Review Article

A review and analysis on knowledge discovery and data mining techniques

Bhagawan Singh[1] , Vivek Dubey[2] and Jitendra Sheetlani[2]
PhD Scholar, Department of Computer Science, School of Engineering, Sri Satya Sai University of Technology & Medical Sciences, Sehore, Madhya Pradesh, India[1]
Professor, Department of Computer Science, School of Engineering, Sri Satya Sai University of Technology & Medical Sciences, Sehore, Madhya Pradesh, India[2]

Abstract
Data mining is used for the knowledge discovery in the area of engineering, medical diagnosis, business analytics, etc. The main aim of this paper is to explore the technological findings in the several fields suggested above and analysis the methods on the basis of the capability of knowledge discovery. In this regard several methodologies have been discussed which are previously published for the analysis. This analysis provides us a proper insight regarding the gaps, advantages and future implications and directions.

Keywords
Data mining, Apriori, FP-Growth, SPADE, ECLAT.

1.Introduction

Data mining algorithms are used in nearly every field. It is helpful in meaningful extraction in different areas of engineering, medical and business application. There are several algorithms and there extended version are already proposed for the betterment. Some of the commonly used algorithms are Apriori, frequent-pattern (FP)-growth, Equivalence class clustering and bottom up lattice traversal (ECLAT), tree-projection, hyper-structure mining (H-Mine), direct hashing and pruning (DHP), sequential pattern discovery using equivalence classes (SPADE), etc [1-7].

In the present situation in everyday life the database is becoming quicker. So that to limiting data likelihood for prune is the best alternative in information mining [8]. The previous approach gives a bearing along these lines, so we can diminish the exchanges and furthermore need with dynamic least help. The data mining (DM) and knowledge in databases (KDD) developments have been founded on the way that the genuine esteem isn't in putting away the information, but instead in our capacity to extricate helpful reports and to discover intriguing patterns and connections [9-12].

The arrangement of DM forms used to separate and confirm designs in information is the center of the learning disclosure process. These procedures include information choice, information preprocessing, information change, DM, and translation and assessment of examples. Different specialists have made recommendations that area information should lead the DM procedure [9].

High-utility data mining] is a prominent assignment in the field on learning revelation. Conventional incessant example digging calculations that scans for gathering of incessant happened things has been reached out by utilizing the approach introduced on [13]. Notwithstanding the helpfulness of incessant example mining, it accepts that everything has comparable significance and has single event in every exchange [13-17]. High-utility example mining settles this impediment by considering that everything may have a weight that will include some helpful data in scanning for those things. A few applications can get helpful data by mining the high utility itemsets in value-based databases, for example, showcase bin investigation, click stream examination, and natural applications [16].

The FP-development calculation is quicker than the Apriori algorithm around a request of extent, the recursive FPgrowth calculation for mining incessant itemsets needs to over and over build restrictive example bases and contingent example tree during

---

*Author for correspondence

the time spent mining, so the recursive mining model not just poor execution of the time, and the storage room is expansive [15]. Notwithstanding, in the previous decades, there are just a few applications in light of FP-development calculation. What's more, despite the fact that a great deal of new calculations were proposed, the change of these calculations depended on parallelization of FP-development calculation [18-21].

The main objectives of our paper are as follows:
1. Investigating and analyzing different data mining techniques for the knowledge discover and pattern identification process.
2. List the factors which needs the improvement and can be helpful in the data mining explorations.
3. Identify and discuss the problem statements along with the suggestive majors for the future analysis and use.

2.Related work
In 2016, Shrivastava et al. [22] suggested that the mining of frequent itemset in association rule mining (ARM) is important. It can be applied in several domains including remedial, biology, banking, retail, market basket analysis, etc. Occasional itemsets finds the shrouded a relationship among the information things. The uncommon solidification of the itemsets can be fascinating and more productive. They have analyzed on the high utility infrequent itemsets using utility pattern rare itemset (UPRI) algorithm. They successfully found the high utility rare itemset patterns and explore the comparisons.

