Volume 6, Issue 5, May 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Fuzzy Logic in Online Banking System Based on Customer Classification

Tanwy Barua
MEEE, AIUB, Dhaka,
Bangladesh

Abstract*— In the last decades, information systems have revolutionized the way information can be stored and processed. As a result, the information volume has significantly increased. It becomes difficult to analyze the large amounts of available data and to generate appropriate management decisions. In practice, Information systems mostly use relational databases in order to store these data collections. This paper makes a comparison between traditional or classical classification and fuzzy classification. Experimental results demonstrate that the proposed intelligent fuzzy query is more effective than the conventional query and it provides the user the flexibility to query the database using natural language. Online Banking system refers to manage the whole banking system through web. In banking system, it is necessary to retrieve important information of customers from their historical data to give them some opportunities, such as discount on loan interest. To do this it is very necessary to classify them. I have classified the customers based on classical logic and fuzzy logic.

Keywords*— Fuzzy, Set Theory, KDD, Data Base, Transformed Database.

I. INTRODUCTION

Fuzzy Logic is a mathematical tool to represent the imprecise, uncertain and vague information. It is based on fuzzy set theory. It is multi-valued logic. It is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. It can be implemented in hardware, software, or a combination of both. Fuzzy Logic (FL) provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information [1]. FL's approach to control problems mimics how a person would make decisions, only much faster. There are two types of classification. One is classical classification and the other is fuzzy classification. Classical classification is based on classical set theory and fuzzy classification is based on fuzzy set theory.
Difference between classical set theory and fuzzy set theory:
The membership degree of each element in classical set is either 0 or 1.
For example, A= {(1,1),(2,0),(3,1),(4,1)}
The membership degree of each element in fuzzy set is in [0,1].
For example, A= {(1,0),(2,0.2),(3,0.5),(4,1)}
Difference between Classical classification and Fuzzy classification:

In classical classification a customer is classified into only one class, whereas in fuzzy classification a customer may be involved into different classes. In this work, I have classified the customers based on classical logic and fuzzy logic to give some discount on their loan interest and compare both approaches. It is necessary to retrieve important information of customers from their historical data to give them some opportunities, such as discount on loan interest. To do this it is very necessary to classify them. I have seen fuzzy classification provides better performance than classical classification. Fuzzy Logic is based on fuzzy set theory. It is multi-valued logic. It is a problem-solving control system methodology. It can be implemented in hardware, software, or a combination of both [2]. Fuzzy Logic (FL) provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information.

Types of classification:

1. Classical Classification.
2. Fuzzy Classification

- Classical classification is based on classical set theory.
- Fuzzy classification is based on fuzzy set theory.

II. KNOWLEDGE DISCOVERY IN DATABASE

Frawley et al states that "Knowledge discovery is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data." In order to get this information, I try to find patterns in the given data set. To know if a pattern is valuable, the assessment of its interestingness and certainty is crucial. Patterns that are interesting and certain enough according to the user's measures are called knowledge. The output of a program that discovers such useful patterns is called discovered knowledge.

According to KDD exhibits four main characteristics:

- High-Level Language (HLL): The discovered knowledge is represented in a language that does not necessarily have to be directly used by humans, but its expression should be comprehensible [3].
- Accuracy: The measure of certainty implies whether the discovered patterns portray the contents of a database properly or not.
- Interestingness: Discovered knowledge is considered interesting if it fulfils the predefined biases. By denoting a pattern interesting, I mean that it is novel, potentially useful and the discovery process is nontrivial.
- Efficiency: Even for large Datasets, the running time of the algorithm is acceptable and predictable.

