Suggested Marketing Strategy Using Apriori and FP-Growth Algorithms in retail sales in Egypt

Amira H. SHalaby¹, Dr.Mohmoud M.Abd el letif², Dr.Amira Rezk³, Prof.Mohamed Hagag⁴
¹,³Mansours University, college of CS & IS, Dept of IS, Egypt
²,⁴Helwan University, college of CS & IS, Dept of IS, Egypt

Abstract

Due to the increase of retail sales in Egypt and all over the world, came the importance for the managers of supermarkets to develop marketing strategy to maximize their profits, by getting rid of inactive products. Product bundling is one of the most important marketing strategies used to get rid of stock by making integrated bundles of inactive products and demanded products with discount prices. We can do that through our recommendation system and also increase customers' faith by keeping up with their purchase habits changes in low prices. In this study, first association rules are applied to find the best-integrated bundles with optimal suggested bundle size according to customer habits. Second testing these resulting product bundles to eliminate bundles didn't contain Products aim to get rid of them. Finally the given suggested bundles’ elements replaced with stagnant products that are the same kind of product in the bundle but with another trade. During that study algorithms (Apriori and FP-growth) were studied. Although the two algorithms give strong association rules, and the results were so near. But FP-Growth algorithm was more efficiency as Apriori algorithm caused problems with minimum support parameter as, in a small system transaction, big minimum support did not work using it. Also, it makes the memory PC faces memory hangs when minimum support was very low. On the other hand, all of this didn't appear with FP-Growth algorithm, and it was faster in dealing data.

Indexing terms/Keywords

Product Bundle; Marketing Strategy; Association Rule; Apriori Algorithm; FP-growth Algorithm; Weka
1. Introduction

There is no doubt that Retail sales in all over the world have an observed increase. Fig (1) shows the regional retail sales by % share 2008 compared to Regional 2015f. From that figure we can notice the increase of retail sales percentage in four countries from five countries as examples.

![Fig (1) regional retail sales by % share 2008 compared to Regional 2015f](image)

Egypt have a noticing increase in last years. As reported by United Nations that 63% of the Egyptian population has an economically active (in 2005), the economically active percentage increased to be 65.1% (in 2010). Additionally, United Nation expected retail sales in Egypt will grow from EGP164.52 billion (US$30.30bn) in 2011 to be (EGP 262.26 billion) by 2015 with growth of 159.4%. (Source Business Monitor – 2014).

By 2014 supermarkets growth in Cairo to be 700 supermarkets. Also, hypermarkets are almost ten hyper. Top supermarkets in 2014 were Metro Supermarkets spent USD 110 million in sales across 40 stores. Carrefour has 17 hypermarkets and 70 supermarkets in Egypt. B-TECH began in 1997 with three retail branches, now B-TECH has 57 branches (Source: Planet Retail - 2014).

From that increase in retail sales came the need to product bundles. As it gives satisfactions to customers by providing them products, they need in one or more bundles in law price and satisfactions of retailer by getting rid of stock.

The problem of our paper is how to generate a system that is giving a suggested bundle that some of its products need to get rid of them and also are compatible with other products. And how to determine the size of the bundle that gets a satisfaction of customers and retailer at the same time. In our work, we mine the database of customer's transactions in a supermarket by performing an association rule technique to find the products that have relations with each other’s. As we deal with logical quality relationships in our system, the most suitable algorithms from association rule algorithms will be Apriori algorithm and FP-Growth algorithm. As we know Apriori algorithm is excellent in mining nominal transactions as it's simple to write a program with it, but it take a long time to compile. While FP-Growth algorithm take less time in compiling than Apriori. But it's much more difficult in writing a program than Apriori because of needing to implement the tree and list structure [1]. Weka tool was used to find association rules of items to generate bundles. Once by using Apriori algorithm and another by using FP-Growth Algorithm to compare the result and time be taken, to discover the most suitable algorithm. The remainder of our paper divided as follows:-

Section 2 introduces relevant definitions like product bundle and its relation with marketing and define association rule with it algorithms A priori algorithm and FP – Growth algorithm. Section 3 presents literature survey of product bundle problems using different algorithms. Section 4 describes suggestion of product generation system using A priori and FP-Growth algorithms comparing between each result. Section 5 outlines conclusion and future work.