In 2016, Vanahalli et al. [23] suggested that the bioinformatics has contributed to high dimensional datasets. It is the collaboration of large number of features and a small number of samples. The conventional calculations exhaust the majority of the running time in mining extensive number of little and fair size things which does not encase important and noteworthy data. The current research concentrated on mining extensive cardinality itemsets called as gigantic itemsets which are huge to numerous applications, particularly in the field of bioinformatics. The current incessant giant itemset mining calculations are unsuccessful in finding complete arrangement of noteworthy incessant monster itemsets. The mined enormous itemsets from existing calculations give wrong help data which influences affiliation examination. Mining noteworthy regular huge itemsets with precise help data helps in accomplishing an abnormal state exactness of affiliation examination. The proposed

work features a novel pre-preparing method and base up push specification calculation to mine huge successive gigantic itemsets with exact help data. A novel pre-handling strategy proficiently uses least help limit and least cardinality limit to prune superfluous examples and highlights. The investigation comes about show that the proposed calculation has high precision over existing calculations. Execution contemplate demonstrates the effectiveness of the pre-preparing method.

In 2016, Li et al. [24] suggested that the valuable information mining can be performed through the infrequent itemsets. 2L-XMMMS model has the specialty in assigning two different minimum supports to every item. It can be able to mine frequent and infrequent itemset. The author suggested that the efficiency of this model is low as it is based on Apriori algorithm. The authors have used FP-Growth algorithm and named it 2LMS_FP. Their result suggest that the proposed method is fast than 2L-XMMS.

In 2017, Ghorbani [25] suggested that the conventional strategies for finding regular itemsets accept that datasets are static also, the instigated rules are important over the whole dataset. In any case, this isn't the situation when information is worldly. Their work main aim is to improve the efficiency of mining frequent itemsets on temporal data. Since examples can hold in either all or then again a portion of the interims, we propose another calculation to confine time interims, which is called visit itemset mining with time 3D squares. Our concentration is building up a productive calculation for this mining issue by expanding the outstanding from the earlier calculation. The idea of time solid shapes is proposed to deal with time progressive systems. Since examples can hold in either all or on the other hand a portion of the interims, they proposed another calculation to confine time interims, which is called visit itemset mining with time solid shapes. Our concentration is building up a productive calculation for this mining issue by broadening the notable from the earlier calculation. The thought of time 3D shapes is proposed to deal with time chains of command.

In 2017, He et al. [26] suggested that the data mining plays an important role in Big Data. They have proposed MAFIM algorithm which is based on mapreduce and FP-tree for improving the mining efficiency. The data distribution is done by mapreduce. FP-tree has been used for frequent itemset computation. Then the results obtained by

mining were combined by the center node and achieve the global frequent itemsets by mapreduce. Their results suggest that that the MAFIM algorithm is fast and effective.

In 2017, Phuong and Duy [27] proposed average-utility item sets (EHAUI-Tree) algorithm for improving HUUI-Tree algorithm to apply for adding new database transactions without restart. At to begin with, the estimation of refreshed information is figured. At that point, thing sets which roll out improvements will be figured and refreshed relying on the refreshed information esteem what's more, the past High Average-utility Upper-bound (HAUUB). This calculation utilizes the descending conclusion property of a normal utility thing set and a list table structure. In expansion, an information structure for thing sets is proposed to limit memory use and boost ascertaining proficiency. The trial result demonstrates that EHAUI-Tree is more powerful than HAUl-Tree while including new exchanges for the past database. The strategy applies the descending conclusion properties of HAUUB Item set and Index Table. Moreover, the Bit-Array-structure thing set is additionally proposed to diminish utilizing memory and figure all the more successfully. The consequence of this calculation is superior to anything HAUl-Tree on refreshing new exchanges.

In 2017, Zulkurnain and Shah [28] suggested that the data flooding is in different sectors including banking, telecom, scientific experiments, etc. The useful information extraction is possible through data mining from this flooded data. It helps in decision making system from large database by extracting meaningful information. Frequent itemset mining is one of the concentration inquire about regions and an essential advance to balance affiliation rules. Time and space necessities for creating visit itemsets are of absolute significance. Calculations to mine visit itemsets viably help in discovering affiliation rules and furthermore help in numerous other information mining errands. In this paper, a proficient half breed calculation was composed utilizing a binding together procedure of the calculations improved Apriori and FP-Growth. Results show that the proposed half breed calculation, but more mind boggling, expends less memory assets and quicker execution time.

