

Enabling Customer Reward with a Hybrid Intelligent System (Case Study: J&J Shopping Mall)

Okon E. Uko
Department of Computer
Science,
University of Port Harcourt,
Choba,
Rivers State, Nigeria.

Enyindah P.
Department of Computer
Science,
University of Port Harcourt,
Choba,
Rivers State, Nigeria.

Enefiok A. Etuk
Department of computer
Science
Michael Opara University of
Agriculture,
Umudike, Abia State, Nigeria.

ABSTRACT

A reward program (also known as a loyalty program) is a marketing technique often adopted by some companies to appreciate customers who frequently make purchases with such companies. Normally, this kind of program leads to giving a loyal customer gifts in forms of customer free merchandise, coupons, rewards and advance released products. In most situations, reward programs in these companies are often biased due to the mechanism employed in determining a loyal customer. Hence, this article introduces a method that uses two efficient artificial intelligent techniques namely; fuzzy logic and expert system (fuzzy expert system) to tackle the challenge of biasness in loyal customer selection. These Hybrid system worked efficiently when executed on Jane and Juliet (J&J) Shopping Mall's data and sales manager (domain expert) 's rule. It was able to fuzzify the different linguistic variables and also aggregated the firing rules (from the expert rule-base) to generate a crisp loyalty output after applying a fuzzy inference model. Object Oriented and Design (OOAD) Methodology was used in the design of the system and JAVA was used to implement it. Matlab's Simulink was also used in the simulation of the results.

General Terms

Artificial Intelligence, Expert System, Fuzzy Logic, Hybrid Intelligent System, Fuzzy Expert System.

Keywords

Customer reward, loyalty, customer relationship management, hybrid intelligence, fuzzy expert system, artificial intelligence, systems integration, rule-base, fuzzification, defuzzification, crisp output, rule-firing, loyalty analysis.

1. INTRODUCTION

Customer reward (loyalty) program is a popular strategy used by businesses to keep track of their most frequent customers and provide incentives for customers to return to their businesses. A repeat customer spends 67 percent more on a given purchase than someone new to a business will [1]. So, managing a good customer reward system is worth it.

Different technologies have been used in recent times to implement customer reward programs. These technologies happen to be faced with a lot of challenges such as poor data integration, etc. Many customer reward programs focus on rewarding customers for spending a specific amount at the store, and do not collect detailed information about customers and their purchase behaviors. This often leads to loss of

valuable opportunities to analyze trends and insights in some specific customer variables.

Fuzzy logic provides a very convenient way to dig-out very important customer data into groups (in form of input parameters) and fuzzify this information internally to aid proper and informed decision making on customer reward pattern. Fuzzy logic, when used singularly, is not often very efficient in solving problems of this nature. Hence, this paper marries the power of two artificial intelligence technique - fuzzy logic and expert system - to build a hybrid intelligent system that is more efficient for customer rewarding.

The following terms will be defined for a better understanding of the implemented system:

1.1 Hybrid Intelligent System:

This describes a software system that applies, in parallel, two or more methods and techniques from artificial intelligence to solve a particular problem.

1.2 Fuzzy Logic:

Fuzzy logic is a field in computing (a sub-field in Artificial Intelligence) that allows solving difficult simulated problems with many inputs and output variables. Fuzzy logic is able to give results in the form of recommendation for a specific interval of output state, so it is essential that this model is strictly distinguished from the more familiar logics, such as Boolean algebra.

1.3 Expert System:

A piece of software programmed using artificial intelligence techniques. Expert systems use databases of expert knowledge to offer advice or make decisions in specific areas.

1.4 Input Parameters:

These are input variables defined for a given fuzzy system based on which the output result will be calculated. For example, in this customer reward program, our input parameters will be; turnover, payment behavior, proximity, etc.

1.5 Membership Function:

A membership function (MF) sometimes called fuzzifier is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the universe of discourse.

1.6 Fuzzification:

Fuzzification is the process of changing a real scalar value into a fuzzy value. This is achieved with the different types of fuzzifier (membership functions).

1.7 Defuzzification:

Defuzzification is the process of generating a quantifiable result in Crisp logic, in the presence of fuzzy sets and corresponding membership degrees. It is the process that maps a fuzzy set to a crisp set.