2. Background

2.1 Marketing Strategies

Marketing examines the need and purchase habits of customers to produce products they need in suitable price and place, promoting for all of this to the customer. This what can we name marketing mix or 4 Ps are the product – place – price and promotion. Marketing strategy is the foundation of marketing plan.

To make marketing strategies, marketing experts need first to make three analysis:-

1) Market analysis (the size of the market and the expecting growth of it, customers purchase habits).

2) Competitor analysis (competitor customer, brands, prices, and so on).

3) Company analysis (business objectives, strengths and weaknesses).
After these analyses, experts must identify target customers. Then decided what strategy to use, putting in their mind the 4Ps [2].

Promotion can be said to be the most important one in the marketing mix, as it is the only way to get way for selling organization’s product. Promotion can be one of the following methods (Public Relations, Sales promotion, Interactive or Internet Marketing, Direct Marketing, Advertising, Personal Selling) the previous methods are called promotion mix [3].

In order the achieve all promotion mix, product bundle appears to be a substantial aid because it is attracting customer with more than one product, product bundle gives the chance to choose integrated product with each other.

2.2 Product bundle

Product bundle means buying two or more product as one product, as products are purchased together in one package always at discount price than buying those singles. The retailer gets the advantage of buying old stored products and the customer get the advantage of discount [4]. So when generating product bundle we have two problems 1) Integrated product bundling 2) price bundling.

1) Integrated Product bundling means how to collect two or more separate products together to form the bundle, without paying attention to bundle price. The only important thing is choosing products that must integrated with each other.

2) Price bundling means buying products with discount in order to get rid of overstock or old product. Price bundling problem can be solved by selling two or more product at less price without paying any attention to the integration between them, or by selling one and get another. The only important thing here is deciding the discount ratio [5].

For both two problems, solutions must satisfy user requirements (implicit requirements), which can be known by knowing their habits according to their previous purchases. And retailer requirements (explicit requirements) [6]. To do this, we need massive database, which is having data about customer habits and products amount in stores. That data is used to extract our knowledge and constraint. From this point of view we need to use data mining.

2.3 Association rule

Reference [7] defines Data mining as a technology that deal with data. Data mining is not data search, but it’s used in data statistic, analysis, extract relations between event, finding rules and future predictions.

When we think about data mining methods, we should have an open mind in inputting data. Data mining specialists classified data mining methods into two sets according to their techniques: -1) Descriptive, which focus on models that summarize data for inference aim. -2) predictive which concentrate on models’ creation to be able to generate prediction results for future cases [8].

In data mining, we have these methods. Estimation, Prediction, association rules, Classification, Clustering, Sequence-based, Fractal-based algorithms, Description and visualization, Fuzzy logic, neural networks. You can use one or more of these methods according to data and purpose of data mining

As our system seeks to find the relations between products to make product bundle, so we will use association rule. As association rule is a method that is looking for Relations between items. Association rule often does (if, then) for example (if A then B or B and C). Association rule depends on min support and min confidence as parameters when generating rules [9].

2.4 Association Rule Algorithms

Association rule algorithms are the way by which data be executed to find a relation between items or item sets. It can be one or more of “classification – A priori algorithm – predictive A priori- clustering – FP-growth algorithm”.

We can choose from these algorithms according to the purpose of the output of the system, and size and type of the database we deal with [10 ]-[13].

2.5 Weka tool

Weka tool grants a common application environment for automatic solving data mining problems in many fields especially in marketing scope. It includes a set of data processing techniques and machine learning algorithms. Like A priori and FP-growth algorithms that can identify relations between data sets (attributes) through generating rules. Weka is an easy tool because it has GUI interface for data exploration and the experimental comparison of different machine learning techniques on the same problem. It has various types of models (like decision trees, rule sets, linear Discriminates). Also, Weka can be installed on almost all operating systems [14].

