AN INSIGHT INTO ASSOCIATION RULE MINING BASED ON GENETIC ALGORITHM

[1] Aniket Patel, [2] Jaimin Dave, , [3] Swati Koshti

[1] Assistant Professor IT, [2] Assistant Professor IT [3] Student

[1,2,3] Silver Oak College of Engineering & Technology

Abstract: Data mining is the process that results in the discovery of new patterns in large data sets. It utilizes methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract knowledge from an existing data set and transform it into a human-understandable structure. In data mining, association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases. Association rules are usually required to satisfy a user-specified minimum support and a user-specified minimum confidence at the same time. Genetic Algorithm (GA) is a search heuristic that is routinely used to generate useful solutions to optimization and search problems. In previous, many researchers have proposed Genetic Algorithms for mining interesting association rules from quantitative data. In this paper we represent a survey of Association Rule Mining Using Genetic Algorithm. This paper provides the major advancement in the approaches for association rule mining using genetic algorithms.

***Index Terms* - Data Mining, Association Rule, Genetic Algorithm**

1. INTRODUCTION

Data mining techniques operate on structured data such as corporate databases; this has been an active area of research for many years. The main tasks of Data mining are generally divided in two categories:

Predictive and Descriptive. The objective of the predictive tasks is to predict the value of a particular attribute based on the values of other attributes, while for the descriptive ones, is to extract previously unknown and useful information such as patterns, associations, changes, anomalies and significant structures, from large databases. There are several techniques satisfying these objectives of data mining. Mining Associations is one of the techniques involved in the process mentioned above and among the data mining problems it might be the most studied ones. Discovering association rules is at the heart of data mining. Mining for association rules between items in large database of sales transactions has been recognized as an important area of database research. These rules can be effectively used to uncover unknown relationships, producing results that can provide a basis for forecasting and decision making. The original problem addressed by association rule mining was to find a correlation among sales of different products from the analysis of a large set of data. Genetic Algorithms are a family of computational models inspired by evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure and apply recombination operators to these structures as to preserve critical information. Genetic algorithms are often viewed as function optimizer, although the ranges of problems to which genetic algorithms have been applied are quite broad.

An implementation of genetic algorithm begins with a population of chromosomes. One then evaluates these structures and allocated reproductive opportunities in such a way that these chromosomes which represent a better solution to the target problem are given more chances to reproduce than those chromosomes which are poorer solutions. The goodness of a solution is typically defined with respect to the current population [1] and [3].

Genetic Algorithms are used for a number of different application areas. An example of this would be multidimensional optimization problems in which the character string of the chromosome can be used to encode the values for the different parameters being optimized [4].

The genetic algorithms are important when discovering association rules because they work with global search to discover the set of items frequency and they are less complex than other algorithms often used in data mining. The genetic algorithms for discovery of association rules have been put into practice in real problems such as commercial databases, biology and fraud detection event sequential analysis. In this paper we discussed a survey of Association Rule Mining Using Genetic algorithms.

2. BACKGROUND

Association Rule

The task of mining association rules over market basket data is considered a core knowledge discovery activity. Association rule mining provides a useful mechanism for discovering correlations among items belonging to customer transactions in a market baske database. Let D be the database of transactions and J = {J1... Jn} be the set of items. A transaction T includes one or more items in J (i.e., T ⊆ J). An Association Rule has the form X=>Y, Where X a``nd Y are non-empty sets of items (i.e. X ⊆ J, Y ⊆ J) such that X ∩ Y=Ø. A set of items is called an item set, while X is called the antecedent. The support of an item (or item set) x is the percentage of transactions from D in which that item or item set occurs in the database. In other words, the support of an association rule X=>Y is the percentage of transactions T in a database where X ∪ Y ⊆ T. The confidence or strength c for an association rule X ⇒ Y is the ratio of the number of transactions that contain X ∪ Y to the number of transactions that contain X. An item set X ⊆ J is frequent if at least a fraction support of the transaction in a database contains X. Frequent item sets are important because they are the building blocks to obtain association rules with a given confidence and support.

Essentially, association mining is about discovering a set of rules that is shared among a large percentage of the data. Association rules mining tend to produce a large number of rules. The goal is to find the rules that are useful to users. There are two ways of measuring usefulness, being objectively and subjectively. Objective measures involve statistical analysis of the data, such as support and confidence.

