A Positive-Negative Frequent Itemsets mining based on PSO

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

Received: May 15, 2018 Accepted: June 20, 2018

ABSTRACT *An efficient PSO has been proposed for frequent itemsets mining. It is based on random bias removal in each iteration. Numeric and character both data can be accepted in our framework.Mushroom T10I4D100K dataset have been used for experimentation. Our approach also provides the facility of artificial dataset creation. Support value has been calculated based on the flexibility of the multi-minimum support in one time. Positive as well as negative association has been considered. Then PSO algorithm has been applied and the results are compared based on the positive (Max-Min) and Negative (Max-Min). The results are compared from the previous approaches.It clearly shows that the produced results in case of minimization and maximization are better in comparison to the previous method in case of support threshold as well as time.*

Keywords: *Data mining, Association rules, Positive and negative associations, PSO*

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 andpruning (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.

Traditional data mining algorithm 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 Apriorialgorithm 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 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 this paper are to mine positive and negative association rule mining based on PSO. It also includes the Max-Min values association in the positive and negative associations.

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 lw 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 HAUl-Tree.

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 improvedApriori 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 theproductive high-utility occasional examples from client related information. The arrangement of high benefit corresponded gathering of things.

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 frequentregular itemsets miner (WFRIM) has been proposed for miningWFRIs. 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 datastructure. 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 itemsetsfromtransactional databases which has been named mining maximalitemsets 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 itemsetsmining 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 frequentitemsets. 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 memoryrequired 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 TopKRules. 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 inrun 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. Material and Methods

For the appropriate data reduction and selection apriorialgorithm is used. The high support itemset is found and their support count is evaluated. Steps involve in apriori algorithm is as follows:

Step 1: Dataset selection as the input.

Step 2: Determine the number of itemset.

Step 3: Support allocation has been done for each itemset.

Support = (Total number of transactions for A) / (Complete transactions)

Step 4: Only those itemset are retained which passes the accepted threshold otherwise the itemset are rejected for the further process.

Step 5: Generate the possibility from the outcome itemset till the dataset selection.

Step 6: The above steps are repeated till the values from the set are not completed.

Binary PN Particle Swarm Optimization

ARM–association rule measures
I_1, I_2, I_3….–Iterations
RN_x–Random number
AR–Association rules
RN_{x-1}–Random number of previous step
Input:

- Rules (R_1,R_2….R_n)

Output:

- RS_1, RS_2,……RS_n

Step 1: Association rules values are the input.
Step 2: The values are used either as the minimization or maximization.
Step 3: The below process are same for maximization and minimization.
Step 4: Particle assignment and updations
i=1 to 5
if i==1

AR_i= (R_1 + R_2 + R_3 + …….+R_n)/n
If (RN_{xi+1}>RN_{xi})
RN_{xi+1} = RN_{xi}
else
No change

If i==2
AR_i= AR_{i-1} + (R_1 X RN_{xi} + R_2 X RN_{xi} + R_3 X RN_{xi} + …….+R_n X RN_{xi})/n
If (RS_{ti+1}>RS_{ti})
RS_{ti+1} = RS_{ti}
else
No change
RV_{x-1} = RV_{xi}

else
AR_i= AR_{i-1} + (R_1 X RN_{xi} + R_2 X RN_{xi} + R_3 X RN_{xi} + …….+R_n X RN_{xi})/n-RN_{x-1}/n
If (RS_{ti+1}>RS_{ti})
RS_{ti+1} = RS_{ti}
else
No change
RV_{x-1} = RV_{xi}

Step 5: Association rule accuracy has been obtained.
Step 6: Finish

In this paper an efficient PSO has been proposed based on random bias removal in each iteration. Figure 1 shows the procedural flowchart of our approach. In our approach first the dataset is selected. Numeric and character based data has been accepted in our approach. Then separation has been checked. Space separation has been accepted otherwise rejected. Then support count of the data has been calculated for each item presented in the dataset. In our approach there is flexibility for the multi-minimum support in one time. Positive as well as negative association has been considered in our approach. As negative association may have some impact in overall accuracy also. Then PSO algorithm has been applied and the results are compared based on the positive (Max-Min) and Negative (Max-Min). The results are compared from the simple rules based association and impacts are compared.