2. RELATED WORKS, RESEARCH PURPOSE AND CONTRIBUTION

2.1 Related Works

A review of the literatures on customer reward programs have brought to the fore different techniques which mostly includes artificial intelligence and statistical analysis techniques. This paper reviews some of these related literatures to provide an overview of the current situation on customer reward systems. [7] proposed a data mining technique for customer relationship management through the implementation of an approach called NeoVista Solutions. This approach [7] combined elements of clustering and neural network technology to enable retailers to automatically review their point-of-sale history and equate store groupings to sales patterns. This combination helps management to receive exact data which assisted them to continually fine-tune their replenishment system. [5] proposed a model for detecting customer level intentions to purchase in Business-to-Customer (B2C) websites using the technique of order preference by similarities to ideal solutions (TOPSIS) and Fuzzy Logic rule-base. This TOPSIS method [5] was used to rank the most important features of attributes and fuzzy logic was used to build model to show real levels of the influencing customer satisfaction and purchase intention. [3] proposed a new procedure, based on expanded RFM model by including one additional parameter, joining WRFM-based method to K-means algorithm applied in DM with K-optimum according to Davies–Bouldin Index, and then classifying customer product loyalty in under B2B concept.

[9] applied concurrent reinforcement learning for customer interactions in a company. In his solution [9], the company has an objective function, such as maximizing revenue, customer satisfaction, or customer loyalty, which depends primarily on the sequence of interactions between company and customer. A key aspect of this setting is that interactions with different customers occur in parallel. As a result, it is imperative to learn online from partial interaction sequences, so that information acquired from one customer is efficiently assimilated and applied in subsequent interactions with other customers. [12] applied an improved balanced random forest (IBRF) learning method to predict customer churn in Chinese banks [12]. Sampling techniques and cost-sensitive learning were integrated into the approach to achieve a better performance than most existing algorithms.

[4] proposed a solution that integrates various techniques of data mining, such as feature selection via induction, deviation analysis, and mining multiple concept-level association rules to form an intuitive and novel approach to gauging customer loyalty and predicting their likelihood of defection. Immediate action triggered by these "early-warnings" resulting from data mining was considered as the key to eventual customer retention [4]. Also, [6] proposed an ID3 Decision tree Algorithm to determine customer loyalty factor at Semarang

Ceramics Company. The research process combined ID 3 algorithm and 5 SERVQUAL attributes; reliability, assurance, tangible, empathy, and responsiveness. The result of the research was Responsiveness attribute and it indicated that quick service to the customer is the main factor which influences the customer loyalty. Hence, this customer loyalty information can be used to reward customers appropriately.

[8] applied generalized sequential pattern (GSP) algorithm to mine data from PT. Armada international motor company, Indonesia which consist of customer behavior and profile [8]. This approach was able to find customer service pattern both sequentially and simultaneously. These results were used as promotion planning and customer reward strategy base.

2.2 Research Purpose

In this study, a hybrid intelligent method that utilizes fuzzy logic and expert system techniques (in parallel) is employed to fuzzify and defuzzify customer information such as turnover, payment behavior, proximity, specialty and loyalty and at the same time formulate a rule-base based on expert knowledge to enable a more efficient determination of the level of loyalty of customer in a business outfit (J&J Shopping Mall) and thereby enhancing customer reward decision.

2.3 Research Contribution

For the case study, the use of fuzzy expert system will enhance an unbiased reward of customer loyalty efforts. It will help to effectively classify customer payment behavior, turnover, specialty, proximity and loyalty.

Academically, this study will serve as a reference material for further researches on the application of Hybrid intelligent (fuzzy expert) system to different problem areas today. This study also presents a major modification in the design of fuzzy expert systems.

Furthermore, this study will contribute to the society by providing a more efficient ways to run small and large businesses by reducing the business failure rate (often caused by poor customer relationship management).

3. FUZZY-EXPERT SYSTEM MODEL

Fuzzy expert system (FES) model is a hybrid intelligence model that meshes, in parallel, two artificial intelligence (Fuzzy Logic and Expert System) approaches to solve problems in a particular area of concern. This model is made up of the crisp input, fuzzification interface, inference engine, domain expert, knowledge base, defuzzification interface and the crisp output.

Based on fuzzification process, the first step in the implementation of the design is to prepare the data to be sent into the hybrid intelligent model by determining the different input parameters to be used. These input parameters which are about five of them in our case study will be collected at the fuzzification interface and broken down into different linguistic variable with the help of a membership function that capitalizes on the data stored in the knowledge base. The knowledge base serves the fuzzification interface with membership function data, the inference engine with rules to be fired at any point based on input values and the defuzzification interface with output membership function data. The architectural design of this model is as shown in Fig. 1 below.

