Enhancing Customer Relationship Management Using Fuzzy Association Rules and the Evolutionary Genetic Algorithm

Ahmed Abu-Al Dahab[1], Riham M Haggag[2], Samir Abu-Al Fotouh[3]

Business Information Systems, Helwan University, Helwan, Egypt[1, 2]
Accounting Information Systems, Mansoura University, Mansoura, Egypt[3]

Abstract—The importance of Customer Relationship Management (CRM) has never been higher. Thus, companies are forced to adopt new strategies to focus on customers, given the competitive climate in which they operate. Also, companies have been able to maintain customer data within large databases that contain all information related to customers, thanks to the tremendous technological development seen recently. Multilevel quantitative association mining is a significant field for achieving motivational associations between data components with multiple abstraction levels. This paper develops a methodology to support CRM to improve the relationship between retail companies and their customers in the retail sector to retain existing customers and attract more new customers, by applying data mining techniques using the genetic algorithm through which an integrated search is performed. The proposed model can be implemented because the proposed model does not need the minimum levels of support and trust required by the user, and it has been confirmed that the algorithm proposed in this research can powerfully create non-redundant fuzzy multi-level association rules, according to the results of these experiments.

Keywords—*Customer relationship management (CRM); fuzzy association rule mining; multilevel association rule; quantitative data mining*

I. INTRODUCTION

Customer Relationship Management (CRM) is a critical component of an organization's information system architecture on which organizations fully rely to improve business relationships with customers. Depending on information technology, including the Internet, data warehouses, data mining, etc. After-sales service achieves a lot of the results that are obtained, which enable companies to talk, link, and exchange goods and services with their customers through various of media, including the Internet, call centers, phone calls, faxes, sales personnel, and surveys to attract new customers and retain customers and increase the happiness of customers, and gain the trust and loyalty of customers, which mainly contribute to achieving the greatest return on profitability.

One of the most important factors in the survival of companies is the relationship that binds the company to their customers. It is important to understand that when companies collect a large volume of unmanaged data, such as consumer transactions or sales, decision-makers will not benefit much from this type of data. Companies must use modern technologies that help analyze a huge amount of data, process it, and convert it into useful information that can be understood and analyzed, which will give them good management skills and ideal basic competitiveness in the market [1].

In the information age, data mining is one of the newest and fastest-growing sub-solutions in machine learning. Knowledge discovery from data (information) is a variety of techniques for extracting common patterns from large or high-dimensional data.

Sets are technologies that provide us with accurate information that we may apply in a variety of fields, including business, engineering, and medical sciences. To find new knowledge, different strategies can be used to produce interesting rules known as association rule mining [2].

The Genetic Algorithm (GA) is a representative model that draws inspiration from the theory of natural evolution. Genetic algorithms are used because they are the most advanced and least complex algorithms when compared to other algorithms, for the variety of applications that they can be used for, and because they employ the scientific research method to identify the set of repetitive elements. This approach is better and easier to use than other genetic algorithms this heuristic method is used routinely to maintain high-quality responses in order to simplify the research of problems, dealing with them, the ability to address them, and the selection of the best solutions and results. These algorithms are widely used in mining important data to determine association rules. They are also used to find association rules in practical issues such as business databases and fraud detection [3].

Data mining efforts aim to identify association rules at the individual concept level, finding more comprehensive and meaningful knowledge by processing data and extracting the required information at the same time by exploring the different levels of ideas. Classifications of related elements can be thought of as hierarchical trees, pre-mapped for real-world uses. Inner nodes define classes or concepts that are created from lower-level nodes; [4] the leaf nodes of the hierarchical tree represent the real elements that the transactions are searching for. Fig. 1 provides a straightforward illustration of this. There are many reasons to mine multilevel association rules, including (a) the fact that these rules are more logical and user interpretable. (b) We can find answers to unwanted and

undesirable rules using the rules of the multilevel association. Applications that use spatial data analysis are encouraged [5].
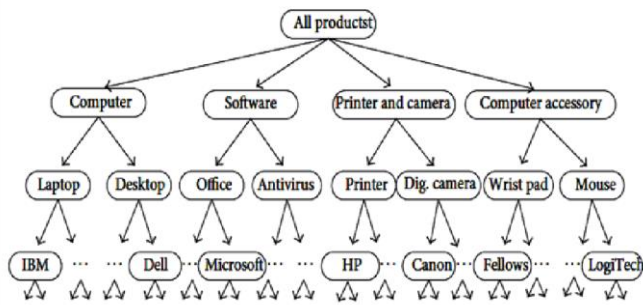

Fig. 1.   The established taxonomy.

A. Motivation and Rationale

Customer contact and after-sales services have become major components of corporate strategies due to the increasing competition for retail activities. In the past, companies used to focus more on selling products and services than on looking into the details of the people who made those purchases and neglecting their wants and needs. And because of the abundance of competition, companies needed to intensify efforts to retain their existing customers, because it was difficult to obtain new ones. In addition, when economic and social conditions were adjusted for, customers were less likely to respond positively to any marketing communications from companies due to their lifestyles [6]. Companies have evolved from product/service-focused strategies to customer-focused strategies as a result of this environment. Thus, developing lasting relationships with clients has become a major strategic objective. In fact, to maintain a positive business relationship with customers, companies that want to stay ahead of the curve must constantly improve service standards. Some companies have invested in building large databases, which enable them to keep a large amount of data related to customers. Numerous data are collected for each customer, allowing analysis of the consumer's complete purchase history. But the knowledge gathered is rarely used to build enterprise processes such as Customer Relationship Management (CRM). In fact, most companies do not incorporate knowledge into their decision-support processes. Information overload and knowledge starvation are recurring problems caused by massive amounts of data. The rate at which analysts can process data has lagged behind in terms of whether the data is usable and relevant to the application [6].

Several trials have been directed at effectiveness (number of bases) and performance (speed). Quality has been neglected in the context of mining the rules of quantitative correlation. The mining algorithms in modern multilevel quantitative correlation rules completely rely on intensive looks at the database to obtain regular models that transcend all the different levels of abstraction [7]. Only users can plan the minimum support related to personal databases and this is based on the assumption of mining algorithms [8] Mining quantitative association rules is not a manageable expansion of mining categorical association rules. Since the search space is unlimited, we aim to detect a measurable set of exciting

solutions (quantitative rules), near the optimal answers. This illustrates why we have decided to solve this search problem with meta-heuristic routines, mainly genetic algorithms [9].