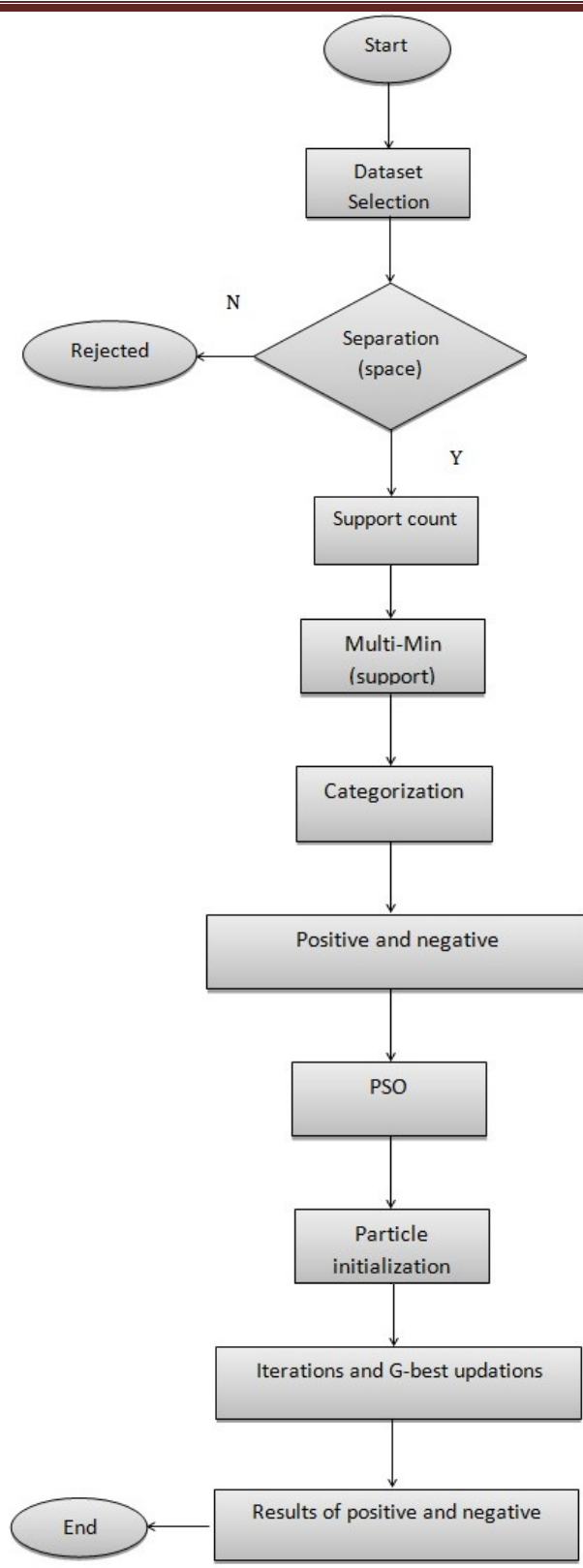

Figure 1: Flowchart of the working process

4. Result and Discussion

In this section results are compared based on different experimentation performed and compared with the previous methods. T10I4D100K, Mushroom and artificial created dataset from our approach. IBM Almaden Quest research group has developed the T10I4D100K dataset. In this dataset there are 1000 items and 100000 transactions. The mushroomdatasetis the collection of different species of mushrooms. It contains 8124 instances and 119 attributes. It is derived from 24 species. Both of the dataset are dense. For the first experiment T10I4D100K dataset has been considered. The minimum supports considered are 30%, 40%, 50% and 60%. Here P shows positive association and N shows negative association. Min and max shows the minimum and maximum values. B shows before PSO and A shows after PSO.Figure 2 shows the result for minimum support (30) for before and after PSO. The negative and positive rule counts are 722 and 73 respectively. Figure 3 shows the result for minimum support (40) for before and after PSO.Th e negative and positive rule counts are 767 and 28 respectively.Figure 4shows the result for minimum support (50) for before and after PSO. The negative and positive rule counts are 783 and 12 respectively.Figure 5shows the result for minimum support (60) for before and after PSO. The negative and positive rule counts are 792 and 3 respectively.