Support

The rule X ⇒ Y holds with support s if s% of transactions in D Contains X ∪ Y. Rules that have a s greater than a user-specified support is said to have minimum support.

Confidence

The rule X ⇒ Y holds with confidence c if c% of the transactions in D that contain X also contain Y. Rules that have a c greater than a user-specified confidence is said to have minimum confidence.

The rule X ⇒ Y holds with confidence c if c% of the transactions in D that contain X also contain Y. Rules that have a c greater than a user-specified confidence is said to have minimum confidence.

Almost all association algorithms are objective and use some form of statistical analysis to determine the usefulness of a rule.Thus the set of all transactions used in the analysis must be sufficiently large in order for association rules to be concluded from it. Therefore, for the rest of this article, the term large will be used to describe a set of data with enough transactions to obtain association rules.

There are a few commonly used terms that must be defined:

- ***Itemset:*** An item set is a set of items. A k-item set is an item set that contains k number of items.
- ***Frequent itemset:*** This is an item set that has minimum support.
- ***Candidate set:*** This is the name given to a set of item sets that require testing to see if they fit a certain requirement [1] and [5].

Algorithms for mining association rules

The "classical" associations rule problem deals with the generation of association rules by defining a minimum level of confidence and support that the generated rules should meet. This is the case since first of all we want the rule to contain items that are purchased often - if we know that whoever buys product A buys product B as well but product A occurs only in two out of ten million transactions then it is of low interest unless the profit margin is high. Furthermore, within the set of transactions that contain item A, we want to know how often they contain product B as well; this is the role of rule's confidence.

If we introduce the term *frequent* for an item set X that meets the criterion that its support is greater than the minimum value set- for example, we might say that we want all items or set of items that were bought by more than 70% of our customers then our problem is restricted to finding all frequent item sets from the database. If we know these, then we can derive all association rules by following a simple strategy. This strategy involves the calculation for every frequent item set X and very subset Y of it- which is neither the null neither set nor X - of the confidence level of all rules of the form X\Y=>Y; the latter is essential so that we will comply with the part of the definition that demands the intersection set to be the null.

As an example, if they consider an item set that consists of pork steaks, coke and beer, the previous definition suggests that we should look for the confidence level of, say, the rule "pork steak => coke and beer". They, then, drop those that do not meet the minimum confidence level criterion. The problem, however, is that a small number of items is able of generating a large search space. This space, on the other hand, has a very interesting property that facilitates our work;there is a border that separates the frequent item sets from the infrequent ones- thus, the problem is restricted on finding that border.

This is done through a mapping procedure between the Set items and the set of natural numbers and the use of special classes. Each algorithm presented below will be characterized by how it looks for the border between frequent and infrequent item sets and how it calculates the support value for each of them. The first can be done either by using the breadth-first search (BFS) or the depth-first search (DFS). In the breadth-first algorithm given a tree and a goal state they try initially all paths- ways of reaching our goal state - of length one, then all paths of length two and so on till they reach the goal state. In the depth-first search they try a path first till they get to a dead-end; then they return to the top and look for alternatives. As far as the calculation of support value is concerned, it can be done using either the number of the subset's occurrences in the database or by using set intersections [5-6].

Strengths and Weaknesses of Association Rules

The first and maybe, one of the biggest advantages of association rules is that the results produced are very easy to be understood. This is the case since they are only statements in plain English, which combine one or many items with other items from the database; for example a customer who buys chips buys Coke as well. This way, the end user can get a string insight of his business and decide what actions he should take.

Furthermore, it is a technique that can be used when we do not really know where to start from but they want to examine the trends in our business. It is probably the best technique to use when they do not have something particular in mind; when they do not want to predict the value of a specific attribute.

The computations that need to be performed are not very complex; that is another strong point of the method. In the classical association rule problem that was described above all they need to calculate is the support and the confidence of a number of rules.

On the other hand, it has drawbacks associated with it. The first is that despite the fact that the computations involved are simple, the search space can be very large with the presence of a relatively small number of items; in fact the relationship between the number of items and the search space is of exponential magnitude.

Moreover, it does not support 'rare' items; these are items that are not purchased very often. As a result of that, the support of any rule, which incorporates those items, will be very low resulting in it being thrown out of the class of frequent item sets; this is the case since the support of the rule would be very low.