2.5 Literature survey

References [10], [11], [15], [16] and others treated with an effective breadth-first (level-wise) method for generating frequent sets. Its A priori algorithm that was the core of all algorithms.

While references [13], [14] introduced the problem of management massive discovered rules produced by association rule and theoretical analysis by using the FP-Growth algorithm. But all these researchers focus on generating rules, but with a
little attention to advantage and disadvantage of each algorithm. As association rules was known first as market basket analysis, because it was used for commercial purpose before using it in many fields. Many studies appear in considering how to use it in marketing. We will focus on studies on product bundle. We found that researchers in reference [17] produced engine to find the product bundle of products that satisfy user requirements & seller needs at the same time. This system was called Intelligent Bundle Suggestion & Generation (IBSAG). They use Genetic Algorithm (GA) as a way to discover product bundle that has that structure they generate a bundle of 4 product. But using GA applications have the disadvantage of limitation of performing control in real time because of random solutions and convergence - many issues cannot be settled by genetic algorithm techniques. Limitation occurs due to poorly known suitability functions that generate severe chromosome blocks despite the fact that only good chromosome blocks cross-over [18].

We can use traditional algorithms that can be executed on association rule (A priori algorithm – Filtered Associator – FP-Growth – predictive a priori and Tertius) as traditional algorithms can be made easy by using Weka tool (Frank, 2014). Researchers in reference [19] produced collaborative filtering mechanism for firm's product bundling strategy as they supposed product bundling go through three steps:-

(I) clustering customers. By Using Adaptive Resonance Theory (ART) to organize different kind of customers’ groups according to their habits

(II) Using association rule to find relations between two types of products

(III) Applying Collaborative Filtering to generate personal product bundling list after calculating recommendation system value. But their system produces just two types of products in each bundle as they Depending on finding relation between two products only. In their experiment, the successful probability hit was about 0.6748.

In reference [20] researchers modelled a framework for product bundle system, in which an association rule method used to find relations between the customers and the products they purchase in a specific place after clustering it and their system also helps in finding suggestions for a new places. Their methodology was

The integration of clustering and association rule mining. They concentrate on products marketing according to customers around market place. The purpose of their framework discovering group-based rules of customer if a retailer want to open a new location.

In reference [21], researcher identifies associated products using Apriori algorithm method. He chooses Weka tool to develop the rules. The outputs then be used in designing layout of the Products in the supermarket. The researcher has few transactions (1049 transactions), so the results was 5 category associations. This is the reason why the Aprior didn't make problems with his output.

Many researchers talk about the tool of Weka like in reference [22], They studied comparative case studies (of the utility of association rule mining using Weka tool by using Three association algorithms (Apriori, Predictive Apriori, and Tertius). They found that the three algorithms give a strong association rules, but they have problems with the number of cycles taken to generate the frequent item sets, memory used, Nominal data type and minimum support needed [23].

But they didn’t generate a system using any tool, they just make Comparative analysis between previous association rule algorithms by using ready medical and economic data set. Also their comparative study didn’t deal with FP-growth algorithm.

3. Methods

In our system market basket analysis was used to get association rules in order to discover the relations between products in customer transactional database. The Apriori algorithm is used in the data mining process.

Two algorithms were used, Apriori and FP-growth on the same data under the same conditions. To build comprehension between the two algorithms and discover which will be more suitable.

If conditions statements was applied on the resulted items on

The product bundle. In order to decide which product bundle doesn't contain products like the same type of inactive product to be eliminated.

After that another if statements was applied on the resulted bundle to replace product items with the same type inactive product but with another trade.