Figure 6 shows the result for minimum support (30) for before and after PSO (created dataset). Figure 7 shows the result for minimum support (30) for before and after PSO (Mushroom). Figure 8 shows the result for minimum support (50) for before and after PSO (Mushroom). Figure 9 shows the result for minimum support (60) for before and after PSO (Mushroom).Figure 10 shows the support utilization comparison from the previous method. It clearly shows that our results in case of minimization and maximization are better in comparison to the previous method. Figure 11 shows the time comparison for Mushroom dataset and figure 12 shows the time comparison for T10I4D100K dataset. It has been compared from [32] and [33] approach. The time comparison also suggested that our results outperform from the previous approaches.

In our approach both positive and negative associations are considered and the performance has been improved in both of the cases. It is clear from different experimentation and results discuss above. It also suggests that the PSO approach is beneficial in the association rules for the improvement in the impact of the final support thresholds considering with different cases.

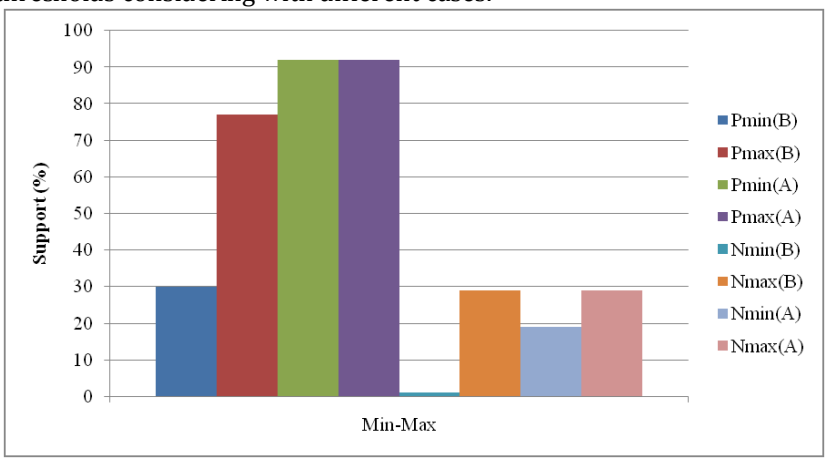

Figure 2 Result for minimum support (30) for before and after PSO (T10I4D100K)

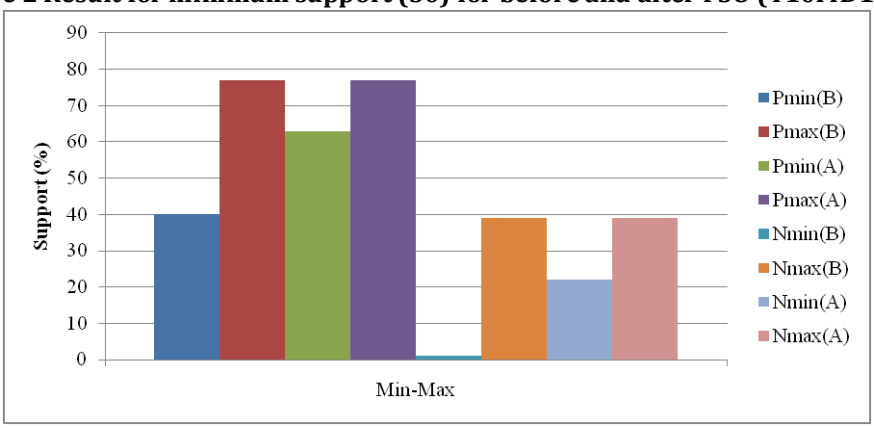

Figure 3Result for minimum support (40) for before and after PSO(T10I4D100K)