Each of the rules can be regularly defined or inferred and this can confuse the user when managing quantitative correlation rules mining. But more importantly, and some of these rules do not produce new knowledge and these rules may be redundant; some attempts to sell redundant controls in flat datasets. However, datasets can have multiple concept levels or a hierarchy/taxonomy, so the redundancy in these datasets requires modification. This topic is one of the stages of this research. Currently, the approach here is to identify all redundant rules and eliminate them immediately, as a result of which the number of rules that the user has to deal with will be reduced and the information content will not be reduced [10]. To locate familiar elements at different levels of abstraction, this paper presents a modified version of the Apriori algorithm with the aim of mining fuzzy multilevel association rules in large databases [11].

B. Research Problems

Understanding and meeting consumer needs are essential responsibilities for a business to compete in this day of intense competition. As a result, customer relationship management (CRM) now ranks highly among company concerns. How businesses attract new clients and retain existing ones is crucial in today's cutthroat culture. CRM is thought to use appropriate analytical devices with limited materials to attract the most valuable consumers and heighten their desire to increase purchases. The manager can search and examine a vast amount of data using data mining methods to find significant patterns and rules. This approach may also be seen as a means of acquiring knowledge to locate ambiguous association rules to carry out CRM since interactions between the expert system and users can be facilitated by fuzzy knowledge representation. The purpose of this study is to introduce the field of data mining, explicitly mining multi-level quantitative fuzzy correlation rules, to support CRM managers. To achieve precisely defined excellence, it is imperative that by answering these questions, it is probable to be provided in a basic manner. For our research, there are several dominant difficulties that we determined to gather around and seek to crack these challenges and they are:

- Boolean features can be studied as an exceptional example of categorical features and is an important student for generalizing Boolean data mining algorithms for quantitative features, however, we have to either obtain different algorithms or somehow transform the problem of quantitative correlation rules into a Boolean problem. Accordingly, we will use a novel approach to discover quantitative correlation rules emerging from a dataset with multiple concept levels.

- Due to the number of bases, this expands greatly with the number of elements. But this complication will be done by using some advanced algorithms that can cut the search space very efficiently [18]. Choosing this issue mainly helps the user when examining the ruleset. However, the process of developing more valuable quality measures for the rules through the use of the

genetic algorithm with the fitness function (relative confidence of the association rule) and the aim to confirm the most intriguing association rules is indicative of the advanced goals to solve this problem [22].

- Without prior knowledge, the purpose of these valid (estimated) intervals is to extract quantitative data which can be a very complex task. Moreover, these time intervals may not be sufficiently agreeable for experts to quickly gain non-intuitive knowledge from these generated rules. This fuzzy organic function can help with these problems.

- We can completely and effectively specify that there are no redundant association rules in hierarchical datasets.

C. Problem Statement

Market information in the actual world typically contains measurable quantities; therefore developing a sophisticated data-mining algorithm capable of working with quantitative data presents a difficulty to researchers in that field [7] The multilevel association rules mining problem can be defined as follows:

The following elements represent I = {$i_1, i_2, ..., i_n$} and Γ is a classification tree through which the multilevel taxonomic relationships between these elements can be clarified as an example of that field awareness. Element i_1 is the parent of i_2 and element i_2 is the descendant of i_1 if and there are some advantages of element i_1 over i_2. Through which only leaf nodes are displayed within the database. The symbols *DD* represent a database of transactions where each transaction in the database *DD* is a set of elements as T $\subseteq$ I. Each of the transactions is associated with an identifier *TTIIDD*. The symbol P represents the set of positive integers, and the symbol IV denotes the set I × p. Couple $_< x,v _>{\in}I_v$ which means the quantitative attribute x, with the respective values and that $v.I_v$ {$\forall_< x,l,u _>{\in}I _\times p_\times p/l _\leq u$}, l represents the lower bound and u stands for the upper bound of *p*. The next trio $_< x,l,u _>{\in}I_v$ and denotes the quantitative x with a value within the interval [l, u]. Note that the coefficient *TT* contains the item $x_{\in}I$ if *x* if x in *TT* or x is the parent of some of these elements in *TT*. In addition, the transaction X includes $\subseteq$ I with the condition that *TT* holds each element of X.

The multilevel association rule is a product of the $X \Rightarrow Y$ model, where $X \subseteq I$, $Y \subseteq I$, and $X \cap Y =$. No element in *YY* is the source of any element in *X*; This is $Y \cap C$ estors (x) = $_\phi$ because the pattern rule "x$\Rightarrow$ ancestors (x) = $_\phi$ " is fairly true with *100%* confidence, and is redundant. X and Y may each contain Γ elements of any level of Γ [12][13][14].

Quantitative association rule mining still has some restrictions, such as [11]: (1) the design's adoption of a separation of the quantitative attribute prevents it from being usable by all users and attributes. (2) Users, and even specialists, frequently find it challenging to supply certain thresholds, such as the minimal amount of support, curiosity, and confidence. (3) If we use quantitative features, the search space may be very big. Fourth, the algorithm's specified rules may be too numerous to handle.

II. LITERATURE REVIEW

In this section, we compare quantum correlation rule mining algorithms taking into account the shape of these rules, and discuss the drawbacks and advantages of each technique, and the type of database that can be used. [15] [16]:

1) Discretization: The primary objective of this is to convert quantities of data into Booleans by examining the separation of numerical features into groups of intervals. Then this algorithm can be processed to detect the logical association rules to prepare the rules of quantities. Two main representations of the sections are included. Fixed partition, in which the groups of spacers are separated, and another type, where the ends of the spacers are overlapped with each other.

Beyond being the first effort on this topic, the main advantage of this approach is that it can manipulate numerical and categorical data in the same way. Disjoint sets incur harm from the Min_Sup and Min_Conf thresholds, whereas overlapping sets experience the cutting boundary problem. Situations (disjointed or overlapped), however, produce complications. Information will always be lost if intervals are used instead of the actual continuous data. The guidelines we provide will simply be an assessment of the ideal outcomes. Another issue is the expansion of the characteristics dimension; here, the issue is the requirement for additional memory and processing time for these data.

2) Adjusted difference analysis: This approach is based on using discretization and adjusted difference analysis to identify relationships between two properties. Any combination of numbers and categories might be used for the two qualities. This method may distinguish between positive and negative association rules without the requirement for user support or confidence criteria. The fact that it does not require any user criteria and can acquire a new substantial objective measure of the association rules are two of its benefits. Similar to the first strategy, this one has discretization issues that are a drawback. Additionally, this method is unmistakably thought of as producing a special case rule because the rules are always between just two characteristics.

