# Comparative Study of Gaussian and Zshaped curve membership function for Fuzzy Classification Approach

Pooja Agarwal<sup>1\*</sup>, Arti Arya<sup>2</sup>, Suryaprasad J.<sup>3</sup>

<sup>1\*</sup>Dept. of CSE, PESIT Bangalore South Campus, Bangalore, India <sup>2</sup> Dept. of MCA, PESIT Bangalore South Campus, Bangalore, India <sup>1</sup> Dept. of ISE, PESIT Bangalore South Campus, Bangalore, India e-mail: poojaagarwal@pes.edu, artiarya@pes.edu, surya@pes.edu

Abstract— Customer Retention is a major aspect of the growth of any business. It is always good to retain those customers who are more loyal or can be converted to be loyal. In this paper, a fuzzy decision tree based approach is proposed in combination with Z-shaped curved membership function to find the customer loyalty and has been compared with the approach of fuzzy decision tree with Gaussian membership function. As a result, it has been found that Z-shaped membership function gives 71% accuracy whereas with Gaussian membership function it gives 63.7% accuracy.

**Keywords**— Customer Loyalty, Customer Relationship Management, Membership functions, Decision Tree, Fuzzy Decision Tree.

# I. INTRODUCTION

Customer retention is crux of any business growth. For retaining the customer, it is required to be known that which customer need to be retained based on their loyalty towards the organization. The current work focuses on finding the loyalty level of the customers. After an intense literature survey [5,6,7,12], the four most important attributes that really influence the loyalty of customers are taken into consideration. These four attributes are:

- (a) Total Expenditure (TE),
- (b) Frequency of visits (FV),
- (c) Life Time of a customer (LT),
- (d) Mode of Payment (MP).

There can be many other attributes contributing towards identifying customer loyalty, but the above stated attributes are most influential ones [5,6,7,12]. In the previous work [6], fuzzy decision tree using Gaussian membership function had achieved the accuracy up to 63.7%. In this paper, an approach has been proposed to improve the accuracy using Z-shaped curve membership function. The accuracy so achieved is 71%.

The classification of a customer as 'loyal' or 'not loyal' can be considered in fuzzy sense, as it may not be a firm decision for an individual customer. Fuzzy logic is employed to handle the concept of partial truth, where the truth value may lie between 0 and 1 that is false and completely true. As an example, the weather is cold or hot should not be a rigid decision. It is better to find the temperature level of cold and hot. Usually it is told that weather is slightly cold/hot or weather is too cold/hot. It might vary based on the individual preference to tolerate cold and hot weather. Here slightly cold or too hot are the soft boundaries and this decision is called partially true or fuzzy decision.

Fuzzy decision tree [1] keeps the boundary of the result soft. Soft boundaries help in classifying a customer, so that he can belong to two classes at the same time. So, the result generated by fuzzy decision tree for the customer loyalty is more flexible.

This paper proposes a Fuzzy decision tree based approach to identify the customer loyalty in four different classes as Super Premium Loyal customer (SPL), Premium Loyal customers (PL), valued customers (VC), Normal customers (NC). Class label NC is assigned to a customer who is not a very regular customer and SPL customer is of great value to the organization, whereas other two loyalty class labels lies in between them.

#### II. RELATED WORK

Jose et al. [11] has defined the different impurity measures to find out the impurity level of the different attributes which is used for the generation of decision tree and called as 'impurity level', which helps to identify that which impurity measure can be used to reduce the CPU usage.

Qiaohong Zu et al. [12] constructed a simple and intuitive extended Bayes model, which gives a better classification effect. The customer classification model uses these three factors to classify customers: Customer lifetime value, Customer credit, Customer loyalty.

The customers were first clustered in k-means algorithm [4] (a partition technique that finds a specific number of clusters which is represented by the average of all the points in the cluster (centroid)) then this cluster was used in customer classification prediction by weighted Bayes algorithm thereby improving the accuracy of the classification.

Assigning weights [5] to the attributes shows how the attribute affects the decision in current data environment in through theory this does not reflect the prior information of the decision-maker.