3.1 Proposed Model

Suggestion system aims to develop Intelligent Bundle system to support the merchant to produce compatible product bundle that satisfy merchant and user requirements.
The work will split in four steps: (I) preparing data to extract explicit and implicit constraints, (II) Bundle Generator, (III) testing bundle according to explicit constraint in "I", (IV) replace bundle products with the same type inactive bundle product but with a different trademark. (Fig 2 shows these four steps).

We will use dataset of customers' transactions of Sama supermarket (it is a new supermarket locates in Giza – Haram Street). We have transactions for three months. As our system also aims to know the most suitable association rule algorithm we will examine every week transactions, in order to change numbers of attributes and transactions. Table (1) shows the numbers of transactions and attributes every week of three months.

| Week No. | No. of Transactions (customers) | No. of Attributes (product) |
|----------|---------------------------------|----------------------------|
| 1st week | 9998                            | 158                        |
| 2nd week | 5796                            | 79                         |
| 3rd week | 9999                            | 20                         |
| 4th week | 4896                            | 169                        |
| 5th week | 4896                            | 79                         |
| 6th week | 4597                            | 20                         |
| 7th week | 9798                            | 259                        |
| 8th week | 9990                            | 79                         |
| 9th week | 753                             | 20                         |
| 10th week| 4893                            | 59                         |
| 11th week| 9648                            | 159                        |
| 12th week| 4886                            | 79                         |

3.2 Preparing Data

Preparing the data is the first step in our system. In order to manipulate data into a suitable form for analysis and processing, numerous data types are obtainable in association rule. As we have a database of customer transactions with numbers of quantity and price of each item in each transaction. In our system price doesn't matter. But as
we used Weka tool to product bundle process data by using Apriori and FP-growth algorithms. (Fig 3 shows part of the table of transaction after preparing data)

To prepare our retail data we went into four steps:
1) - decide what concentrate questions to solve (what products will be taken into consideration – what type of customer we deal with….etc). In our system, we decided to take all data in customer transaction except calculated fields like (No. of products, etc.).
2) - cleaning data (data normalization and data smoothing).
3) - New data construction. In our system, we skipped that step as there is no new data will be obtained.
4) - data formatting, the most important step in our preparing steps as we added commas instead of tabs (i.e., in Weka formatting).

3.3 Bundle Generation

The second step we used Weka tool on the prepared purchasing dataset. First, we open the first data set (1st-week data set) and choose to compile it using Apriori algorithm. Then we determined min support and confidence. We put them 0.1 and 0.9. To examine the effect of changing minimum support on Apriori algorithm we change it to be 0.3 and 0.05 (we choose these values randomly). After that, we save the three product bundle.

Second, we repeat the first step for the rest of dataset (the eleven data set).
We do the first and second step mentioned above, but by using FP-Growth algorithm.

3.4 Testing Bundle According to Explicit Constraints

After generating bundles, we must compare products in product bundle and quantity of stored products and retailer requirements (explicit constraint).

According to these constraints, we eliminate bundles that didn’t subject to conditions. (Fig 4 shows the flowchart of testing algorithm) that testing was done as a separate step because stored product quantity and retailer requirement may change at any time.

4. result

After executing Apriori Algorithm and FP – growth Algorithm on our supermarket data set by using Weka tool. We get the result appeared in the next sections under different condition listed in the tables below.

4.1 Execution of A priori Algorithm

While we execute our database, we change the number of instances times and the number of attributes other times. Also, we change min support parameter under different conditions. Table 2 (in the next page) shows the result of execution with various conditions.