3) A fuzzy approach based on integrating the concepts of fuzzy logic and fuzzy sets with the Apriori algorithm: It reforms numerical data into fuzzy member between [0,1] with a membership function; then operates with the fuzzy member with an adjusted Apriori technique that can comfortably extract the rules, which are stated in linguistic terms. These approaches are based on the fuzzy additions to the classical association rules mining by establishing support and confidence in the fuzzy rule. While the mining results are straightforward to interpret by human operators, two shortcomings still insist on implementing such fuzzy approaches to the original problems. One is the computational time for mining from the database, and the other is the precision of deduced rules. A more formal description, as well as a survey of the existing methods of quantitative association rule mining, can be found in [17].

Numerous scholars have absorbed fuzzy multilevel association rules mining in the literature [13–17]. Without specifying the actual minimum support, certain of these solutions suggested multilayer membership functions via systems of ant colonies and genetic algorithms. Setting the functions for each item and then determining the minimal supports are used to increase computing performance. Other projects benefited from the efficiency and versatility that were done using OLAP and data mining techniques [12].

Up-to-date, there exist only a few algorithms for quantitative multilevel fuzzy association rule mining (QMLFRL). For example, in [25] the authors advised a QMLFRL based on the idea that the minimum support for an item at a higher taxonomic concept is valued as the minimum of the minimum supports of the items pertaining to it, and an item minimum support for an itemset is established as the maximum of the minimum supports of the items enclosed in the itemset. Under this limitation, the characteristic of downward closure is conserved, such that the original Apriori algorithm can be simply prolonged to find fuzzy large item sets..

The authors of [26] provide a brand-new, very innovative genetic-based strategy for choosing criterion values for common item sets. That is the method, an advanced coding technique is chosen, and both the fitness functions have some assurance. The user-specified minimum support is not necessary for our model. Using the genetic algorithm, a thorough search can be carried out. The experiment's findings show that the suggested approach is able to generate fuzzy multilevel association rules that are non-redundant.

The authors of [5] proposed a new technique for extracting quantitative association rules that may concurrently learn rules and quantize an attribute by using a clustering algorithm. They carried out clustering utilizing all qualities simultaneously ahead of time, and they extracted the clusters in the rules from the "association" feature. The authors' technique is superior to the traditional Cartesian product-type quantization technique in terms of total Rule extraction and quantization accuracy, which was supported by the numerical experiments.

The concept threshold - common item groups that are generated by the evolutionary algorithm - is used in the additional pertinent work provided in [3] to develop the quantitative dataset-based rules. In this example, crossover and mutation are used to unify the rule in various ways and can detect the co-occurrence of item sets. Here, comprehensibility, interestingness, and confidence are the three objectives being investigated. As a result, the created laws have developed the principle of multi-objective association. By achieving these goals, the search space for fitness functions is reduced. Finally, distribution-based optimal criteria are developed regarding the numerically valued attribute (A rule's right side displays the distribution of the values for numerical qualities like the mean or variance.)

The advantage of the previous systems is that they contain language terms that make established rules seem much more natural to human specialists; nonetheless, they may produce a significant amount of fascinating association rules. However, since the sparseness of data in three dimensions, makes it not always easy to establish efficient association rules (meeting the lowest level of confidence and assistance) between data points at low (basic) variations in abstraction. Other related issues include (1) inadequate support for hierarchies that are needed to change over time (2) Real-world application requirements cannot be satisfied by algorithm efficiency; (3) the potential to eliminate the connection between several notion levels; (4) Their approach allowed users to specify various things with differing minimal supports[20][27].

To the book [18] that it is necessary to have a customer relationship management system from the requirement of a comprehensive analysis of the market and the study of consumer needs, product developments, and product life cycles. Information mining is the collection of data on the conduct of product sales over time to analyze different market trends. With this information, a product life cycle can be established and new products with some new improvements can be developed according to market trends and customer desires. Although it is usually a good idea to keep arbitrators global, fuzzy arbitrators operate on a group parameter that can be either general or class-specific. They base their choices on a comprehensive review of the product's sales figures in addition to analyzing other data.

Customer relationship management (CRM) seeks to create a "Learning Relation" with customers to help businesses concentrate on their needs, which are the cornerstone of all corporate operations. Businesses put their consumers at the center of management and operations by monitoring client reactions to specific goods and services. Businesses gain the knowledge necessary to raise the caliber of their goods and services by doing this. In other words, businesses find innovative strategies to keep their current clients, attract new customers, and encourage customers' contributions and loyalty to businesses through regular contact and complete knowledge [19].

A. Research Contribution

By examining several analytical characteristics of customer relationship management in the retail industry, the thesis idea presented in this study contributes to the marketing literature. In addition, it provides recommendations for companies on how they can benefit from using CRM analytics to help customers achieve understanding, which enhances customer connection. This paper incorporates the idea that was developed and inspired in part by work on QMLFRL, but that a genetic algorithm is used to calculate a minimum level of confidence and a minimum level of support for each level in this classification regardless of the nature of this data; this makes the system automated. Previous studies, it has fully investigated single-level association rule mining with GA, such as multi-target mining rules and single-level association mining rules. However, when setting up big data analysis, multilevel forms of correlation rules are regularly powerful. Multilevel association rules require big data mining to be a more efficient and effective method. GA-based multilevel association rule mining in this paper is only one attempt to discover multilevel association rules in big data with high efficiency.

III. THE PROPOSED MODEL

This project's main goal was to create practical Heuristic techniques used to mine multilevel association rules in huge datasets, with the primary goal of determining the minimal support and minimum confidence levels for each taxonomic level automatically. By utilizing the genetic algorithm's capacity to swiftly identify without completing exhaustive searches, numerous solutions can be found simultaneously in a huge multidimensional problem. our proposed technique can improve mining efficiency while maintaining the desired precision but avoiding the exhaustive list of potential association rules The definitions listed below are those connected to multilevel association rules [12][13][14]:

The first definition: the element set defines that X is the set of data objects known as the set of data elements $\{X_i, X_j\}$, where $X_i, X_j \in I$. which supports this set X in set S, $\sigma(X / S)$ is the number of transactions (in S) c` overnight X versus the total number of transactions in S. The confidence of $X \Rightarrow Y$ in S, ϕ ($X \Rightarrow Y/S$), is the fraction of $\sigma(X \cup Y / S)$ is in competition with σ (X/S), i.e., the probability that element set Y occurs in S when element mining occurs that X occurs in S.

The second definition: is that the group of elements X is very large in the group s at level L if the size of the support X is not less than the corresponding minimum of the support σ_L' Confidence in the $X \Rightarrow Y/S$ rule is very high at L if its confidence ratio is not lower than the equivalent minimum confidence threshold for ϕ_L'

Third identification: The rule $X \Rightarrow Y/S$ is very strong if $X \cup Y/S$ is large at the current level and the confidence of $X \Rightarrow Y/S$ is high at the current level.