The rough set theory is used to find the solution of attribute important weights [5]. A rough set is used when one has to represent incomplete knowledge. They are sets with fuzzy boundaries-sets which cannot be precisely characterized using available set of attributes [12]. It has the following characteristics: the solution does not need any prior data information [5], simplify data and obtain the minimum knowledge representation. The extended Baye's algorithm [12] was constrained under both rough set attribute importance theory and expert prior knowledge and it also combined cluster preprocessing with classification prediction to effectively and efficiently classify the customers with multifactor based on customer value and customer behavior.

# III. BASIC TERMINOLOGY

#### A. Decision Tree

Decision tree classification technique is used for a wide range of classification problems and is an important tool for classification. A decision tree is a prediction model where each internal node denotes a test on an attribute, each outgoing branch represents an outcome of the test and each leaf node is labeled with a class.

The model produced by the decision tree is represented in the form of a tree structure. Learning of decision tree involves deciding which split to make at each node and how deep the tree should be. A leaf node indicates the target class labels. The examples are classified by sorting them down the tree from the root node to some leaf node.

#### B. Overfitting in Decision Tree

Let us consider, that h is error of hypothesis. Consider the error on training data as  $\text{error}_{\text{train}}$  (h) and error over the entire distribution D of data be  $\text{error}_{D}$  (h). Then a hypothesis h "overfits" the training data [2,3] if there is an alternative hypothesis,  $h_{a}$ , such that:

error  $_{train}(h) < error _{train}(h_a)$ 

error  $_{D}(h) < error _{D}(h_{a})$ 

When a decision tree is built, many of the branches will reflect anomalies in the training data due to noise or outliers. As a result, poor accuracies would be obtained for unseen data. Tree pruning methods address this problem of overfitting the data [2]. Such methods typically use statistical measures to remove the least reliable branches, generally resulting in faster classification and an improvement in the ability of the tree to correctly classify independent test data. The designer may not be aware of the existence of outliers at the time of the tree design, even though outliers are usually easily detected in a training set [3].

#### IV. PROPOSED METHOD

#### A. Fuzzy Decision Tree

There have been several extensions to the decision trees to deal with classification. One such extension is Fuzzy decision tree (FDT) which combines the ability of decision tree with fuzzy representation. There are several heuristic algorithms available to generate FDT [9]. In this work, FDT is generated by considering the leaf nodes as decision making nodes after applying Z-shaped curve membership function. The weight of the attributes is computed using Information Gain and Entropy [4].

Fuzzy Logic as implemented in Decision trees will be employed for the automatic classification of customers. Using Fuzzy decision tree technique for classification, the resulted value may belong to either side of the tree or may fall in the overlapping (Fuzzy) region. Fuzzy Decision Tress (FDT) uses soft boundaries for classification where the customer can belong to two classes at the same time, thus generating a more flexible result.

By introducing uncertainty, this method proves to be beneficial in identifying the potential loyal customers for an organization who may not be categorized high on the loyalty ladder. Fuzzy decision trees improve the robustness and generalize the classification because of fuzziness [7]. Fuzzy decision tree gives the result within a range of [0, 1], where this range signifies the degree of membership of a data object belonging to a particular class label. Here Z-shaped membership function [8] is used to determine the magnitude of participation of each input. The construction of the FDT has been implemented by assigning weights to the different contributing factors or attributes using Gini index values, as all the attributes may not affect the construction of FDT in a similar way.

The objective of Fuzzy Decision Tree is to indicate that a customer may belong to more than one class at the same time but with different membership grades. The membership function [8] is a graphical representation of the degree of participation of each customer.

Once the functions are inferred, they are defuzzified into a crisp output which drives the system. Defuzzification [8] is the process of producing a crisp result in fuzzy logic [8], given fuzzy sets and corresponding membership degrees.