Finally, the implementation of the algorithms tends to lead to a very big number of association rules. The most likely case is that they will be confusing and we will have to examine each one of them very carefully to determine if it is of any interest to us. They might, for example, have association rules like whoever buys milk brand A byes serials brand B. It might be a true rule but it might be the result of a promotion. If the latter is the case then they would need to think whether there would be any changes once the promotion is over or if a partnership is formed between another milk brand and a serial's manufacturer [2] and [6].

3. GENETIC ALGORITHM

In the computer science field of artificial intelligence, a genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.

- GAs are inspired by Darwin's Theory about Evolution "Survival of Fittest".
- GAs are adaptive heuristic search based on the evolutionary ideas of natural selection and genetics.
- GAs are intelligent exploitation of random search used in optimization problem.
- GAs, although randomized, exploit historical information to direct the search into the region of better performance within the search space.
- Genetic algorithms find application in bioinformatics, phylogenetic, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics and other fields [3] and [8].

Definitions

Chromosome: A chromosome (also sometimes called a genome) is a set of parameters which define a proposed solution to the problem that the genetic algorithm is trying to solve. The chromosome is often represented as a simple string; although a wide variety of other data structures are also used. We have to redefine the Chromosome representation for each particular problem, along with its fitness, mutate and reproduce methods.

Gene: A Gene is a part of chromosome. A gene contains a part of solution. For example if 162759 is a chromosome then 1, 6, 2, 7, 5 and 9 are its genes.

Fitness: Fitness (often denoted ω in population genetics models) is a central idea in evolutionary theory. It can be defined either with respect to a genotype or to a phenotype in a given environment. In either case, it describes the ability to both survive and reproduce, and is equal to the average contribution to the gene pool of the next generation that is made by an average individual of the specified genotype or phenotype. If differences between alleles at a given gene affect fitness, then the frequencies of the alleles will change over generations; the alleles with higher fitness become more common [1] and [7].

Survey of Association Rule Using GA

a. Optimization of Association Rule Mining through Genetic Algorithm

Rupali Haldulakar and Prof. Jitendra Agrawal et. al. designed a novel method for generation of strong rule. In which a general Apriori algorithm is used to generate the rules after that they used the optimization techniques. Genetic algorithm is one of the best ways to optimize the rules .In this direction for the optimization of the rule set the design a new fitness function that uses the concept of supervised learning then the GA will be able to generate the stronger rule set.

Here they have used Abolan Dataset basically these dataset used for the classification. In this Database 8 attributes are exited. It is provided by the MCI. Genetic Algorithm directly not work on the raw data then whole data had encoded in the form of Binary representation technique (0 and 1). The most important part of Genetic Algorithm is a design of Fitness Function:

$$f(x) = \frac{\text{Support (x)}}{\text{Minsupport}}$$

Support is the Support of New rule generated through genetic operation. Normal case the value of q(Support(x) < min support) is rejected for the better performance of genetic algorithm. We have used class learned classifier for the prediction for rejected those value near to the Maximum value.

The value of q class is divided into two parts C1 and C2. q = {C1, C2}

C1 = {those value or Data min Support less than 0.5}

C2 = {those value or Data min Support greater than 0.5}

Now

$$f(q) = \frac{\text{Support (C2)}}{\text{Minsupport}}$$

The selection strategy based on the basis of individual fitness and concentration pi is the probably of selection of individual whose fitness value is greater than one and f(α) is a those value whose fitness is less than one but near to the value of 1.

Now pi=

$$= \frac{f(xi).e^{-\alpha f(\alpha)}}{\sum f(xj)}$$

Where α is an adjustment factor.

Genetic Operation

The Genetic operators determine the search capability and convergence of the algorithm. Genetic operators hold the selection crossover and mutation on the population and generate the new population.

Select operation*:* In this algorithm it restores each chromosome in the population to the corresponding rule, and then calculate selection probability pi for each rule based on above formula.

Crossover operation*:* In which multi point crossover are used. It classifies the domain of each attribute into a group and classifies the cut point of each continuous attributes into one group and the crossover carried out between the corresponding groups of two individuals by a certain rate.

Mutation Operation: Any bit in the chromosomes is mutated by a certain rate that is, changing “0”to”1”,”1”to”0”.

The algorithm finally extracts the rules that meet the confidence threshold given by users, so the final output is not one optimal rule, but rather a set of rules that meet the threshold.

Rules extraction

The frequent rules are generated according to the fitness function and genetic operators. In order to mine the strong association rules finally, these rules must be extracted again. Extraction criteria are: output the rule which meets the minimum confidence given by users, otherwise abandon it.