The fourth definition: This mysterious transaction, denoted by the symbol T, is presented by that

$$\overline{\overline{T}} = \{(x, \mu(x)) \mid \forall x \in I\}, 0 \le \mu(x) \le 1, \mu: I \to [0,1], \overline{\overline{T}} \subseteq T$$

Where T is a general set of those coefficients, and $\mu(x)$ is a degree of membership in x.

Fifth definition: A The set of soft quantitative coefficients is denoted by symbols T_q' . Let (F, E) is the soft set over the universe U and $X \subseteq E$, F stands for the fuzzy energy set of U, and E is the set of parameters. It is defined as the set of X attributes that support a transaction if:

$$T_q' = \{(< x,l,u >,e) \mid\mid \forall < x,l,u > \in I \times p \times p \mid l \le u, e \in E\}$$

In general, the focus is always on digging into association rules on a single conceptual level. And some applications cannot locate the link in the multiple abstract levels in the large databases of transactions, where each of those transactions consists of a set of elements and its classification (hierarchy) on these elements, it is expected that the links between these elements will be discovered in any level of this classification. To investigate the process of searching for multi-level association rules, everyone wants to bear the cost of this data in a multi-level association and that on the multiple levels of this abstraction and these effective methods of mining in multi-level rules. We can achieve the first specifications by producing classifications of concepts, from the primitive level of concepts to the higher level of those concepts, or we can find more efficient and effective methods for the process of exploration and research in the rules of the multi-level association.[13].

Only one modification of the Apriori algorithm, which deals with datasets within databases, particularly transaction records, or records including a certain number of fields, or which uses a multi-level "bottom-up approach" has been made to *ML_T2L1* [21] [30]. Which runs the *ML_T2L1* algorithm from the transaction table and these tables contain the hierarchical information in which it is encoded. Each of these levels deals with the data set completely separately. The next first: level *1* search (which is the highest level in the levels of the hierarchy) is done for very large groups of *1* item using the Apriori algorithm. Secondly, we then use the list of large group1 items from level 1 to correct, revise and cut the data set of this transaction for any element that does not have any predecessor or recurrence in the large group1 list of level *1*, which eliminates any transaction that may It has no items in common (so it only has rare items when grading using the Level 1 Big List of 1 Item). This is through the large *1-item* level *1* list, after which the large *2-item* level *1* list is completed (this using the dataset cleaned from before). Then large item sets are inferred from level 1 and this process is repeated until there are no duplicate item groups to find in level 1. This is because *ML_T2L1* only selects items that are offshoots of these duplicate items in level *1* (essentially they must be descended). From this level this set can be for large *1* element) and is always recursive by itself, and the collections of elements are completed at level *2* of the iterative transaction table. For level *2* the large 1-item groups are built, from which the 2 large groups are determined and then the 3 large item groups, and so on. From the same well-filtered datasets) and so on. The fixes for *ML_T2L1* so that all these levels are explored using the Apriori algorithm or the large 1itemsets are not displayed at the level of the proposed system core steps are as follows: [13][14][23][20] [27] [30-32] :
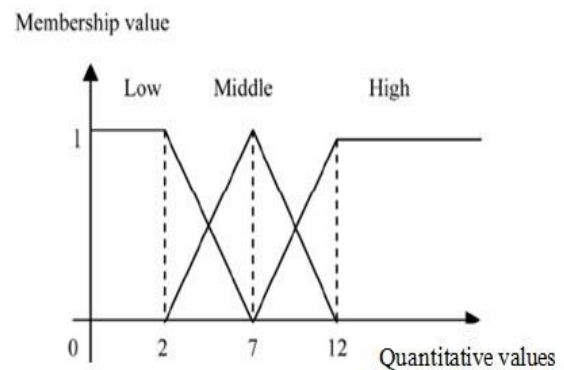

Fig. 2. Shows how the elements in perform as members.

Input: A collection of *N* transactional quantitative data D, a predetermined Γ with the sum of the parent element groups $\{i_1, i_2, ..., i_n\}$, the set of membership functions for each of the elements in the different levels. In our example, all the organic functions have the same form as we can see in Fig. 2; but the primary x-axis for each element is determined based on the

highest quantitative value of each element with which it is associated. Finally, the parameter f is determined by the minimum support α_k and the minimum confidence λ_k that is gained by the genetic algorithm.

The Output: a set of fuzzy multi-level association rules without the constraints of minimum support and optimal trust.

Step.1: All previously defined labels are compiled using their order symbol "*" and numbers based on the formula, $C = \rho * 10 + i$, where i is the position number of the node at the current level l, C denotes the symbol of the node i^{th} at this current level and ρ is the parent symbol of node it i^{th} at that current level.

Step 2: All clause terms are interpreted in each transaction statement and the encryption scheme is approved. Then we define $k = 1$ and $r = 1$ where k, $1 \leq k \leq x$ is the number of the last level, x is the number of the level in a specific classification and r indicates the number of elements that have been saved in the set of recurring elements.

Step 3: We group all elements with the same first k number in each D_i transaction, and enter all quantities of elements in these similar groups into D_i. This stands for the j-th group I^k_j of D_i as v_{ij}^k.

Step4: We have explored a lot of membership functions for each of the different data items where each data item has its features and each has its membership function, and then we convert the v_{ij}^k value of each transaction D_i of all encrypted group I^k_j to fuzzy set f_{ijl}^k (Eq. 1) by plotting a v_{ij}^k over all the given membership functions, where I^k_j is the *j-th* element in level k, v_{ij}^k is the quantum value of I^k_j in D_i and h^k_j is the number of fuzzy regions of I^k_j, R^k_{jl} $(1 \leq l \leq h^k_j)$ is the *l*-th fuzzy region of I^k_j, f_{ijl}^k is the fuzzy organic value of R^k_{jl}

$$\left(\frac{f^k_{ij1}}{R^K_{j1}} + \frac{f^k_{ij2}}{R^K_{j2}} + \cdots + \frac{f^k_{ijh}}{R^K_{jh}} \right) \tag{1}$$

Step 5: We construct the candidate set C_1^k by summing all the fuzzy regions (through linguistic terms) with organic values greater than zero. Where the numerical principal values k_{jl} for each fuzzy region are $R^k_{jl\ n}$ in the

Transactional data as $\quad s^k_{jl} = \sum_{i=1}^{n} f^k_{ijl}$

Step 6: Check to see if the value S k_{jl} for each R^k_{jl} region in C_1^k is greater than or equal to the lower bound of α_k which represents the lower bound of the best support for level k that can be obtained by applying this genetic algorithm to the set of coefficients Included at all this level according to Γ (see. Genetic algorithm. 1). Whether R^k_{jl} is boundary matched, then we put it into the large *1-items* L_1^k collection of level k. that:

$$L_1^k = \{ R^k_{jl} \mid S^k_{jl} \geq \alpha_k, R^k_{jl} \in C_1^k \} \tag{2}$$

Step 7: If L_1^k is empty, we can go to step 3, where $k = k + 1$; If it doesn't, then we create a C_2^K requester array of L_1^1, L_1^2,..., L_1^K to catch the "level crossing" group of elements. Therefore, the pool of C_2^K applicants created must fulfill the following conditions: (1) each pool must consist of two items

in the C_2^K *and* contain at least one item only. L_1^k. (2) the two regions in a group consisting of two elements may not have the same name as this element. (3) that in a hierarchical relationship in this classification, the names of the elements in the set may not be made up of only two elements. (4) the two bulk element sets each containing a candidate element set of 2 must have support values greater than or equal to the minimum

From the support, $\alpha_{k=2}$, in each case.

