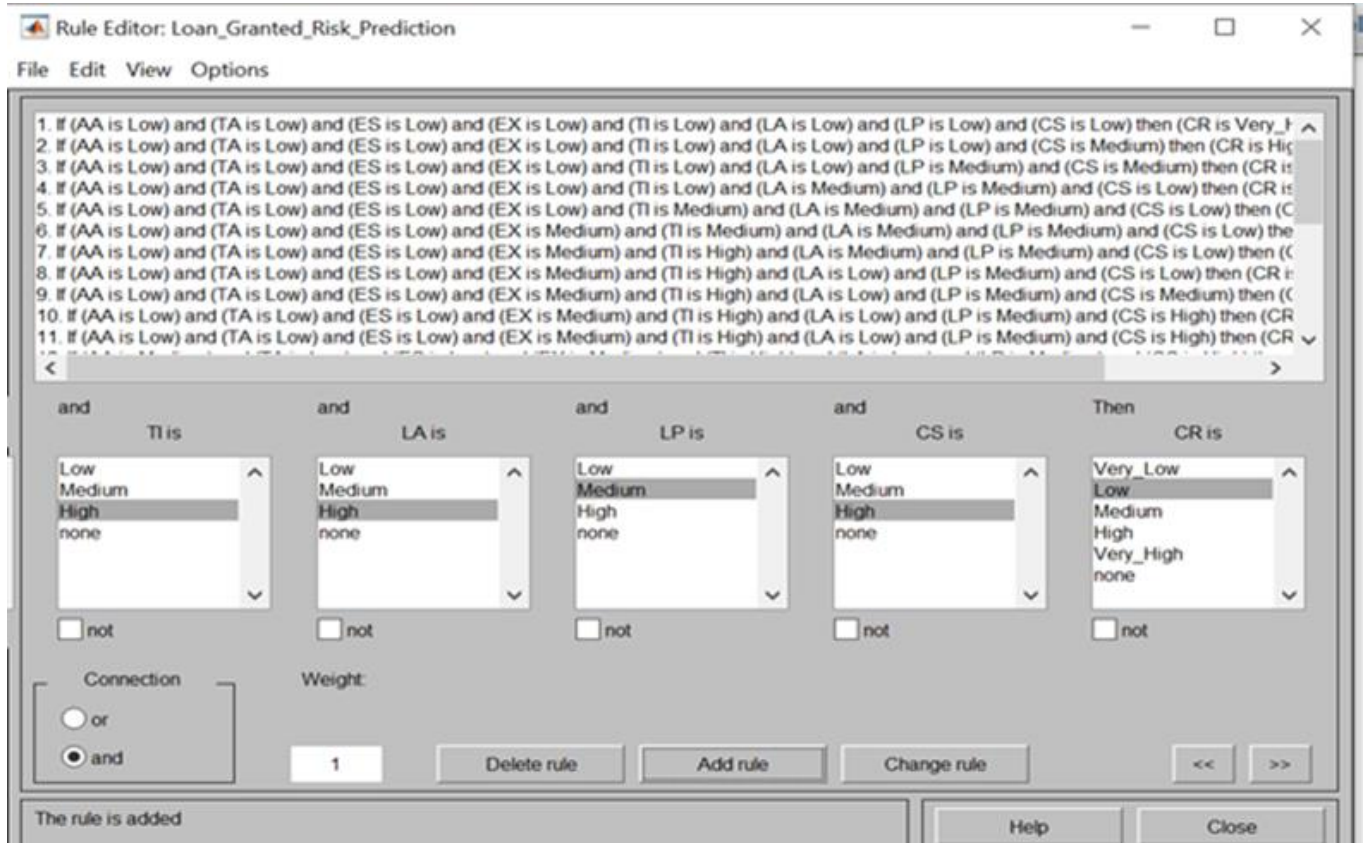Figure 2: Output variable membership functions plot in Matlab editor