| id | # of instance | # of attribute | Min support | Min confidence | # of cycles | # of large frequent itemsets | Size of bundle | # of best rules found |
|----|---------------|----------------|--------------|----------------|-------------|-----------------------------|----------------|----------------------|
| 1  | 9998          | 159            | 0.1          | 0.9            | 18          | 8                           | 4,5,3           | 11205                |
| 2  | 9998          | 79             | 0.1          | 0.9            | 18          | 3                           | No rules found |                      |
| 3  | 9998          | 20             | 0.1          | 0.9            | No rules found | No rules found               | No rules found |                      |
| 4  | 4895          | 159            | 0.1          | 0.9            | 18          | 8                           | 4,3,5           | 11880                |
| 5  | 4896          | 79             | 0.1          | 0.9            | 3           | No rules found               | No rules found |                      |
| 6  | 4896          | 20             | 0.1          | 0.9            | No rules found | No rules found               | No rules found |                      |
| 7  | 9998          | 159            | 0.3          | 0.9            | 13          | 4                           | 1,2,3           | 11                   |
| 8  | 9998          | 79             | 0.3          | 0.9            | No rules found | No rules found               | No rules found |                      |
| 9  | 9998          | 20             | 0.3          | 0.9            | No rules found | No rules found               | No rules found |                      |
| 10 | 4896          | 79             | 0.3          | 0.9            | No rules found | No rules found               | No rules found |                      |
| 11 | 9998          | 159            | 0.05         | 0.9            | pc faces memory hangs | No rules found |                      |
| 12 | 4896          | 79             | 0.05         | 0.9            | 19          | 9                           | 3,4,5           | 45679                |

4.2 Execution of FP – growth Algorithm

As we execute our database with Apriori algorithm, and as we change the number of instances times, the number of attributes other times and min support parameter under different conditions we do this typically by using FP-Growth algorithm. Table 3 (in the next page) shows the result of execution with various conditions.
### Table 3: Results of Execution with FP – Growth Algorithm

| id | # of instance | # of attribute | Run Information |          |          |          |          |          |          |          |          |
|----|--------------|----------------|----------------|----------|----------|----------|----------|----------|----------|----------|----------|
|    |              |                | Min support    | Min confidence | Min lift | Min lev | Min conv | Size of bundle | # of best rules found |
| 1  | 9998         | 159            | 0.1            | 0.9        | 1.1      | 0.01     | 1.78     | 4,5,6    | 11164    |
| 2  | 9998         | 79             | 0.1            | 0.9        |          |          |          |          | No rules found |
| 3  | 9998         | 20             | 0.1            | 0.9        |          |          |          |          | No rules found |
| 4  | 4896         | 159            | 0.1            | 0.9        | 1.09     | 0.03     | 1.72     | 5,6      | 11823    |
| 5  | 4896         | 79             | 0.1            | 0.9        |          |          |          |          | No rules found |
| 6  | 4896         | 20             | 0.1            | 0.9        |          |          |          |          | No rules found |
| 7  | 9998         | 159            | 0.3            | 0.9        | 1.44     | 0.12     | 3.74     | 3,4      | 49       |
| 8  | 9998         | 79             | 0.3            | 0.9        |          |          |          |          | No rules found |
| 9  | 9998         | 20             | 0.3            | 0.9        |          |          |          |          | No rules found |
| 10 | 4896         | 79             | 0.3            | 0.9        |          |          |          |          | No rules found |
| 11 | 9998         | 159            | 0.05           | 0.9        | 1.15     | 0.01     | 2.19     | 5,6,7    | 213045   |
| 12 | 4896         | 79             | 0.05           | 0.9        | 1.44     | 0.05     | 3.73     | 5,6,7    | 45452    |

### 4.3 Execution of Testing Bundle According to Explicit Constraints

Based on the scientific analysis in section 4.1, the FP-growth algorithm gives better results than Apriori algorithm when generating our bundles.

So we passed rules that produced by using FP-Growth algorithm to our explicit constraint. Table 4 show comparing between result before and after applying explicit constraint.

![Flowchart of testing algorithm](Fig 5: Flowchart of testing algorithm)

- **Accept the bundle**
- **Eliminate the bundle**
- **Suggested product bundle**
- **Products in retailer product list to get rid**
- **Eliminate the bundle**
- **Products amount > amount average**

---------

*Fig 5: Flowchart of testing algorithm*
4.3.1 Execution of Testing Algorithm

Based on the scientific analysis in section “4.4.1”, the FP-growth algorithm gives better results than Apriori algorithm when generating our bundles.

