User centric approach to itemset utility mining in Market Basket Analysis

Jyothi Pillai
Associate Professor
Bhilai Institute Of Technology,
Durg-491 001, Chhattisgarh,India
jyothi_rpillai@rediffmail.com

ABSTRACT

Business intelligence is information about a company's past performance that is used to help predict the company's future performance. It can reveal emerging trends from which the company might profit [31]. Data mining allows users to sift through the enormous amount of information available in data warehouses; it is from this sifting process that business intelligence gems may be found [31]. Within the area of data mining, the problem of deriving associations from data has received a great deal of attention. This problem is referred as "market-basket problem".

Association Rule Mining (ARM), a well-studied technique in the data mining field, identifies frequent itemsets from databases and generates association rules by assuming that all items have the same significance and frequency of occurrence in a record. However, items are actually different in many aspects in a number of real applications such as retail marketing, nutritional pattern mining, etc [26]. Rare items are less frequent items [32]. For many real world applications, however, utility of rare itemsets based on cost, profit or revenue is of importance. For extracting rare itemsets, the equal frequency based approaches like Apriori approach suffer from "rare item problem dilemma". Utility mining aims at identifying rare itemsets with high utility. The main objective of Utility Mining is to identify the itemsets with highest utilities, by considering profit, quantity, cost or other user preferences [40].

Also valuable patterns cannot be discovered by traditional non-temporal data mining approaches that treat all the data as one large segment, with no attention paid to utilizing the time information of transactions. Now, as increasingly complex real-world problems are addressed, temporal rare itemset utility problem, are taking center stage. In many real-life applications, high-utility itemsets consist of rare items. Rare itemsets provide useful information in different decision-making domains such as business transactions, medical, security, fraudulent transactions, and retail communities. For example, in a supermarket, customers purchase microwave ovens or frying pans rarely as compared to bread, washing powder, soap. But the former transactions yield more profit for the supermarket. A retail business may be interested in identifying its most valuable customers i.e. who contribute a major fraction of overall company profit [40].

In this paper, these problems of analyzing market-basket data are considered and important contributions are presented. It is assumed that the utilities of itemsets may differ and determine the high utility itemsets based on both internal (transaction) and external utilities.

Keywords : Business intelligence, Association rule Mining, Utility Mining, Apriori, Market Basket

1. Introduction

1.1 Data Mining in Business

The Supermarket revolution, first sparked off during 1920s in United states, won acclaim and acceptance almost globally by 1950s. Since last two decades the computerization processes have generated large data repositories, paving way to new intelligent approaches of exploiting these data repositories for Business Planning and forecasting [27].

There is a transition of the economy from an era of competitive advantage based on information to one based on knowledge creation. The earlier era was characterized by relatively slow and predictable change that could be deciphered by most formal information systems[33]. During this period, information systems based on programmable recipes for success were able to deliver their promises of efficiency based on optimization for given business contexts [33].

The new business environment is characterized by radical and discontinuous change, which overwhelms the traditional organizational response of predicting, and *reacting* based on pre-programmed heuristics. Instead, it demands *anticipatory* response from organization members who need to carry out the mandate of a faster cycle of knowledge-creation and action based on the new knowledge [33].

There is an increasing hype about the wonders delivered by newest information technologies in an era characterized by *knowledge* as the critical resource for business activity. With the advent of new technologies, such as data mining, intranets, videoconferencing and web casting, several technologists are offering such solutions as a panacea for meeting the business challenges of the knowledge era.

Data warehousing and business intelligence provide a method for users to anticipate future trends from analyzing past patterns in organizational data [31]. Data mining is more intuitive, allowing for increased insight beyond data warehousing and allow for statistical predictions, groupings and classifications of data.

Most companies collect, refine and deduce massive quantities of data [31]. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources and can be integrated with new products and systems, as they become part of the system. When implemented on high performance client/server or parallel processing computers, data mining tools can analyze massive databases to deliver answers to many different types of predictive questions [31], [34]. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions [31].