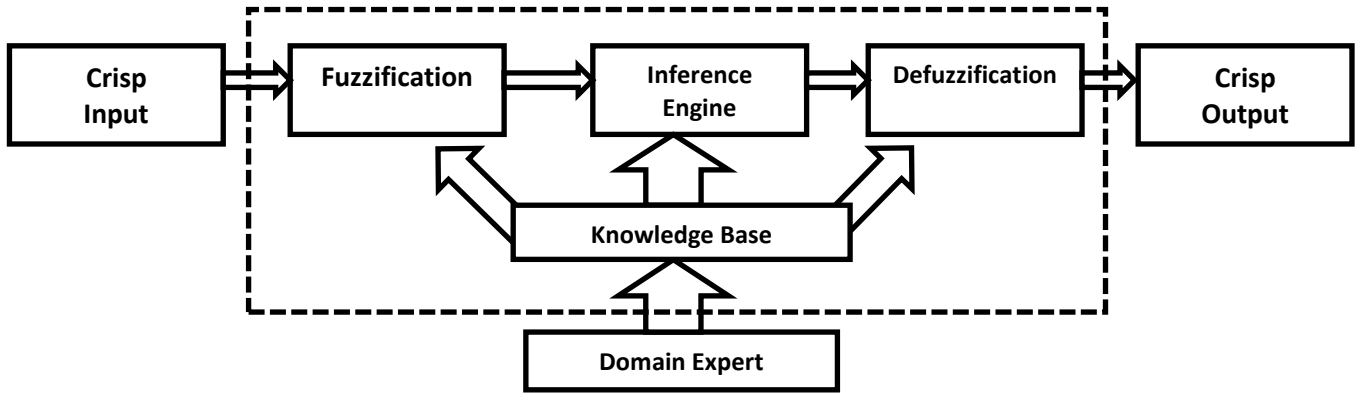


Fig. 1 Hybrid Intelligent (Fuzzy Expert System) Model

So, the model can be divided into three main stages which are:

3.1 The front-end stage:

At the front end of the fuzzy expert system model, crisp inputs are supplied and fuzzification is carried out. Crisp inputs are sent into the system. Here, in our case study, the system is fed with customer values such as payment behavior, turnover, proximity, specialty and loyalty. In the cause of carrying out fuzzification at this stage, input membership functions are generated with information obtained from the knowledge base as well as the crisp input. Domain expert knowledge are also organized to form a knowledge base that comprises of rules.

3.2 The processing stage:

This involves the aggregation of the functions of the knowledge base and the inference engine. At this stage, the knowledge base actually feeds the inference engine with a collection of expert rules. And the inference engine uses these rules to infer the fuzzy output value to be used in the defuzzification stage.

3.3 The back-end stage:

The activities carried out at the back-end of the fuzzy expert system involve basically defuzzification and generation of crisp output. The process of defuzzification involves the use of a mathematical model to transform a fuzzy set value into a crisp output that can be used for proper decision making.

4. MATERIALS AND METHODS

The methodology employed in this paper is the object oriented and design methodology. This methodology breaks down the entire system into subsystems and modules. One of the limitations to this research study was that business managers found it difficult to disclose information about their businesses. The research, however, is a fuzzy-expert system and is not data intensive. The knowledge base has one hundred and twenty rules that could possibly be fired by fuzzified inputs. The fuzzification method that was applied at the front end of this hybrid intelligent system is called The Triangular membership function. The triangular curve of this function depends on three values namely; a_1 , a_2 and a_3 as showcased in equation (1) below. The endpoint of the triangular membership function is determined by a_2 and a_3 while the peak location is determined by a_1 . Crisp input supplied at the frontend of the system are evaluated into linguistic expressions for Loyalty (Turnover, Payment Behavior, Proximity and Specialty) as shown in equations (2) – (5).

$$\mu(x) = \begin{cases} 0, & x < a_1 \\ (x - a_1)/(a_2 - a_1), & a_1 \leq x < a_2 \\ (a_3 - x)/(a_3 - a_2), & a_2 \leq x < a_3 \\ 0, & x > a_3 \end{cases} \quad (1)$$

$$TO(x) = \begin{cases} 25 \leq x < 50, & \text{"Low Turnover"} \\ 50 \leq x < 75, & \text{"Medium Turnover"} \\ 75 \leq x \leq 100, & \text{"High Turnover"} \end{cases} \quad (2)$$

$$PB(x) = \begin{cases} 25 \leq x < 50, & \text{"Less Attractive"} \\ 50 \leq x < 75, & \text{"Attractive"} \\ 75 \leq x \leq 100, & \text{"More Attractive"} \end{cases} \quad (3)$$

$$PR(x) = \begin{cases} 25 \leq x < 50, & \text{"Close"} \\ 50 \leq x < 75, & \text{"Closer"} \\ 75 \leq x \leq 100, & \text{"Closest"} \end{cases} \quad (4)$$

$$SP(x) = \begin{cases} 25 \leq x < 50, & \text{"Low"} \\ 50 \leq x < 75, & \text{"Medium"} \\ 75 \leq x \leq 100, & \text{"High"} \end{cases} \quad (5)$$

$$LO(x) = \begin{cases} 25 \leq x < 50, & \text{"Pseudo"} \\ 50 \leq x < 75, & \text{"Latent"} \\ 75 \leq x \leq 100, & \text{"True"} \end{cases} \quad (6)$$

$$\mu_{low-To}(x) = \begin{cases} 0, & x < 0 \\ \frac{x-0}{25}, & 0 \leq x < 25 \\ (50-x)/25, & 25 \leq x < 50 \\ 0, & x > 50 \end{cases} \quad (7)$$

$$\mu_{med-To}(x) = \begin{cases} 0, & x < 25 \\ \frac{x-0}{25}, & 25 \leq x < 50 \\ (50-x)/25, & 50 \leq x < 75 \\ 0, & x > 75 \end{cases} \quad (8)$$

$$\mu_{high-To}(x) = \begin{cases} 0, & x < 50 \\ \frac{x-0}{25}, & 50 \leq x < 75 \\ (50-x)/25, & 75 \leq x < 100 \\ 0, & x > 100 \end{cases} \quad (9)$$