Some advantages of using fuzzy decision tree over decision tree are:

- 1. By using Fuzzy decision tree for classification, we can maintain transparency as well as a high accuracy rate.
- 2. It is more flexible than Decision tree.
- 3. It improves the robustness and generalizes the classification due to the fuzzy reasoning.
- 4. It can handle uncertainty appropriately.

Fuzzy Decision Tress (FDT) uses soft boundaries for classification where the customer can belong to two classes at the same time, thus generating a more flexible result. Here Z-shaped curve membership function [10] is used to determine the magnitude of participation of each input.

#### B. Z-shaped curve membership function

The Z-shaped curve is spline-based function of x. Because of its shape like Z it has been named as Z-shaped curve membership function. The parameters a and b are the extremes of sloped portion of the curve given by:



From literature [10,12] on the similar lines, it has been found that the space of the customer loyalty factor lies in between [0.25, 0.75]. The Z-Shaped membership function of customer loyalty is represented by  $Z_{cl}(x)$  and the values of a and b are 0.25 and 0.75 respectively. The four class labels are depicted as four different values of Z-shaped membership function, where 1 for "Normal customer", 0 for "SPL customer" and rest class labels "Valued customer" and "PL customer" lies in between them.

$$Z_{CL}(x) = \begin{cases} 1 & x \le 0.25 \\ 1-8(x-a)^2 & 0.25 \le x \le 0.5 \\ 8(b-x)^2 & 0.5 \le x \le 0.75 \\ 0 & x \ge 0.75 \end{cases}$$

Fig 1 represents the graphical representation of the membership function of customer loyalty.



Fig 1: Graphical Representation of membership function of customer loyalty

#### C. Entropy and Information Gain

A Fuzzy Decision Tree(FDT) classifiers are typically a top-down greedy approach, which provides a rapid and effective method for classifying data instances. The root node of DT is chosen based on the highest information gain of the attribute. Given a training dataset, D, the expected information gain to correctly classify an instance,  $x_i = D$ , is given in following equation, where  $p_i$  is the probability that  $x_i = D$  is a part of a class  $C_i$ .

$$Info(D) = -\sum_{i=1}^{m} p_i log_2 p_i$$

Where Info(D) is the average amount of information needed to identify  $C_i$  of an instance  $x_i$  D. The goal of FDT is to iteratively partition D into subsets  $\{D_1, D_2, ..., D_n\}$ , where all instances in each  $D_i$  belong to the same class  $C_i$ . Info<sub>A</sub>(D) is the expected information required to correctly classify an instance  $x_i$  from D, based on the

partitioning by attributes A. The following equation shows  $Info_A(D)$  calculation, where  $\begin{bmatrix} D_1 \\ D \end{bmatrix}$  acts as the weight of the j<sup>th</sup> partition.

$$info_{A}(D) = \sum_{i=1}^{n} \frac{|D_{i}|}{|D|} info_{A}(D_{i})$$

Information gain is defined as the difference between the original information requirement and the new requirement that is shown in the following equation

#### $Gatn (A) = Info(D) - Info_A(D)$

The Gain Ratio is an extension to the information gain approach, also used in Decision Tree. It applies a kind of normalization to information gain using a "split information" value defined analogously with Info(D) as shown in the following equation

$$Splittinfo_{A}(D) = -\sum_{l=1}^{n} \frac{|D_{l}|}{|D|} \log_{2}(\frac{D_{l}}{D})$$

The attribute with the maximum Gain Ratio is selected as the splitting attribute, which is defined in following equation

$$GeinRetio (A) = \frac{Gain(A)}{\text{SplitInfo}(A)}$$

Following equation defines the gini index value for a dataset D, where p<sub>i</sub> is the frequency of class Cj D.

$$dint(D) = 1 - \sum_{i=1}^{m} p_j \circ p_i$$

In this way, the attribute to split with the least gini index value is chosen.

#### V. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

The proposed approach is tested on customer dataset (collected through an online survey, https://goo.gl/p4LZjK), this contains around 650 records and online retail dataset, this dataset contains around 5.5 lakh records (https://archive.ics. uci.edu/ml/datasets/Online+Retail) with each example representing items bought in a single transaction.