Step 8: If L_l^k is null, then increase k by one, $r = 1$, and go to step 3 else set $r=r+1$.

(A) create the candidate set C_2^K in the case of $r = 2$, where C_2^K is the set of elements that have been nominated with the two elements at the k level of the set.

L^1_1, L_1^2, L_1^3,..., L_l^k to learn the "level intersection" of the set of repeating elements. Therefore, each of these groups must contain or contain at least two elements in C_2^K; but at the same time L_l^k contains only at least one element; Note that the first element must not be the same as the next in the classification. The two possible combinations are made in C_2^K.

(B) create the candidate group C_r^k if $r > 2$, where C_r^k is the group of elements that were nominated with *r*-items at the k level of $L^k_r - 1$ and this is done in the same way and the same steps mentioned in the previous steps.

Step 9: For each of the obtained *r-itemset* s with element set $(S_1, S_2,....S_r)$ in C_r^k:

The value of fuzzy S is computed in all data for each D_i transaction by using the minimum operator

as $f_{is} = \min (f_{is1}, f_{is2},..., f_{isr})$

B) *Estimate the scalar cardinality of S in all of the* Transactional data as $\text{count}_s = \sum_{i=1}^{n} \int is$

If those numbers are greater than or equal to the set of elements you predetermined including $S = (S_1, S_2,...., S_r)$, $r > 2$ as follows: minimum support α_k place S into L_r^k.

Step 10: In the case of L_r^k is equal to nothing, then we increase the K by only one and move on to the next step for that; if you did not increase r by one, we move to step 8 immediately.

step11: If $k > x$ the next STEP, else set r $=1$ go to stop 3

step12: produced fuzzy association rules for every common r- Itemset, including S = (S1, S2... Sr), r >2 as follows:

- Catch all the rules A$\rightarrow$ B where A$\subset$ S, B $\subset$ S and A$\cap$ B $=\varphi$, A$\cup$B = S.

- Calculate the confidence level for each association rule using by $\frac{\sum_{i=1}^{n} \min (f_{iS})}{\sum_{i=1}^{n} \min (f_{iA})}$.

Step 13: Bases that have high confidence values λ_k and are not less than a pre-determined confidence threshold are selected, where λ_k represents the pre-determined least

confidence value of the *k-level* that was found from the application of the genetic algorithm.

Step 14: We delete all redundant rules in multilevel data sets, and this rule R_1 is considered redundant in rule R_2 if (1) the set of elements X_1 consists of at least two elements and at least one of them consists of a descendant of On the contrary, (2) the group of elements in X_2 consists of at least two elements, provided that at least one of them is a descendant of these elements in X_1, and (3) all other non-original elements in X_2 are all It is in the X_1 element group. In the additional case (4) the confidence interval for R_1 (C_1) is less than or equal to the confidence interval for R_2 (C_2).

A. The Extraction of Parameters using the Genetic Algorithm

We define a genetic algorithm as a type of deep search algorithm that is used to mine the minimum support and minimum support for each taxonomic level. It searches a variety of options to deal with a given problem [26]. The algorithm generates a "big set". One of the scientific solutions to each problem that allows this algorithm to "evolve" over many generations, to find the right optimal solution for each problem. This algorithm begins by selecting a set of solutions (known as chromosomes) on which to start the algorithm. To create a new population, solutions from a single population are selected and maintained. The structure of the basic genetic algorithm is as follows: (see Fig. 3).

Procedure genetic algorithm

```
begin (1)
      t = 0;
      initialize P(t);
      evaluate P(t);
      While (Not termination-condition) do
      begin (2)
            t = t + 1:
            select P(t) from P(t -1);
            recombine P(t);
             evaluate P(t):
      end (2)
  end (1)
```

Fig. 3.   Structure of the genetic algorithm [30].

1) [*Start*] Make arbitrary n-chromosome samples (relevant solutions to the issue).

2) [*Fitness*] Examine the population's fitness (qualifying) function f(x) for each chromosome x.

3) [*The new population*] evolves frequently and generates a new population once the new population is complete, by repeating these following steps.

Selection: Through the population, two of the parents' chromosomes are selected based on their fitness (the higher the level of fitness to the better, the higher the possibility of selection).

Crossover: The crossing over of parents to produce a new generation (children) if there is a possibility of a crossover, and the offspring is the exact reflection of the parents in the case of no crossover possibility.

Mutation: The GA creates a fresh generation of mutations at each spot with a mutation probability (site on the chromosome).

Accepting: In the new society, the privileged new generation is stored.

1) [*Replace*] control freshly produced population to improve the algorithm's path.

2) [*TEST*] If the final condition is met, the program terminates and gives the best solution in this current set to them.

3) [*Loop*] Go the step 2.

The GA maintains a population of results p (t) during repetition t. r_t^1, r_N^t, where r_t^1 refers to the rule set that is produced at random for each level. The function is used to evaluate each answer. E (•) and E (r_t^1) is an indicators of how suitable a solution is. A person's fitness value determines whether they have the necessary ability to live and procreate in the next generation. The next iteration is *(t+1)*. A new resident is formed based on the procedures (2) and (3) [30].

B. Data Encoding

A population often consists of several sets of rules. The coding method used by the system is called the Michigan approach, in which each chromosome is treated as a set of all applicable rules. For each level, with this approach, randomly generated association rules are introduced into each level. Michigan technology uses binary coding for its encryption, which is (0, 1), where the number "1" indicates that the knowledge base is included in the knowledge base, while the number "0" indicates that it is excluded, that is, it does not use it. The main advantage of this technique is that all the rules of the rule are codified, or in other words encrypted, and accordingly, there is no necessity to conduct a quantitative analysis of these rules that cannot be dispensed with, to know and determine that the method in which it is dealt with works as planned or not because it is quite unlike the Bates method, which compiles all possible rules at runtime Genetic algorithm. The size of the chromosome is inversely related to the size of the base of the base, meaning that if the size of the base of the bases increases, the size of the number of groups of elements also increases. Noting that the large size of the chromosome is a major defect [8][29] [31].