Variants by researchers were mainly focusing to improve the efficiency of mining process but business managers were looking more for business rules rather than frequent item sets or association rules, so that the results are directly convertible to their profitability and productivity. Data mining is often performed on an ad-hoc basis where fairly complex business problems are analyzed.

It is since 1993 that Association Rule Mining was first discovered and serves as trigger to the researching activities in the field of Data Mining with large application potential in Market Basket Analysis. Despite the hyperbole the large scale implementations of data mining are not yet visible.

The problem of deriving associations from data was first formulated by Agrawal et al and is often referred as the “market basket problem”. Market Basket Analysis is the analysis of the data with a view to improve the performance of the retail outlet. The canonical example from which the problem gets its name is a supermarket. The items are products and the baskets are customer purchases. Determining what products customers are likely to purchase together can be very useful for planning and marketing [25].

Association Rule Mining (ARM) is an important area of data mining research which deals with the discovery of association rules. ARM is the mining of Association rules for finding the relationships between data items in large datasets. The frequency of itemset is not sufficient to reflect the actual utility of an itemset. For example, the sales manager may not be interested in frequent itemsets that do not generate significant profit. Recently, one of the most challenging data mining tasks is the mining of high utility itemsets efficiently. High utility frequent itemsets contribute the most to a predefined utility, objective function or performance metric [41]. For example, from marketing strategy perspective, it is important to identify product combinations that have a significant impact on company’s bottom line i.e. having highest revenue generating power [41].

1.2 User-Centric Approach to Market Basket Analysis

Most of the ARM approaches discover the association rules pertaining to frequently occurring entities, by considering the utilities of the itemsets to be equal [1]. However, the real world datasets contain both frequent and relatively infrequent or rarely occurring entities. Knowledge pertaining to rare entities may contain interesting knowledge useful in decision-making process. Research efforts are being made to investigate efficient approaches to extract rare knowledge patterns, by realizing the importance of knowledge pertaining to rare entities. The retail market can be more effectively studied by using the Association Rule mining on supermarket business data but it is also needful that Association mining be explored from a more realistic viewpoint.

Standard methods for mining association rules like Apriori and Frequent Pattern growth (FP-growth), are based on the support–confidence model, involving

two steps. First find all frequent itemsets (or frequent patterns) that satisfy minimum support (minsup) constraint. Second, generate all association rules that satisfy the minimum confidence (minconf) constraint. The frequency of an itemset may not be a sufficient indicator of interestingness because it does not reveal the utility of an itemset, which can be measured in terms of cost, profit, or other expressions of user preference. In many applications, some items appear very frequently in the data, while others rarely appear [35]. The traditional approaches suffer from dilemma called "rare item problem dilemma". If frequencies of items vary, two problems encountered are (1) At high minsup value, then rules or patterns of rare items will not be found, as rare items fail to satisfy the minsup value (2) To find rules that involve both frequent and rare items, minsup has to be set very low. This may cause combinatorial explosion, that is too many rules or frequent patterns may be generated.

Although standard ARM algorithms are capable of identifying distinct patterns from a dataset, they sometimes fail to associate user objectives and business values with the outcomes of the ARM analysis. For example, in a retail mining application, frequent itemsets identified by the standard ARM algorithms may contribute only a small portion of the overall company profit because high profit items are rare and do not appear in rules with high support count values [26].

The necessity to develop methods for discovering association patterns to increase business utility of an enterprise has long been recognized in data mining community [36]. This requires modeling specific association patterns that are both statistically (based on support and confidence) and semantically (based on objective utility) relating to a given objective that a user wants to achieve or is interested in [37].

Another problem that arises during the data mining process is treating data that contains temporal information. The business data collected in realistic situations have shown that it can be Simple or Complex, Sparse or Dense, Time-independent or time-sensitive. Temporal association rule mining is to discover the valuable relationship among the items in the temporal database [23]. This incorporation is especially necessary if we want to extract useful knowledge from dynamic domains, which are time varying in nature[37]. Discovering sequential relationships in a time sequence is often important to many application domains, including financial, manufacturing, etc. However, in a lot of cases is practically a computationally intractable problem and therefore it poses more challenges on efficient processing than non-temporal techniques. The temporal high utility itemsets are itemsets whose support is larger than a pre-specified threshold in current time window of the data stream.

