iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VI **Month of publication:** June 2023

DOI: https://doi.org/10.22214/ijraset.2023.54067

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

A Comprehensive Review on Association Rule Mining and its Algorithms: Apriori, FP Growth, and SS-FIM

Priyankaben Vashi
Dept. of CE/IT, P. P. Savani University, Kosamba-394125, Gujarat, India

Abstract: ***The technique of extracting usable information from databases is known as data mining. It is employed to manage massive amounts of data. In order to forecast future trends and behaviours, data mining is performed. Today, data mining is a technique that is growing helpful since it has many uses.***
Massive amounts of data are produced daily in many different fields. Data mining is utilised since it is challenging to analyse such a big volume of data.
It is often referred to as database knowledge discovery (KDD). Data mining technologies are used to forecast the behaviour and pattern of the data, assisting organisations in decision-making [1]. Association Rule Mining is one of the data mining techniques discussed in the present article. It has three algorithms. 1. Apriori, 2. FP-Growth, and 3. SS-FIM. In the current work, all three algorithms are discussed. After examining each algorithm, it was determined that the SS-FIM algorithm had the best combination of features.
Keywords: Association Rule Mining; Apriori; Data Mining; FP Growth; SS-FIM

I. INTRODUCTION

The technique of extracting usable information from databases is known as data mining. It is employed to manage massive amounts of data. In order to forecast future trends and behaviours, data mining is performed. Today, data mining is a technique that is growing helpful since it has many uses. Massive amounts of data are produced daily in many different fields. Data mining is utilised since it is challenging to analyse such a big volume of data. It is often referred to as database knowledge discovery (KDD). Data mining technologies are used to forecast the behaviour and pattern of the data, assisting organisations in decision-making [1]. Due to its success in several fields, data mining is becoming more and more popular. It has uses in many different fields, including healthcare, finance, telecommunication, business, and education [2]. There are two types of data mining: descriptive mining and prescriptive mining.

It was found and summarised the already-existing data in descriptive mining. Predictions are made from historical data in prescriptive mining [3]. There are several approaches used in data mining, including classification, clustering, and association rule mining. One of the most essential data mining methods is association rule mining. It is employed to identify the database's recurrent patterns. Finding intriguing associations and correlations between the various database elements is the primary goal of association rule mining [4]. In several fields, association rules are employed to identify patterns in the data. It can be determined how many combinations of occurrences happen at once with the use of patterns. A significant amount of data is subjected to association rules. Association rules are employed in the corporate world to find recurring patterns that aid in decision-making and marketing [5]. Additionally, association norms are relevant in a wide range of different professions, such as those in medicine, market basket analysis, libraries, etc. [6]. For instance, if a customer purchases shampoo, he might also purchase conditioner. These kinds of linkages can be used as information to aid in decision-making for marketing-related activities [7]. Association rules are typically represented as X=>Y. The database item sets, in this case, are X and Y. X and Y are referred to as antecedent and consequent, respectively. Two fundamental metrics make up association rule mining: support (s) and confidence (c).

Support is the likelihood that X and Y will occur simultaneously in a transaction. Confidence is likewise a probability, but confidence has a condition attached to it. Confidence, if a transaction includes X, it also includes Y [8-10]. In association rule mining, the following two steps are used to generate rules:

1) All frequent item sets are located with the least amount of assistance.
2) Strong association rules with confidence c exceeding a predetermined threshold value are created using these frequent item sets.

It takes greater concentration to complete the first stage because it is challenging to identify all the commonly occurring item sets [11].

©IJRASET: All Rights are Reserved

A. Association Rule Mining in various areas

Due to its many benefits, association rule mining is applied in many different fields. The next section discusses a few of the fields that have used association rule mining.