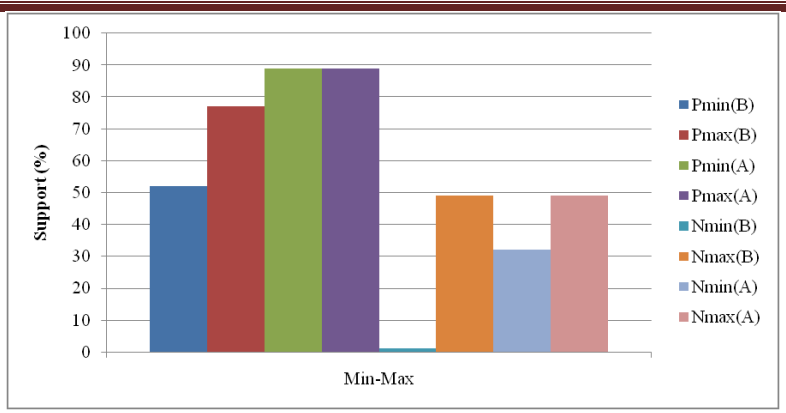

Figure 4 Result for minimum support (50) for before and after PSO(T10I4D100K)

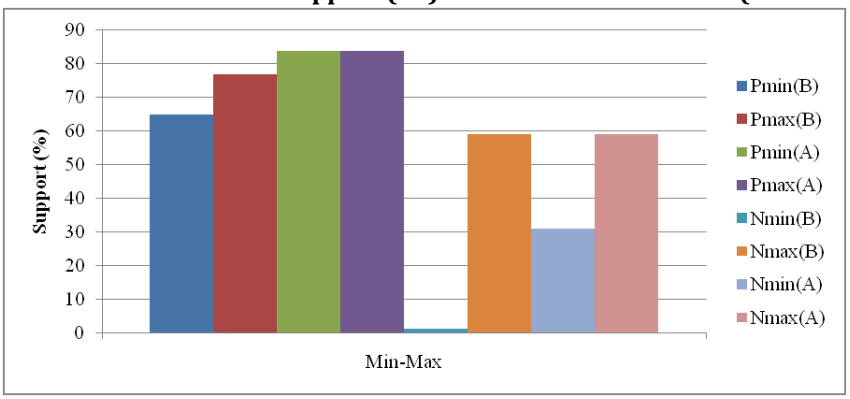

Figure 5 Result for minimum support (60) for before and after PSO(T10I4D100K)

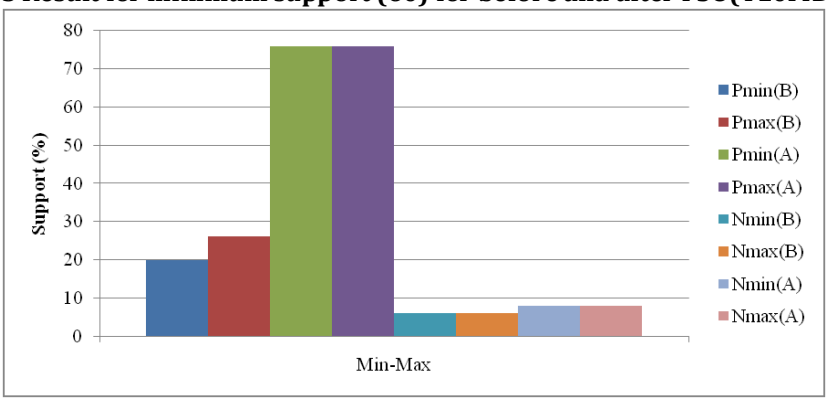

Figure 6Result for minimum support (30) for before and after PSO(Created dataset)