The 'wicked environment' of the new world of business imposes the need for variety and complexity of interpretations of information outputs generated by computer systems. Such variety is necessary for deciphering the multiple worldviews of the uncertain and unpredictable future. Non-linear change imposes upon organizations the need for devising non-linear strategies. Such strategies cannot be 'predicted' based on a static picture of information residing in the company's databases. Business analysts still require more precise knowledge from data miners to facilitate their business understanding, providing new insights and ultimately leading to Business Intelligence. The modeling of imprecise and qualitative knowledge, as well as the transmission and handling of uncertainty at various stages are possible through the use of fuzzy sets. Fuzzy logic is capable of supporting, to a reasonable extent, human type reasoning in natural form [24]**.**

An effective Data Mining Approach shall be required for business analyst to consider these issues. A user-centric approach to data mining shall be thus more desirable rather than the current procedure-centric ones.

In this paper, a more user centric approach rather than traditional procedure centric approach of data miners is proposed. In this proposal a theoretical approach to Temporal Rare Itemset Utility Mining with fuzzy approach is presented, which allows item utility values to be dynamic over time for finding profitable rare itemsets.

In section 2, some related works are discussed. In section 3, the proposed theoretical concept is detailed. Section 4 shows conclusion & future work.

2 Literature review

2.1 Related research work

To address the "rare item problem" persisting in "single minsup framework", efforts have been made in Multiple Support Apriori (M.S.Apriori) (B.Liu and Ma, 1999), CFP growth (Ya-Han Hu, 2004) and improved Multiple Support Apriori(Kiran and Reddy, 2004) Utility mining is now an important association rule mining paradigm. In [1], a good foundational and theoretical model to utility itemset mining is introduced where a utility table UT<I,U> is defined by items I and their utilities U computed for each transaction and locally. This approach is improved in [5].

Chu, C.-J. et al proposed a Novel method, namely THUI (Temporal High Utility Itemsets)- Mine in [2], for mining temporal high utility itemsets from data streams efficiently and effectively. The Novel contribution of THUI-Mine is that it can effectively identify the temporal high utility itemsets by generating fewer candidate itemsets such that the execution time can be reduced substantially in mining all high utility itemsets in data streams. In this way, the process of discovering all temporal high utility itemsets under all time windows of data streams can be achieved effectively with less memory space and execution time.

Keshri Verma et al. have proposed algorithm H-mine, in [13], which takes advantage of H-struct data structure and dynamically adjust link in the mining process. This algorithm gives an efficient time sensitive approach for mining frequent item in the dataset. This approach reduces the size of dataset and increases the performance & efficiency of algorithm. Temporal FP-tree, uses divide & conquer technique for construction & traversing of tree, which is used to decompose the mining task into a set of smaller tasks for mining confined pattern in conditional database which reduce the search space on specific time interval when the data is sparse [13].

Ranjana Vyas et al, have proposed Temporal data mining, implemented and compared on performance issues such as temporal dimension in existing Associative Classifiers CBA, CMAR AND CPAR. In [27], conclusion is that Temporal Associative Classifier performs better in terms of classifier accuracy as compared to their non-temporal counterparts.

In [6], Ying Liu et al. present a Two-Phase algorithm to efficiently prune down the number of candidates and precisely obtain the complete set of high utility itemsets. It performs very efficiently in terms of speed and memory cost both on synthetic and real databases, even on large databases that are difficult for existing algorithms to handle. **G.C.Lan et al** proposed a new kind of patterns, named Rare Utility Itemsets in **[**42], which consider not only individual profits and quantities but also common existing periods and branches of items in a multi-database environment.

Many researchers have suggested that the items have the dynamic characteristics in terms of transaction, which have seasonal selling rate and it holds time-based association ship with another item. In [1], a foundational approach is given as a static ARM approach (without consideration of temporal or fuzzy features). In many applications, for example stock markets or data streams, use of discrete-valued utilities alone is inadequate. In cases where the values are uncertain, a fuzzy representation may be more appropriate. This motivates our exploration of the issue of efficiently mining high utility rare itemsets in temporal databases.