Market basket analysis, such as that found in supermarkets, is one of the greatest and most prevalent examples of association rule mining. The owners of these businesses hope to both pique the interest of current patrons and draw in new ones. These stores contain numerous databases and a substantial volume of transactional records. The managers might be curious to discover if certain things are regularly bought. Analysing clients' purchasing patterns is the key goal. To generate the rules to determine the frequently occurring item sets in a store, association rule mining is used. For instance, if customers are purchasing bread, the majority of them will also be purchasing milk. Therefore, managers will benefit from placing milk close to the bread. As a result, the association rule is useful for creating a store layout. Therefore, a market basket is a collection of several things bought by a customer in a single transaction [12-16].

A bank uses a system called CRM, or customer relationship management, to manage its relationships with customers. Typically, banks track consumer behaviour to learn about their preferences and interests. Banks can improve the coherence between credit card clients and the bank in this way. Here, association rules allow banks to get to know their clients better so they can offer them high-quality services. [17, 18].

The study of medicine makes use of association rules. It is beneficial for doctors to heal their patients. To determine the likelihood of a condition, association rules are used. A condition needs to be well explained. So, diagnosing is not a simple task in and of itself. The doctors will benefit from adding a few additional symptoms to an existing ailment and then determining how the symptoms relate to one another. For instance, when a doctor is examining a patient, it is evident that he will need all of the patient's information. Only then can he make the best choice for the patient. As one of the most significant data mining research techniques, the association rule is helpful in this situation [19-25].

In census data, association rule mining has a lot of potential. Association law supports the advancement of business and good public policy. The census produces a lot of statistical data. Planning for public services and businesses can benefit from the information because it relates to the economic and population census. It comprises health, education, transportation, and other services. In the corporate world, it has built new banks, factories, and shopping centres [26-28].

Any cell in an organism is made up primarily of proteins. There are a variety of DNA technologies and instruments available for quickly determining DNA sequences. Essentially, amino acid sequences made up of over 20 different sorts make up proteins. Each protein has a three-dimensional structure that is determined by the amino acid sequence. As a result, proteins' ability to function is overly dependent on amino acids. Different amino acids in a protein are given association rules. Understanding the protein composition will be helpful [29, 30]. There are numerous techniques and algorithms available for producing association rules in databases. Those algorithms are discussed in the subsequent sections.

II. APRIORI ALGORITHM

It uses a mining technique for Boolean association rules. The entire database is examined after each set of frequent item sets is formed, and the association rules between the data are mined from the frequent item sets to provide us with decision assistance. A frequent item-set is an item-set that has more support than the specified minimum support count. An item-set is a group of 0 or more items. Typical evaluation standards for frequently used item sets [31-34]:

One of the two fundamental criteria of association rules is support. It is the proportion of transactions in the All sample of dataset D that contain both x and y to all transactions. If having two data x and y that need to be analysed for correlation., The corresponding support degree is:

$$Support(x, y) = P(xy) = \frac{num(xy)}{num(All\ Samples)} \quad (1)$$

$$Support(x \rightarrow y) = P(x \cup y) = \frac{count(X \cup Y)}{|D|} \quad (2)$$

For instance, "there is a 28% probability that an individual in the population will contain both X and Y," according to a support rating of 28%.

Confidence: Conditional probability is the ratio of the number of transactions including x and y to the number of transactions including y.

$$Confidence(x \rightarrow y) = P(x|y) = \frac{P(xy)}{P(y)} \quad (3)$$

$$Confidence(x \rightarrow y) = P(X|Y) = \frac{Support(X \cup Y)}{Support(Y)} \quad (4)$$

The confidence is 52% if assuming that "52% of the terms containing X contain Y".

The frequent item-sets in the database can only be identified with custom minimum support or a combination of custom support and confidence. By setting the minimum support count and iterating continually, its main goal is to recover all frequent item-sets and identify all item-sets that are larger than or equal to the support [35].

The two characteristics of the Apriori algorithm can efficiently decrease the algorithm's search complexity, decrease the amount of processing, and increase the effectiveness of mining common item-sets. First, all subsets of X are frequent item-sets if X is a frequent item-set. Second, all supersets of X are non-frequent item-sets if X is a non-frequent item-set.