Values are assigned to the linguistic variables; Turnover, Payment Behavior, Proximity and Specialty towards loyalty determination. Labels were assigned to each linguistic stating the degree of the assigned value as shown in Equation (7)-(9). Equation (2) evaluates the degree

of highness of the Turnover, if the value of Turnover is for instance, 50, the degree of influence for medium evaluates to 0.5 (50%) severity, whereas, 75 evaluates to 0.75 (75%) for high Turnover. MatLab is employed in this paper for the membership function plots, the graphical formats which show the fuzzy membership curves for the Turnover, Payment Behavior, Proximity, Specialty and Loyalty and presented in Figures (4 – 8) respectively with overlap parameters of 0.5 based on the defined ranges. Triangular membership functions are used to describe the variables. The degree of membership for a “Turnover” of 65 for instance, projects up to the middle of the overlapping part of the "medium" and "high" function so the result is "medium" membership = 0.40 and "high" membership = 0.60, while low is zero. Only rules associated with "medium" and "high" turnover will actually apply to the output response. We obtain our rule base from derivation based on the supermarket’s records, expert experience, control engineering knowledge and the experience of a manager who is working at Jane and Juliet supermarket for over 5 years. The expert also assisted in defining the fuzzy rules and the fuzzy set. There are 4 inputs in the knowledge base namely; turnover, payment behavior, proximity and specialty, with 3 fuzzy sets each as antecedent parameters and 3 fuzzy sets each as consequent parameters.

Also, being that this research used an object oriented analysis and design (OOAD) methodology in its implementation; a Unified Modelling Language (UML) Use Case diagram (in figure 2) is used to model the user interactions with the system as well as interactions with external components like the security, fuzzifier, inference engine and the defuzzifier. The use case description for figure 2 is shown below.

4.1 Description:

The user enters customer’s Turnover, followed by Payment Behavior, Proximity and Specialty in numbers. The system verifies if those inputs are valid inputs. If inputs are not valid system prompts re-entry or eventually reports malicious input attempt. If inputs are valid, Fuzzifier receives this input and transforms them into linguistic variables with the help of a triangular input membership function. These linguistic variables are transformed, with the help of a domain expert into inference rules which is selected by the inference engine on receiving inputs. It is based on the rule that eventually fires that the defuzzifier aggregates with the help of a mathematical defuzzification model to obtain defuzzified and crisp loyalty value. Finally, the user obtains this crisp loyalty value and determines which customer has the highest loyalty score based on previous evaluations and effectively decide to reward the customer with the highest loyalty.

4.2 Trigger:

1. Operator enters customer’s Turnover, Payment Behavior, Proximity and Specialty.

4.3 Actors:

1. J&J Shopping Mall Operators.
2. Fuzzifiers
3. Inference Engine
4. Defuzzifier

4.4 Pre-Conditions:

1. Connection to user database
2. Connection to knowledge base

4.5 Goals (Successful Conclusions):

1. Generate customer final/actual loyalty
2. Decide on customer to reward

4.6 Failed Conclusions:

1. Incomplete Input
2. Invalid Input

4.7 Extensions:

1. If input is still invalid after three (3) attempts (delay for 5 minute)
2. If Input is incomplete (emphasize field as required)

4.8 Steps of Execution:

1. Operator Inputs customer data
 - A. Customer data is not complete.
 - B. Prompt user to complete remaining fields
2. Validate Input.
3. Compute Loyalty.
4. Generate report.
5. Operator marks loyal customers in descending order.

The mathematical model that illustrates the computation carried out at the inference and defuzzification stage can be described as follows:

4.9 Inference Model:

- α_i represents the matching degree of a given input which satisfies the condition of the i th rule and $i=1, 2, \dots, 120$.
- \wedge represents the fuzzy logical AND operator.
- $T_i(v_0)$ represents the turnover value of a particular rule numbered i (ranging from 1 to 120 in the rule base).
- $PB_i(x_0)$ represents the Payment Behavior value of a particular rule numbered i (ranging from 1 to 120 in the rule base).
- $PR_i(y_0)$ represents the Proximity value of a particular rule numbered i (ranging from 1 to 120 in the rule base).
- $S_i(z_0)$ represents the Proximity value of a particular rule numbered i (ranging from 1 to 120 in the rule base).
- $L_i(w)$ represents the rule consequence of a particular rule i in the rule base ($i=1 \dots 120$).

