

F-PNWAR: Fuzzy-based Positive and Negative Weighted Association Rule Mining Algorithm

K. Mangayarkkarasi^{#1}, M. Chidambaram^{*2}

[#] Research Scholar,

Research and Development Center, Bharathiar University, Coimbatore, Tamilnadu, India

¹ kanthimangai@gmail.com

^{*} Assistant Professor

Rajah Serfoji College, Bharathidasan University, Thanjavur, Tamilnadu, India

² chidsuba@gmail.com

Abstract—Association Rule Mining (ARM) algorithm motivates on mining of the Positive Association Rules (PARs). In recent times, the researchers focused on mining the Negative Association Rules (NARs) by finding the interesting infrequent itemsets. Existing ARM algorithms discovers only the PARs and treat each item with same significance. But, the significance of each item may differ from each other. This paper proposes a Fuzzy-based Positive and Negative Weighted Association Rule (F-PNWAR) mining algorithm for the market-based data analysis. The itemsets are ranked and weight is assigned to the itemsets based on the rank. The positive and negative weighted itemsets are extracted and rule is generated. The proposed F-PNWAR algorithm is compared with the existing weighted ARM (WARM), Fuzzy WARM (FWARM), Enhanced FWARM (E-FWARM), traditional K-means and Adaptive K-means algorithms. The comparative analysis shows that the proposed F-PNWAR algorithm achieves maximum frequency item rate, association rule rate, accuracy and minimum execution time than the existing algorithms.

Keyword - Fuzzy Association Rules, Negative Association Rule (NAR), Positive Association Rule (PAR), Weighted Association Rule Mining (WARM)

I. INTRODUCTION

Association Rule Mining (ARM) [1, 2] is the highly significant research area in the data mining and knowledge discovery applications. Association rules provide an efficient way for identifying the certain dependencies between the attributes in the database [3]. An association rule is formed as $P \Rightarrow Q$, where 'P' and 'Q' are disjoint itemsets. The support value of the itemsets is not less than the user-specified minimum support. As the correlation between two itemsets is positive, it is called as PAR. Savasere et al. and Wu et al. [4, 5] proposed mining of the NARs $P \Rightarrow \neg Q$, $\neg P \Rightarrow Q$ and $\neg P \Rightarrow \neg Q$. This means that if the itemset 'P' is in the transaction, then the itemset 'Q' will not be in the same transaction. There has been a limited research work for finding NARs among the infrequent itemsets. The ARM algorithms are rarely designed for mining the NARs. However, the current researchers have focused more on the extraction of NARs.

Hong [6] developed a method for mining the fuzzy association rules from the quantitative transactions by integrating the fuzzy set concepts and modified AprioriAll algorithm. The quantitative values in the transactions are transformed into linguistic terms and filtered to find association rules. Liu [7] extended the conventional association rule model to specify different minimum item support values for each item. But, it suffers from the rule missing or explosion problems, as the minimum support for each item is calculated by multiplying the support with a constant percentage.

The existing ARM approach considered the PAR while defining the degree of support and confidence. The accuracy of the NARs is high while handling the huge unstructured data [8]. Hence, it is not easy to find the interesting frequent itemsets[9]. The existing ARM algorithms aimed at the mining of frequent itemsets in a transaction database. However, these algorithms neglect the important itemsets with the minimum support value. These itemsets can generate NARs with the high degree of confidence. Hence, the discovery of both PARs and NARs plays a significant role in building a reliable decision support system. This paper presents an F-PNWAR algorithm for mining the PARs and NARs based on the rank. The input data obtained from the OneDrive is filtered and analyzed based on the data types. The zero-mean normalization is applied to convert the string type data into integer-type data. Then, the items in the dataset are ranked. The weight is assigned to the items based on the rank. The positive and negative weighted itemsets are extracted and rule is generated. The data received from the cloud is analyzed based on the generated rule.

The remaining sections in the paper are arranged in the following way: Section II describes the existing ARM approaches. Section III explains the proposed F-PNWAR algorithm including the mining of the PARs and NARs. Section IV illustrates the experimental analysis of the proposed F-PNWAR algorithm. The conclusion of the proposed F-PNWAR algorithm is described in Section V.

II. RELATED WORKS

Mallik et al. [10] proposed a weighted rule mining approach for ranking the association rules using the interestingness measures. The proposed algorithm generated a few number of frequent itemsets than the existing mining algorithms. Hence, it saved the execution time. Pears et al. [11] automated the weight assignment process by formulating a linear model that obtains the relationships between items. The Valency model is extended by increasing the field of interaction beyond the immediate neighborhoods. The experimental results show that the rules are mined efficiently at a much lower level of support than the basic model. However, the computational cost is high, while recomputing the entire set of weights.