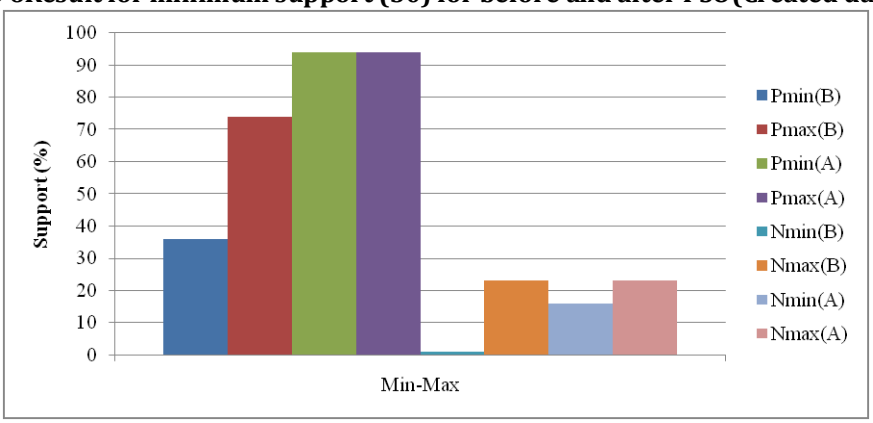

Figure 7Result for minimum support (30) for before and after PSO (Mushroom)

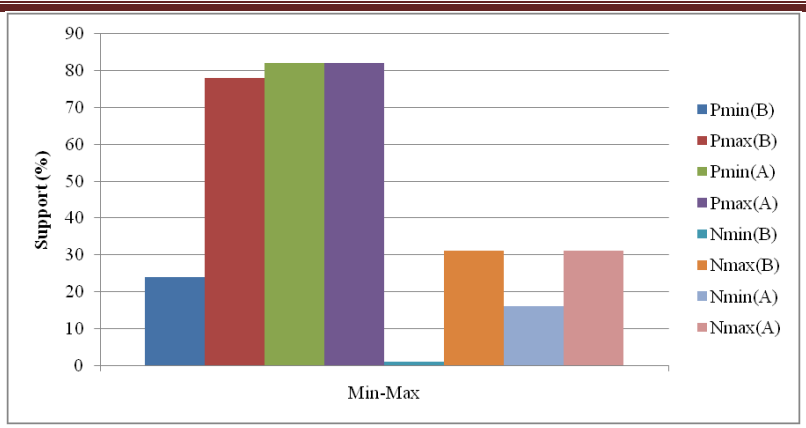

Figure 8 Result for minimum support (50) for before and after PSO (Mushroom)

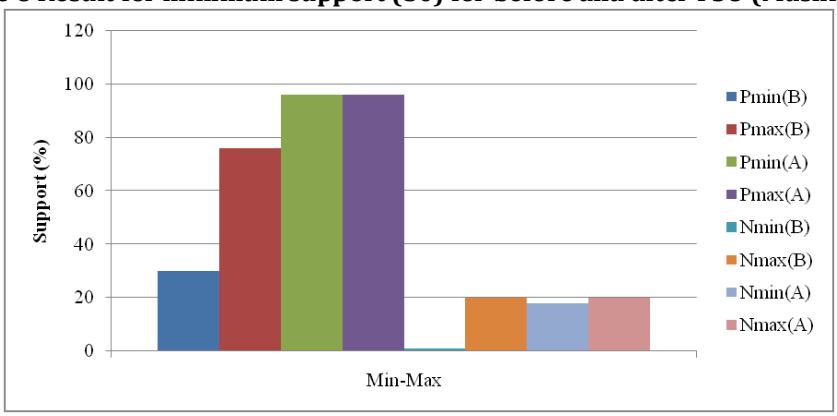

Figure 9 Result for minimum support (60) for before and after PSO (Mushroom)