III. FP-GROWTH ALGORITHM

Association rules (X→Y) are the most frequently used method in typical transactions. The goal is to look for any relationships between database entries. The relationship shows that other items (Y) are likely to arise when items (X) do [36]. For instance, a consumer may receive milk in addition to their bread order. As a result, association rules can help decision-makers identify potential products that consumers are likely to buy. It also makes it easier to plan marketing tactics [37]. The traditional association rule algorithm scans databases quite slowly, especially when using the Apriori algorithm, which frequently reduces the effectiveness of data mining. Han et al.'s [38] Frequent-Pattern growth (FP-growth) mining method, which is thought to be more effective and avoids the requirement to develop candidate itemsets, was offered as a solution to the aforementioned problem. Reading the data set one transaction at a time and assigning each one to a path in an FPtree (frequent pattern tree) is how FP-Growth is created. Different transactions may cross paths since they may share many goods at once. It can be used the FP-tree structure to compress data more efficiently when paths overlap more. It can be extracted frequent itemsets straight from the structure in memory, as opposed to repeatedly traversing the data saved on a disc, if the size of the FP-tree is small enough to fit into the main memory [39]. As a result, it only needs to scan the database twice instead of creating a candidate itemset.

Additionally, one must take into account the user's impression and the cognitive uncertainty of subjective judgements when making decisions. In 1965, Zadeh put up the fuzzy set theory [40] to address the problem of cognitive uncertainty caused by vagueness and ambiguity. They are useful in conducting studies of decision-making because linguistic variables and linguistic values [41-43] can be defined with hazy notions to subjectively match the potential cognition of a decision-maker. As a result, fuzzy data mining has recently gained prominence as a research topic.

Each transaction database item is treated as a linguistic variable through fuzzy frequent pattern growth, and each linguistic variable is partitioned according to its linguistic value. By doing this, fuzzy association rules can be properly explained using natural language. The suggested method is divided into two steps.

The first is to create a fuzzy FP-tree by scanning the database twice, and the second is to locate huge 1-itemsets by scanning the database once. To create the fuzzy association rules, one conditional pattern base and one conditional fuzzy FP-tree will then be derived from each node in the fuzzy FP-tree [44].

The use of a small data structure and the elimination of repetitive data are the main benefits of the FP-Growth algorithm. Faster than other association mining methods and faster than tree research is FP-growth database analysis. According to [45], the health of the economy depends on our ability to determine "What classes or groupings of products are likely to give great services to customers." Market basket circles include all significant areas of market analysis that may be conducted on consumer transaction retail data. It is expensive to manage a large number of candidate sets, and it is tiresome to search the database repeatedly and test a large number of candidates by matching patterns, which is especially true for mining long patterns, according to [46]. This is true for situations where there are a lot of frequent patterns, long patterns, or low minimum support thresholds. This is the underlying cost of candidate generation, regardless of the implementation method utilised.

IV. SS-FIM

Frequent Itemset Mining (FIM) tries to uncover strongly linked hidden patterns in a database that is represented using a transactional form. FIM is widely employed in numerous applications in the real world where it is used as a preprocessing step to extract common patterns. Let's take the information retrieval problem as an example, where a collection of documents is converted into a transactional database, where each document is treated as a transaction and each term as an item. Studying the various connections between the terms of the papers is made possible by FIM. For instance, inferring a high reliance between the terms "computational" and "intelligence" from the pattern "computational, intelligence" is typical. Over the past 20 years, the FIM problem has been thoroughly investigated in the literature. Google has utilised apriori and fp-growth [47] extensively.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

However, the newest and most advanced technologies are producing enormous amounts of data (big data), such as social networks, sensors, and the Internet of Things, which are then stored in the cloud and processed using HPC technology. Because of this, the algorithm is unable to handle the steadily growing volume of databases. For instance, running the Apriori algorithm on the Webdocs database, which has more than 1,500,000 transactions and more than 500,000,000 objects, takes several months and days. Numerous metaheuristic-based solutions have been developed to address this issue utilising methods like genetic algorithms [48] and swarm intelligence [49].

