ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 8(33), DOI: 10.17485/ijst/2015/v8i33/80081, December 2015 ISSN (Online) : 0974-5645

A Comprehensive Analysis on Associative Classification in Medical Datasets

D. Sasirekha[1*] and A. Punitha[2]

[1]Department of Computer Science, Bharathiyar University, Coimbatore - 641046, Tamil Nadu, India; sasirekhadheeraj@gmail.com
[2]Department of Computer Applications, Queen Mary's College, Chennai - 600004, Tamil Nadu, India; apunithaganesh@yahoo.com

Abstract

Association rule mining along with classification technique is capable of finding informative patterns from large data sets. Output of this technique is of the form if-then which is easy to understand for the end users and also for prediction. Termed as Associative Classification, it is having wide application on medical domain in diagnosing diseases and to analyze medical datasets which are unstructured, heterogeneous, incomprehensible and voluminous. Analysing these data allows physicians to predict the diseases as well as to take vital decisions. This paper presents a detailed study on associative classification and the phases of associative classification procedure. Several associative classification methods viz., CPAR, MMAC, CAR, CBA, CMAR, etc. along with their merits and demerits are also presented in a lucid manner.

Keywords: Association, Associative Classification, Classification, Knowledge Discovery, Medical Datasets

1. Introduction

Medical data contains huge volume of information in an unstructured format. Data mining discovers insightful, interesting and novel patterns which are descriptive, understandable and predicative from large amount of data[1]. Data mining includes core techniques such as association, classification, clustering, prediction, combination, etc., that are used for different type of mining and data recovery operations. Association is a technique used to know the correlation between two or more items. Association rule mining defines all the rules that exist in the data set that satisfies some amount of support and confidence. The best application of association rule mining is market basket analysis which investigates the purchase behaviour of customers[2]. Association rule mining tries to find the group of items that are frequently purchased in the presence of other items in the shopping cart. This group of items is said to be frequent itemset.

On the other hand, classification is also one of the most important techniques in data mining. Constructing fast and accurate classifier for large data is an important task in knowledge discovery. There are several classification techniques such as Naive Bayesian classifier, C4.5, PART, Prism, IREP, etc. Normally, classification techniques produce small subsets of rules and there is a great chance that detailed rules that are vital in decision making may get skipped[3].

In association rule mining, the target of knowledge discovery is not predetermined, whereas classification rule mining works by fixing one and only one predetermined target[4]. In order to make it convenient and to save time in knowledge discovery process, these two techniques can be integrated. A new approach that integrates association rule mining and classification is said to be Association Classification[5]. This fusion is done by considering with minimum subset of association rules called Class Association Rules (CARs). This new approach yields higher accuracy[5]. There exist many associative classification techniques.

* Author for correspondence

X. Yin and Jiawei Han[6] presented Classification based on Predictive Association Rules (CPAR), which combines the advantages of both associative classification and rule based classification. CPAR adopts greedy method to generate rules from large training data. CPAR generates and tests more rules than ordinary rule-based classification. Wenmin Li et al.[7] proposed accurate and efficient classification based on multiple class association rules. The method extends Frequent Pattern (FP) growth mining method, constructs FP-tree and mines large databases in an efficient way. CPAR is consistent and highly effective.

Experimental results carried out by many studies show that classification based association rule mining is highly effective and constructs strong predictive and accurate classifiers than traditional classification methods[6,7].

2. Association Rule Mining

Association Rule Mining (ARM) is a tool used to find interesting associations and correlations among large set of data items which is a strong mechanism used in market basket analysis. Association rules shows attribute value conditions that occur frequently together in a given data set. Association rule mining finds group of items that are frequently sold together along with the presence of other items in customers shopping cart. Let D be a database which stores transaction such that $D = \{t_1, t_2, t_3 \ldots t_n\}$ and I be the set of items. Set of items, called as itemset and an itemset with k-items is said to be k-itemset. The support of an itemset is the number of transactions in D where itemset occurs as subset and it is denoted by σ. Itemset is said to be frequent or large if support value is greater than user specified minimum support[8]. Association rule is of the form X→Y, where X and Y are disjoint itemsets. The support for the rule X→Y, is the probability of transactions containing both X and Y. The confidence of the association rule X→Y is the conditional probability that atransaction contain Y, given that it contains X. An association rule is frequent and strong if its support is greater than minimum support and its confidence is greater than minimum confidence respectively,

$$\text{support}(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N}$$

$$\text{confidence}(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}$$

The ARM process works in two phases, with finding all frequent itemsets having minimum support and confidence in the first phase, and generating strong association rules from the frequent itemsets in the second phase. Apriori is the most classic association rule mining algorithm which achieves good level of performance. However, ARM is having certain drawbacks such as