Azadnia et al. [12] developed a new approach by integrating the Genetic Algorithm (GA) with the ARM algorithm such as Traveling Salesman Problem algorithm to find the best travel path. The GA is applied for sequencing the batches to reduce the tardiness. Nithya and Duraiswamy[13] applied average ranking feature selection approach and Fuzzy Weighted ARM (FWARM) classifier to diagnose the medical dataset. The classification accuracy is improved and the number of rules is reduced by ranking the appropriate potential attribute. Hence, the computation time is minimized. Galárraga et al. [14] developed a model for mining the Horn rules on large Resource Description Framework (RDF) knowledge base (KB) and supporting the Open World Assumption (OWA) scenario. The precision and coverage of the proposed model are improved. The rules can be mined quickly than the existing approaches.

Lee et al. [15] proposed a utility-based ARM method for evaluating the association rules by measuring the business benefits of the firms. Vo et al. [16] developed new algorithms for efficient mining of the Frequent Weighted Itemsets (FWI) from the transaction databases. The proposed algorithm achieved a significant reduction in the mining time than the Apriori-based algorithms. Tew et al. [17] concentrated on the behavior-based clustering and study of the interestingness measures. The domain knowledge is crucial to select a proper interestingness measure for a specific task and business objective. Wanaskar et al. [18] presented and investigated a novel approach based on WARM algorithm and text mining. The algorithm is improved by adding semantic knowledge to the results. Better web recommendation performance is achieved. Babashzadeh et al. [19] proposed a new approach for modeling the medical query contexts by mining the semantic-based association rules. The clinical data retrieval performance is improved.

Savasere et al. [20] developed an algorithm for mining NARs for statistically dependent items by integrating the frequent itemsets and domain knowledge. But, this approach required a set of predefined hierarchical classification structure. This makes it difficult to generalize. Morzy[21] introduced the dissociation rule concept. This algorithm maintains the number of generated patterns low. Antonie and Zaiane[22] proposed an algorithm that discovers NARs with high negative correlation between the antecedents and consequents. But, there is a need to continuously update the coefficients and there is no guarantee of all NARs. This paper presented the Fuzzy-based algorithm for extracting the PARs and NARs.

III. PROPOSED F-PNWAR ALGORITHM

The input data is obtained from the OneDrive. OneDrive is a file hosting service that allows the user to store the files. Initially, the pre-filtering method is applied on the input data to remove the data redundancy. The filtered data is analyzed based on the data types. The zero-mean normalization is applied so that all the data are made to slide vertically. Thus, the average value of the data is zero. The string type data is converted into integer-type data. Then, the items in the dataset are ranked. The weight is assigned to the items based on the rank. The F-PNWAR mining algorithm is applied to find the positive and negative weighted itemsets and rule is generated. The data is received from the cloud. Finally, the data analysis is performed based on the generated rule. Fig.1 shows the overall flow diagram of the proposed F-PNWAR algorithm.

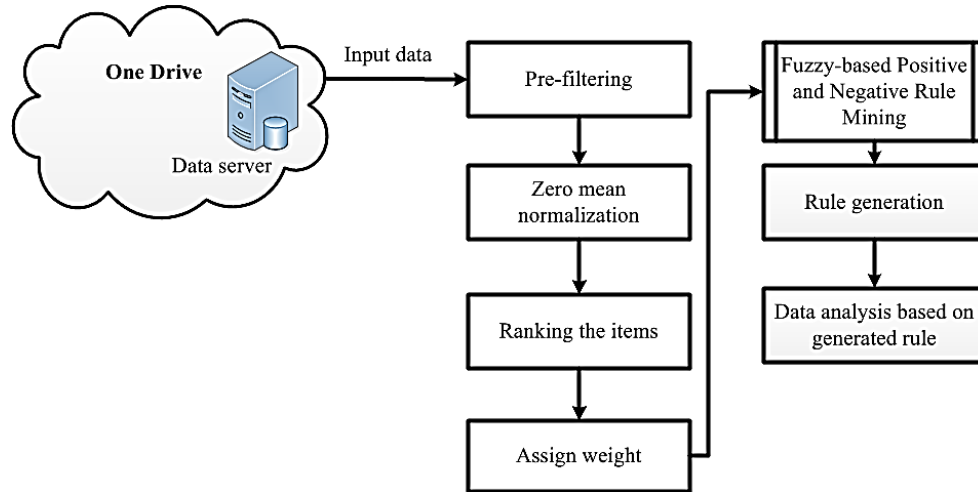


Fig.1 Overall flow diagram of the proposed F-PNWAR algorithm

To find valuable association rules, Shapiro [23] presented interestingness measurement of association rules. If $\sup(P \cup Q) \approx \sup(P) \times \sup(Q)$, $P \Rightarrow Q$ is considered as uninteresting rules. The association rule $P \Rightarrow Q$ is interesting, only if the $\sup(P \cup Q) - \sup(P) \times \sup(Q)$ is not less than a specified minimum interesting value, $\min_interest$. The same method is adopted to measure the interestingness of NARs [24].