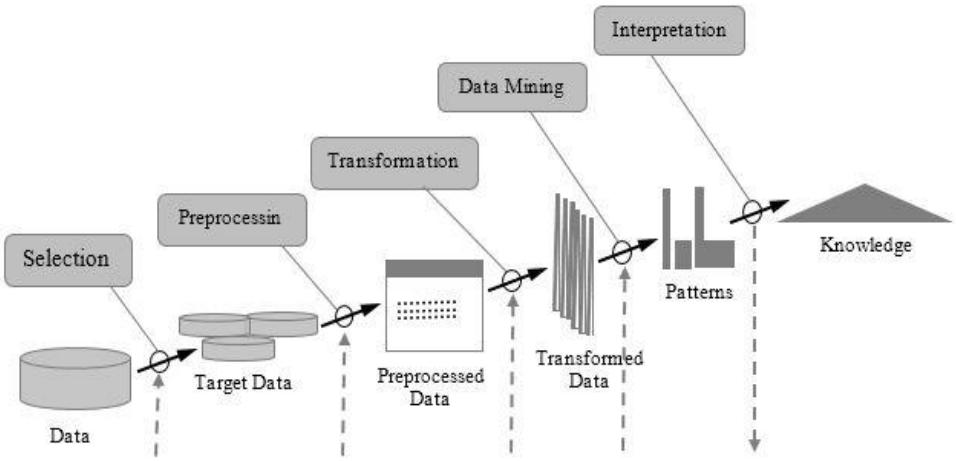

Fig 1: KDD Process

The KDD Process consists of several steps that are in place to achieve the defined goals for knowledge discovery. The KDD Process is interactive and iterative and requires decisions made by the user. Proposes nine basic steps.

1. Data Understanding: learning the application domain for prior knowledge and goals of the application.
2. Creating a target data set: selecting the subset of the data on which the data mining will be performed.
3. Data cleaning and pre-processing: removing noise or outliers, developing strategies for handling missing data.
4. Data reduction: reduce dimensionality of the data set in order to get rid of data that is unnecessary for completing the mining task and thereby keep the computing time low.
5. Selecting the data mining method: the most important task here is to find the method that will best suit the completion of the KDD goals.
6. Choosing the data mining algorithm: there are many different data mining algorithms. Deciding on an efficient one to search for patterns in data is critical and includes decisions about appropriate models and parameters.
7. Data mining: applying the previously chosen algorithm to the data set and searching for interesting patterns in a particular representational form.
8. Interpreting: mined patterns include the visualization of mined patterns and a possible return to any of the steps 1-7 if the results are unsatisfactory.
9. Consolidating discovered knowledge: documenting the results and incorporating them into another system.

III. IMPLEMENTATION

Fuzzy Classification is a well-established technique for discovering knowledge from data. We have seen that in classical classification a customer is classified into only one class, whereas in fuzzy classification a customer may be involved into different classes [4]. The Classical Classification leads to the underestimation- or overestimation of values due to the hard boundary of the intervals. To overcome this problem the approach of Fuzzy Classification has been developed. In this chapter, firstly I will implement the Classical Classification and then Fuzzy Classification and finally I will compare between those two approaches.

In my research work, I will provide some discount on loan interest over all customers of our banking system. For this reason I have to classify them. I have used four classes in my work. They are

C1: Very very committed customer
C2: Very committed customer
C3: Committed customer
C4: Non-committed customer

I offer discounts 5%, 4%, 3% and 0% on loan interest for C1, C2, C3 and C4 classes respectively.

IV. OPERATIONS

To perform this approach I have to follow the following steps:

1. Creating Target Database

There are huge amount of data in our database, all data are not required to retrieve knowledge, so I have to select only three attributes, which are Customer Name, Number of Loan Payment per year, Number of Transaction per year in his/her Current account and this is my target database [5].

2. Creating Pre-processed Database

By removing some unexpected data from target database I have got processed database. In my research work the unexpected data means the number of loan payment which have not been paid in due time. If number of loan payment is greater than 12, then the number of loan payment will be 12. After removing the unexpected data from my assumed database, the processed database is shown in table 1.

Table 1: Processed Database for Classical Classification

Customer Name	No of Loan payment per year	No of Transaction per year
Kalam	7	4
Jamal	12	11
Rahim	7	8
Karim	5	9
Shohag	12	12
Rafique	10	3
Sumon	0	0

According to classical classification approach, we know that the membership degree of a customer in regular payer and regular customer is either 0 or 1. The scenario of each customer of processed database with their membership degree is shown in fig 2.