- Holding large number of candidate itemsets which is costly.
- Scanning the database multiple times to check number of candidate itemsets, which is again costly during runtime.
- Items with high confidence and low support may get ignored.
- Setting appropriate parameter for mining algorithm.
- ARM discovers too many rules.
- Discovery of poorly understandable rules.

3. Classification

It is the process of predicting a class label for a given unlabeled point. It involves examining the features of new objects and trying to assign one predefined set of classes. Building a model from classified objects in order to classify previously unseen objects as accurately as possible is the goal of classification[9]. Devising a procedure for classification in which exact classes are known in advance is termed as pattern recognition or supervised learning. There are many cases of unsupervised learning in which classification class is not known in advance. Three main standards of classification techniques include, statistical based classification, machine learning based classification and neural network based classification[16].

All the above classification approaches work by adopting divide-and-conquer, separate-and-conquer and by statistical approaches. Various classification algorithms have been developed such as PART, RIPPER, Prism, etc. A traditional classification problem can be defined as follows: Let D denotes the domain of all possible training instances, let Y denotes list of class labels and H denotes set of classifiers. The objective is to find a classifier h such that $h \in H$ which minimizes the probability that

$h(d) = y$ for each test case (d, y)

Apart from various advantages, classification techniques contain certain drawbacks also. Few classification algorithms generate empty branches or insignificant branches[10]. Few algorithms require long searching time and more memory. Few classification algorithms may not perform well with very few records as they need large number of data to achieve good results. Moreover, majority of the classification algorithms are costly.

4. Associative Classification

Associative classification is a type of classification approach that adopts association rule mining and classification rule mining to build classification models[11]. It is a special kind of ARM where target attribute is considered in rule's right hand side. In associative classification, the entire data set is divided into two parts with 70% of data used for training and 30% for testing the accuracy of the classifier[11]. Associative classification contains various advantages such as

- Training the data is very efficient regardless of the size of training set.
- High dimensionality training sets can be handled easily and assumptions on dependent and independent attributes are not needed.
- Classification is done in quick manner.
- The generated rules are easy to understand by humans.

Associative classification works in three steps,

- Generate a set of association rules from training set with certain support and confidence threshold as candidate rules.
- Prune the discovered rules that may introduce over fitting.
- Perform classification to make prediction for test data and measure the accuracy of the classifier.

Associative classification uses mainly minimum support and minimum confidence[13]. The data in associative classification can be represented in three ways such as horizontal, vertical and set theory data representation. Horizontal representation is like a format adopted in ARM. In this format, a database consists of transactions, with each transaction being identified with transaction identifier and list of items in the transaction[12]. The data is represented in rows, with each row representing a transaction. Classification Based Association (CBA) uses this type of representation.

Vertical data representation transforms training data into a table form which contains transaction identifier and simple intersections are employed to discover frequent itemsets to produce CARs[12]. Associative classification algorithms such as Multi-Class Multi-Label Associative Classification (MMAC) and Multi-Class Classification based Association Rules (MCAR) use vertical data representation. During rule discovery, frequent items of N-1 size discovers the possible frequent items of size N. Rough set theory data representation is based on discarding redundant attributes from training data sets and selects reducts of attribute-value pairs that can represent the complete training data set in a decision table context[14]. ROSETTA selects reducts of attribute-value pairs. Reduct is a subset of attribute values and class attribute that represents the whole table.

4.1 Associative Classification Methods

4.1.1 CBA

Classification Based on Association Rules (CBA) selects a sub set of rules to form the classifier model. CBA works in two phases viz., rule generation and classifier building. The perfect classifier is built using Class Association Rules (CARs). CBA is a simple algorithm which prefers database to be inside the main memory. There are several associative classification algorithms that use CBA properties to find frequent itemsets and to generate class association rules.