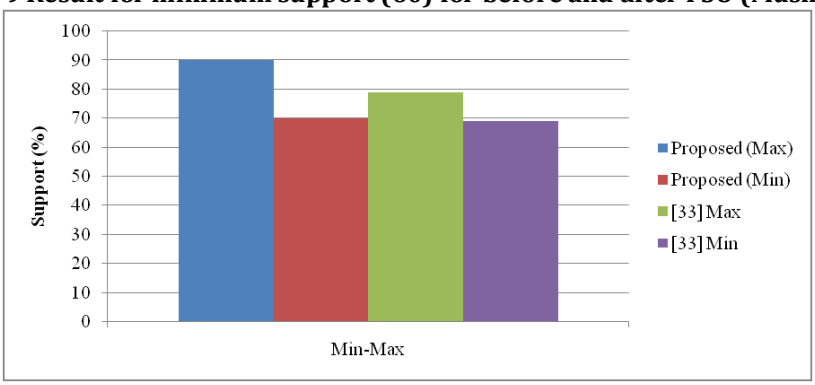

Figure 10 Support utilization comparison

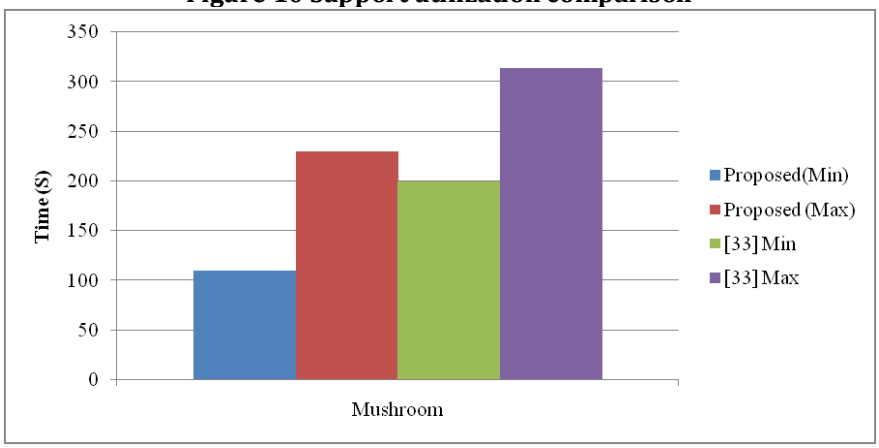

Figure 11 Time comparison for Mushroom dataset

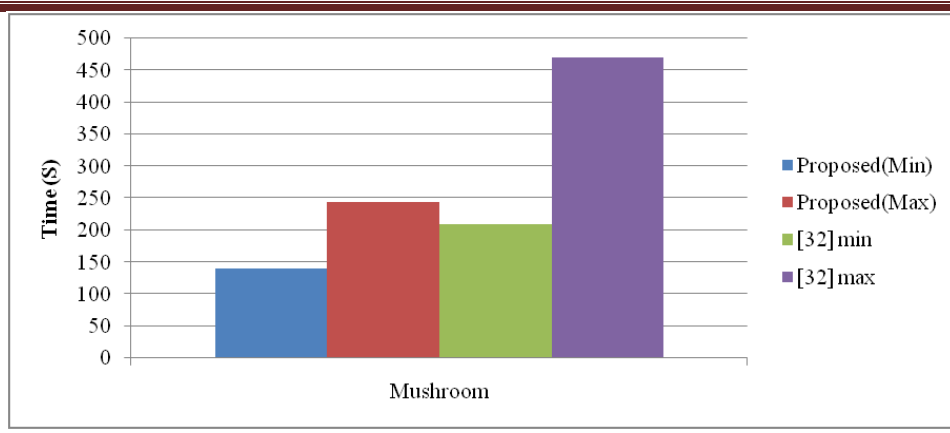

Figure 12 Time comparison for T10I4D100K dataset

5. Conclusions

This paper provides the systematic analysis of the approaches used in frequent itemset mining.This paper proposed an efficient frequent itemsets mining based on PSO for the positive (Max-Min) and negative (Max-Min). This approach provides the flexibility of the multi minimum support selection with numeric and character based data support. The results obtained from the proposed approach in case of support threshold values and times are superior from the previous approach. The results also indicate that the PSO approach is beneficial in the association rules for the improvement in the impact of the final support thresholds considering with different cases.