2.2 Companies using Data Mining

Most of the companies in the retail business are using Data Mining in one-way or the other. Applications of Data Mining in some companies are discussed in [30].

Wal-Mart, a pioneering leader in data mining and data management, captures point-of-sale transactions from over 2,900 stores in six countries and continuously transmits this data to its massive 7.5 terabyte data warehouse. Wal-Mart allows more than 3,500 suppliers to access data on their products and perform data analyses. These suppliers use this data to identify customer-buying patterns at the store display level. They use this information to manage local store inventory and identify new merchandising opportunities.

Burger King that has restaurants in 11,435 locations in 50 states and 58 countries with annual sales of $11.3 billion "is using Business Objects of Business Intelligence to help them achieve operational excellence throughout their organization, and maintain their leadership in a competitive industry", says Dave Kellogg, Senior Group Vice President Worldwide Marketing of Business Object, a leading provider of enterprise business intelligence. Business Objects enables organizations to track, understand, and manage enterprise performance. The company's solutions leverage the information that is stored in an array of corporate databases, enterprise resource planning (ERP), and customer relationship.

Other companies supplement their customers' transactional information with external data such as postal codes to do a market basket analysis. Practically every retailer now records all the details of each POS (Point of Sale) transaction for stock keeping purposes. Sometimes these are supplemented by customer information. Home Depot, for example

supplements the data with ZIP or postal code of the purchaser. Sometimes the cashier may also enter the sex and appropriate age of the customer into the cash register. Affinity cards and credit card numbers can be used to track repeat customers.

Another example of what data mining can do involves the directed targeting of customers for new products, at a fraction of the cost. A credit card company can leverage its vast warehouse of customer transaction data to identify customers most likely to be interested in a new credit product. Using a small test mailing, the attributes of customers with an affinity for the product can be identified. Recent projects have indicated more than a 20-fold decrease in costs for targeted mailing campaigns over conventional approaches.

3. Problem Definition

3.1 Business Intelligence

The advent of computing technology has significantly influenced our lives and two major impacts of this effect, are Business Data Processing and Scientific Computing. The knowledge based information system is a solution to manage the changing demand of the customers. As a result, Business Intelligence (BI) covers the process of transforming data from various data sources into meaningful information that can provide insights into business.

Data mining offers organizations an indispensable decision-enhancing environment to exploit new opportunities by transforming data into strategic weapon. Knowledge Management systems integrate data mining results along with Decision Support Systems in an organization while incorporating domain knowledge. Association Rules may be converted into business rules while incorporating domain knowledge and experience in general, similarly Retail market knowledge, Customer knowledge and related experience may be combined with Association rules for decision making in the Market Basket Analysis.

The traditional paradigm of information systems is based on seeking a consensual interpretation of information based on socially dictated norms or the mandate of the company bosses. This has resulted in the confusion between 'knowledge' and 'information'. However, knowledge and information are distinct entities!! While information generated by the computer systems is not a very rich carrier of human interpretation for potential action, 'knowledge' resides in the user's subjective context of action based on that information. Hence, it may not be incorrect to state that knowledge resides in the user and not in the collection of information.

Association rule mining needs to be treated as semi-automatic process of finding hidden useful patterns in Market Basket Analysis., which coupled with other supportive mechanism like DSS and using Knowledge Management approach helps the Business analysts in taking informed business decisions.

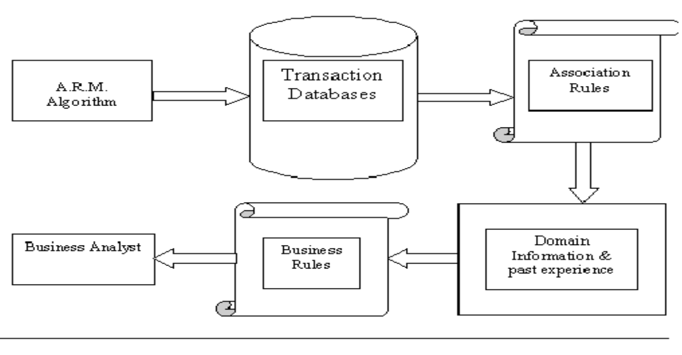

Figure 1: From Association rules to Business rules

We propose an approach as shown in figure 2 as more user-centric and has more usability by business analyst.