The Gini index or Information Gain value calculated for each of the attribute is as follows:

Information Gain<sub>Total Expenditure</sub>= 0.502

Information Gain Frequency of visit = 0.627

Information Gain Life time = 0.652

The procedure to build decision tree using z-shaped curve membership function is given in algorithm 1.

The combined gini values of each of the 4 attributes are compared, the attribute having least value is considered to give the best split in decision tree and therefore it is considered to be at the root node.

In the above calculation of Information Gain, it is clear the attribute **"Total Expenditure"** is the least, hence it is the root node. The value of information gain is not calculated for mode of payment, as it contains binary value.

# Nomenclature:

 $\beta$  – Total expenditure of individual customer (TE)

 $\alpha$  – Average expenditure per customers (Agv(TE))

m – Median of Total expenditure

**n** – Total number of customers

**mode** – mode of payment (1 for payment through credit/ debit card; 0 for payment through cash/ other mode of payment)

 $\tau$  – Frequency of visit (FV)

- $\delta$  average expenditure per visit  $(\beta/\tau)$
- $\phi$  average expenditure over all transactions (Avg( $\delta)$ )
- lt lifetime of customer
- $\Upsilon$  average no. of visits per customer (Avg( $\tau$ ))

 $\mu$  – z-shaped curve membership function

d - standard deviation of individual total bill amount

# Algorithm1: Fuzzy Decision Tree Classifier for Customer loyalty using Z-Shaped Membership function

**Input:** A transactional data set having the attributes Total Expenditure, Frequency of visit, Life time of a customer and Mode of payment for each customer id and the root node is selected based on the Gini index value of attributes i.e. "Total Expenditure".

#### **Procedure:**

```
If ((\beta > (\alpha + m/n)))
  If ((mode = = 1) \parallel (mode = = 0 \&\& \beta \ge (0.31 N))) then
     If (\delta \ge \phi) then
         If (( lt \ge 365 days && \tau \ge \Upsilon) ||( \delta \ge (N))) then
                                                                    "SPL Customer"
         else if( lt <=183 days || \delta > (0.75*N)) then
            "PL Customer"
         else Z = Z(a, \alpha, b)
            If Z < 0.5 then "PL Customer"
            else "SPL Customer"
     else "Valued Customer"
  "Valued Customer"
  else
       "Normal Customer"
else if(\beta < (\alpha - m/n)) then
   if (\delta \ge \varphi \parallel \delta \ge (0.38*N)) then
       if(\beta \ge (2*N))
                           then
         "Valued Customer"
       else "Normal Customer"
   else "Valued Customer"
else
   Z = Z(a,\alpha,b);
         if (Z< 0.25) then "Normal Customer"
         else if (Z \ge 0.25 \&\& Z < 0.5) then
                  "Valued Customer"
         else if (Z \ge 0.5 \&\& Z < 0.75) then
                  then "PL Customer"
         else "SPL Customer"
end if
```

N is used as a parameter based on the minimum threshold limit.

A sample fuzzy decision tree is represented Figure 2. Each of the tree node and leaf node represent the value of the Z-shaped membership function of the, which represents the degree of membership. In fig 2  $F_{ij}$  represents the degree of membership of each attribute.



#### Fig2: Fuzzy Reasoning

In fuzzy decision tree, a customer can be the part of more than one type of customer categories, which leads to a better decision for the customer loyalty. The algorithm is executed for Customer dataset of 650 records and online retail dataset of 5.5 lakhs records over Gaussian and Z-shaped curved membership functions. It is observed that Customer dataset achieves 61.2% accuracy with Gaussian membership function and 68.7% accuracy with Z-shaped curved membership function. It has been represented in table 1.

	Customer dataset	Online dataset
FDT with Gaussian MF	61.2%	63.7%
FDT with Z-shape MF	68.7%	71%

Table 1: Accuracy achieved from FDT by different datasets using different membership functions

The algorithm is developed in a Linux environment with ubuntu 16.04.1 on desktop-amd64.iso configuration using R version 3.2.3.