References

[1] Agrawal R, Srikant R. Fast algorithms for mining association rules. InProc. 20th int. conf. very large data bases, VLDB 1994 Sep 12 (Vol. 1215, pp. 487-499).

[2] Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. In ACM SIGMOD record 2000 May 16 (Vol. 29, No. 2, pp. 1-12). ACM.

[3] Zaki MJ, Parthasarathy S, Ogihara M, Li W. New Algorithms for Fast Discovery of Association Rules. InKDD 1997 Aug 14 (Vol. 97, pp. 283-286).

[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 Sep 1;7(32):176.

[5] Agarwal RC, Aggarwal CC, Prasad VV. A tree projection algorithm for generation of frequent item sets. Journal of parallel and Distributed Computing. 2001 Mar 1;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. InData Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on 2001 (pp. 441-448). IEEE.

[7] Dubey AK, Dubey AK, Agarwal V, Khandagre Y. Knowledge discovery with a subset-superset approach for Mining Heterogeneous Data with dynamic support. InSoftware Engineering (CONSEG), 2012 CSI Sixth International Conference on 2012 Sep 5 (pp. 1-6). IEEE.

[8] Babu DB, Prasad RS, Umamaheswararao Y. Efficient frequent pattern tree construction. International Journal of Advanced Computer Research. 2014 Mar 1;4(1):331.

[9] Li K, Cui L. A kernel fuzzy clustering algorithm with generalized entropy based on weighted sample. International Journal of Advanced Computer Research. 2014 Jun 1;4(2):596.

[10] Horeis T, Sick B. Collaborative knowledge discovery & data mining: From knowledge to experience. InComputational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on 2007 Mar 1 (pp. 421-428). IEEE.

[11] Feng Y, Wu Z, Zhou Z. Enhancing reliability throughout knowledge discovery process. InData Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on 2006 Dec (pp. 754-758). IEEE.

[12] Vityaev EE, Kovalerchuk BY. Relational methodology for data mining and knowledge discovery. Intelligent Data Analysis. 2008 Jan 1;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. InInternational symposium on methodologies for intelligent systems 2014 Jun 25 (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 Jan 1;38(1):85-107.