3.2 Association Rule Mining Approaches

Mining Association rules is one of the research problems in data mining. Given a set of transactions where each transaction is a set of items, an association rule is an expression of the form X $\Rightarrow$ Y, where X and Y are sets of items. Association rule mining (ARM) is a popular technique for finding co-occurrences of items in a set of transactions or a database. Rules with confidence and support above user-defined thresholds (minconf and minsup) were found. As data continues to grow and its complexity increases, newer data structures and algorithms are being developed to match this development. For example, frequent itemsets can be found out by analyzing market basket data and then association rules can be generated by predicting the purchase of other items by conditional probability [40].

Most utility items are linked with temporal constraints i.e. an itemsets utility is a temporal feature that may vary over a given time period. An important research issue extended from the mining of association rules is the discovery of temporal association patterns in data streams due to the wide applications in various domains. Temporal Data Mining can be defined as the activity of discovering

interesting correlations or patterns in large data sets of temporal data accumulated for other purposes. For time-variant data streams, there is a strong demand to develop an efficient and effective method to mine various temporal patterns. However, most methods designed for traditional databases cannot be directly applied to the mining of temporal patterns in data streams because of their high complexity.

Temporal Association Rule Mining- Temporal Association rule adds time constraint on association rule.
Temporal Association Rule Mining can be divided into 4 steps: [23]

(1) Data preprocessing. It includes data cleaning, data integration, data exchange and data reduction.
(2) To find frequent itemsets which have the support not less than min_sup.
(3) To generate association rules with frequent itemsets with time. The rules generated are temporal ones.
(4) To generate rule sets and output.

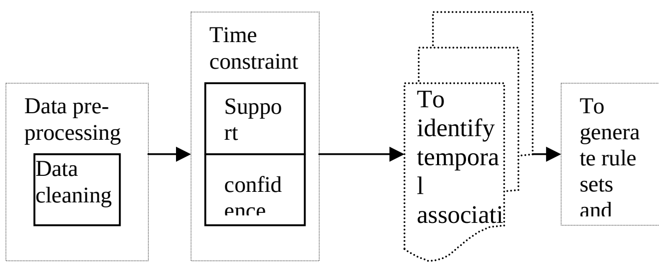

Figure 2: Temporal Association Rule Mining

Almost all Association rule mining algorithms treat each itemset to be of equal value (or importance) and only use support and confidence heuristics to measure rule quality. In real life applications, an itemset can be valued more highly, not only because of its frequency, but also of its importance or utility. The frequency of itemset is not sufficient to reflect the actual utility of an itemset. For example, the sales manager may not be interested in frequent itemsets that do not generate significant profit. A Utility Mining model was defined by Yao et al, (2004).

The term *utility* is commonly used to mean “the quality of being useful” and utilities are widely used in decision making processes to express user’s preferences over decision objects towards decision objectives. In decision theory, we have the well-known equation “Decision =probability + utility,” which says that a decision object is chosen based on its probability and utility. Since association mining can be viewed as a special decision problem where decision objects are patterns, we may have, correspondingly, an equation “Interestingness (of a pattern) = probability + utility.”

For Example, in business situations a manager may use Utility mining to discover the best business strategies by specifying his/her objective as “high profit and low risk of loss”[36]. Another example is in medical field. A doctor may use Utility mining to find the best treatments for a disease by specifying an objective “high effectiveness and low side-effects” [38].