An interesting NAR is defined as

$$P \cap Q = \emptyset; \quad (1)$$

$$\sup(P) \geq \min_sup, \sup(Q) \geq \min_sup, \sup(P \cup \neg Q) \geq \min_sup; \quad (2)$$

$$\sup(P \cup \neg Q) - \sup(P) \times \sup(\neg Q) \geq \min_interest \quad (3)$$

The condition $\sup(Q) \geq \min_sup$ should be satisfied, due to the interest in the mining of frequent itemsets in association rules. Similarly, the NAR conditions are defined as $\neg P \Rightarrow Q$ and $\neg P \Rightarrow \neg Q$. If $P \Rightarrow \neg Q$ is a NAR, $P \cup Q$ will be an interesting infrequent itemset. If 'i' is an interesting infrequent itemset, one expression $i = P \cup Q$ will exist. This makes one of the interesting NAR $P \Rightarrow \neg Q$, $\neg P \Rightarrow Q$ and $\neg P \Rightarrow \neg Q$ to hold.

A. Fuzzy Association Rules

The support of an itemset can be computed by finding the fuzzy logic AND of the membership values of the items, for each transaction and adding these values. Let the transaction database be 'D' and itemset $X = \{x_1, x_2, x_3, \dots, x_k\} \subseteq I$. The support of the transaction to the itemset 'X' is defined as

$$\sup(X, t) = \bigcap_{i=1}^k \mu_{x_i}(t) \quad (4)$$

If the fuzzy logic AND is obtained as the result, the support of the itemset from the transaction database is defined as

$$\sup(P) = \sum_{t \in D} \sup(P, t) = \sum_{t \in D} \prod_{i=1}^k \mu_{x_i}(t) \quad (5)$$

B. Positive and Negative Weighted Fuzzy Association Rules

Let us assume μ_x is the membership function of x for all $x \in I$. For each transaction $t \in D$, $\mu_x(t)$ represents the degree that 't' contains the item 'x'.

1) Positive weighted fuzzy association rules

The support of itemset 'P' $\sup(P)$ is considered as the number of transactions in the database that contains the itemset. The weighted minimum support is indicated as $wmin_sup$. Let 'P' and 'Q' be two itemsets. $P \Rightarrow Q$ is the positive weighted fuzzy association rule, if the following conditions are satisfied

1. $P \cap Q = \emptyset$;
2. $wsup(P \cup Q) - wsup(P) \times wsup(Q) \geq \min_interest$
3. $wsup(P \cup Q) = \max_{x \in P \cup Q} w(x) \sum_{t \in D} \prod_{x \in P} \mu_x(t) \prod_{y \in Q} \mu_y(t) \geq wmin_sup$
4. $Conf(P \Rightarrow Q) \geq wmin_sup$

2) Negative weighted fuzzy association rules

Let 'P' and 'Q' are two itemsets, if $P \Rightarrow \neg Q$ is a NAR, both 'P' and 'Q' are frequent. This means that the support value of these itemsets should not be less than the support threshold, while $P \cup Q$ should be infrequent.

The three types of negative fuzzy association rules are defined as follows

$P \Rightarrow \neg Q$ is a negative fuzzy association rule, if the following conditions are satisfied

1. $P \cap Q = \emptyset$;
2. $wsup(P) \geq \min_sup, wsup(Q) \geq \min_sup$;
3. $wsup(P \cup \neg Q) - wsup(P) \times wsup(\neg Q) \geq \min_interest$;
4. $wsup(P \cup \neg Q) \max_{x \in P \cup Q} w(x) \sum_{t \in D} \prod_{x \in P} \mu_x(t) \prod_{y \in Q} (1 - \mu_y(t)) \geq \min_sup$;
5. $Conf(P \Rightarrow \neg Q) = wsup(P \cup \neg Q) / wsup(P) = wsup(P) - wsup(P \cup Q) / wsup(P) \geq \min_conf$.

$\neg P \Rightarrow Q$ is a negative fuzzy association rule, if the following conditions are satisfied.

1. $P \cap Q = \emptyset$;
2. $wsup(P) \geq \min_sup, wsup(Q) \geq \min_sup$;
3. $wsup(\neg P \cup Q) - wsup(\neg P) \times wsup(Q) \geq \min_interest$;
4. $wsup(\neg P \cup Q) \max_{x \in P \cup Q} w(x) \sum_{t \in D} \prod_{x \in P} (1 - \mu_x(t)) \prod_{y \in Q} \mu_y(t) \geq \min_sup$;
5. $Conf(\neg P \Rightarrow Q) = wsup(\neg P \cup Q) / wsup(P) = wsup(Q) - wsup(P \cup Q) / (1 - wsup(P)) \geq \min_conf$.