#### VI. CONCLUSION

In current times, efforts to reduce customer churning has been in limelight as customer retention has become one of the major motives of business organizations. For reducing customer churning, Customer Loyalty Classification becomes vital for the growth of business organizations and thereby assisting in the enhancement of the economy of a country. Although many researchers have explored this area of customer classification for their loyalties towards different organizations, but there was a good scope of improvement over accuracy. In this paper, Fuzzy decision tree approach is proposed for loyalty classification and 71% accuracy is achieved, which is comparable with the accuracy obtained by gaussian membership function

#### References

- Olaru, Cristina, and Louis Wehenkel. "A complete fuzzy decision tree technique." Fuzzy sets and systems 138.2 (2003): 221-254. Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.
   Bramer, Max. "Avoiding overfitting of decision trees." Principles of Data Mining. Springer London, 2016. 121-136.
- [2] Branet, Max. Avoiding over-fitting of decision decs. Thirdpies of Data Mining. Springer Eondon, 2010. 121-150.
   [3] Wang, Tao, et al. "Handling over-fitting in test cost-sensitive decision tree learning by feature selection, smoothing and pruning." Journal of Systems and Software 83.7 (2010): 1137-1147.
- [4] V. Kumar, T. Pang-Ning, S. Michael, "Introduction to Data Mining", Pearson Education ,2009
- [5] http://www.cec.uchile.cl/~his02/index\_files/ rough\_tutorial.pdf.
- [6] Arya, Arti, Agarwal, Pooja et al. "Automatic Fuzzy Classification tool for Customer Loyalty using Gaussian Membership Function." Data Mining and Knowledge Engineering 2.7 (2010): 168-173.
- [7] Arya, Arti, and Agarwal Pooja. "Fuzzy Decision Tree based Automatic Classifier for Customer Loyalty." In proc. of Intl. Conf. on Data Management, 2010.
- [8] John, Yen, and Reza Langari. "Fuzzy logic: intelligence, control, and information." Publish by Dorling Kindersley, India (1999): 379-383.
- [9] [Bartczuk, Łukasz, Piotr Dziwiński, and JanuszStarczewski. "New method for generation type-2 fuzzy partition for FDT." Artificial Intelligence and Soft Computing. Springer Berlin/Heidelberg, 2010.
- [10] Zhu, A-Xing, et al. "An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic." Geomorphology 214 (2014): 128-138.
- [11] J. Ranilla, O. Luaces, and A. Bahamonde, "A Heuristic for Learning Decision Trees and Pruning them into Classification Rules," Artificial Intelligence Comm., vol. 16, no. 2,2003, pp.71-87.
- [12] Z. Qiaohong and L. Wenfeng,"The Research of Customer Classification Based on Extended Bayes Model"(IEEE)('08).

#### Authors Profile

Ms. Pooja Agarwal is Associate professor in the Dept of CSE in PESIT Bangalore South Campus. She has completed her M.Tech in computer science in 2001 from Banasthali Vidyapith. She has 14years of experience in academics and research.

Her research interest includes Machine learning, and Data mining, Automata Theory and Algorithms. Prof Agarwal is a life member of IAENG. Mail id: poojaagarwal@pes.edu

Ms. Arti Arya is professor and HOD in the Dept of MCA in PESIT Bangalore South Campus. She has completed her Ph.D. in computer science in 2009 from MD University. She has 16+ years of experience in academics and research.

Her research interest includes Data Analytics, Machine learning, Artificial Intelligence, Knowledge Based Systems and Data Mining, Data Analytics etc.

Dr, Arya is senior member of IEEE and life member of CSI & IAENG.

Mail id: artiarya@pes.edu

Suryaprasad J. is professor and Director in PESIT Bangalore South Campus. He has completed his Ph.D. in computer science in 2003 from Florida Atlantic University. He has 27+ years of experience in industry, academics and research.

His research interest includes Embedded Systems, System level Design Methodology, Verification Methodology, Machine learning etc.

Mail id: surya@pes.edu