Utility is considered as a measure of how “useful” (e.g. profitable”) an itemset is [2]. Utility u(V) of an itemset X is the sum of the utilities of X in all transactions containing X. Utility mining is the problem of finding all itemsets in a transactional database above a minimum utility threshold satisfying a given a minimum support and confidence. Utility mining aims at identifying the itemsets with high utilities [2]. The temporal high utility itemsets are the itemsets which appear infrequently in current time window of large databases but whose support is larger than a pre-specified threshold in current time window of the data stream. Utility Mining gives additional constraints to items to help generate rules with sufficient conditions such as revealing the utility of an itemset and thus help generate more informed association rule. For example, low support itemsets may be more profitable than high support itemsets and hence pruning techniques in normal association rule mining may miss such items.

A need to associate value to mined item sets is especially important in business analytics applications, such as retail analysis, targeted marketing, or client segmentation, since as pointed out recently [1,2], the utility of extracted “patterns” in decision-making can only be addressed within the micro-economic framework of the enterprise. In other words, client segments or product combinations identified through the frequent item mining, are interesting only to the extent in which they can be used in the decision-making process to increase some objective function, business metric, or utility (e.g., maximize revenue, minimize marketing or inventory cost, etc.) [39].

Transaction Utility in a transaction is the value of an item in a transaction. The transaction utility reflects the utility in a transaction database and is transaction dependent. However, frequency of an itemset alone does not assure its interestingness because it does not contain information on its subjectively defined utility such as profit in euros or some other variety of utility.**External Utility** of an item is a numerical value associated with an item defined by user. It is transaction independent [2]. User- defined utility (such as quantity of an item or profit) is based on information not available in the transaction dataset. It often reflects user preference and can be represented by an external utility table or utility function. **Utility of item** is the quantitative measure of utility for item in transaction, defined as product of Transaction Utility and External Utility.

The goal of Utility Mining is to identify high utility itemsets which derive a large portion of the total utility. For example, for stock exchange data, share prices of one particular product may remain constant in same period of time that one other product’s share price changes ten times. Utility mining frameworks ought to reflect such dynamicity within the time partitions being considered.

In most business applications, frequent itemsets may not generate much profit while rare itemsets may generate a very high profit. Rare itemsets are very important and can be further promoted together because they possess high associations and can bring some acceptable profits. Rare itemsets provide very useful information in the real-life applications such as security, business strategies, biology, medicine and super market shelf-management.

4. Conclusions and Future Work

The paper chronicles the related issues to Classical Association Rule Mining approaches to Market Basket Analysis and proposes a new theoretical approach for making business data mining more realistic and usable to business analyst. In the proposed approach, modifcation to data mining is done by incorporating time information in it and also an approach for converting Association rules into business rules have been proposed. This approach bridges the gap between technology developers and end users.

Marketers are interested in knowing how various marketing programs affect the discovery of subtle relationships. The outcome of systematic business data mining would also enable the business analyst in taking crucial business decisions such as finalizing discounting policy and organizing shelf space in supermarket scenario based on temporal high utility item set scenario.The proposed user-centric approach provides managers with systematic guidelines and suggestions regarding rare itemset utility mining where item utility values are allowed to be dynamic over time. The proposed approach can generate insights which can lead effective planning for retail marketing in Supermarket and online stream mining. The future work will include development and implementation of an efficient algorithm that treat data within Supermarket transactional database in more realistic manner.

5. References

[1] Yao, Hong, Hamilton, H., and Butz, C. J. 2004. *A Foundational Approach to Mining Itemset Utilities from Databases*, Proceedings of the Third SIAM International Conference on Data Mining, Orlando, Florida, pp. 482-486.

[2] Chu, C., Tseng, V. S., and Liang, T. 2008. *An efficient algorithm for mining temporal high utility itemsets from data streams*. *J. Syst. Softw.* 81, 7 (Jul. 2008), 1105-1117

[3] Hu, J., Mojsilovic, A. *High-utility Pattern Mining: A Method for Discovery of High-utility Item Sets, Pattern Recognition*, Vol. 40, 3317-3324.

[4] Ale, J. M. and Rossi, G. H. (2000). *An Approach to Discovering Temporal Association Rules*. In Proceedings of the 2000 ACM Symposium on Applied Computing, Vol.1, J. Carroll, E. Damiani, H. Haddad, and D. Oppenheim, Eds. SAC '00. ACM Press, New York, NY, pp 294-300.