They only mined such association rules X⇒Y that Y was age. The setting of parameter: the size of evolutionary population N=100, crossover rate=0.006, mutation rate=0.001.The experiment was executed on Pentium IV CPU 2.58GHZ machine and software was MATLAB (2007). The below table represent the rule generation with genetic and without genetic algorithm for particular support and confidence.

The proposed optimize association rule mining using new fitness function. In which fitness function divide into two classes‟ c1 and c2 one class for discrete rule and another class for continuous rule. Through this direction it got a better result.The genetic algorithm does not sufficient effective and it can’t incorporate with other techniques, so it will need to improve in future work [1].

b. Optimization of Association Rule Mining using Improved Genetic Algorithms

Manish Saggar, Ashish Kumar Agarwal and Abhimunya Lad et.al. proposed to optimize the rules generated by Association Rule Mining (apriori method), using Genetic Algorithms. In general the rule generated by Association Rule Mining technique do not consider the negative occurrences of attributes in them, but by using Genetic Algorithms (GAs) over these rules the system can predict the rules which contains negative attributes. The main motivation for using GAs in the discovery of high-level prediction rules is that they perform a global search and cope better with attribute interaction than the greedy rule induction algorithms often used in data mining. The improvements applied in GAs are definitely going to help the rule based systems used for classification.

The genetic algorithms are applied over the rules fetched from association rule mining. Now for demonstration its utility, the database is produced synthetically. Now, the authors firstly implemented Association Rule mining (using a-priori technique) by the help of their toolkit. And then the GAs are applied to evolve the rules which contains the negations in attributes and are of richer quality.

In this paper the authors have tried to use the enormous robustness of GAs in mining the Association Rules. The results generated when the technique applied on the synthetic database, includes the desired rules, i.e. rules containing the negation of the attributes as well as the general rules evolved from the Association Rule Mining. This technique needs major modifications to improve the complexity reduction of Association rule mining and Genetic Algorithms by using distributed computing [2].

C. Optimized association rule mining using genetic algorithm

Anandhavalli M, Suraj Kumar Sudhanshu, Ayush Kumar and Ghose M.K et. al. is to find all the possible optimized rules From given data set using genetic algorithm. The rule generated by association rule mining algorithms like priori, partition, pincer-search, incremental, border algorithm etc., does not consider negation occurrence of the attribute in them and also these rules have only one attribute in the consequent part. By using Genetic Algorithm (GAs) the system can predict the rules which contain negative attributes in the generated rules along with more than one attribute in consequent part. The major advantage of using GAs in the discovery of prediction rules is that they perform global search and its complexity is less compared to other algorithms as the genetic algorithm is based on the greedy approach.

In this paper the genetic algorithm is applied over the rules fetched from Apriori association rule mining. The proposed method for generating association rule by genetic algorithm is as follows:

- Load a sample of records from the database that fits in the memory.
- Apply Apriori algorithm to find the frequent itemsets with the minimum support. Suppose A is set of the frequent item set generated by Apriori algorithm.
- The output set, which contains the association rule.
- Input the termination condition of genetic algorithm.
- Represent each frequent item set of A as binary string.
- Select the two members from the frequent item set using roulette wheel sampling method.

Apply the crossover and mutation on the selected members to generate the association rules.

- Find the fitness function for each rule
- if (fitness function > min confidence)

- If the desired number of generations is not completed, Apply Apriori algorithm to find the frequent itemsets with the minimum support again.

They have dealt with a challenging association rule mining problem of finding optimized association rules. The frequent Itemsets are generated using the Apriori association rule mining algorithm. The genetic algorithm has been applied on the generated frequent itemsets to generate the rules containing positive attributes, the negation of the attributes with the consequent part consists of single attribute and more than one attribute.

The results reported in this paper are very promising since the discovered rules are of optimized rules. This technique need to minimize the complexity of the genetic algorithm and scanning of database by applying theorem on the generated rule [3].

d. An Improved Algorithm for Mining Association Rules in Large Databases

Farah Hanna AL-Zawaidah, Yosef Hasan Jbara and Marwan AL-Abed Abu-Zanona et. al. Present a novel association rule mining approach that can efficiently discover the association rules in large databases. The proposed approach is derived from the conventional Apriori approach with features added to improve data mining performance. They had performed extensive experiments and compared the performance of the algorithm with existing algorithms found in the literature. Experimental results show that the approach outperforms other approaches and show that approach can quickly discover frequent item sets and effectively mine potential association rules.