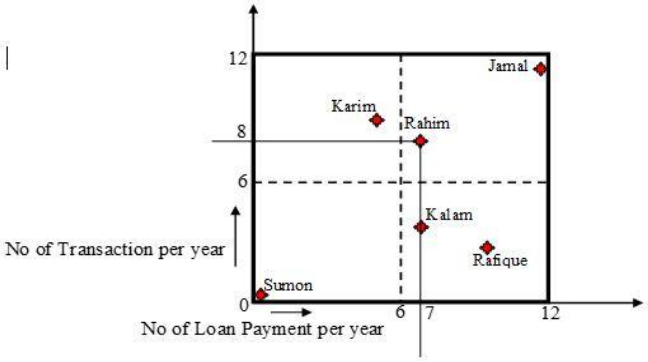

Fig 2: Customer over Different Classes

In fig 2, we have seen Jamal belongs to Class C_1 and Rahim belongs to Class C_1, i.e in C_1 the membership degree of Jamal and Rahim is 1 and in other classes the membership degree for Jamal and Rahim is 0. This is one of the problems in Classical Classification. Here we have seen that the number of loan payment for Jamal and Rahim are not same, but they belong to same classes with same membership degree. That is the main drawback of classical classification. After computing the membership degree of each customer, I have got the database, which is shown in table 2.

Table 2: Processed Database for CC with the value of different classes.

Customer Name	No of Loan Payment Per Year	No of Transaction Per Year	Values In different classes
Kalam	7	4	C1:0, C2:1, C3:0, C4:0
Jamal	12	11	C1: 1, C2:0, C3:0, C4:0
Rahim	7	8	C1: 1, C2:0, C3:0, C4:0
Karim	5	9	C1: 0, C2:0, C3:1, C4:0
Shohag	12	12	C1: 1, C2:0, C3:0, C4:0
Rafique	10	3	C1:0 , C2:1, C3:0, C4:0
Sumon	0	0	C1:0 , C2:0, C3:0, C4:1

In my research work, my main goal is to provide some discount on loan interest based on their status. Since the discount over classes C_1, C_2, C_3 and C_4 are 5%, 4%, 3% and 0% respectively, the general expression to calculate the total discount of a customer is shown in equation (1).

Discount(customer_name)=C1*5+C2*4+C3*3+C4*0…….Equation (1)

For example, the membership degrees in C1, C2, C3 and C4 classes for Kalam customer are respectively 0, 1, 0 and 0. So the overall discount on loan interest of kalam will be by using equation (1).

Discount(Kalam)= 0 * 5 + 1 * 4 + 0 * 3 + 0 * 0 = 4%

In the same way, after computing the discount on loan interest for all customer of processed database I have got the table 3.

V. DATABASE

The database in which the membership degree of any numerical attribute is included is called fuzzy database. In my research work, I have applied fuzzy membership degree on number of loan payment and the number of transactions. I used S-shaped and Z-shaped membership functions that match with this data. The used membership function for Regular Loan Payer and Regular Customer is S-shaped membership function, which is shown in fig-3

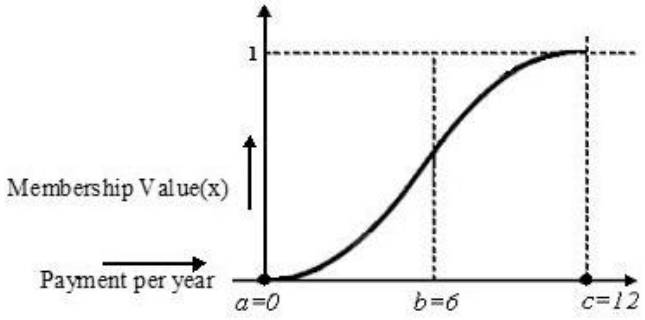

Fig 3: S-Shaped Membership function

$$\mu(x:a,b,c) = \begin{cases} 0 & \text{if } x <= a \\ (x-a)^2/((b-a)*(c-a)) & \text{if } a<x<= b \\ 1-((x-c)^2/((c-b)*(c-a))) & \text{if } b<x<= c \\ 1 & \text{if } x>c \end{cases}$$

For Non Regular Loan Payer and for Non Regular Customer I have used Z-shaped membership function.

$$\mu(x:a,b,c) = \begin{cases} 1 & \text{if } x <= a \\ 1-2\left(\frac{x-a}{c-a}\right)^2 & \text{if } a <= x <= b \\ 2\left(\frac{x-c}{c-a}\right)^2 & \text{if } b <= x <= c \\ 0 & \text{if } x >= c \end{cases}$$