C. Generic Operators

It is common to use genetic factors for reproduction, crossover, and mutation. To achieve genetic operators, it is not necessary to select a single member of the population to work on. The aggregation plan depends primarily on the level of physical fitness of the entire population. The system controls the elections in the style of the roulette wheel model. Multiple parents and mutants are selected depending on their fitness, that is, the higher the value of the fitness function a candidate has, the higher the probability of being selected. To implement the sampling of the roulette wheel, first, all applicant values are normalized so that they have the same chance range, then a random number between 0 and 1 is estimated using the chance number function, and a candidate is tested based on a match between this value and an adjustment value The value is based on suitability and, accordingly, the candidate is elected [31].

After an individual has been selected, cloning operators do not move the individual as it is selected; instead, they just replicated it using the existing population through the new population without transfer. Starting with two selected and selected individuals, the intersection operator moves to the transition point (a number among *1* and *L-1* being an integer, anywhere *L* is the length of the strings) chosen at random. The mutation is the third genetic factor. Random shifts in population composition and arrangement can intermittently or irregularly produce beneficial results by deviating from the local optimum. In GA, the sole purpose of this mutation is to oppose each leg of this chain, *i.e.* change 0 to 1 and vice versa change 1 to 0 with the chance of p_m [31].

The algorithm stops working when it reaches a state of deterioration or decay - if the chromosome produced changes between the best and the worst in the population, and this is by less than 0.1%. The implementation is completed when the final number of these generations that the user previously specified reaches. In addition, the algorithm avoids randomly generating the initial set because it can appear in rules that exclude training data in cases where the fitness is too low. Accordingly, a population of the same rules that are locked and contain at least one training instance may result in overfitting in the relevance of this data. The evidence showed that initializing non-random methods can improve the accuracy of the solution and can significantly reduce the running time. [24]. As a result, we developed a new approach to adapting practice case selection to serve as a "seed and foundation" for generating these rules based on a change in the groups of elements within each level. [32].

In general, the genetic factor helps to manage population heterogeneity and prevent early alignment with the local optimum [27]. The primary goal is the exploration of fascinating association rules. As a result, the fitness function is very necessary to check the significance of the chromosome, and it also greatly influences the convergence of the genetic algorithm. In this case, the proposed system checks two different fitness functions. The initial function of fitness is the congruent confidence of the corresponding correlation rule as in Eq. 3, while the connection between the confirmation (conf) and support (sup) properties of the second fitness function, which is necessary to determine the correlation rule (see Equation 4) [8] [12] [27]. The parameters $\alpha\alpha$ and are the important factors for the confidence balance, support in the fitness function, $\beta+\alpha=1$. Using the GA approach to mine confirmed association rules from the huge database, the threshold for the fitness function had to be determined in advance; In this case. $\alpha=\beta=0.5$.

$$f_1: rconf(X \rightarrow Y) = \frac{sup(X \cup Y) - sup(X) \times sup(Y)}{sup(X)(1 - sup(Y))} \quad (3)$$

$$f_2 \ (x \rightarrow y) = \alpha \times sup(x \rightarrow y) + \beta \times conf(x \rightarrow y) \quad (4)$$

By implementing the suggested approach, only the most interesting rules are announced by the fitness function-defined interestingness measure, as opposed to standard mining models that produce an infinite number of interesting rules. Because GA performs a find and deals globally with attribute cooperation better than greedy rule selection algorithms, they

are the preferred method for learning high-level prediction rules [27].

Very briefly, the evolutionary method that has been proposed here for mining quantum correlation rules is driven particularly by the following: (1) the passed rules may be too many to handle; (2) the search space may be too large when we encounter quantitative attributes; (3) users, even experts, usually get bored when specifying the minimum support; and (iv) quantitative attribute splitting is not available for every attribute and user [22]. However, mining association rules fall short in terms of benefits; they also have several flaws, most notably a sophisticated algorithm. With the number of things, the number of rules grows exponentially. However, certain sophisticated algorithms are used to tackle this complexity and effectively reduce the search space. The second challenge is choosing intriguing rules from a set of rules, or attaining rules from rules.

The proposed work tackles the second issue, which essentially helps the user scan the rule set, and useful quality controls on the rules are implemented based on genetic algorithms. Usually, managing association rule mining results in a large number of rules being discovered or inferred, confusing the user. More importantly, some of these criteria might not be necessary and produce no new information. Some efforts have been made to address duplicate rules in flat datasets, but redundancy in these datasets needs to be focused on because they can contain a hierarchy/taxonomy or compound idea levels. One of the characteristics of this study is this problem.

IV. EXPERIMENTS AND RESULTS

We perform various experiments in this section to evaluate the performance of the proposed strategy and validate the improvements. Experiments are performed on a computer Intel ® Core™ i5-245oM CPU @ 2.50 GHz, running on a Windows 7, 64-bit operating system, x64-based processor, and 6 GB of memory. All codes are implemented under MATLAB version 7.8.0.

A. Dataset

We have relied on the information contained in [28] because it is considered an important standard to be used for comparison. And this incoming information consists of the data of the goods, their quantities, and their elements declared in each purchase container that was marketed, which is known as the market basket data. Each set of data contains more than 1,000 sales receipts for sales in a food warehouse, and each asset contains 7 contracts (10,000 transactions).). The pre-determined classification in the first level contains 7 nodes that describe the items that were made in this test, and the second level consists of 14 nodes related to flavor, taste, taste or different types of other stocks and comes in the third level, which consists of 48 nodes that express information about factories and production companies. These transactions contain a database that contains some data such as the name of the product, its shape, and the quantity of the goods that were purchased. However, it is not permissible in any way to use the same element more than once in one transaction.

B. Methodology

The methodology proposed by investigating various analytical CRM characteristics in the retail industry, this thesis contributes to the marketing literature. Additionally, it offers recommendations for businesses on how analytical CRM is used to facilitate consumers' acquisition of information, which in turn improves the relationship with customers. This thesis also emphasizes the potential of large-scale database data mining techniques for CRM.

It was compared to the method typically used in [28] which proposes a method for mining multi-level fuzzy quantitative association rules that applies the GA optimization approach. The goal was to (1) enhance multi-level optimum support and confidence employed to obtain interesting rules to unearth fresh in-depth data. (2) Removing unnecessary rules that were present in the conventional method. Both approaches consolidate fuzzy boundaries rather than explicit border intervals by inferring large item collections top-down and with increasing depth.