In 2017, Hong et al. [29] suggested erasable-itemset (EI) mining that is helpful in finding the itemsets which also not affect the factory's profit. They have proposed an incremental mining algorithm for erasable itemset is proposed. This basically based on fast-update (FUP). Their results show that the proposed algorithm executes faster than the batch approach in the intermittent data environment.

In 2017, Ismail et al. [30] suggested that the mining high-utility examples is the method for finding sets of valuable things that can give a high benefit in a client exchange database. Finding High-utility itemsets give valuable data that can help in basic leadership by obviously distinguish sets of lucrative things that clients purchased in retail location. Finding client beneficial things in retail location utilizing conventional high-utility strategies is wrong to discover intermittent client practices and furthermore in what way those things identified with each other's do. They have resolved the limitations by providing new method for discovering the productive high-utility occasional examples from client related information. The arrangement of high benefit corresponded gathering of things. They characterize another example development calculation with new tree structure. Test assessments demonstrate that our calculation can uncover valuable data.

In 2017, Klangwisan et al. [31] proposed frequent-regular itemsets mining for finding the interesting itemsets based on the occurrence behavior. Generally, an itemset is distinguished as fascinating, on the off chance that it happens every now and again and consistently in a database. In any case, this undertaking just thinks about things without characterizing distinction or on the other hand centrality of everything which may influence the missing of vital/fascinating learning in true applications. They have introduced mining weighted frequent regular itemsets (WFRIs). Weighted frequent regular itemsets miner (WFRIM) has been proposed for mining WFRIs. A FP-tree like structure named WFRI-tree is outlined to productively keep up competitor itemsets amid mining process. The idea of overestimated-weighted-recurrence of things/itemsets under worldwide/neighborhood most extreme weight is likewise connected to early prune look space. Exploratory outcomes on manufactured and genuine datasets demonstrate effectiveness of WFRIM in the terms of computational time, memory utilization and capacity to discover significant itemsets.

In 2017, Jiang and He [32] proposed a more efficient non-recursive FPNR-growth algorithm and corresponding data structure. Their results from the experiment show the FPNR-growth algorithm performance superiority than the FP-growth

algorithm. It is efficient in storage and mining time both.

In 2017, Mohammed et al. [33] suggested that the maximal frequent itemset is the biggest incessant itemset in a database which isn't secured by different itemsets. All visit itemsets can be developed from maximal one. Additionally, it is conceivable to center on any piece of the maximal continuous itemset to direct Data Mining. Honey bees' Algorithm is basic, powerful and populace based stochastic enhancement calculation which is in view of honey bees' regular searching propensities. It plays out a neighborhood look joined with irregular inquiry. They have presented maximal frequent itemset-oriented bees'algorithm. It is used to mine maximal frequent itemsets from transactional databases which has been named mining maximal itemsets bees' algorithm (MMIBA). The wellness, coding, scout honey bees' obligations, kind of collected data, and end criteria have been situated to the issue of maximal successive itemset mining. MMIBA was connected on genuine databases accessible freely on the Internet which are chess, Mushroom, Malignancy Cells, Census information, and Dense Census. The analyses were proficient relying upon three levels of least bolster limit, which are twenty five, fifty, and seventy five rates of the databases' sizes, to approve the effectiveness of MMIBA. The level twenty five rate was portrayed to explain the capacity of MMIBA in mining the MFIs with low estimations of least help, while the level seventy five rates to approve its capacity in high least help values.

In 2017, Subbulakshmi et al. [34] suggested frequent itemsets as the major task in data mining. Be that as it may, in genuine, each thing can't be given same importance. And furthermore things that are visit as of late are critical ones to be mined.