This paper they attack the association rule mining by an apriori based approach specifically designed for the optimization in very large transactional databases. The developed mining approach called Feature Based Association Rule Mining Algorithm.

The developed approach adopts the philosophy of Apriori approach with some modifications in order to reduce the time execution of the algorithm. First, the idea of generating the feature of items is used and; second, the weight for each candidate itemset is calculated to be used during processing. The feature array data structure is built by storing the decimal equivalent of the location of the item in the transaction. Transforming here means reorganizing and transforming a large database into manageable structure to fulfill two objectives: (a) reducing the number of I/O accesses in data mining, and (b) speeding up the mining process. There is one mandatory requirements for the transforming technique, that the transaction database should be read only once within the whole life cycle of data mining. By storing the appearing feature of each interested item as a compressed vector separately, the size of the database to be accessed can be reduced greatly.

Using minimum leverage thresholds at the same time incorporates an implicit frequency constraint. e.g., for setting a min. leverage thresholds to 0.01% (corresponds to 10 occurrence in a data set with 100,000 transactions) one first can use an algorithm to find all itemsets with minimum support of 0.01% and then filter the found item sets using the leverage constraint. By using Leverage measure they reduce the generation of candidate's itemsets and thus reduce the memory requirements to store a huge number of useless candidates. This is one of the main contributions of this paper.

This paper is to improve the performance of the conventional Apriori algorithm that mines association rules by presenting fast and scalable algorithm for discovering association rules in large databases. The approach to attain the desired improvement is to create a more efficient new algorithm out of the conventional one by adding new features to the Apriori approach. The proposed mining algorithm can efficiently discover the association rules between the data items in large databases. In particular, at most one scan of the whole database is needed during the run of the algorithm. Hence, the high repeated disk overhead incurred in other mining algorithms can be reduced significantly. They compared our algorithm to the previously proposed algorithms found in literature. The findings from different experiments have confirmed that the proposed approach is the most efficient among the others. It can speed up the data mining process significantly as demonstrated in the performance comparison. Furthermore, gives long maximal large itemsets, which are better, suited to the requirements of practical applications. They demonstrated the effectiveness of the algorithm using real and synthetic datasets. They developed a visualization module to provide users the useful information regarding the database to be mined and to help the user manage and understand the association rules.

The proposed technique need to improve in the mining multidimensional association rules from relational databases and data warehouses and also in mining multilevel association rules from transaction databases [4].

e. Extraction of Interesting Association Rules Using Genetic Algorithms

Peter P. Wakabi-Waiswa and Dr. Venansius Baryamureeba et. al. present a Pareto based multi objective evolutionary algorithm rule mining method based on genetic algorithms. They used confidence, comprehensibility, interestingness, surprise as objectives of the association rule mining problem. Specific mechanisms for mutations and crossover operators together with elitism have been designed to extract interesting rules from a transaction database. Empirical results of experiments carried out indicate high predictive accuracy of the rules generated.

In this paper deal with the ARM problem as a multi-objective problem rather than as a single one and try to solve it using multi-objective evolutionary algorithms with emphasis on genetic algorithms (GA). The main motivation for using GAs is that they perform a global search and cope better with attribute interaction than the greedy rule induction algorithms often used in data mining tasks. Multi-objective optimization with evolutionary algorithms is well discussed.

They proposed to use multi-point crossover operator. There were some difficulties to use the standard multi-objective GAs for association rule mining problems. In case of rule mining problems, need to store a set of better rules found from the database. If they follow the standard genetic operations only, then the final population may not contain some rules that are better and were generated at some intermediate generations. These rules should be kept. For this task, an external population is used. In this population no genetic operation is performed. It will simply contain only the non-dominated chromosomes of the previous generation. At the end of first generation, it will contain the non-dominated chromosomes of the first generation. After the next generation, it will contain those chromosomes, which are non-dominated among the current population as well as among the non-dominated solutions till the previous generation.

The scheme applied here for encoding/decoding the rules to/from binary chromosomes is that the different values of the attributes were encoded and the attribute names are not. For encoding a categorical valued attribute, the market basket encoding scheme is used. For a real valued attribute their binary representation can be used as the encoded value. The range of values of that attribute will control the number of bits used for it. The archive size is fixed, i.e., whenever the number of non-dominated individuals is less than the predefined archive size, the archive is filled up by dominated individuals. Additionally, the clustering technique used does not loose boundary points.