Zhitong Su[24] adopted CBA for discovering informative association rules. To select and rank small subset of high quality rules, entropy based associative classifier is built. This classifies uses information gain parameter to select and rank the rules. Shekhawat and Dhande[25] adopted CBA methodology and Neural Network based Associative Classification (NNAC) system is presented. NNAC improves the efficiency of the classifier.

4.1.2 CMAR

CMAR stands for classification based on multiple association rules. CMAR performs weight analysis using multiple strong association rules. The weight analysis is performed with Chi-Square (χ^2) method. This determines the strength of the association rule under both support and class distribution[28]. CMAR also works in two phases with rule generation in first phase and class distribution in second phase.

In order to improve efficiency and accuracy, Classification Rule Tree (CR-Tree) data structure is used. The new data structure is an extension of Frequent

Pattern Growth (FP-tree), and it compactly stores and retrieves large number of rules for classification. In order to speed up the classification process, CMAR uses variant of FP-growth algorithm, which is faster than apriori-like methods[28].

4.1.3 CARGBA

It stands for classification based on association rule generated in a bidirectional approach. CARGBA works in two phases. In the first step, it generates a set of high confidence rules of smaller length[27]. Then this set is increased by adding high confidence rules of higher length. In the second phase, specific rules are generated without any support and pruning. This result is a better mixture of class association rules. Since all the rules are not used in classification, the second phase of CARGBA builds a classifier called CARGBA classifier.

4.1.4 CPAR

CPAR stands for classification based on predictive association rules. CPAR uses expected accuracy measure to evaluate each rule to avoid over fitting problem. CPAR follows the basic idea of First Order Inductive Learner (FOIL) algorithm in rule generation[29]. It seeks for the best rule condition that can bring maximum gain among available dataset. Once the condition is identified, weights of positive examples associated with it will be deteriorated by a multiplying factor.

During classifier building process, a common error occurs, called misclassification penalty. This error can be minimized by using modified CPAR (M-CPAR)[26]. CPAR combines the advantages of associative classification and traditional rule based classification. CPAR generates rules directly from training data by using greedy algorithm.

4.1.5 CARPT

Classification Algorithm based on association rules. This method is more preferable if frequent rules need to be generated directly with generating frequent itemsets. Here, the data of any form is converted into two-dimensional array. The horizontal position of two dimensional array represents item number and property types. The vertical position represents transaction number[30].

CARPT uses Trie-tree to build accurate classifier. Trie-tree influences two properties for generating frequent rules. First property states that "if a sub tree takes a non-frequent bucket for root nodes, then all the buckets of the sub tree are not frequent". Second property states that "if $<i_1, i_2, \ldots, i_n>$ is the frequent item set. Then there cannot be frequent item set which contains two or more items taken for a prefix". Usage of trie-tree reduces the cost of the query time thereby improving efficiency[30].

4.1.6 HMAC

Hierarchical Multi-label Associative Classification (HMAC) is another classification technique performed on hierarchically arranged data. Sawinee Sangsuriyun et al.[31] presented a paper on HMAC using negative rules. Here, the Rules of Positive Set (RPS) are generated first. The generated rules are pruned so that rules without minimum support and confidence are removed.

Pruning is done using two methods such as Pearson's correlation coefficient pruning and redundant rules pruning. For a given set of coefficient, φ is used to select rules with user specified threshold. HMAC classifier is built, which selects a sub set of high quality rules. HMAC classifier also predicts the class label for the object by analyzing the subset rules[31]. The efficiency and set similarity of this algorithm is tested with F-measure and Jaccard's coefficient.

4.1.7 ACCF

ACCF stands for Associative Classification based on Closed Frequent itemsets. It was proposed by Li et al.[32] and it works in two phases viz., rule generation and classifier construction. ACCF is the extended version of CHARM algorithm. ACCF discovers a set of Closed Frequent Itemsets (CFIs) along with their tid (transaction identifier) sets and class labels.

The support and confidence for each rule is calculated by using tidsets for both items and corresponding class labels. A set of CARs is generated by evaluating the rules and selecting the rules with minimum support and minimum confidence. The generated CARs are ranked according to their support, confidence and rule length. ACCF starts with first ranked rule and applies to tidset for classification[32].