So we passed rules that produced by using FP-growth algorithm to our explicit constraint. (Table six show comparing between result before and after applying explicit constraint).

Table (4) number of rules (bundles) generated by FP-Growth algorithm

| id | Number of rules (bundles) before passing explicit constraint | Number of rules (bundles) before passing explicit constraint | Number of eliminated rules (bundles) |
|----|-------------------------------------------------------------|-------------------------------------------------------------|-------------------------------------|
| 1  | 11164                                                      | 6015                                                       | 5149                                |
| 2  | 11823                                                      | 6994                                                       | 4829                                |
| 3  | 11823                                                      | 6140                                                       | 5683                                |
| 4  | 213045                                                    | 139487                                                     | 73558                               |
| 5  | 45452                                                      | 26218                                                      | 19234                               |

5. Discussion

As we see in the previous section, the result of execution the two algorithms on our data sets. In the next section, we will analysis each the result of each table to discover our conclusion.

5.1 Result Analysis

The most significant result we found from result illustrated in tables two by using Apriori algorithm:

When min support increase, the number of rules found decreased and may occur no rules found.

The most observed changed can be found, generating rules be so slowly when reducing min support. And if the number of instances is small the memory was hang

But by using FP-Growth algorithm we found, when decreasing the number of attributes causes no rules found, we can solve it by reducing min support to get rules.

5.2 Comparative Analysis Between both results from the result of applying Apriori and FP-growth Algorithms, we found that there is not big difference between two algorithms except when minimum support increase. As it works better with FP-Growth. For example, when it’s 0.05 and 9998 instances and 159 attributes, it didn’t give a result with Apriori algorithm while giving result with FP–Growth. It’s clear that FP-Growth is faster than Apriori algorithm when dealing with data.

The problem of decreasing attribute can be solved by decreasing minimum support parameter. Fig (5) show comparative graph between the results of the two algorithms.

Fig (5) a comparative graph between the result of the two algorithms.
CONCLUSIONS AND FUTURE WORK

This paper introduces a new system platform that has additional features more than previous systems. As we present a suggestion for a system that produce not only integrated products bundle, but also it gives choices of the size of bundles. Keeping into consideration retailer, and customer requirements our system can be used in making an active promotion or even in organizing items in a supermarket. The most important issue of our system is to get rid of inactive products and increase retailer sales. Also, we discovered that using the FP-Growth algorithm in retail data mining be more efficient. As we see when min support decrease than 0.6, instances be more than 5000 transactions and number of attribute increase than 79 products the system be so slow when compiling by using Apriori algorithm. But under the same numbers it worked by using FP- algorithm faster than Apriori. As in retail data mining, there are a lot of related transactions. In our future work, we will focus on finding new methods using data mining also, to solve the problem of product bundling price on integrated products in bundles. And connecting products with small Min support and product with height Min support, having height confidence in the same bundle.

Acknowledgments

The authors would like to thank many thanks to the responsible for Sama market for giving me data I need.