Figure 3: Some Fuzzy if-then rules

V. IMPLEMENTATION DETAILS

To implement the loan granting risk prediction we have used fuzzy tool box in Matlab. For common high level issues the Fuzzy inference systems (FIS) Editor can be used. Total number of inputs and outputs can be used to deal it with their names. Here, we have used 8 input variables and one output variable. For each input variable we have taken into consideration of three triangular membership functions with a range of 0 to 100. For output variable we have consider five triangular membership functions with a range of 0 to 100. At the MATLAB prompt, type fuzzy to start the system from the beginning. The generic unnamed FIS Editor appears, with one input, input1, and one output, output1. We modified and added 8 inputs and one output. All unnamed inputs and output are renamed as AA-applicant's age, TA-total assets, ES-employment stability, EX-no of years working experience, TI-total income, LA-loan amount, LP-loan period, CS-among other things(credit score), CR- risk to grant a loan shown in figure 1. When a rule's antecedent is made up of multiple inputs, we must use proper fuzzy operators (AND, OR) to get a single integer that represents the result of the complete antecedent.

R: IF x is A^1 and A^2 and A^3 and A^4 …………and A^L THEN y is B^s ………………(1)

For multiple conjunctive antecedents, we considered the min rules to find the membership values of the antecedent part as shown in equation (2).

$\mu_{A^S}(x) = \min \{\mu_{A^1}(x), \mu_{A^2}(x), \mu_{A^3}(x) ……… \mu_{A^L}(x)\}$……………………(2)

For multiple disjunctive antecedents, we considered the max rules to find the membership values of the antecedent part. For evaluation of the fuzzy rules, we have considered Mamdani if-then implication rules. Figure 2 depicts the membership function editor,

which displays input and output variables, as well as their ranges and probability of occurrence. Figure 3 shows a rule editor that may be used to add, amend, and delete rules.

As a result, the numerous fuzzy sets acquired during rule evaluation must be consolidated into a single resultant fuzzy set in some way. For this aggregation process, we used max fuzzy rules.

To get the final crisp set defuzzification is used. It is used for each and every fuzzy inference variable. It is mainly converts aggregated fuzzy output to crisp number. This is the last step in the designing of fuzzy controller. The final desired output for each variable is generally a single number. Since the aggregate of

the number of fuzzy sets is itself, a fuzzy set. Here, we used centroid defuzzification method.

VI. RESULT ANALYSIS

For each candidate, we consider a new set of circumstances. Table 3 shown the results of different circumstances of input values. We have calculated percentage of predicted risk for loan granting to a customer. Figure 4, shows a sample case where if applicant age is 20, the total asset is 50, employment stability is 20, working experience is 20, total income is 50, the loan amount is 20, the loan repayment period is 15, and the applicant credit score is 80 then the percentage of risk to grant a loan is 50.