4.1.8 GARC

Gain based Association Rule mining (GARC) uses extended association rule mining to construct a classifier[33]. It works in three steps. First it discovers candidate itemsets by applying information gain measure. The candidate itemsets with best-split attribute value are generated. By this way, redundant and unwanted candidate itemsets are removed. In the second step, it combines rules generation and frequent itemset generation process. In the third step, it produces compact set of rules that are short and easy to understand.

The classifier built using GARC can be applied to discrete and continuous data. The performance of GARC can be enhanced by applying pruning strategies[33]. GARC classifier produces better results in terms of efficiency, accuracy and understandability.

4.1.9 ACN

Associative Classifier with Negative rules (ACN) is a different associative classification method that mines relatively large set of negative association rules first. Then positive and negative association rules are used to build a classifier. The advantage of generating negative rules is that, good number of negative rules can be replaced with weak positive rules. Hence, there will be a reduction in number of inaccurate positive results in the final classifier thereby increasing the accuracy[34].

Generally computational cost of mining negative association rules is high and it will become still higher if positive association rules are also need to be mined. ACN uses the variant of Apriori algorithm to address the cost issue and tries to generate both forms of rules with low overhead. ACN works in two phases. First it designs an ACN rule generator and then it builds classifier. ACN uses several pruning strategies to cut down the number of rules generated[34]. ACN is highly consistent, accurate and effective on various kinds of databases.

4.1.10 Comparison of Various Associative Classification Methods

Table 1. Comparison of associative classification methods

Method	Advantages	Disadvantages
CMAR	• It finds frequent patterns and generates association rules in one step. • Since CR-tree data structure is used, both accuracy and efficiency is improved. • By pruning process, CMAR selects only high quality rules for classification. • CMAR is superior to C4.5 and CBA interms of accuracy and it is also scalable.	• Since it searches for only high quality rules, it is slower. • Performing weighted analysis adds substantial computational load to the algorithm.
CPAR	• Redundancy is avoided since each is generated by comparing with the set of already generated rules. • It generates high quality predictive rules directly from dataset.	• CPAR is more complex to understand as well as to implement. • Usage of greedy algorithm to train the dataset adds additional computational overhead to the algorithm.
CBA	• Simple algorithm, finds valuable rules. • Capable of handling data in table form as well as in transaction form. • It does not require the whole dataset to be fetched into main memory.	• Training the dataset often generates huge set of rules leading to redundancy. • Classification is based on parameter called confidence, which can be biased sometimes.
CARG-BA	• Capable of handling various kinds of databases, and the method is highly consistent and effective. • Rules are produced without any support pruning.	• Multiple scans are needed to generate rules which makes the algorithm slow. • Overall accuracy of the algorithms is less.

CARPT	• Storing data in vertical data format reduces number of scans to the database. • Time and space is saved effectively.	• The data should be in two dimensional array of vertical data format.
GARC	• It gets rid of many candidate itemsets in the training phase itself.	• It removes frequent items. • Only few rules are generated.
ACCF	• Handles both continuous and discrete datasets. • No redundant rules are generated.	• Taking additional time for learning the frequent itemset.
ACN	• Usage of coverage pruning algorithm discards rules that does not cover atleast one training case. • It is time efficient and achieves better accuracy. • Too many negative rules are generated with low cost and these negative rules control the inclusion of inaccurate positive rules.	• It calculates support and confidence for negative rules which is unnecessary and adds computation overhead.
HMAC	• Improved accuracy is achieved since negative redundant rule pruning is used. • Capable of handling complex data.	• Rule generation is based on combination of two scoring methods viz., F-measure and Jaccard's measure which leads to additional computational overhead.

4.2 Rule Ranking Procedures

The generated class association rules are well-organized using support and confidence, and rules with high support and confidence is ranked on top. There are several methods available to rank the generated class association rules. CBA algorithm uses support, confidence and antecedent length to arrange the discovered rules[15]. Rules are arranged based on confidence. If two or more rules are having same confidence, then they are arranged by their support. If the support values are also same, then their antecedent lengths are used for arranging. If the antecedent lengths are also identical, then randomly rules are arranged[15]. Other algorithms such as CARGBA, Associative Classification based on Closed Frequent itemsets (ACCF) are also using this approach for arranging rules. It is also possible to order the generated rules without the above three parameters. Instead, a new parameter is used which is called as class distribution, which reflects the number of items in a class, is used to arrange the generated rules.