These methods produce results in a decent amount of time, however, when working with highly huge and complex solution spaces, they may only extract a very small number of frequent itemsets. Figure 1 compares, for certain FIM benchmark databases, the runtime and the number of frequent itemsets found using various metaheuristic-based techniques, including GA [50], PSO [51], BSO [21] and Apriori (the correct algorithm). Initial minimum support is set at 10%. It is evident that although the metaheuristic-based techniques extract fewer frequent itemsets than Apriori, they are faster. This can be explained by the fact that Apriori analyses the intrinsic qualities of the FIM issue to enumerate all frequent itemsets, whereas metaheuristic-based FIM techniques randomly search the itemset space. According to the explanation above, it is preferable to employ metaheuristic-based FIM algorithms to find the necessary information. However, one significant drawback is that they might make only a few common items sets discoverable.

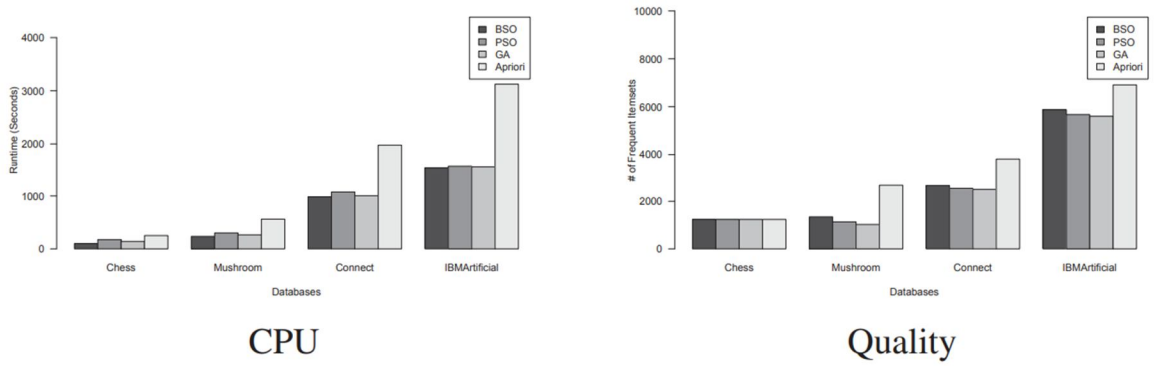


Figure 1 Comparison of the Apriori algorithm and metaheuristic-based FIM's runtime and number of patterns

Numerous exact-based techniques have been put out in the literature [33]. These methods incur a high cost in terms of memory and CPU runtime since they repeatedly scan the database to build all frequent item sets. To address this issue, the Single Scan for Frequent Itemsets Mining (SSFIM) method, in which transactions are processed one at a time to produce candidate frequent itemsets, has recently been proposed [34]. The support of generated candidate itemsets is updated as transactions are processed, resulting in the generation of all frequent itemsets at the conclusion of a single database scan. The main drawback of this strategy is the creation of itemsets for every transaction, where all potential combinations are looked into. The maximum number of items allowed in a transaction affects the temporal complexity of this significantly. As a result, for dense transactional databases, SSFIM performance is significantly decreased. Metaheuristic-based FIM has been created to address the shortcomings of the current accurate FIM algorithms. These methods change the search into an itemsets space, where every point is seen as a potential remedy for the issue.