$\neg P \Rightarrow \neg Q$ is a NAR, if these conditions are satisfied.

1. $P \cap Q = \emptyset$;
2. $wsup(P) \geq \min_sup, wsup(Q) \geq \min_sup$;
3. $wsup(\neg P \cup \neg Q) - wsup(\neg P) \times wsup(\neg Q) \geq \min_interest$;
4. $swsup(\neg P \cup \neg Q) \max_{x \in P \cup Q} w(x) \sum_{t \in D} \prod_{x \in P} (1 - \mu_x(t)) \prod_{y \in Q} (1 - \mu_y(t)) \geq \min_sup$;
5. $Conf(\neg P \Rightarrow \neg Q) = wsup(\neg P \cup \neg Q) / wsup(\neg P) = 1 - wsup(P) - wsup(Q) + wsup(P \cup Q) / (1 - wsup(P)) \geq \min_conf$.

C. Algorithm for mining positive and negative weighted fuzzy association rules

The fuzzy ARM algorithm [4] transforms quantitative value into a fuzzy set with the linguistic terms by using the membership functions. The scalar count of each linguistic term is estimated. The support value of the itemsets is computed. An iterative search method is applied to find the large itemset. Each item uses the linguistic term with the maximum count. The number of fuzzy regions will become identical to the number of original items. This algorithm focuses on the important linguistic terms, the time complexity is minimized. Table I shows the symbols and descriptions used in the mining algorithm[25].

TABLE I. Symbols and Descriptions

Notation	Descriptions
n	Total number of transactions in database
m	Total number of items
D_i	D_i is the i^{th} transaction in D , $1 \leq i \leq n$
I^g	g^{th} item, $1 \leq g \leq m$
R^{gk}	k^{th} region of I^g , $1 \leq k \leq I^g $, where $ I^g $ is the number of fuzzy regions for item I^g
v_i^g	Quantitative value of item I^g in D_i
f_i^g	Fuzzy set converted from v_i^g
f_i^{gk}	Membership value of v_i^g in region R^{gk}
$count^{gk}$	Scalar cardinality of region R^{gk}
$maxcount^g$	Maximum count value among $count^{gk}$ values
$maxR^g$	Fuzzy region of item I^g with $maxcount^g$
C_k	Set of candidate itemsets with 'k' items
L_k	Set of large fuzzy itemsets
WL_k	Set of large weighted fuzzy itemsets
NL_k	Set of non-large weighted fuzzy itemsets

F-PNWAR Algorithm

Input: 'n' number of transactions consisting of customer identity (ID), number of purchased items with their quantities, a set of membership functions, minimum weighted fuzzy support threshold $wmin_sup$, minimum weighted fuzzy confidence threshold $wmin_conf$ and minimum interest threshold $min_interest$. 'D' is the transaction database and D_i is the i^{th} transaction in D.

Output: A set of PAR and NAR.

Step 1: Convert the quantitative value v_i^g of each itemset I^g into a fuzzy set f_i^g using the Fuzzy membership functions.

The f_i^g is denoted as $\left(\frac{f_i^{g1}}{R^{g1}} + \frac{f_i^{g2}}{R^{g2}} + \dots + \frac{f_i^{gl}}{R^{gl}}\right)$. R^{gk} is the k^{th} fuzzy region of the itemset I^g and f_i^{gk} is the fuzzy membership value of the quantitative value in the fuzzy region.

Step 2: Compute the scalar cardinality of each R^{gk} defined as $count^{gk} = \sum_{i=1}^n f_i^{gk}$.

Step 3: Find the max $count^g = MAX_{k=1}^l(count^{gk})$, such that $1 \leq g \leq m$. Let $maxR^g$ be the region with the maximum count for the item. The region with the maximum count value represents the fuzzy characteristic of the item.

Step 4: Calculate the fuzzy support of $maxR^g$ $wfsup(maxR^g) = max_count^g$. Verify if the fuzzy support of the region is greater than or equal to the predefined minimum weighted support threshold sup , for $g = 1$ to m . If the maximum count value is greater than or equal to sup , arrange $maxR^g$ in the large 1-itemsets L_1 . $L_1 = \{maxR^g | max_count^g \geq sup, 1 \leq g \leq m\}$.

Step 5: If the large itemset is null, the algorithm is stopped. Otherwise, move to the next step.

Step 6: Set $k=1$, where 'k' represents the number of items in the current large itemsets.

Step 7: Create the candidate set C_{k+1} from the large itemset.