Baralis and Torino[11] proposed lazy associative classification algorithms which uses antecedent lengths, which is the opposite of CBA rule ranking practise. Here the rules which hold more attributes are ranked on top and these rules are called as specific rules. Information gain, another parameter used to rank the generated rules. Information gain is a mathematical measure that represents how well a given attribute is used for classification. To calculate information gain of an attribute, it is important to calculate entropy.

$$D=\sum -P_k \log_2 P_k$$

Where P_k is the probability of class k belongs to D. The information gain of an attribute is calculated by

$$G = Entropy(D) - \sum\left(\left(|D_a| / |D|\right) * Entropy(D_a)\right)$$

Where D_a is the subset of an attribute which has the value a, modulus of D_a is the number of data cases in D_a and modulus of D is the number of data cases in D.

4.3 Building the Classifier and Pruning

Classifier is the set of rules that are built from training set. Since a classifier builds many rules, result is delayed in classification process. Therefore it is necessary to remove useless and redundant classifier[16]. And also this removal will increase the speed of the classification process.

Database coverage is a pruning method which checks for the rules covering at least an object. If so, it is added in the classifier and its corresponding training object is deleted from the training dataset. This procedure is repeated until all training objects are deleted or all the rules are examined. If the training data set is not empty at the end of pruning, all the remaining cases are put under default class rule. This procedure is adopted by several associative classification algorithms such as CBA, CMAR, CAAR, ACN and ACCF.

High precedence method iterates over the ranked rules starting with highest rank. All the training cases covered are discarded and the rule is inserted into the classifier. Rule that does not cover training case are removed. High classification pruning method selects a rule that partially covers a training case and has a common class to that of training case[17]. Based on this, a training case is inserted or deleted from the classifier. The whole procedure is repeated until training data becomes empty.

Pruning can also be done using mathematical calculations. The decision to replace a sub-tree with a leaf node or to keep as such is done by calculating pessimistic error estimation measure over training data set. CBA uses this approach. Chi-Square test can also be used to decide a rule relevancy for its inclusion in class association rules set.

CMAR uses this approach for pruning the rules set. Redundant rules are removed by using long rules pruning which eliminates long rules with confidence values larger than their subset rules[18]. The rules that may lead to misclassification on the training data set can be removed by using lazy method. In this method, every rule that covers a training data case and have both the same class, the rule is inserted in the primary rule set else deleted. Conflicting rules, which may cause the classifier to classify the test case, must be removed.

5. Associative Classification in Medical Datasets

Medical data mining is an emerging technology in medical field that solves the traditional medical problems such as congestion, long wait time and delayed patient case. This technique helps physicians to make accurate diagnosis of diseases. Generally medical data set are widely distributed, in unstructured format, heterogeneous in nature and in huge volume. These data need to be organized in a form which is understandable. Advantage of using data mining in medical domain is to improve the accuracy of the output with huge amount of data. Medical data mining has great potential for exploring hidden patterns among medical data sets. Among various data mining techniques, associative classification technique is having a very good application on medical domain. Most of the attributes in medical data sets are often associated with quantitative domains such as Body Mass Index (BMI), age, Blood Pressure (BP), etc.

Rafel Rak et al.[19] proposed multi-label associative classification on medical documents from Medical Literature Analysis and Retrieval System Online (MEDLINE). This method performs classification on medical repository MEDLINE. It is based on associative classification which considers multi label features. Akhil Jabbar et al.[20] proposed heart disease prediction system using associative classification and genetic algorithm. This method performs high level rules that are accurate and comprehensible, and contains high interesting value. K. Ruth Ramya et al.[21] proposed class based approach for medical classification of chest pain. The class label is used in classification to minimize the searching space. This method also synchronizes the rule generation and classifier building phase.