Frequent itemsets mining is fused with these two variables to give another strategy called recent weighted frequent itemsets mining. Late weighted frequent itemsets mining is an expansion of frequent itemsets mining errand that is utilized for mining the incessant itemsets that fulfill the weight and regency limitations. This paper centers around mining maximal later weighted successive itemsets. maximal recent weighted frequent itemsets is a minimal portrayal of recent weighted frequent itemsets. This portrayal is especially helpful when the database estimate is substantial and memory is an issue. Number of maximal late weighted regular itemsets will be in particular lesser than the quantity of late

weighted incessant itemsets for a given dataset. The time required for mining and memory required for putting away is less for maximal successive itemsets at the point when contrasted and visit itemsets. The consequences of the proposed technique demonstrate a change in the time multifaceted nature and furthermore the memory space required for mining.

In 2017, Khode and Mohod [35] suggested that the ining high utility itemsets from a value-based database alludes to the revelation of itemsets with high utility like benefits. In spite of the fact that a number of applicable methodologies have been proposed as of late, however they bring about the issue of creating an expansive number of competitor itemsets for high utility itemsets. Such countless applicant itemsets corrupts the mining execution regarding execution time and space prerequisite. The circumstance may turn out to be more regrettable when the database contains loads of long exchanges or long high utility itemsets. A rising point in the field of information mining is utility mining which not just considers the recurrence of the itemsets yet in addition thinks about the utility related with the itemsets. The primary target of high utility itemset mining is to recognize itemsets that have utility qualities over a given utility edge. Therefore Utility mining assumes an essential part in numerous constant applications and is an essential research theme in information mining framework to discover the itemsets with high benefit. They have presented the implementation of first module. It covers pre-processing and product base dataset by using Top K Rules. For this a new framework has been presented. Top-k high utility web get to designs, where k is the coveted number of HUIs to be mined. Two sorts of productive calculations named TKU and Technical knockout are proposed for mining such itemsets.

In 2017, Wang et al. [36] suggested that the data stream frequent pattern mining is an important aspect in the field of data mining. Their paper discusses their definitions, relationship in terms of frequent itemsets and sliding windows. It also analyzes the classifies sliding windows from data processing models then analyzes the use of sliding windows in run of the mill visit itemsets mining calculations, and outlines the mining methods and proficiency of normal incessant itemsets mining calculations.

In 2018, Bai et al. [37] suggested that the high-utility itemset mining (HUIM) is an emerging area of data mining. Authors suggested that it differs from the frequent itemset mining (FIM). As the last considers

just the recurrence factor though the previous has been intended to address both amount and benefit components to uncover the most gainful items. The difficulties of producing the HUI incorporate exponential unpredictability in both time and space. Also, the pruning procedures of decreasing the hunt space which is accessible in FIM in light of their monotonic and hostile to monotonic properties can't be utilized as a part of HUIM. They proposed a novel selective database projection based high-utility itemset mining algorithm (SPHUI-Miner). They have also introduced HUI-RTPL data format. They also proposed two novel data structures. These are selective database projection utility list (SPUList) and tail-count list to prune the search space for HUI mining. Particular projections of the database lessen the examining time of the database making our proposed approach more productive. It makes one of a kind information cases and new projections for information having fewer measurements in this manner bringing about quicker HUI mining. They demonstrated upper limits on the measure of memory devoured by these projections. Their results suggest that the SPHUI-Miner algorithm outperforms in terms of computation time, memory usage, scalability, and candidates generation.

3.Problem identification

The following problem statements have been found by the research and analysis:

- Positive and negative rule mining along with the ranking based frequent item periodization is missing.
- Less exploration is done in the direction of selective database projections with upper bounds and lower bound pruning techniques with average utility frequency.
- The number of candidate set, frequency generation set and accuracy combined missing in the previous work.
- Better algorithms are needed in case of knowledge discovery in the negative rule mining and threshold level assignment.
- The candidate reduction to speed up the process is also missing in the previous literature.
- There is the need of exploration in mining weighted frequent itemsets in the data streams.

4.Discussion and analysis

Table 1 shows the results analysis from the previous published literature. It shows the comparisons from different methods and results are compared. Based on this gaps are highlighted.