Hence, the matching degree of input was given as:

$$\alpha_i = T_i(v_0) \wedge PB_i(x_0) \wedge PR_i(y_0) \wedge S_i(z_0), T_i(v_0) \wedge PB_i(x_0) \wedge PR_i(y_0) \wedge S_i(z_0), \dots, T_i(v_0) \wedge PB_i(x_0) \wedge PR_i(y_0) \wedge S_i(z_0)$$

And Mamdani Max-Min inference model is used for the final aggregation of the firing rules is given below.

$$L(w) = L'_1(w) \quad L'_2(w) \quad L'_3(w) \quad \dots \quad L'_n(w)$$

where $L_i(w) = \alpha_i$
and

$$L'_i(w) = (\alpha_{i1} \quad L'_{i1}(w)), (\alpha_{i2} \quad L'_{i2}(w)), \dots, (\alpha_{in} \quad L'_{in}(w))$$

4.10 Defuzzification Model:

At the defuzzification stage, the centroid of gravity method was adopted. Below is the model and representations;

$\mu_A(x)$ represents membership value in the membership function

x represents the center of membership function.

Finally, this model is used to compute the crisp output as shown below;

$$\text{Crisp output} = (\sum \mu_A(x).x) / \sum \mu(x)$$

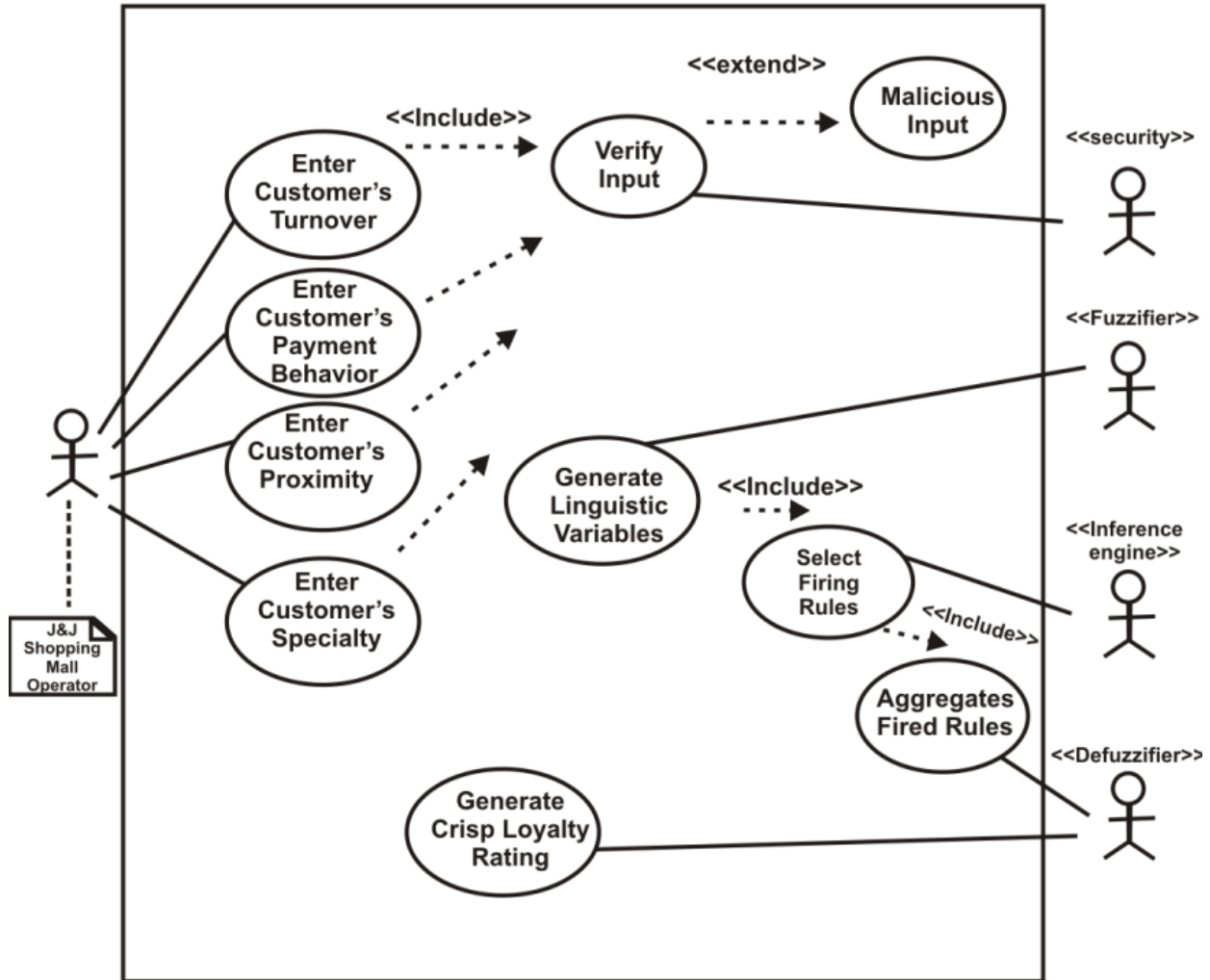


Fig. 2 UML-Use Case diagram of the Hybrid Intelligent Customer Reward system

5. EXPERIMENTATION AND RESULT DISCUSSIONS

The research study adopted Matlab Simulink toolbox to simulate user interface and fuzzy inference and to carry out the experimentation of the architectural design.

Here is the Result of evaluation of fuzzy rule base inference for two ranges of inputs; Turnover, Payment Behavior, proximity and specialty at 25 50 60 80 and 25 40 60 70 as shown in table 1 and 2 respectively.

Table 1. Rule base evaluation for Turnover, Payment Behavior, Proximity and Specialty at 25 50 60 and 80