Genetic algorithms are adaptive procedures that maintains a population of potential solutions to candidate set problem. The genetic algorithm contains three operators such as selection, crossover and mutation. These operators allow algorithm to explore search space. Genetic algorithms can be implemented in two ways: through external support and direct application over analysis. The reason for preferring genetic algorithm in data mining is that more attributes and more observations can be handled. However, genetic algorithm can be implemented only when the data is in discrete structures.

Fitness function used by selection operator in genetic algorithm provides support notion for an association. It defines figure of merit divided by utility measure of a sample in numerical format. The fitness function works in two phases; with finding associations of small support in the first phase and finding associations of large support in the second phase. The fitness function can be expressed using the below formula

$$F=\left((1-S)\mathrm{x}\frac{\frac{\mathrm{T}}{10}-10\mathrm{xSF}}{\mathrm{T}}\right)+2\mathrm{x}\left(S\,\mathrm{x}\frac{\frac{\mathrm{T}}{10}+10\mathrm{xSF}}{\mathrm{T}}\right)$$

Where S represents the support, T represents total number of features and SF represents number of selected significant features. Fitness function is a flexible way of expressing model criteria and tradeoffs among multiple objectives.

Genetic algorithms is having wide applications in medical domain. Asha Gowda Karegowda et al.[21] presented a paper on applications of genetic algorithm for medical diagnosis of PIMA Indian diabetes. This model integrates genetic algorithm and back propagation network to diagnose diabetes. Erhan Elveren and Nejat Yumusak[22] proposed tuberculosis diagnosis system using genetic algorithm. The method works with neural network trained data and genetic algorithm. Prem Pal Singh Tomar and Ranjit Singh[23] presented evolutionary continuous genetic algorithm for clinical decision support systems. This method, called Medical Multimedia based Clinical Decision Support System (MM-CDSS), supports diagnosis for four major heart diseases using patient's symptoms.

6. Conclusion

This paper presents ananalysis on associative classification and various traditional methods of associative classification. Performed review outlines the merits and demerits of traditional associative classification algorithms. For example, CARPT method saves time and space, whereas it removes frequent items during classification process. It is concluded from the study that if the objective of the classification is to generate high quality rules without redundancy, CPAR and CBA can be used. If the objective is to generate high quality accurate rules, CMAR is the best option. For performing classification on heterogeneous data sets, it is better to use CARGBA. Associative classification can also be performed by generating negative rules. ACN and ACCF methods use this concept and they are accurate in generating association rules. The variations of CBA algorithm such as HMAC and GARC, performs associative classification with higher accuracy. This paper reviews about associative classification on medical data sets. It is obvious from the study that inclusion of genetic algorithms while performing associative classification on medical data sets will produce better results.

7. References

1. Kononenko, Igor, Kukar M. Machine learning and data mining: Introduction to principles and algorithms. Horwood Publishing; 2007.
2. Chen MC, Chiu AL, Chang HH. Mining changes in customer behavior in retail marketing. Expert Systems with Applications. 2005; 28(4):773–81.
3. Thabtah F, Cowling P, Peng Y. MCAR: Multi-Class Classification based on Association Rule. The 3rd ACS/IEEE International Conference on Computer Systems and Applications; 2005.
4. Rygielski C, Wang JC, Yen DC. Data mining techniques for customer relationship management. Technology in Society. 2002 Nov; 24(4):483–502.
5. Weiyang L, Alvarez SA, Ruiz C. Efficient adaptive-support association rule mining for recommender systems. Data Mining and Knowledge Discovery. 2002; 6(1):83–105.
6. Yin X, Han J. CPAR: Classification based on Predictive Association Rules. SDM. 2003; 3:331–5.
7. Li W, Han J, Pei J. CMAR: Accurate and efficient classification based on multiple class-association rules. Proceedings IEEE International Conference on Data Mining-ICDM; San Jose, CA. 2001. p. 369–76.
8. Tao F, Murtagh F, Farid M. Weighted association rule mining using weighted support and significance framework. Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM; 2003. p. 661–6.
9. Kirkos E, Spathis C, Manolopoulos Y. Data mining techniques for the detection of fraudulent financial statements. Expert Systems with Applications. 2007; 32(4):995–1003.
10. Zheng Y, Liu P, Lei L, Yin J. R-C4. 5 decision tree model and its applications to health care dataset. 2005 Proceedings of International Conference on Services Systems and Services Management (ICSSSM'05); 2005. p. 1099–103.
11. Baralis E, Chiusano S, Garza P. A lazy approach to associative classification. IEEE Transactions on Knowledge and Data Engineering. 2008; 20(2):156–71.
12. Yu G, Li K, Shao S. Mining high utility itemsets in large high dimensional data. Proceedings of the 1st International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia and Workshop. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering); 2008. p. 47.
13. Lin W, Alvarez SA, Ruiz C. Efficient adaptive-support association rule mining for recommender systems. Data Mining and Knowledge Discovery. 2008; 6(1):83–105.
14. Thabtah F. A review of associative classification mining. The Knowledge Engineering Review. 2007; 22(01):37–65.
15. Zaki MJ. Mining non-redundant association rules. Data-Mining and Knowledge Discovery. 2004; 9(3):223–48.
16. Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Emerging Artificial Intelligence Applications in Computer Engineering; 2007. p. 3–24.
17. Nauck D, Kruse R. Obtaining interpretable fuzzy classification rules from medical data. Artificial Intelligence in Medicine. 1999; 16(2):149–69.
18. Scheffer T. Finding association rules that trade support optimally against confidence. Principles of Data Mining and Knowledge Discovery. Heidelberg: Springer Berlin; 2001. p. 424–35.
19. Rak R, Kurgan L, Reformat M. Multi-label associative classification of medical documents from medline. IEEE Proceedings Fourth International Conference on Machine Learning and Applications; 2005. p. 177–86.