Table 1 Result analysis

S.NO	Authors	Year	Work	GAP
1	Nan et al. [38]	2018	In this work a safety systematics model have been developed for accident cases. Similar units analysis and similarity degree evaluation are the two main components of this method. The advantages of this model are it is capable in enabling qualitative and quantitative data collecting about different accidents.	Relationship between similar data can be developed in the future.
2	Rocha et al. [39]	2018	They have explored the expressive power of the multi-relational inductive logic programming (ILP) data mining approach. They explored the capability of ILP to handle and combine (implicit) multiple relationships in the E-COMPARED data. They focused on the scores of the PHQ-8 self-report questionnaire. All the measures were considered during treatment, which is in between 9 to 16 weeks.	Different cutoff criteria may be examined in the future.
3	Lu et al. [40]	2017	They have introduced materials data mining (MDM) process. Both subjective and quantitative strategies in machine learning can be embraced in the MDM procedure to achieve distinctive errands in materials disclosure, outline, and improvement.	More exploration is needed in candidate generation and optimization.
4	Rojas et al. [41]	2015	They proposed augmented model visualization for data mining (AMV-DM), based on models of visual perception and interaction to represent and operate with data mining models. The plan has at its center the utilization of reciprocal representations connected to DM demonstrated amid the change stage. This correlative sees relate to: a moment elucidating strategy of information mining, and a proper arrangement of graphical curios.	Different combination of DM techniques hybridization has been suggested.
5	Vadim [42]	2018	The main aim of this work is to analysis the large amounts of information and comparison of methods. Due to of the immense assortment of information writes and types of sorting out data	Practical situations and implementations are missing.

			genuine information may not generally be investigated by machine learning instruments. For the change of "crude" information to the information, which can work proficiently Data Mining systems, take care of the issue of pre-handling. They suggested K-nearest neighbor and decision trees for such problems.	
6	Darrab et al. [43]	2017	They presented multiple item support-ECLAT; MIS-ECLAT algorithm. It is helpful in mining frequent patterns including rare itemsets under multiple support thresholds (MIS) by utilizing a vertical representation of data. They compared MISeclat to tree based algorithm, MISFP-growth28 and another recent algorithm, CFP-growth++22 in terms of execution time, memory usage and scalability on both sparse and dense databases and found to be better in comparison to previous methods.	This method may be used on other dataset for checking the effectiveness along with the parameter variability.
7	Jabbour et al. [44]	2018	To mine the most relevant negative rules, the mining task of negative association rules is regularly combined with new measure, for example, lift or conviction to restrict the arrangement of separated association rules have been suggested in this work. They have addressed the problem of mining strong negative rules by extending the previous SAT-based approach [45].	They have suggested to mine other kind of associations using a constraints approach such as exception and rares ones in the future.

5.Conclusions and future directions

This paper addresses and explores different data mining methods in the process of achieving efficient knowledge discovery system creation and mining frequent itemset. It covers the positive and negative association rule mining along with the prospective, impact and insights in the future. It also covers the results which are published currently along with the future suggestions. From the above analysis and discussion it is clear that an efficient data mining knowledge discovery model is needed which can capable of handling candidate generation, negative and positive association rule mining.

Conflicts of interest
The authors have no conflicts of interest to declare.