V. CONCLUSION

As discussed in the present article, the algorithms of the Association rule mining are discussed i.e., Aprori, FP Growth, and SS-FIM. For Apriori, The algorithm's spatial complexity is low, which eases the programmer's coding burden. Numerous repeated operations can be avoided thanks to the algorithm's design and pruning operation, which also increases the algorithm's processing speed. The quantity of database scans will be drastically decreased if a set minimum support criterion is high. It is appropriate for sparse data and has a high degree of ductility and exploitability. The candidate set's size will grow exponentially as more data is added to the transactional database, using more RAM and adding to the workload. However, pruning the candidate set necessitates scanning the database, which adds time and decreases algorithm performance. As a result, the algorithm's effectiveness decreases exponentially as the size of the database grows. Numerous methods are developed based on Apriori to address these issues, such as the FP-growth algorithm, SS-FIM, hash-based technology, thing reduction technology, partition technology, etc.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Han et al. offered the new idea of the FP-growth algorithm, which can be one of the representations of the itemsets that do not require candidate generations. Phases of the Apriori process that create candidate itemsets can move on without regard to association length. However, because the database may need to be scanned repeatedly to produce a large number of candidate itemsets, mining with the Apriori algorithm does not archive the goal efficiently. As a result, FP-growth does the initial scan of the transaction database, filters out the popular itemsets, and then steadily expands support. Create an FP-tree structure using the transaction database in the second scan after that.

Then, assign each item node in the FPtree using a header table such that each branch of the tree is connected to the others.

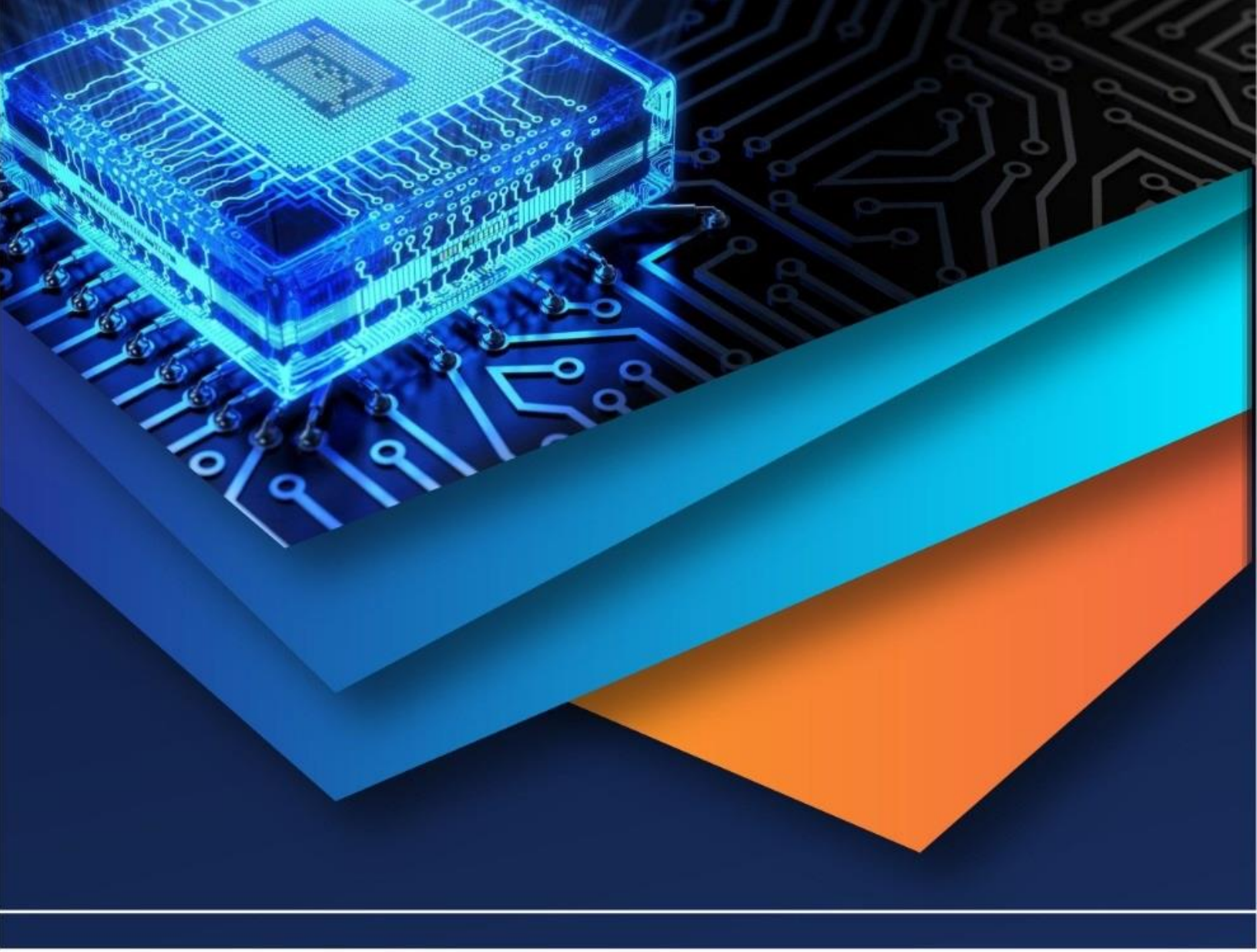
The objective of SS-FIM is to find frequent itemsets while reducing the number of database scans and generated candidates. This is done to get around the Apriori heuristic's drawbacks. Creating every conceivable itemset for each transaction is the core goal of SS-FIM. A generated itemset t's support is increased by one if it was already created when handling a previous transaction. If not, its support is created and set to one. Up until all of the database's transactions have been handled, the process is repeated. With SS-FIM, all frequent item sets can be located by doing a single scan of the transactional database. Because the frequent itemsets are taken straight from the transactional database and a certain itemset is frequent, SS-FIM is also comprehensive. As a result, nothing is lost during the itemset generating procedure. Moreover, each transaction first generates prospective itemsets, and as transactions are processed, a hash table is utilised to keep track of the partial frequency of occurrence of candidate itemsets. Owing to the above discussion, SS-FIM eliminated many of the limitations offered by Apriori and FP-Growth.