Rule No.	Premise Variables				Conclusion Part of rule	Min. Value (non-zero)	Max. Value
	Turnover	Payment Behavior	Proximity	Specialty			
15	0.75	0.10	0.65	1.0	Latent	0.10	0.10

Table 2. Table Rule base evaluation for Turnover and Payment Behavior at 25 40 60 and 70

Rule No.	Premise Variables				Conclusion Part of rule	Min. value (non-zero)	Max.Value
	Turnover	Payment Behavior	Proximity	Specialty			
14	0.75	1.0	0.65	0.15	No Loyalty	0.15	0.35
15	0.75	0.50	0.65	0.35	Pseudo Loyalty	0.35	

Finally, a defuzzification strategy is used to evaluate crisp output.

The values of the turnover, payment behavior, proximity and specialty indicated in Table 2 are inserted into the rule base under the view rule editor and the outputs computed for all the cases are recorded. The graphical construction of inference mechanism of fuzzy sets in Table 2 is shown in Figure 3 and Figure 4 shows the performance surface for Loyalty (Pseudo=0.4, Latent=0.4 and True= 0.55), generated in the Matlab Fuzzy Logic Toolbox.

Given the inputs [150 225], the defuzzification result leads to the following calculation;

Crisp output= (16*0.75)/0.75=16

This is illustrated in the output membership function in fig 5. After the strength of each membership function linguistic variables (in this case, only “No Loyalty”) has been calculated using Max-Min method. The value is mapped to x-axis of the graph ($\mu_A(x)$) for each. The center of the linguistic variable is determined and clipped to the x-axis ($\mu_A(0.75)$). This gives us $\mu_A(0.75)= 16$, $x=0.75$ which when substituted in the crisp output formula, gives a crisp value of 16.