References

[1] Chen, Ch. and Liao, Sh. 2008. ‘Association Rule Algorithms for Logical Equality Relationships’, IEEE 8th International Conference on Computer and Information Technology Workshops. 2008 IEEE Workshops 116. PP: 26-30.
[2] Belch, G. and Belch, M. 2003. ‘An introduction to integrated marketing communication, the role of IMC disclosure of information by consumers to commercial organizations. In: advertising and promotion, six edition, © The McGraw−Hill Companies. pp: 2-65.
[3] Hall, N. 2009. ‘How to achieve an effective promotional mix’. © The chartered institute of marketing. UK.
[4] Chandrashekar, G. 2009. Product Bundling - A Banking Perspective. ©2009 Infosys Technologies Limited, Bangalore.
[5] Roth, S. 2007. Understanding Pricing Objectives and Strategies for the Value-Added Ag Producer. Produced by Information and Communication Technologies in the College of Agricultural Sciences. © The Pennsylvania State University 2007.
[6] Themistocleous, Ch., Smith, A. And Wagner, Ch. 2014. The ethical dilemma of implicit vs explicit data collection: Examining the factors that influence the voluntary disclosure of information by consumers to commercial organizations. 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering. pp: 1-6.
[7] Guoxiang, L. and Zhiheng, Q. 2013. Data Mining Applications in Marketing Strategy. 2013 Third International Conference on Intelligent System Design and Engineering Applications. pp: 518-520.
[8] Setiabudi, D., Budhi, G., Purnama, W. and Noertjahyana, A. 2011. Data Mining Market Basket Analysis’ Using Hybrid-Dimension Association Rules Case Study in Minimarket X’, 2011 International Conference on Uncertainty Reasoning and Knowledge Engineering. 978-1-4244-9983-0/11/$26.00 ©2011 IEEE. Pp: 196-199.
[9] Danso, S. 2006. An Exploration of Classification Prediction Techniques in Data Mining: The insurance domain. [Online]. Available from: http://www.comp.leeds.ac.uk/sccd/MSc%20Dissertation.pdf (Accessed: 2015-07-11).
[10] Gupta, S. and Mamta, R. 2014. A Survey on Association Rule Mining in Market Basket Analysis’, International Journal of Information and Computation Technology ISSN 0974-2239 Volume 4, Number 4. pp: 409-414.
[11] Savi, G., Roopal, M. 2014. A Survey on Association Rule Mining in Market Basket Analysis. International Journal of Information and Computation Technology. ISSN 0974-2239 Volume 4, Number 4 (2014). pp: 409-414.
[12] Shengwei, L., Lingsheng, L. and Chong H. 2009. Mining closed frequent itemset based on FP-Tree. Granular Computing. GRC ’09. IEEE. pp: 354-357.
[13] Verhein, f. January 2008. Frequent Pattern Growth (FP-Growth) Algorithm an Introduction. ©Copyright 2008 Florian Verhein, School of Information Technologies, the University of Sydney, Australia. pp: 3-10.
[14] Frank, E., Hall, H., Trigg, L., Holmes, L. and Ian H., (2014). Data Mining in Bioinformatics using Weka2014. [Online]. Available from: http://www.cs.waikato.ac.nz/ml/weka (Accessed: 2015-07-11).
[15] Roopesh, D. 2007. Data mining techniques for dynamically classifying and analysing library database. 5th International CALIBER -2007, Panjab University, Chandigarh. PP: 08-10
[16] ZHANG, Sh. and YANG, Q. 2003. Data preparation for data mining. Copyright # 2003 Taylor & Francis. pp: 375-381.
[17] Birtolo, C., Chiara, D., Losito S., Ritrovato, P. and Veniero, M. 2013. Searching optimal product bundles by means of GA-based engine and market basket analysis. 978-1-4799-0348-1/13/$31.00 ©2013 IEEE. pp: 448-453.
[18] Reeves, C. and ROWE, J. 2002. Genetic algorithms: principles and perspectives a guide to ga theory. © School of Computer Science University of Birmingham. pp: 58-61.
[19] Guo-Rong, L. and Xi-Zheng, Z. 2006. Collaborative Filtering Based Recommendation System for Product Bundling. Management Science and Engineering. 2006 ICMSE ’06. pp: 251-254.

[20] Bhanu, S. and Pavai, M. December 2009. Retail Market analysis in targeting sales based on Consumer Behaviour using Fuzzy Clustering – A Rule Based Model. Journal of computing. volume 1. Issue 1. issn: 2151-9617. pp: 92-99.

[21] Isti, S. 2005. Design of product placement layout in retail shop using market basket analysis. MAKARA, TEKNOLOGI. VOL. 9. NO. 2, pp: 43-47

[22] Lekhal, A., Srikrishna, C. and Vinod, J. 2013. Utility of Association Rule Mining: a Case Study using weka tool. 978-1-4673-5301-4/13/$31.00 ©2013 IEEE.