Step 8: For the newly generated (k+1) itemsets with the items $(s_1, s_2, \dots, s_{k+1})$ in the candidate set,

- Calculate the fuzzy value f_i^s for the item as $f_i^s = MAX_{i=1}^n(f_i^s)$ and $f_i^s = f_i^{s1} \wedge f_i^{s2} \wedge \dots \wedge f_i^{s_{k+1}}$.
- Calculate the scalar cardinality of the item as $count^s = \sum_{i=1}^c f_i^s$.
- If the count of the item is not less than sup , then
 - Set the item 's' in L_{k+1} .
 - If the weighted fuzzy support of the item 's' $\sum_{i=1}^{k+1} w(s_i)count^s \geq wmin_sup$, set 's' in WL_{k+1} ;
 - Else set 's' in S_{k+1} ;

Step 9: If L_{k+1} is null, then perform the next step. Otherwise, set $k = k + 1$ and repeat steps 7 to 9.

Step 10: $WL = \cup_k WL_k$; $S = \cup_k S_k$;

Step 11: For each itemset 'i' in WL Do {

Step 12: For any $P \cup Q = i$ and $P \cap Q = \emptyset$ Do {

Step 13: If $wsup(P \cup Q) - wsup(P) \times wsup(Q) \geq min_interest \wedge$

Step 14: $wsup(P \cup Q) \geq wmin_sup \wedge Conf(P \Rightarrow Q) \geq wmin_conf$

Step 15: Then $PAR = PAR \cup \{P \Rightarrow Q\}$;

Step 16: }

Step 17: }

Step 18: For each itemset 'j' in 'S' Do {

Step 19: For any $P \cup Q = j$ and $P \cap Q = \emptyset$ Do {

Step 20: If $wsup(P) \geq wmin_sup \wedge wsup(Q) \geq wmin_sup \wedge wsup(P \cup \neg Q) \geq wmin_sup \wedge$

Step 21: $wsup(P \cup \neg Q) - wsup(P) \times wsup(\neg Q) \geq min_interest \wedge$

Step 22: $wsup(P \cup \neg Q) \geq wmin_sup \wedge Conf(P \Rightarrow \neg Q) \geq wmin_conf$

Step 23: Then $NAR = NAR \cup \{P \Rightarrow \neg Q\}$;

Step 24: If $wsup(P) \geq wmin_sup \wedge wsup(Q) \geq wmin_sup \wedge wsup(\neg P \cup Q) \geq wmin_sup \wedge$

Step 25: $wsup(\neg P \cup Q) - wsup(\neg P) \times wsup(Q) \geq min_interest \wedge$

Step 26: $wsup(\neg P \cup Q) \geq wmin_sup \wedge Conf(\neg P \Rightarrow Q) \geq wmin_conf$

Step 27: Then $NAR = NAR \cup \{\neg P \Rightarrow Q\}$;

Step 28: If $wsup(P) \geq wmin_sup \wedge wsup(Q) \geq wmin_sup \wedge wsup(\neg P \cup \neg Q) \geq wmin_sup \wedge$

Step 29: $wsup(\neg P \cup \neg Q) - wsup(\neg P) \times wsup(\neg Q) \geq \min_interest \wedge$
 Step 30: $wsup(\neg P \cup \neg Q) \geq wmin_sup \wedge Conf(\neg P \Rightarrow \neg Q) \geq wmin_conf$
 Step 31: Then $NAR = NAR \cup \{\neg P \Rightarrow \neg Q\}$;
 Step 32: }
 Step 33: }

IV. PERFORMANCE ANALYSIS

The performance of the proposed work is evaluated by applying it in the groceries dataset [26] on a system with Intel(R) Core i3-3220 x64-based processor and 8 GB capacity. The proposed E-FWARM algorithm is compared with the WARM and FWARM [27] and traditional K-means and Adaptive K-means algorithms [28]. Fig.2 shows the frequent item rate analysis of the proposed F-PNWAR and existing WARM, FWARM and E-FWARM algorithms. The proposed F-PNWAR algorithm achieved maximum frequent items than the WARM, FWARM and E-FWARM algorithms. There is a linear decrease in the number of frequent items with respect to the increase in the support value. Fig.3 illustrates the association rule rate analysis of the proposed F-PNWAR and existing WARM, FWARM and E-FWARM algorithms. The proposed F-PNWAR algorithm extracts more number of association rules than the existing WARM, FWARM and E-FWARM algorithms. There is a gradual decrease in the number of association rule with respect to the increase in the weighted confidence value.