[5] Yao, H. and Hamilton, H, J. 2006. *Mining Itemset Utilities from Transaction Databases*, Data and Knowledge Engineering, 59(3): 603-626

[6] Liu, Y., Liao, W., and Choudhary, A. 2005. *A Fast High Utility Itemsets Mining Algorithm*. Proceedings of the Utility-Based Data Mining Workshop.

[7] Teng, W. G., Chen, M. S., and Yu, P. S. 2003. *A Regression-Based Temporal Pattern Mining Scheme for Data Streams*. Proceedings of the 29th International Conference on Very Large Databases, pp 93-104.

[8] Ahmed, C. F., Tanbeer, S. K., Jeong, B-S, and Lee, Y-K. 2008. *Handling Dynamic Weights in Weighted Frequent Pattern Mining*, IEICE Trans. Information and Systems, Vol. E91-D:2578-2588.

[9] Han, J., Pei, J. and Yiwen, Y. 2000. *Mining Frequent Patterns Without Candidate Generation*. Proceedings ACM-SIGMOD International Conference on Management of Data, ACM Press, pp1-12.

[10] Coenen, F., Leng, P. and Ahmed, S. 2004. *Data Structures for association Rule Mining: T-trees and P-trees*. IEEE Transactions on Data and Knowledge Engineering, Vol 16, No 6, pp774-778.

[11] Yun, U. 2007. *Mining lossless closed frequent patterns with weight constraints*. Know.-Based Syst. 20, 1 (Feb. 2007), 86-97.

[12] Ning, H. and Yuan, S-C. 2006. *Temporal Association Rules in Mining Method*, First International Multi-Symposiums on Computer and Computational Sciences - Volume 2 (IMSCCS'06) pp. 739-742.

[13] Verma, K., Vyas, O. P. and Vyas,R. 2005. *Temporal Approach to Association Rule Mining Using T-Tree and P-Tree, Machine Learning and Data Mining in Pattern Recognition*, 651-659, LNS Volume 3587.

[14] Cheng-Yue Chang, Ming-Syan Chen, Chang-Hung Lee, *Mining General Temporal Association Rules for Items with Different Exhibition Periods* , Second IEEE International Conference on Data Mining (ICDM'02),

[15] Yo-Ping Huang; Li-Jen Kao; Sandnes, F.-E. , 2005. *A prefix tree-based model for mining association rules from quantitative temporal data, Systems, Man and Cybernetics*, 2005 IEEE International Conference on Volume 1, Issue , 10-12 Oct. 2005 Page(s): 158 - 163 Vol. 1

[16] Edi Winarko and John F. Roddick,. 2005. *Discovering Richer Temporal Association Rules from Interval-Based Data, Data Warehousing and Knowledge Discovery*, LNCS 3589, 315-325.

[17] Kriegel, H-P et al. 2007. *Future Trends in Data Mining*, *Data Mining and Knowledge Discovery*, 15:87–97

[18] Gaber, M. M,, Zaslavsky, A. and Krishnaswamy, S. 2005. *Mining data streams: a review*. SIGMODRecords

[19] Lu, S., Hu, H. and Li, F. 2005. *Mining weighted association rules*. Intelligent Data Analysis, 5(3):211–225.

[20] Agrawal ,R. Srikant, R. 1994. *Fast Algorithms for Mining Association Rules*, In: Proceedings of 20th International Conference on Very Large Databases, Santiago, Chile, pp. 487-499.

[21] Yingjiu Li, Peng Ning, X. Sean Wang , Sushil Jajodia R. 2003. *Discovering calendar- based temporal association rules*, Data & Knowledge Engineering volume 4, Elesvier publisher, Volume 44, pp 193-214.

[22] Agrawal R., Imielinski T., and Swami A., "*Mining association rules between sets of items in large databases*", Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, pages 207–216, 1993.