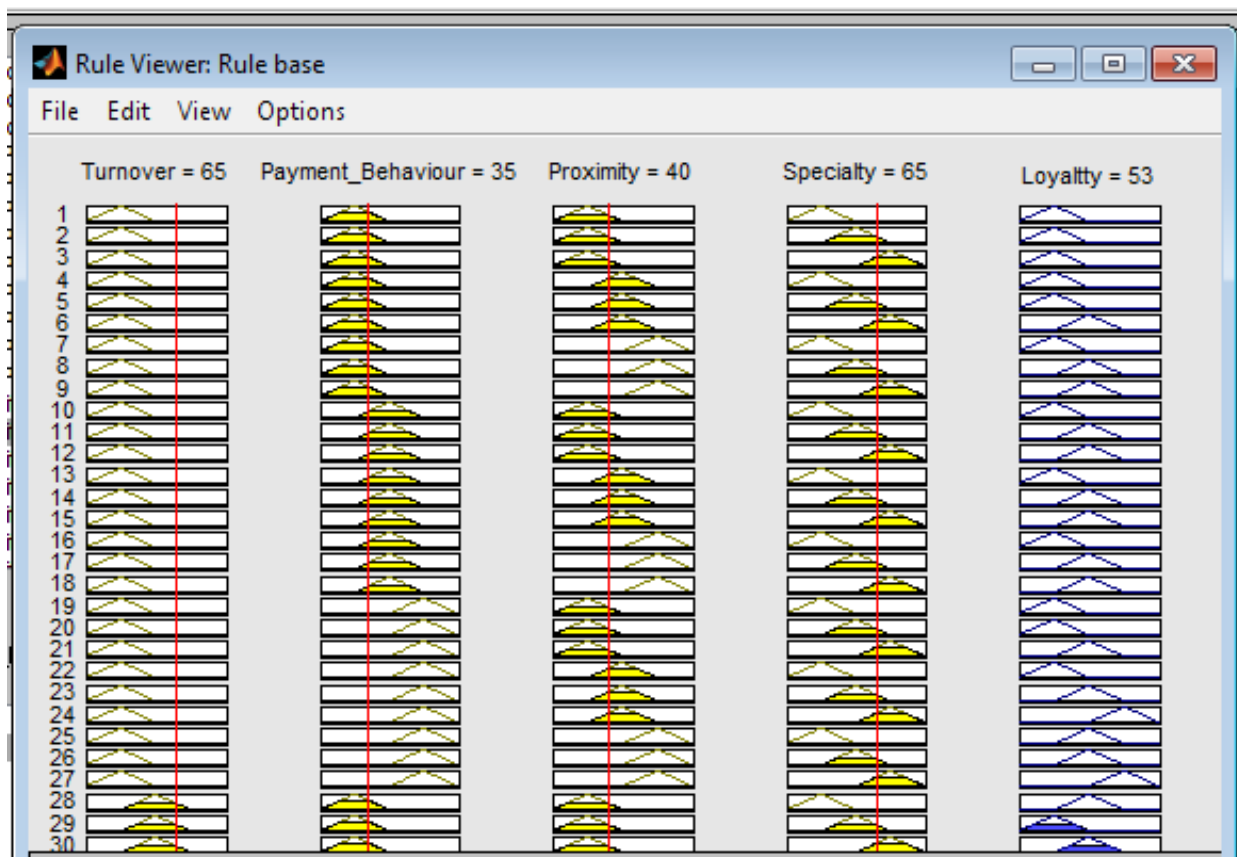


Fig. 3 Visualization of the inference mechanism for the fuzzy expert system

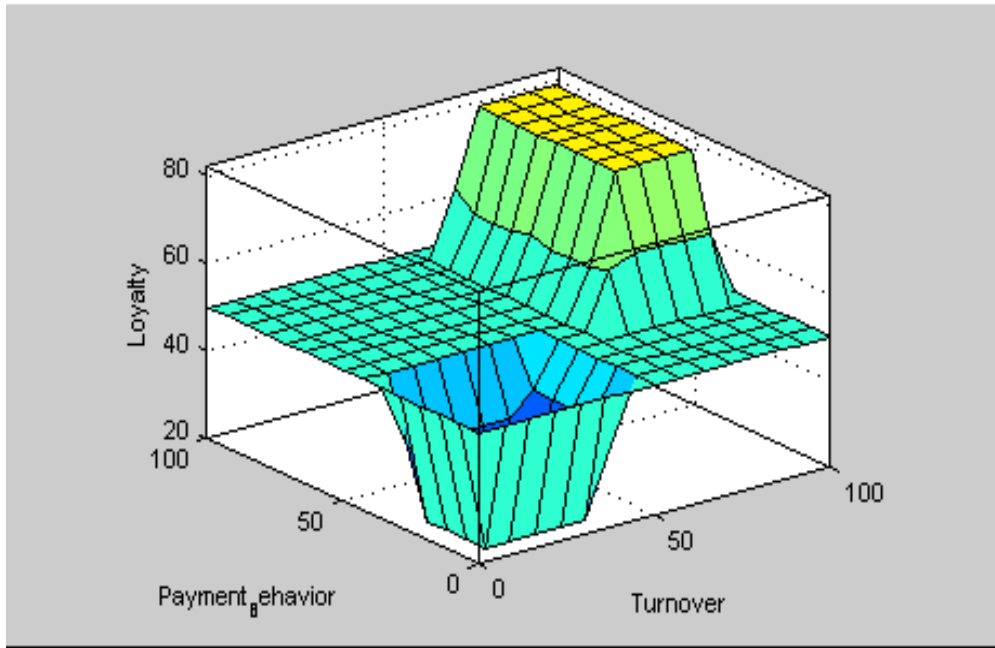


Fig. 4 Fuzzy expert system inference result in Matlab surface view

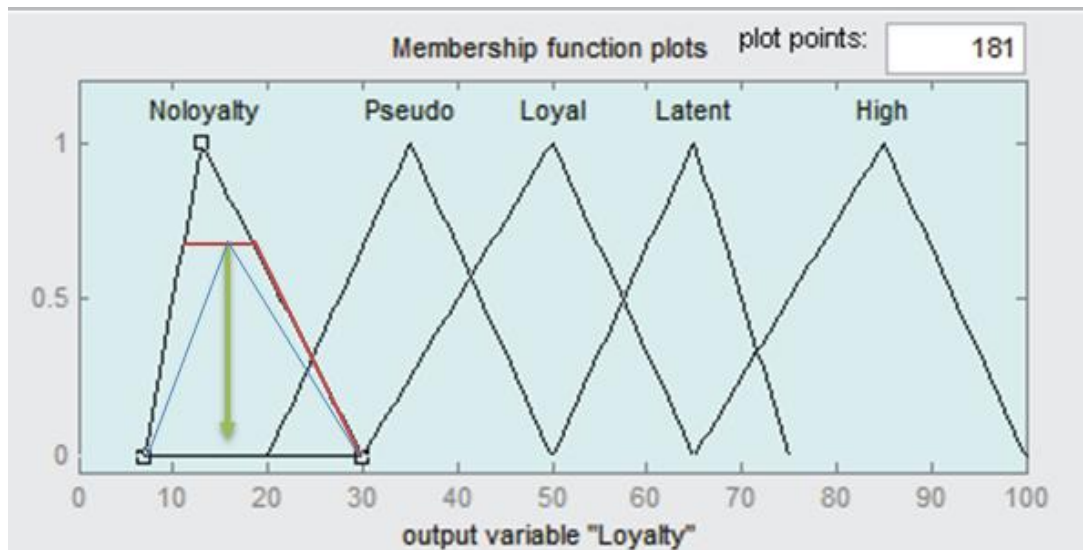


Fig. 5 Fuzzy Centroid of Gravity Mapping the Max-MIN Fuzzification to the Crisp Output (on the x-axis).

These particular input conditions indicates value of 75% (No Loyalty), therefore No Loyalty is expected by such a customer with 75% possibility which is the required system response. Also from the final output, a rating of 16 out of 100 is obtained for the customer's loyalty. Finally, linguistic variables are attached to the crisp output using the scaled function in Eqn. (10).

$$LO(x) = \begin{cases} \text{No Loyalty,} & x < 20 \\ \text{Pseudo Loyalty,} & 20 \leq x < 50 \\ \text{Loyal,} & 50 \leq x < 65 \\ \text{Latent Loyalty,} & 65 \leq x < 75 \\ \text{High Loyalty,} & x \geq 75 \end{cases} \quad (10)$$

6. CONCLUSION

This paper has succeeded in applying two artificial intelligence techniques to build an improved hybrid intelligent customer loyalty and reward system. The Hybrid System from its implementation and experimentation has been confirmed to deliver 92% improvement when compare to other systems that used fuzzy logic model alone to solve this particular problem. With the help of this new system, customer relationship level within organizations has been noted to increase drastically. However, other artificial intelligence methods can as well be combined to improve systems that were originally designed and developed with a single AI technique. For instance, in subsequent researches, a Neuro-Fuzzy System or a Genetic Fuzzy System can be employed in the same problem area to achieve optimized solutions.

7. REFERENCES

- [1] Bia/Kesley (2015). U.S. CRM; Key to business successes. *Bia/Kelsey report*. U. S.