References
[1] Agrawal R, Srikant R. Fast algorithms for mining association rules. In proceeding of international conference on very large data bases, VLDB 1994 (pp. 487-99).
[2] Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. In ACM SIGMOD record 2000 (pp. 1-12). ACM.
[3] Zaki MJ, Parthasarathy S, Ogihara M, Li W. New algorithms for fast discovery of association rules. In KDD 1997 (pp. 283-6).
[4] Jamil A, Salam A, Amin F. Performance evaluation of top-k sequential mining methods on synthetic and real datasets. International Journal of Advanced Computer Research. 2017; 7(32):176-84.
[5] Agarwal RC, Aggarwal CC, Prasad VV. A tree projection algorithm for generation of frequent item sets. Journal of Parallel and Distributed Computing. 2001; 61(3):350-71.
[6] Pei J, Han J, Lu H, Nishio S, Tang S, Yang D. H-mine: hyper-structure mining of frequent patterns in large databases. In proceedings IEEE international conference on data mining 2001 (pp. 441-8). IEEE.
[7] Dubey AK, Dubey AK, Agarwal V, Khandagre Y. Knowledge discovery with a subset-superset approach for mining heterogeneous data with dynamic support. In CSI sixth international conference on software engineering 2012 (pp. 1-6). IEEE.
[8] Babu DB, Prasad RS, Umamaheswararao Y. Efficient frequent pattern tree construction. International Journal of Advanced Computer Research. 2014; 4(14):331-6.
[9] Li K, Cui L. A kernel fuzzy clustering algorithm with generalized entropy based on weighted sample. International Journal of Advanced Computer Research. 2014; 4(15):596-600.
[10] Horeis T, Sick B. Collaborative knowledge discovery & data mining: from knowledge to experience. In IEEE symposium on computational intelligence and data mining 2007 (pp. 421-8). IEEE.
[11] Feng Y, Wu Z, Zhou Z. Enhancing reliability throughout knowledge discovery process. In international conference on data mining workshop 2006 (pp. 754-8). IEEE.
[12] Vityaev EE, Kovalerchuk BY. Relational methodology for data mining and knowledge discovery. Intelligent Data Analysis. 2008; 12(2):189-210.
[13] Fournier-Viger P, Wu CW, Zida S, Tseng VS. FHM: faster high-utility itemset mining using estimated utility co-occurrence pruning. In international

symposium on methodologies for intelligent systems 2014 (pp. 83-92). Springer, Cham.

[14] Lan GC, Hong TP, Tseng VS. An efficient projection-based indexing approach for mining high utility itemsets. Knowledge and Information Systems. 2014; 38(1):85-107.

[15] Song W, Liu Y, Li J. BAHUI: fast and memory efficient mining of high utility item sets based on bitmap. International Journal of Data Warehousing and Mining. 2014; 10(1):1-15.

[16] Tseng VS, Shie BE, Wu CW, Philip SY. Efficient algorithms for mining high utility item sets from transactional databases. IEEE Transactions on Knowledge and Data Engineering. 2013; 25(8):1772-86.

[17] Rashidi P, Cook DJ. Mining sensor streams for discovering human activity patterns over time. In international conference on data mining 2010 (pp. 431-40). IEEE.

[18] Wang B, Chen D, Shi B, Zhang J, Duan Y, Chen J, et al. Comprehensive association rules mining of health examination data with an extended FP-growth method. Mobile Networks and Applications. 2017; 22(2):267-74.

[19] Xu F, Lu H. The application of FP-Growth algorithm based on distributed intelligence in wisdom medical treatment. International Journal of Pattern Recognition and Artificial Intelligence. 2017; 31(4):1-11.

[20] Pei B, Wang X, Wang F. Parallelization of FP-growth algorithm for mining probabilistic numerical data based on MapReduce. In international symposium on computational intelligence and design 2016 (pp. 223-6). IEEE.

[21] Makanju A, Farzanyar Z, An A, Cercone N, Hu ZZ, Hu Y. Deep parallelization of parallel FP-growth using parent-child Map Reduce. In international conference on Big Data 2016 (pp. 1422-31). IEEE.

[22] Shrivastava S, Johari PK. Analysis on high utility infrequent itemsets mining over transactional database. In international conference on recent trends in electronics, information & communication technology 2016 (pp. 897-902). IEEE.

[23] Vanahalli MK, Patil N. Association analysis of significant frequent colossal itemsets mined from high dimensional datasets. In international conference on electrical, computer and electronics engineering 2016 (pp. 258-63). IEEE.

[24] Li C, Dong X, Dong X, Ren X. FP-growth based method for mining infrequent and frequent itemsets with 2-level minimum support. In international conference on computer science and network technology 2016 (pp. 263-7). IEEE.

[25] Ghorbani M, Abessi M. A new methodology for mining frequent itemsets on temporal data. IEEE Transactions on Engineering Management. 2017; 64(4):566-73.