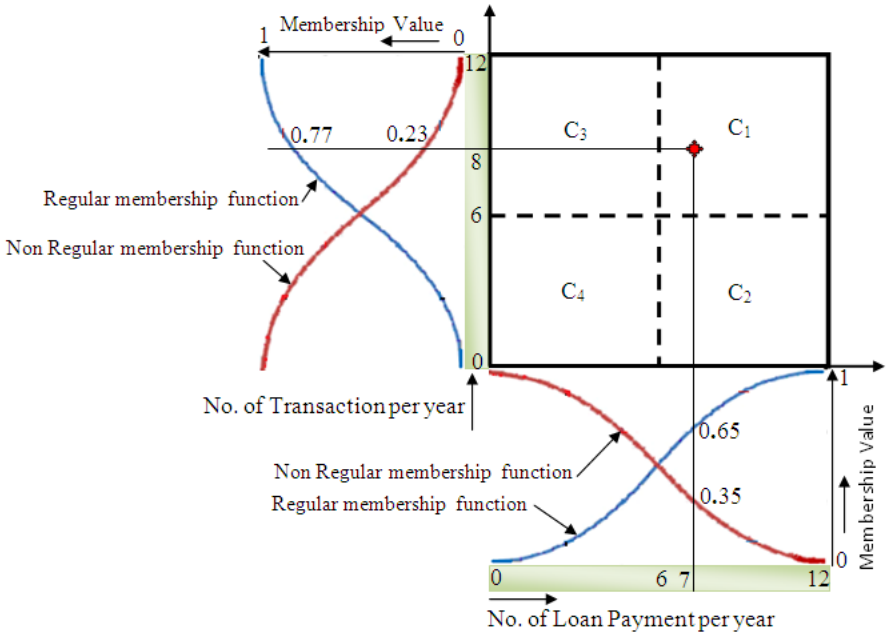

Fig-4: Classes with different membership values

By using eq. (2) in pre-processed database, I have got fuzzy database which is shown in table 5

Table 5: Transformed database (Fuzzy database)

Customer Name	No. of Loan Payment Per Year	Fuzzy Membership Value In regular loan payer	No. of Transaction Per Year	Fuzzy Membership Value In regular customer
Kalam	7	0.652	4	0.222
Jamal	12	1	11	0.980
Rahim	7	0.652	8	0.777
Karim	5	0.347	9	0.875
Shohag	12	1	12	1
Rafique	10	0.944	3	0.125
Sumon	0	0	0	0

VI. CONCLUSION

The most frequent activities conducted online were checking account balances and viewing or paying bills. These features should be easily accessible upon login and include clear and simple terminology, feedback information, and ability to easily identify/fix mistakes. By using this project, one can classify customers by using classical and fuzzy approaches. I have used only classification in the data mining task. One can extend my work by including other tasks of data mining, such as Association rules, Regression, Deviation Detection etc. One can extend our work by using fuzzy modifier, such as very very regular, extremely regular, more or less regular etc.

ACKNOWLEDGMENT

We are earnestly grateful to our mentor, Mihir Barua, Manager-Technical Sales (Siemens Bangladesh Limited) for providing us with his special advice and guidance for this project. Finally, we express our heartiest gratefulness to the Almighty and our parents who have courageous throughout our work of the project.

REFERENCES

[1] Frawley, William J.; Piatetsky-Shapiro, Gregory; Matheus, Christopher J.: Knowledge Discovery in Databases: an Overview. AAAI/MIT Press, 1992.

[2] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic: Knowledge Discovery and Data Mining: Towards a Unifying Framework. In Proceeding of The Second Int. Conference on Knowledge Discovery and Data Mining, pages 82--88, 1996.

[3] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic: The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM, Volume 39, Issue 11 Pages: 27 - 34, 1996.

[4] Fayyad, Piatetsky-Shapiro, Smyth: From Data Mining to Knowledge Discovery in Databases. AI MAgazine, 1996.

[5] Zembowicz, Robert; Zytkow, Jan M.: From Contingency Tables to Various Forms of Knowledge in Databases. American Association for Artificial Intelligence, 1996.