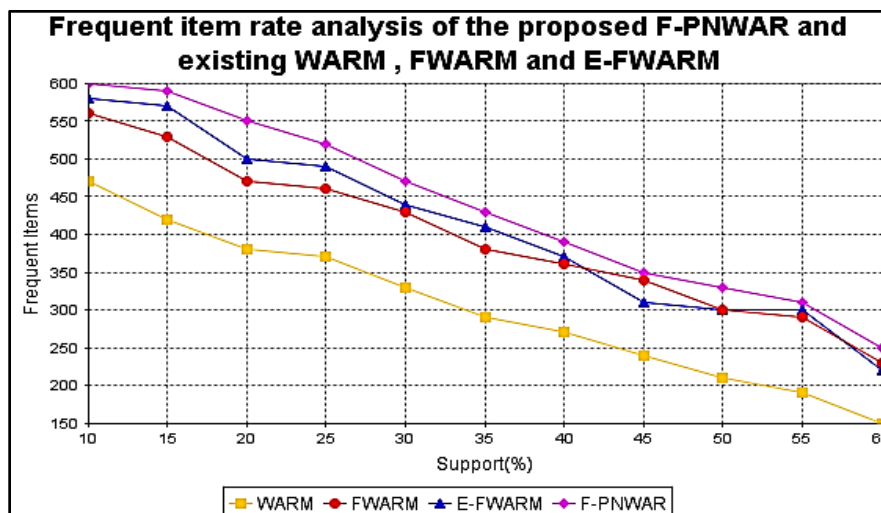


Fig.2 Frequent item rate analysis of the proposed F-PNWAR and existing WARM, FWARM and E-FWARM algorithms

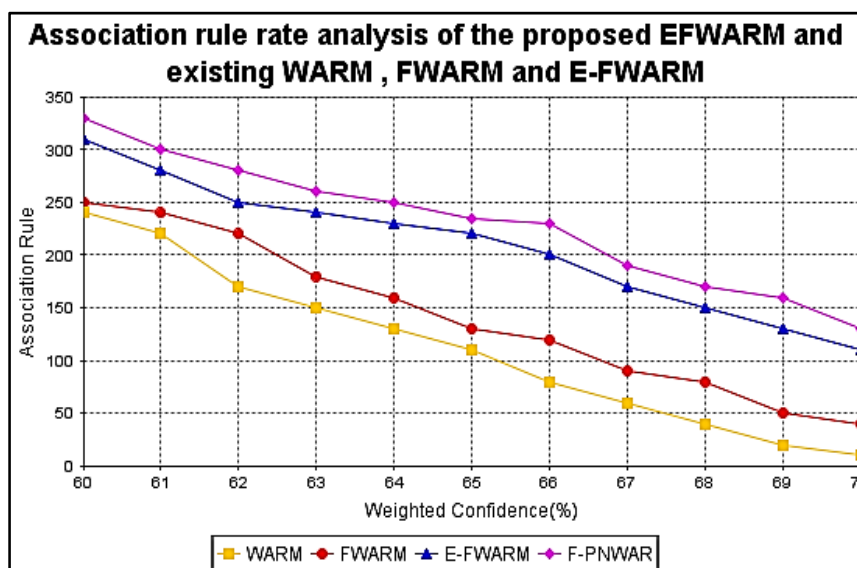


Fig.3 Association rule rate analysis of the proposed F-PNWAR and existing WARM, FWARM and E-FWARM algorithms

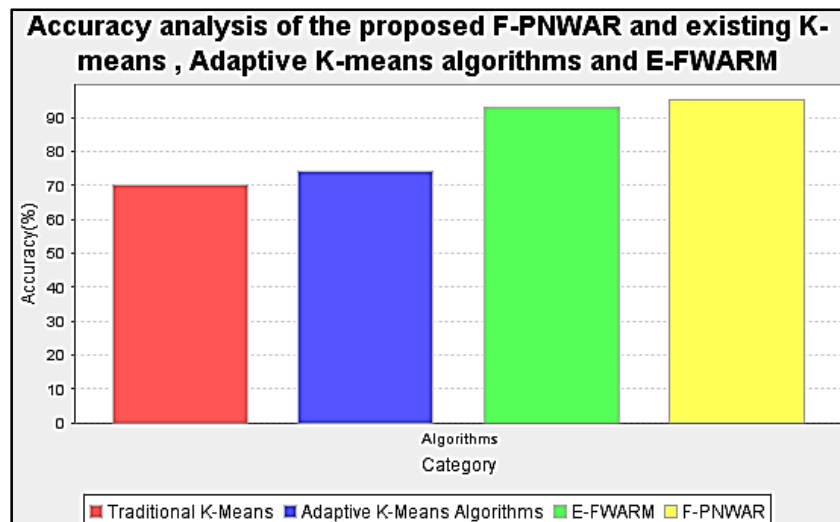


Fig.4 Accuracy analysis of the proposed F-PNWAR and adaptive K-means and traditional K-means algorithms