Table 3: Results of different circumstances of input variables

#Example	Input Details(Low/Medium/High) within a range of 0 to 100								% of risk to grant a loan
	AA	TA	ES	EX	TI	LA	LP	CS	
1	25	50	50	80	80	20	15	80	32
2	25	50	50	80	50	20	15	80	70
3	25	50	50	80	80	70	85	80	76
4	20	50	20	20	50	20	15	80	50
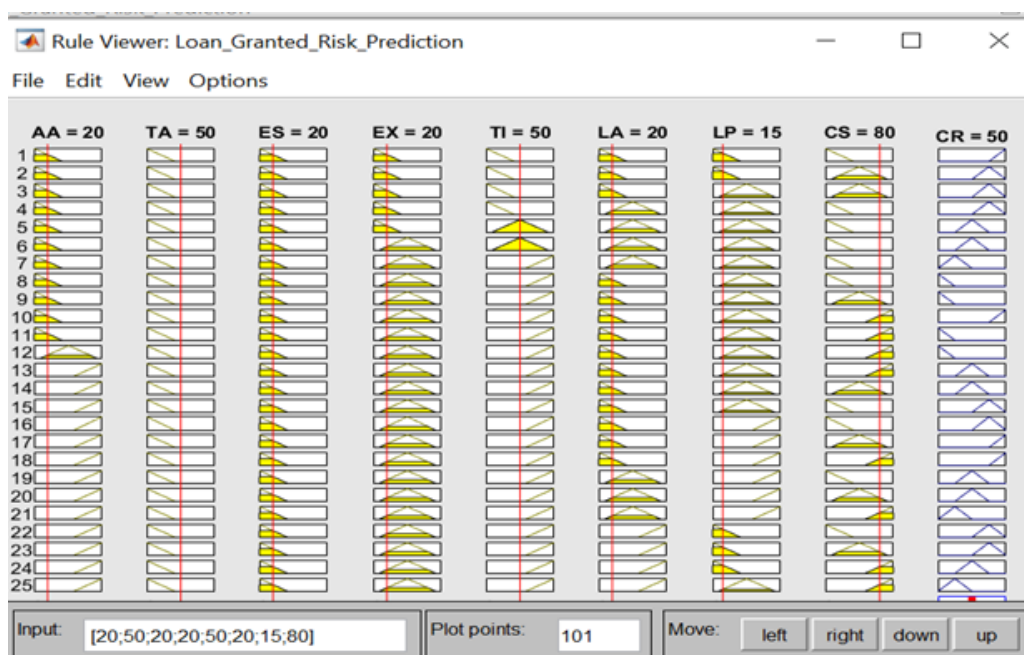
Figure 4: One sample output

case where if applicant age is 20, the total asset is 50, employment stability is 20, working experience is 20, total income is 50, the loan amount is 20, the loan repayment period is 15, and the applicant credit score is 80 then the percentage of risk to grant a loan is 50.

VII. CONCLUSION AND FUTURE SCOPE

A risk prediction model based on fuzzy rules is a good choice for selecting the right candidates for granting a loan. Then there will be no biasedness. In addition, it reduces paperwork, increases trust, cooperation, and understanding among customers, and increases the likelihood of choosing the right, low-risk customer. Further, we want incorporate some machine learning algorithm to investigate customer behaviours and motive and previous loan payment history to minimize the risk for granting a loan.

VIII. REFERENCES

[1]. Lotfi A. Zadeh "Fuzzy sets. Information and control" 8, 338-353 in 1965.

[2]. G. Hayward and V. Davidson "Fuzzy Logic Applications" analyst.128, 1304-1306.

[3]. McCord, A.Z. Fuzzy logic and its applications in hardware, Senior Seminar, CPSC 491-01, March 4, 2009.

[4]. Tohida Rehman "An introduction to Fuzzy Logic Controller and its applications" International Journal of Scientific Research in Science and Technology (www.ijsrst.com),Volume 3, Issue 7 , pages from 540 to 545 with Print ISSN: 2395-6011 and Online ISSN: 2395-602X.

[5]. Tohida Rehman "Fuzzy Rule Based Candidate Selection Evaluator by Political Parties" International Journal of Advanced Research in Computer Science, Volume 8, No. 3, March – April 2017 , pages from 445 to 451 with ISSN No. 0976-5697.

[6]. Ehsan Kamalloo and Mohammad Saniee Abadeh "Credit Risk Prediction Using Fuzzy Immune Learning" Hindawi Publishing Corporation Advances in Fuzzy Systems Volume 2014, Article ID 651324, 11 pages.

[7]. Hamid Eslami Nosratabadi, Ahmad Nadali, and Sanaz Pourdarab "Credit Assessment of Bank Customers by a Fuzzy Expert System Based on Rules Extracted from Association Rules" International Journal of Machine Learning and Computing, Vol. 2, No. 5, October 2012.

[8]. J. A. Fdez, R. Alcalá, M. J. Gacto, and F. Herrera,"Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms", Fuzzy Sets and Systems, 2009, 160(7),pp.905-921.

Cite this article as :