REFERENCES

[1] N. Padhy, P. Mishra and R. Panigrahi, “The Survey of Data mining Applications and Feature Scope”, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), vol. 2, no. 3, (2012**)** June.
[2] N. Paidi, “Data Mining: Future Trends and its Applications”, International Journal of Modern Engineering Research (IJMER), vol. 2, no. 6, November-December (**2012**), pp. 4657-4663.
[3] T. Karthikeyan and N. Ravikumar, “A Survey on Association Rule Mining”, International Journal of Advanced Research in Computer and Communication Engineering, vol. 3, no. 1, (**2014)** January.
[4] Rajak and M. K. Gupta, “Association Rule Mining: Applications in Various Areas”, International Conference on Data Management.
[5] Chandrakar and A. M. Kirthima, “A Survey on Association Rule Mining Algorithms”, International Journal of Mathematics and Computer Research”, vol. 1, no. 10, (**2013),** November.
[6] S. Kotsiantis and D. Kanellopoulos, “Association Rules Mining: A Recent Overview”, GESTS International Transactions on Computer Science and Engineering, vol. 32, no. 1, (**2006**), pp. 71-82.
[7] C. Kaur, “Association Rule Mining Using Apriori Algorithm: A Survey”, International Journal ofAdvanced Research in Computer Engineering & Technology (IJARCET), vol. 2, no. 6, (**2013)** June.
[8] R. Trikha and J. Singh, “Improving the Efficiency of Apriori Algorithm by adding new Parameters” International Journal for Multi Disciplinary Engineering and Business Management, vol. 2, no. 2, (**2014)** June.
[9] S. Maitrey and C. K. Jha, “Association Rule Mining: A Technique for Revolution in Requirement Analysis”, International Journal of Scientific and Research Publications, vol. 4, no. 8, (**2014)** August.
[10] J. K. Jain, N. Tiwari and M. Ramaiya, “A Survey: On Association Rule Mining”, International Journal ofEngineering Research and Applications (IJERA), vol. 3, no. 1, (**2013**) January-February, pp. 2065-2069
[11] S. Soni and O. P. Vyas, “Using Associative Classifiers for Predictive Analysis in Health Care Data Mining” International Journal of Computer Applications, vol. 4, no. 5, (**2010)** July.
[12] T. A. Kumbhare and S. V. Chobe, “An Overview of Association Rule Mining Algorithms”, International Journal of Computer Science and Information Technologies, vol. 5, no. 1, (**2014**), pp. 927-930.
[13] Agarana, M. and A. Ehigbochie, 2015. Optimization of student’s academic performance in a world-class university using operational research model. Intl. J. Math. Comput. Appl. Res., 5: 43-50.
[14] Alos, S.B., L.C. Caranto and J.J.T. David, 2015. Factors affecting the academic performance of the student nurses of BSU. Intl. J. Nurs. Sci., 5: 60-65.
[15] Basil, A.O. and M.A. Yinusa, 2008. Socio-economic factors influencing students academic performance in Nigeria: Some explanation from a local survey. Pak. J. Social Sci., 5: 319-323.
[16] Bayat, A., W. Louw and R. Rena, 2014. The impact of socio-economic factors on the performance of selected high school learners in the Western Cape Province, South Africa. J. Hum. Ecol., 45: 183-196.
[17] Considine, G. and G. Zappala, 2015. The influence of socio and economic disadvantage in the academic performance of school students in Australia. J. Sociology, 38: 129-148.
[18] Erdem, C., I. Senturk and C.K. Arslan, 200°. Factors affecting grade point average of university students. Empirical Econ. Lett., 6: 360-368.
[19] Erdem, H.E., 2013. A cross-sectional survey in progress on factors affecting students academic performance at a Turkish university. Procedia Social Behav. Sci., 0: 691-695.
[20] Kuok, C.M., A. Fu and M.H. Wong, 1998. Mining fuzzy association rules in databases. ACM. Sigmod Rec., 2°: 41 -46.
[21] Michikyan, M., K. Subrahmanyam and J. Dennis, 2015. Facebook use and academic performance among college students: A mixed-methods study with a multi-ethnic sample. Comput. Hum. Behav., 45: 265-2°2.
[22] Okioga, C.K., 2013. The impact of students socio-economic background on academic performance in Universities, a case of students in Kisii University College. Am. Intl. J. Social Sci., 2: 38-46.