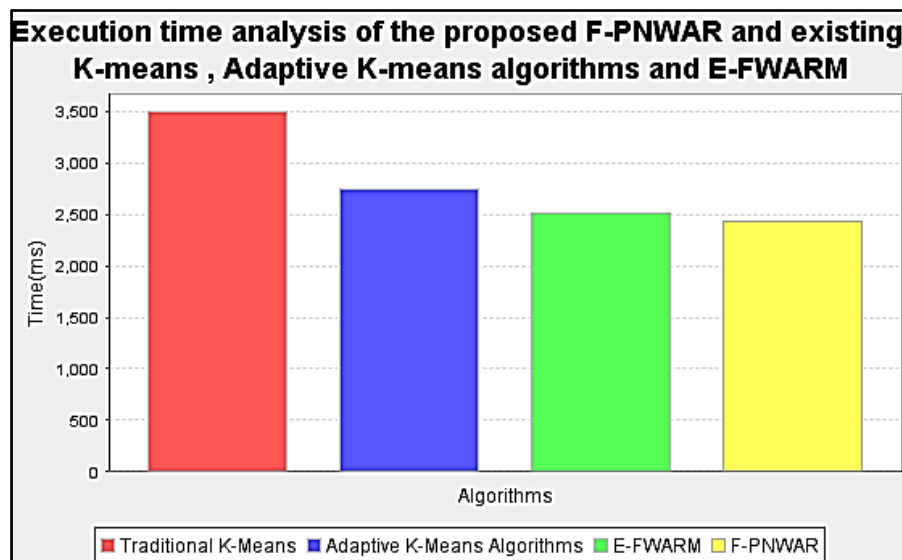


Fig.5 Execution time analysis of the proposed F-PNWAR and adaptive K-means and traditional K-means algorithms

Fig.4 depicts the accuracy analysis of the proposed F-PNWAR and E-FWARM, adaptive K-means and traditional K-means algorithms. The proposed F-PNWAR algorithm yields maximum accuracy of about 97%, while the E-FWARM algorithm yields accuracy of about 93%, traditional K-means and Adaptive K-means algorithms yield accuracy of about 70% and 75% respectively. Fig.5 presents the execution time analysis of the proposed F-PNWAR and E-FWARM, adaptive K-means and traditional K-means algorithms. The proposed F-PNWAR algorithm requires minimum execution time than the E-FWARM, adaptive K-means and traditional K-means algorithms.

V. CONCLUSION

Traditional rule mining methods are accurate, but have very hard and fragile operations. Fuzzy-based mining algorithms provide a robust and efficient approach to explore large search space. This paper presented a Fuzzy-based algorithm for mining both the PARs and NARs. The proposed algorithm efficiently generates NARs, along with the PARs. This algorithm focuses on the significant linguistic terms, the time complexity is minimized. From the performance analysis, it is observed that the proposed F-PNWAR algorithm yields maximum frequency item rate, association rule rate and accuracy than the existing WARM, FWARM and E-FWARM algorithms. The execution time of the proposed F-PNWAR algorithm is lesser than the E-FWARM, adaptive and traditional K-means algorithms.