The proposed algorithm was tested on a dataset drawn from the UCI repository of machine learning databases. For brevity, the data used is of a categorical nature.

In this paper they had dealt with a challenging NP-Hard association rule mining problem of finding interesting association rules. The results of this paper were good since the discovered rules are of a high predictive accuracy and of a high interesting value. This technique does not sufficient reliable for a large dataset, it need to improve for the application in large data set [5].

f. Genetic algorithms for the prioritization of Association Rules

M. Ramesh Kumar and Dr. K. Iyakutti et. al. proposed a novel genetic algorithm based association rule mining algorithm is discussed in this paper. Prioritization of the rules has been discussed with the help of genetic algorithm. Fitness function is designed based on the two measures like all confidence and the collective strength of the rules, other than the classical support and the confidence of the rules generated. The algorithm is been tested for the four data sets like Adult, Chess, Wine, Zoo. They presented a novel algorithm for the rule prioritizing, generated by the apriori algorithm through the application of genetic algorithm. The fitness function is designed based on the user‟s interesting measure and M is the threshold value of the interesting measure considered. The sample data sets have been taken from the UCI data repository for the testing of the algorithm. The environmental measure they had for testing is the population size for performing GA is 200, the selection rate is 10% and the cross over rate is 6% and the mutation rate is fixed as 1%.

The proposed genetic algorithm based association rule mining algorithm for the prioritization of the rules. This approach significantly reduces the number of rules generated in the data sets. The fitness function is designed in such a way that to prioritize the rules based on the user preference.

The technique can be extended by the incorporation of the other interesting measures in the literature to future work [6].

4. CONCLUSION

Traditional rule mining methods, are usually accurate, but have very hard and fragile operations. Genetic algorithms on the other hand provide a robust and efficient approach to explore large search space. In recent years numerous works have been carried out using genetic algorithm for mining association rules. As many works have been carried out on mining association rules with genetic algorithms this paper surveys the existing work on application of Genetic algorithm in mining association rules and analyzes the performance of the methodology adopted. During the survey, we also find some points that can be further improvement in advanced association rules mining with genetic algorithm to achieve more efficient accuracy in result and maintain a high confidence and a good coverage of the database, also providing the user with high quality rules.

5. REFERENCES

[1] Rupali Haldulakar and Prof. Jitendra Agrawal, “Optimization of Association Rule Mining through Genetic Algorithm”, International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 3 Mar 2011, pp. 1252-1259.

[2] Manish Saggar, Ashish Kumar Agarwal and Abhimunya Lad, “Optimization of Association Rule Mining using Improved Genetic Algorithms ”IEEE 2004.

[3] Anandhavalli M, Suraj Kumar Sudhanshu, Ayush Kumar and Ghose M.K., “Optimized association rule mining using genetic algorithm”, Advances in Information Mining, ISSN: 0975–3265, Volume 1, Issue 2, 2009, pp-01-04.

[4] Farah Hanna AL-Zawaidah, Yosef Hasan Jbara and Marwan AL-Abed Abu-Zanona, “An Improved Algorithm for

Mining Association Rules in Large Databases", World of Computer Science and Information Technology Journal (WCSIT) ISSN: 2221- 0741 Vol. 1, No. 7, 2011, pp. 311-316.

[5] Peter P. Wakabi-Waiswa and Dr. Venansius Baryamureeba, "Extraction of Interesting Association Rules Using Genetic Algorithms", Advances in Systems Modelling and ICT Applications, pp. 101-110.

[6] M. Ramesh Kumar and Dr. K. Iyakutti, "Genetic algorithms for the prioritization of Association Rules", IJCA Special Issue on "Artificial Intelligence Techniques - Novel Approaches & Practical Applications" AIT, 2011, pp. 35-38.

[7] 1S. Dehuri, A. K. Jagadev, A. Ghosh and R. Mall, "Multi-objective Genetic Algorithm for Association Rule Mining Using a Homogeneous Dedicated Cluster of Workstations", American Journal of Applied Sciences 3 (11), 2006, pp. 2086 2095.

[8] Ansaf Salleb-Aouissi, Christel Vrain andCyril Nortet, "QuantMiner: A Genetic Algorithm for Mining Quantitative Association Rules", IJCAI-2007, pp. 1035-1040.