[15] Song W, Liu Y, Li J. BAHUI: fast and memory efficient mining of high utility itemsets based on bitmap. International Journal of Data Warehousing and Mining (IJDWM). 2014 Jan 1;10(1):1-5.
[16] Tseng VS, Shie BE, Wu CW, Philip SY. Efficient algorithms for mining high utility itemsets from transactional databases. IEEE transactions on knowledge and data engineering. 2013 Aug;25(8):1772-86.
[17] Rashidi P, Cook DJ. Mining sensor streams for discovering human activity patterns over time. InData Mining (ICDM), 2010 IEEE 10th International Conference on 2010 Dec (pp. 431-440). IEEE.
[18] Wang B, Chen D, Shi B, Zhang J, Duan Y, Chen J, Hu R. Comprehensive association rules mining of health examination data with an extended FP-growth method. Mobile Networks and Applications. 2017 Apr 1;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 Apr;31(04):1759005.
[20] Pei B, Wang X, Wang F. Parallelization of FP-growth Algorithm for Mining Probabilistic Numerical Data based on MapReduce. InComputational Intelligence and Design (ISCID), 2016 9th International Symposium on 2016 Dec 10 (Vol. 2, pp. 223-226). IEEE.
[21] Makanju A, Farzanyar Z, An A, Cercone N, Hu ZZ, Hu Y. Deep parallelization of parallel FP-growth using parent-child MapReduce. InBig Data (Big Data), 2016 IEEE International Conference on 2016 Dec 5 (pp. 1422-1431). IEEE.
[22] Shrivastava S, Johari PK. Analysis on high utility infrequent ItemSets mining over transactional database. InRecent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on 2016 May 20 (pp. 897-902). IEEE.
[23] Vanahalli MK, Patil N. Association analysis of significant frequent colossal itemsets mined from high dimensional datasets. InElectrical, Computer and Electronics Engineering (UPCON), 2016 IEEE Uttar Pradesh Section International Conference on 2016 Dec 9 (pp. 258-263). 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. InComputer Science and Network Technology (ICCSNT), 2016 5th International Conference on 2016 Dec 10 (pp. 263-267). IEEE.
[25] Ghorbani M, Abessi M. A New Methodology for Mining Frequent Itemsets on Temporal Data. IEEE Transactions on Engineering Management. 2017 Nov;64(4):566-73.
[26] He B, Pei J, Zhang H. The Mining Algorithm of Frequent Itemsets based on Mapreduce and FP-tree. In2017 International Conference on Computer Network, Electronic and Automation (ICCNEA) 2017 Sep 1 (pp. 108-111). IEEE.
[27] Phuong N, Duy ND. Constructing a new algorithm for high average utility Itemsets mining. InSystem Science and Engineering (ICSSE), 2017 International Conference on 2017 Jul 21 (pp. 273-278). IEEE.
[28] Zulkurnain NF, Shah A. HYBRID: An efficient unifying process to mine frequent itemsets. InEngineering Technologies and Social Sciences (ICETSS), 2017 IEEE 3rd International Conference on 2017 Aug 7 (pp. 1-5). IEEE.
[29] Hong TP, Lin KY, Lin CW, Vo B. An incremental mining algorithm for erasable itemsets. InINnovations in Intelligent SysTems and Applications (INISTA), 2017 IEEE International Conference on 2017 Jul 3 (pp. 286-289). IEEE.
[30] Ismail W, Hassan MM, Fortino G. Productive-associated Periodic High-utility itemsets mining. InNetworking, Sensing and Control (ICNSC), 2017 IEEE 14th International Conference on 2017 May 16 (pp. 637-642). IEEE.
[31] Klangwisan K, Amphawan K. Mining weighted-frequent-regular itemsets from transactional database. InKnowledge and Smart Technology (KST), 2017 9th International Conference on 2017 Feb 1 (pp. 66-71). IEEE.
[32] Jiang H, He X. An Improved Algorithm for Frequent Itemsets Mining. InAdvanced Cloud and Big Data (CBD), 2017 Fifth International Conference on 2017 Aug 13 (pp. 314-317). IEEE.
[33] Mohammed MA, Al-Khafaji H. Maximal itemsets mining algorithm based on Bees' Algorithm. InNew Trends in Information & Communications Technology Applications (NTICT), 2017 Annual Conference on 2017 Mar 7 (pp. 1-6). IEEE.
[34] Subbulakshmi B, Dharini B, Deisy C. Recent weighted maximal frequent itemsets mining. InI-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2017 International Conference on 2017 Feb 10 (pp. 391-397). IEEE.
[35] Khode S, Mohod S. Mining high utility itemsets using TKO and TKU to find top-k high utility web access patterns. InElectronics, Communication and Aerospace Technology (ICECA), 2017 International conference of 2017 Apr 20 (Vol. 1, pp. 504-509). IEEE.
[36] Wang H, Li F, Tang D, Wang Z. Research on data stream mining algorithm for frequent itemsets based on sliding window model. InBig Data Analysis (ICBDA), 2017 IEEE 2nd International Conference on 2017 Mar 10 (pp. 259-263). IEEE.
[37] Bai A, Deshpande PS, Dhabu M. Selective Database Projections Based Approach for Mining High-Utility Itemsets. IEEE Access. 2018;6:14389-409.