- [2] Cheng C. and Chen Y. (2009) “Classifying the segmentation of customer value via RFM model and RS theory”, *Elsevier Science Ltd*, 36(3), 4176–4184.
- [3] Hosseini S. M. S., Maleki A. and Gholamian M. R. (2010) “Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty”, *Elsevier Science Ltd*, 37(7), 5259–5264.
- [4] Ng K. and Liu H. (2000) “Customer Retention via Data Mining”, *Kluwer Academic Publishers*, 14(6), 569–590.
- [5] Nilashi M. and Ibrahim, O. B. (2014). “A Model for Detecting Customer Level Intentions to Purchase in B2C Websites Using TOPSIS and Fuzzy Logic Rule-Based System”. *Arabian Journal for Science and Engineering*. 1907–1922.
- [6] Nugroho I., Manongga, D. and Utomo W. H. (2013) " ID3 Algorithm to Identify Customer Loyalty Factor at Semarang Ceramics Company", *International Journal of Computer Applications*, 0975 – 8887.
- [7] Rygielski, C., Wang J. and Yen D. C. (2002) “Data mining techniques for customer relationship management”. *Elsevier Science Ltd.*, 483–502.
- [8] Setiyawati N., Utomo W. H. and Manongga, D. (2015). “Analytical Customer Relationship Management for Garage Services Recommendation Using the Generalized Sequential Pattern Method”. *International Journal of Computer Science and Software Engineering (IJCSSE)*, 95-101.
- [9] Silver D., Newnham L., Barker D., Weller S., and McFall J. (2003) “Concurrent Reinforcement Learning from Customer Interactions”. *Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA*. 1-9.
- [10] Tsai C. and Lu Y. (2009) “Customer Churn Prediction by Hybrid Neural Networks”, *Elsevier Science Ltd*, 36(10), 12547–12553.
- [11] Verbeke W., Martens D., Mues C. and Baesens B. (2011) “Building comprehensible customer churn prediction models with advanced rule induction techniques”, *Elsevier Science Ltd*, 38(3), 2354–2364.
- [12] Xie Y., Liu X., Ngai E. W. T. and Ying W. (2009) “Customer churn prediction using improved balanced random forests”, *Elsevier Science Ltd*, 36(3), 5445–5449.