In contrast to the conventional method, which relies on specialists to decide these values manually the suggested approach enables the mining of association rules based on the most advantageous recalculated mining parameters (min sup, min conf). By using GA to determine these characteristics, the suggested system becomes more universal and independent from context. The min-sup and min-conf criteria were chosen at 0.29 and 1.8 for each taxonomic level in the investigations.

In the first experiment, by using the data set we test whether we can provide association rules within a given time frame with a fixed number of first generations. From the results presented in Table I, we see that the initial population size ranges from 30 to 100. We can infer the most robust association rules in the dataset but if the population is too large, and quite the opposite if the population is too small The GA-based algorithm is similar to the stochastic algorithm. Because the computational complexity is rising rapidly, we cannot obtain complete association rules immediately. However, as we can see, there is a high chance that most association rules have already been discovered, even with a small population and a short time frame. As a result, we decided to choose 50 cohorts as the default cohort for the data set because it is compatible with the approach.

TABLE I.   THE CONNECTION BETWEEN THE AMOUNT OF MULTILEVEL ASSOCIATION RULES AND THE NUMBER OF INITIAL GA POPULATION (GENERATION N0. 10 APPLYING MICHIGAN ENCODING USING F1)

No. of Initial population	30	50	70	100
N0. 0f association rule (Redundant)	3709	608	608	608
N0. 0f association rule (non-redundant)	1734	348	348	348

As the Michigan primary coding approach is the proposed system, the primary set used is the chromosome. This is the most important reason for the stability of the number of bases extracted with the initial set, which consists of 50 chromosomes, and each chromosome contains a comprehensive depiction of the rules and laws based on the

Michigan method, and therefore each set will include the least number of these initial sets in the selected bases.

Through the following set of experiments, and using the method based on the GA algorithm, we have verified the applicability of the extracted association rules, instead of the traditional algorithm without GA. And that is to measure its value through the 10,000 transactions, using the f_1 function formula through the initial group = 50, the generation number = 10, and the mutation rate = 0.1. The data set results are shown in Table II.

TABLE II.   CONTRAST STUDY

Methods	No. Of Non-Redundant Rules	Calculated *min-sup*	Calculated *min-conf*	Time (Sec)
Proposed Method with GA	348	L1=0.95 L2=0.67 L3=0.27	L1=1.12 L2=1.45 L3=1.92	450 s
Traditional method without GA [8]	2282	0.28 (L=1 to 3)	1.7 (L=1 to 3)	400 s

TABLE III.   COMPARISON STUDY OF THE 2 FITNESS

Fitness Function	levels	Computed Fitness	No. of association rules (non-.redundant)
f_1 (Eq. 3)	1evel 1	1.90	348
	1eve1 2	1.97	
	1evel 3	2.06	
f_2 (E.q.4) α= β= O.5	1eve1 1	1.07	2248
	1eve1 2	1.26	
	1eve1 3	1.47	

TABLE IV.   EVALUATION OF USING DIFFERENT GA PARAMETERS, THE PROPOSED SYSTEM *F1*

Parameters ratio	No. of association rules (.redundant)	No. of association rules (non-redundant.)
Mutation= 0. 9 Crossover = 0.1	750	418
Mutation=0.8 crossover=0.2	750	418
Mutation= 0.7 crossover= 0.3	608	348

As a result of the above, we find that the approach proposed by us, which is based on the GA algorithm, takes a little more time, by an estimated rate of about 13%, so that it can identify the rules of association with high relevance compared to the traditional method. But concerning quality, the system proposed by us can extract rules of higher quality and more interesting by about 17% of the total rules that were extracted from the comparative system based on the traditional method. In general, the decision maker will be hindered if there are a large number of rules extracted from the market basket analysis. The proposed system based on the GA algorithm

presents us with the most interesting rules according to the fitness function, which is responsible for the evaluation that simulates how perfect the answer is: the higher the number provided by the proposed system, the better the solution, and vice versa.

Through the implementation of the third group of experiments, to compare the suitability of the mined association rule using either the fitness function that takes both support and confidence characteristics into account (Eq. 4). or even a fitness function that takes the relative confidence of the associated association rule into account (Eq. 3). Note that the experiment was performed based on the previous configuration of the GA. Accordingly, and through the results presented in Table III, we find that the use of f_1 improves the mining of association rules and generates an additional improvement in the rate of mined association rules while reducing the amount of extracted rules by 83%. Through this experiment, we conclude that the fitness function is a very crucial issue in the success of GA. This was shown clearly as GA did not get any benefit or advantages from using the f_2 state, and thus we got the same amount of bases extracted as we did use the traditional method (about 2248).

The use of f_1 mainly helps in improving the performance of GA, and this is through the optimal extraction of interesting rules, because of the support of each of the elements separately, as well as the calculation of the union support of the elements in each of these rules, and this is quite the opposite of the function that performs based on use support and confidence for each rule, according to the standard case used in current mining algorithms (such as the Apriori algorithm).

By studying the effect of the GA factor settings on the proposed system, which includes both the mutation rate and the crossover rate, and after it has been compared with several multilevel quantitative mining algorithms. Then we'll just change the setting to just one parameter at a time leaving the setting for another parameter at its default value to keep the number of blend settings small. Table IV shows us the number of (non-redundant) association bases that the system was able to identify in the used dataset with the mutation rate changed from 0.7 to 0.9. The above table shows that decreasing the mutation rate will reduce the number of bases extracted by 17%. This regression is very clear. Managing mutations to maintain genetic heterogeneity from one generation to the next was the main reason for this decline. In GA, intersection operators are used very extensively to supervise the population to focus on one of the best solutions yet. Mutation factors are frequently used to provide exploration (exploitation). As a result, while an intersection tries to focus on a specific area of the scene, a mutant does its best to avoid convergence and explore additional areas.

V. CONCLUSION

The proposed system provides rules based on interesting metrics, and we can easily extend the capabilities of the system if needed by changing the fitness function, and the final results show us the possibility of using quantitative correlation rules to support CRM managers. As the proposed system based on association rule mining maintains a very high accuracy when comparing the proposed algorithm with traditional methods. In addition, the extracted rules are very close to reality. This is because different organic functions are adopted for each unique element. The minimum level of support and confidence is improved. Finally, the non-redundant algorithm was used to improve the quality and application of rules to facilitate interactions between the expert system and the users. This method can be viewed as a knowledge acquisition tool for market basket discovery and analysis (MBA), which helps CRM managers to improve their decision-making process. Future work, involving the use of CRM is supported by an association rule based on genetic taboos that use GA to modify the fuzzy membership function of each element.

REFERENCES

[1] V. Shrivastava, "Customer Relationship Management System Developed using Data Mining and Fuzzy Arbiter", International Journal of Research in Engineering, Science and Management, vol.2, issue 2, pp.437-440, 2019.