[23] Oladipupo, O.O., O.J. Oyelade and D.O. Aborisade, 2012. Application of fuzzy association rule mining for analysing students academic performance. Intl.
[24] J. Comput. Sci. Issues, Vol. 9,
[25] Olayiwola, M.S.O., F.J.O. Oyenuga, A. Oyekunle,
[26] T.O. Ayansola and A. Agboluaj e, 2011. Statistical analysis of impact of socio-economic factors on students academic performance. Intl. J. Res. Rev. Appl. Sci., 8: 395-395.
[27] Myra, S., "Web usage mining for web site evaluation," Communications of the ACM, Vol. 43, pp. 21-30, 1994.
[28] Pedrycz, W., "Why triangular membership functions?," Fuzzy Sets and Systems, Vol. 64, pp. 21-30, 1994.
[29] Tan. P. N., Michael Mteinbach, Vipin Kumar, Introduction to Data Mining, NY: Addison Wesley, 2005
[30] Wang, L. X. and Mendel, J. M. (1992), "Generating Fuzzy Rules by Learning from Examples", IEEE Transactions on Systems, Man, and Cybernetisc, Vol. 22, no.6, pp.1414-1427, 1992.
[31] Zadeh, L. A. (1965), "Fuzzy Sets," Information Control, vol. 8, no.3, pp.338-353, 1965.
[32] Zadeh, L. A. (1975), "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning," Information Science (part 1), Vol. 8, no. 3, pp.199-249, 1975.
[33] Zadeh, L. A. (1975), "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning," Information Science (part 2), Vol. 8, no. 4, pp.301-357, 1975.
[34] Zadeh, L. A. (1976), "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning," Information Science (part 3), Vol. 9, no. 1, pp.43-80, 1976.
[35] Zimmermann, H. -J., Fuzzy sets, Decision making and expert systems, Kluwer, Boston, 1991.
[36] Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: ACM SIGMOD Record, vol. 22, no. 2, pp. 207–216. ACM, June 1993
[37] Han, J., Pei, J., Yin, Y.: Mining frequent patterns without candidate generation. In: ACM SIGMOD Record, vol. 29, no. 2, pp. 1–12. ACM, May 2000
[38] Brin, S., Motwani, R., Ullman, J.D., Tsur, S.: Dynamic itemset counting and impli- cation rules for market basket data. In: ACM SIGMOD Record, vol. 26, no. 2, pp. 255–264. ACM, June 1997
[39] Mueller, A.: Fast sequential and parallel algorithms for association rule mining: a comparison. Technical report CS-TR-3515, University of Maryland, College Park, August 1995
[40] Zaki, M.J., Parthasarathy, S., Ogihara, M., Li, W.: New algorithms for fast discov- ery of association rules. In: Third International Conference Knowledge Discovery and Data Mining (1997)
[41] Amphawan, K., Lenca, P., Surarerks, A.: Efficient mining top-k regular-frequent itemset using compressed tidsets. In: Cao, L., Huang, J.Z., Bailey, J., Koh, Y.S., Luo, J. (eds.) PAKDD 2011. LNCS (LNAI), vol. 7104, pp. 124–135. Springer, Heidelberg (2012). doi:10.1007/978-3-642-28320-8 11
[42] Leung, C.K.-S., Mateo, M.A.F., Brajczuk, D.A.: A tree-based approach for fre- quent pattern mining from uncertain data. In: Washio, T., Suzuki, E., Ting, K.M., Inokuchi, A. (eds.) PAKDD 2008. LNCS (LNAI), vol. 5012, pp. 653–661. Springer, Heidelberg (2008). doi:10.1007/978-3-540-68125-0 61
[43] Cerf, L., Besson, J., Robardet, C., Boulicaut, J.F.: Closed patterns meet n-ary relations. ACM Trans. Knowl. Discov. Data (TKDD) **3**(1), 3 (2009)
[44] Grahne, G., Zhu, J.: Fast algorithms for frequent itemset mining using FP-trees. IEEE Trans. Knowl. Data Eng. **17**(10), 1347–1362 (2005)
[45] Borgelt, C.: Frequent itemset mining. Wiley Interdisc. Rev.: Data Min. Knowl. Discov. **2**(6), 437–456 (2012)
[46] Djenouri, Y., Bendjoudi, A., Mehdi, M., Nouali-Taboudjemat, N., Habbas, Z.: GPU-based bees swarm optimization for association rules mining. J. Supercom- put. **71**(4), 1318–1344 (2015)
[47] Djenouri, Y., Drias, H., Habbas, Z.: Bees swarm optimisation using multiple strate- gies for association rule mining. Int. J. Bio-Inspired Comput. **6**(4), 239–249 (2014)
[48] Gheraibia, Y., Moussaoui, A., Djenouri, Y., Kabir, S., Yin, P.Y.: Penguins search optimisation algorithm for association rules mining. CIT J. Comput. Inf. Technol. **24**(2), 165–179 (2016)
[49] Luna, J.M., Pechenizkiy, M., Ventura, S.: Mining exceptional relationships with grammar-guided genetic programming. Knowl. Inf. Syst. **47**(3), 571–594 (2016)
[50] Hegland, M.: The apriori algorithm tutorial. Math. Comput. imaging Sci. Inf. Process. **11**, 209–262 (2005)
[51] Guvenir, H.A., Uysal, I.: Bilkent university function approximation repository (2000). http://funapp.CS.bilkent.edu.tr/DataSets. Accessed 12 Mar 2012

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)