[26] He B, Pei J, Zhang H. The mining algorithm of frequent itemsets based on mapreduce and FP-tree. In international conference on computer network, electronic and automation 2017(pp. 108-11). IEEE.

[27] Phuong N, Duy ND. Constructing a new algorithm for high average utility itemsets mining. In international conference on system science and engineering 2017 (pp. 273-8). IEEE.

[28] Zulkurnain NF, Shah A. HYBRID: an efficient unifying process to mine frequent itemsets. In 3rd international conference on engineering technologies and social sciences 2017 (pp. 1-5). IEEE.

[29] Hong TP, Lin KY, Lin CW, Vo B. An incremental mining algorithm for erasable itemsets. In international conference on innovations in intelligent systems and applications 2017 (pp. 286-9). IEEE.

[30] Ismail W, Hassan MM, Fortino G. Productive-associated periodic high-utility itemsets mining. In international conference on networking, sensing and control 2017 (pp. 637-42). IEEE.

[31] Klangwisan K, Amphawan K. Mining weighted-frequent-regular itemsets from transactional database. In international conference on knowledge and smart technology 2017 (pp. 66-71). IEEE.

[32] Jiang H, He X. An improved algorithm for frequent itemsets mining. In international conference on advanced cloud and big data 2017 (pp. 314-7). IEEE.

[33] Mohammed MA, Al-Khafaji H. Maximal itemsets mining algorithm based on bees algorithm. In annual conference on new trends in information & communications technology applications 2017 (pp. 1-6). IEEE.

[34] Subbulakshmi B, Dharini B, Deisy C. Recent weighted maximal frequent itemsets mining. In international conference on IoT in social, mobile, analytics and cloud 2017 (pp. 391-7). IEEE.

[35] Khode S, Mohod S. Mining high utility itemsets using TKO and TKU to find top-k high utility web access patterns. In international conference of electronics, communication and aerospace technology 2017 (pp. 504-9). IEEE.

[36] Wang H, Li F, Tang D, Wang Z. Research on data stream mining algorithm for frequent itemsets based on sliding window model. In international conference on big data analysis 2017 (pp. 259-63). IEEE.

[37] Bai A, Deshpande PS, Dhabu M. Selective database projections based approach for mining high-utility itemsets. IEEE Access. 2018; 6:14389-409.

[38] Nan J, Cheng L, Yi L. A similar safety systematics model for accident cases data mining support. Procedia Computer Science. 2018; 131:929-36.

[39] Rocha A, Camacho R, Ruwaard J, Riper H. Using multi-relational data mining to discriminate blended therapy efficiency on patients based on log data. Internet Interventions. 2018; 12:176-80.

[40] Lu W, Xiao R, Yang J, Li H, Zhang W. Data mining-aided materials discovery and optimization. Journal of Materiomics. 2017; 3(3):191-201.

[41] Rojas WC, Quispe FM, Villegas CM. Augmented visualization for data-mining models. Procedia Computer Science. 2015; 55:650-9.

[42] Vadim K. Overview of different approaches to solving problems of data mining. Procedia Computer Science. 2018; 123:234-9.

[43] Darrab S, Ergenç B. Vertical pattern mining algorithm for multiple support thresholds. Procedia Computer Science. 2017; 112 (2017):417–26.

[44] Jabbour S, El Mazouri FE, Sais L. Mining negatives association rules using constraints. Procedia Computer Science. 2018; 127(2018):481-8.

[45] Boudane A, Jabbour S, Sais L, Salhi Y. A sat-based approach for mining association rules. In proceedings of the international joint conference on artificial intelligence 2016   (pp. 2472-8). AAAI Press.

Bhagwan Singh was born in Rajasthan India in 1972. He received his AMIE (ELECTRICAL) engineering from the institution of Engineers (India) in 2000 and AMIE (Electronics & ommunication) engineering from the institution of Engineers (India) in 2010. Master of Computer application from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal in 2010. Currently, He is doing PhD in Computer Science and applications from Sri Satya Sai University of Technology and Medical Sciences Sehore,M.P., India. His research area is data mining.
Email: bhagwansingh@bhel.in