[2] T. Rahman, J. Kabir, and M. Kabir, "Performance Evaluation of Fuzzy Association Rule Mining Algorithms", International Conference on Electrical Information and Communication Technology, pp. 1-4, 2019.

[3] T. Alam, S. Qamar, A. Dixit, and M. Benaida, "Genetic Algorithm: Reviews, Implementations, and Applications", International Journal of Engineering Pedagogy, vol. 10. Issue 6, 2020.

[4] E. Mahmoudi, E. Sabetnia, M. Torshiz, M. Jalal, and G. Tabrizi, "multi-level fuzzy association rules mining via determining minimum supports and membership functions, " the Second International Conference on Intelligent Systems Modeling, and Simulation, Iran, pp.55-61, 2013.

[5] G. Garai, "Application of Genetic Algorithm in Numerous Scientific Fields", Genetic Algorithms, IntechOpen, 2022.

[6] D. Pravinchandra, S. Shah, and M. Vasavada. Analytical CRM for Google Edge Data Mining Framework with Reference to Pharmaceuticals Industry in India, "International Journal of Management, Public Policy and Research", Vol. 2 issue 1, pp.40-61, 2023.

[7] X. Yang, M. Zeng, Q. Liu, and X. Wang, "A genetic algorithm based multilevel association rules mining for big datasets" Mathematical Problems in Engineering, Vol. 2014, 2014, pp.1-10.

[8] X. Yan, C. Zhang, and S. Zhang, "genetic algorithm-based strategy for identifying association rules without specifying actual minimum support, " Expert Systems with Applications vol. 36, no. 2, 2009, pp. 3066-3076.

[9] S. Aouissi, A. Vrain, and C. Nortet, "Quantminer: a genetic algorithm for mining quantitative association rules", the 20th International Conference on Artificial Intelligence , 2007, pp. 1035–1040.

[10] G. Shaw, Y. Xu, and S. Geva, "Utilizing non-redundant association rules from multi-level datasets" the IEEE International Conference on Web Intelligence and Intelligent Agent Technology, vol. 03, Australia, 2008, pp.681-684.

[11] P. Gautam, and K. R. Pardasani, "Algorithm for efficient multilevel association rule mining", International Journal on Computer Science and Engineering, vol. 2, no.5, 2010, pp.1700-1704.

[12] G. Garai,. Application of Genetic Algorithm in Numerous Scientific Fields. In: Genetic Algorithms. IntechOpen, 2022.

[13] V. Ramana, M. Rathnamma, and A. Reddy, "Methods for mining cross level association rule in taxonomy data structures" International Journal of Computer Applications, vol. 7, no.3, 2010, pp.28-35.

[14] S. Saraf, N. Adlakha, and S. Sharma, " Soft set approach for mining quantitative fuzzy association patterns in databases, " International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, issue 11, 2013, pp.359-369.

[15] M. Delgado, N. Manín, M. J. Martín-Bautista, D. Sánchez, and M. -A. Vila" Mining fuzzy association rules: an overview, "Studies in Fuzziness and Soft Computing, Springer Berlin Heidelberg , vol. 164, 2005, pp.351-373.

[16] G. Attila, "A fuzzy approach for mining quantitative association rules, " ACTA CYBERN, vol. 15, no.2, 2001, pp.305-320.

[17] R. Sridevi, and E. Ramaraj, "A general survey on multidimensional and quantitative association rule mining algorithms, " International Journal of Engineering Research and Applications, vol. 3, issue 4, 2013, pp.1442-1448.

[18] ABHILASHA, Vishal Shrivastava. Customer Relationship Management System Developed using Data Mining and Fuzzy Arbiter. (2019)"."Volume-2, Issue-6.

[19] C. Wang, C. Pang, "Applying Fuzzy Data Mining for an Application CRM", Bulletin of Networking, Computing Systems and Software Vol. 1, Num. 1, pp. 46-51, 2012.

[20] P. Gautam, N. Khare, and K. R. Pardasani," A model for mining multilevel fuzzy association rule in database, " Journal of Computing, vol. 2, issue 1, 2010, pp.58-64.

[21] C. Wang, Y. Li, J. Du and G. Corzo, "Mamdani Fuzzy Inference System for Rating the Performance of Sponge City Programme", International Conference on Systems and Informatics, pp. 1-6, China, 2022.

[22] S. Prakash, M. Vijayakumar, and R.Parvathi, " A Novel Method of Mining Association Rule with Multilevel Concept Hierarchy", International Journal of Computer Applications, vol. 5, no. 5, 2011, pp.26-29.

[23] G. Shaw, Y. Xu, and S. Geva, "Eliminating redundant association rules in multi-level datasets, " the 4thP P International Conference on Data Mining, USA, 2008pp.1-8.

[24] E. Ayetiran, and A. Adeyemo, "A data mining-based response model for target selection in direct marketing," International Journal Information Technology and Computer Science, vol. 4, no. 1, 2012, pp. 9-18.

[25] Y. C. Lee, T. P. Hong, and T. C. Wang, "Mining fuzzy multiple-level association rules with multiple minimum supports, " Expert Systems with Applications, vol. 34, no. 1, 2008, pp.459-468.

[26] D. Kanani, and S. Mishra, "An optimize association rule mining using genetic algorithm," International Journal of Computer Applications, vol. 119, issue 14, 2015, pp.11-15.

[27] Y. Wan, Y. Liang, and L. Ding, "mining multilevel association rules from primitive frequent itemsets," Journal of Macau University of Science and Technology, vol.3, issue 1, 2009,pp.10-19.

[28] Q. Gao, F. Zhang, and R. Wang, "Mining Frequent Sets Using Fuzzy Multiple-Level Association Rules", Journal of Electronic Science and Technology, Vol. 16, Issue 2, pp.145-152, 2018.

[29] S. Manish, A. Agrawal, and A. Lad, "Optimization of association rule mining using improved genetic algorithms," the International Conference on Systems, Man and Cybernetics, vol. 4, USA, 2004, pp.3725-3729.

[30] M. Kayaa, R. Alhajj, "Genetic algorithm based framework for mining fuzzy association rules, " Fuzzy Sets and Systems, vol. 152, 2005, pp. 587–601.

[31] S. Tiwari, M. K. Rao" Optimization in association rule mining using distance weight vector and genetic algorithm , "International Journal of Advanced Technology & Engineering Research, vol. 4, Issue 1, 2014, pp.79-84.

[32] P. Wakabi-Waiswa, and V. Baryamureeba, "Mining high quality association rules using genetic algorithms", In Proceedings of the twenty second Midwest Artificial Intelligence and Cognitive Science Conference, USA, 2009, pp. 73-78.