[23] Zhai Liang, Tang Xinming, Li Lin, Jiang Wenliang, *Temporal Association Rule Mining based on T-Apriori Algorithm and its typical application*, Proceedings of International Symposium on Spatio-temporal Modeling, Spatial Reasoning, Analysis, Data Mining and Data Fusion, 2005.

[24] Sushmita Mitra, Sankar K. Pal, Pabitra Mitra, *Data Mining in Soft Computing Framework: A Survey*, IEEE Transactions On Neural Networks, VOL. 13, NO. 1, January 2002.

[25] Sergey Brin, Rajeev Motwani, Jeffrey D. Ullman, Shalom Tsur, *Dynamic* Itemset *Counting and Implication Rules for Market Basket Data* in Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data, volume 26(2) of SIGMOD Record, pp. 255–264. ACM Press

[26] M. Sulaiman Khan, Maybin Muyeba, Frans Coenen, *A Weighted Utility Framework for Mining Association Rules*, Proceedings of the 2008 Second UKSIM European, page(s): 87-92.

[27] O.P.Vyas, Ranjana Vyas, Vivek Ranga, O.P.Vyas, Anne Gutschmidt, *Exploring Temporal Associative Classifiers for Business Analytics*,accepted in Ist International Conference on Advanced Data Analysis, Business Analytics and Intelligence, IIM, Ahmedabad held on 6th-7th June'09.

[28] Kiran, R.U. and Reddy, P.K.(2009), *An improved multiple minimum support based approach to mine rare association rules*, IEEE symposium on Computational Intelligence and Data Mining.

[29] B. Liu, w.H. and Ma, Y(1999), *Mining association rules with multiple minimum supports*, SIGKDD Explorations.

[30] Syed Riaz Ahmed, "*Applications of Data Mining in Retail Business***",** Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04) IEEE 2004.

[31] http://www.aspfree.com/c/a/MS-SQL-Server/Using-Data-Mining-for-Business-Intelligence/

[32] R. Uday kiran, P Krishna Reddy; *An Improved Multiple Minimum Support Based Approach to Mine Rare Association Rules*; Conference: 2009 IEEE Symposium on Computational Intelligence and Data Mining (IEEE CIDM 2009)

[33] Leon, Florin; Atanasiu, Gabriela M; *Integrating artificial intelligence into organizational intelligence Management & Marketing* - Bucharest (2/2008), Issue: 2/2008, Pages: 79-90

[34] http://www.gr1innovations.com/data_mining.html

[35] Bing Liu , Wynne Hsu , Yiming Ma; *Mining Association Rules with Multiple Minimum Supports* ; Publications in Knowledge Discovery and Data Mining(1999).

[36] Yi-Dong Shen, Zhong Zhang, Qiang Yang; Objective-*Oriented Utility-Based Association Mining;* Second IEEE International Conference on Data Mining (ICDM'02),Maebashi City, Japan ,December 09-December 12, ISBN: 0-7695-1754-4

[37] GUIL Francisco, BAILON Antonio, BOSCH Alfonso, MARIN Roque, *An iterative method for mining frequent temporal patterns;* 10th international conference on computer aided systems theory, Las Palmas de Gran Canaria, Spain, February 7-11, 2005, vol. 3643, pp. 189-198, ISBN 3-540-29002-8.

[38] http://news.yahoo.com/s/ap/20091102/ap_on_he_me/us_med_unproven_remedies_med_schools

[39] Jianying Hu· nd Aleksandra Mojsilovic; *High-utility pattern mining: A method for discovery of high-utility item sets,* Pattern Recognition, Volume 40, Issue 11, November 2007, Pages 3317-3324

[40] Jyothi Pillai, O.P.Vyas, *Overview of Itemset Utility Mining and its Applications*, International Journal of Computer Applications (0975 – 8887) Volume 5– No.11, August 2010 Pages 9-13.

[41] J. Hu, A. Mojsilovic, "*High-utility pattern mining: A method for discovery of high-utility item sets*", Pattern Recognition 40 (2007) 3317 – 3324.

[42] G.C.Lan, T.P.Hong and V.S. Tseng, " A Novel Algorithm for Mining Rare-Utility Itemsets in a Multi-Database Environment"