20. Jabbar MA, Deekshatulu BL, Chandra P. Heart disease prediction system using associative classification and genetic algorithm. International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies; 2013. p. 183–92.
21. Karegowda AG, Manjunath AS, Jayaram MA. Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima Indians. International Journal on Soft Computing. 2001; 2(2):17–21.
22. Elveren E, Yumusak N. Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm. Journal of Medical Systems. 2011; 35(3):329–32.
23. Tomar PPS, Singh R. Evolutionary continuous genetic algorithm for clinical decision support system. African Journal of Computing and ICT. 2013; 6(1):127–46.
24. Su Z, Song W, Cao D, Li J. Discovering informative association rules for associative classification. IEEE International Symposium on Knowledge Acquisition and Modeling Workshop; Wuhan. 2008. p. 1060–3.
25. Shekhawat PB, Dhande SS. A classification technique using associative classification. International Journal of Computer Application. 2011; 20(5):20–8.
26. Ganganwar V. An overview of classification algorithms for imbalanced datasets. International Journal of Emerging Technology and Advanced Engineering. 2012; 2(4):42–7.
27. Jiang Y, Shang J, Liu Y. Maximizing customer satisfaction through an online recommendation system: A novel associative classification model. Decision Support Systems. 2010; 48(3):470–9.
28. Ye Y, Li T, Jiang Q, Wang Y. CIMDS: Adapting postprocessing techniques of associative classification for malware detection. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 2010; 40(3):298–307.
29. Dorbala S, DrNayak R. Integrated data mining approach for security alert system. International Journal of Research Computer Communication Technology. 2014; 3(4):464–8.
30. Phua C, Lee V, Smith K, Gayler R. A comprehensive survey of data mining-based fraud detection research. Artificial Intelligence Computer Science. 2010; 14.
31. Sangsuriyun S, Marukatat S, Waiyamai K. Hierarchical Multi-label Associative Classification (HMAC) using negative rules. 9th IEEE International Conference on Cognitive Informatics (ICCI); Beijing. 2010. p. 919–24.
32. Li X, Qin D, Yu C. ACCF: Associative Classification based on Closed Frequent itemsets. Fifth International Conference on Fuzzy Systems and Knowledge Discovery FSKD'08; Shandong. 2008. p. 380–4.
33. Chen G, Liu H, Yu L, Wei Q, Zhang X. A new approach to classification based on association rule mining. Decision Support Systems. 2008; 42(2):674–89.
34. Kundu G. Islam MM. Munir S. Bari MF. ACN: An associative classifier with negative rules. 11th IEEE International Conference on Computational Science and Engineering, CSE'08; Sao Paulo. 2008. p. 369–75.