REFERENCES

- [1] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *Acm sigmod record*, 1993, pp. 207-216.
- [2] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th int. conf. very large data bases, VLDB*, 1994, pp. 487-499.
- [3] N. Suryawansi, S. Jain, and A. Jain, "A review of negative and positive association rule mining with multiple constraint and correlation factor," *Int J Emerg Technol Adv Eng (IJETA)*, vol. 2, pp. 778-781, 2012.
- [4] A. Savasere, E. Omiecinski, and S. Navathe, "Mining for strong negative associations in a large database of customer transactions," in *14th International Conference on Data Engineering*, 1998. *Proceedings.*, 1998, pp. 494-502.
- [5] X. Wu, C. Zhang, and S. Zhang, "Mining both positive and negative association rules," in *19th International Conference on Machine Learning*, 2002.
- [6] T.-P. Hong, K.-Y. Lin, and S.-L. Wang, "Fuzzy data mining for interesting generalized association rules," *Fuzzy sets and systems*, vol. 138, pp. 255-269, 2003.
- [7] B. Liu, W. Hsu, and Y. Ma, "Mining association rules with multiple minimum supports," in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1999, pp. 337-341.
- [8] X. Liu, B. Lang, W. Yu, J. Luo, and L. Huang, "AUDR: an advanced unstructured data repository," in *6th International Conference on Pervasive Computing and Applications (ICPCA)*, 2011, pp. 462-469.
- [9] A. R. Islam and T.-S. Chung, "An improved frequent pattern tree based association rule mining technique," in *International Conference on Information Science and Applications (ICISA)*, 2011, pp. 1-8.
- [10] S. Mallik, A. Mukhopadhyay, and U. Maulik, "RANWAR: rank-based weighted association rule mining from gene expression and methylation data," *IEEE transactions on nanobioscience*, vol. 14, pp. 59-66, 2015.
- [11] R. Pears, Y. S. Koh, G. Dobbie, and W. Yeap, "Weighted association rule mining via a graph based connectivity model," *Information Sciences*, vol. 218, pp. 61-84, 2013.
- [12] A. H. Azadnia, S. Taheri, P. Ghadimi, M. Z. Mat Saman, and K. Y. Wong, "Order batching in warehouses by minimizing total tardiness: a hybrid approach of weighted association rule mining and genetic algorithms," *The Scientific World Journal*, vol. 2013, 2013.
- [13] N. Nithya and K. Duraiswamy, "Correlated gain ratio based fuzzy weighted association rule mining classifier for diagnosis health care data," *Journal of Intelligent & Fuzzy Systems*, vol. 29, pp. 1453-1464, 2015.
- [14] L. A. Galárraga, C. Teflioudi, K. Hose, and F. Suchanek, "AMIE: association rule mining under incomplete evidence in ontological knowledge bases," in *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 413-422.
- [15] D. Lee, S.-H. Park, and S. Moon, "Utility-based association rule mining: A marketing solution for cross-selling," *Expert Systems with applications*, vol. 40, pp. 2715-2725, 2013.
- [16] B. Vo, F. Coenen, and B. Le, "A new method for mining Frequent Weighted Itemsets based on WIT-trees," *Expert Systems with Applications*, vol. 40, pp. 1256-1264, 2013.
- [17] C. Tew, C. Giraud-Carrier, K. Tanner, and S. Burton, "Behavior-based clustering and analysis of interestingness measures for association rule mining," *Data Mining and Knowledge Discovery*, vol. 28, pp. 1004-1045, 2014.
- [18] U. Wanaskar, S. Vij, and D. Mukhopadhyay, "A hybrid web recommendation system based on the improved association rule mining algorithm," *arXiv preprint arXiv:1311.7204*, 2013.
- [19] A. Babashzadeh, M. Daoud, and J. Huang, "Using semantic-based association rule mining for improving clinical text retrieval," in *International Conference on Health Information Science*, 2013, pp. 186-197.
- [20] A. Savasere, E. Omiecinski, and S. Navathe, "Mining for strong negative rules for statistically dependent items," in *Proceedings of the IEEE International Conference on Data Mining (ICDM'02)*, 2002, pp. 442-449.
- [21] M. Morzy, "Efficient mining of dissociation rules," in *International Conference on Data Warehousing and Knowledge Discovery*, 2006, pp. 228-237.
- [22] M.-L. Antonie and O. R. Zaïane, "Mining positive and negative association rules: an approach for confined rules," in *European Conference on Principles of Data Mining and Knowledge Discovery*, 2004, pp. 27-38.
- [23] G. Piatetsky-Shapiro, "Discovery, analysis and presentation of strong rules," *Knowledge discovery in databases*, pp. 229-248, 1991.
- [24] W. Ouyang, "Mining Positive and Negative Weighted Fuzzy Association Rules in Large Transaction Databases," in *Second International Symposium on Knowledge Acquisition and Modeling*, 2009. *KAM'09.*, 2009, pp. 269-272.
- [25] W. Ouyang, "Mining Positive and Negative Weighted Fuzzy Association Rules in Large Transaction Databases," in *Knowledge Acquisition and Modeling*, 2009. *KAM'09. Second International Symposium on*, 2009, pp. 269-272.
- [26] S. Marafi. (2014). Market Basket Analysis with R. Available: <http://www.salemmarafi.com/code/market-basket-analysis-with-r/>
- [27] V. Vidya, "AN EFFICIENT FUZZY WEIGHTED ASSOCIATION RULE MINING WITH ENHANCED HITS ALGORITHM," 2006.
- [28] Deepti Ambaselkar and D. A. B. Bagwan, "Adaptive K-means clustering for association rule mining from gene expression data," *International Journal of Engineering and Technical Research (IJETR)*, vol. 5, pp. 25-28, July 2016.

AUTHOR PROFILE

K.Mangayarkkarasi is currently working towards the PhD degree at the Department of Computer Science Bharatiar University, Coimbatore,Tamil Nadu. She has received her Post Graduate degree M.Sc in A.V.C. College(autonomous) ,Mannampandal affiliated to Bharathidasan University, Tiruchirappalli, TamilNadu. Her research interests include Data mining , parallel processing,Networks and cloud computing.She is working as Assistant Professor in the Department of Computer Science, Dharmapuram Gnanambigai Government arts college for women, Mayiladuthurai, Tamil Nadu since August 2000.

M.Chidambaram is working as Assistant Professor in the department of Computer Science, Sarafoji College, Thanjavur, Tamil Nadu. His research interests include Networks, Grid computing, Parallel processing and Cloud computing.