An Effective Mining of Exception Class Association Rules from Medical Datasets

Abdualkareem Ali Al-Tapan
Department of Compute Science
Faculty of Computer Sciences and Information Systems, Thamar University
Thamar, Yemen
Email : abdualtapan@yahoo.com

Basheer Mohamad Al-Maqaleh
Department of Information Technology
Faculty of Computer Sciences and Information Systems, Thamar University
Thamar, Yemen
Email: basheer.almaqaleh.dm@gmail.com

Abstract — ***Data mining techniques can be applied on medical data to improve quality of decisions in medical field, and to enhance clinician performance. Class Association Rule (CAR) integrates the techniques of association rule mining and classification rules mining. CAR mining is able to find a lot of rules, some important rules are still not discovered. These rules are called Exception Class Association Rules (ECARs), which have high confidence and low support. The goal of ECAR is to discover rare and low support itemsets to generate useful rules from these itemsets and meet the class label (target attribute). Also, ECARs are considered very important by domain's expert in medical field, which lead to improve the performance of analysis and diagnosis. These type of rules cannot be identified easily using traditional algorithms as they focus in frequent itemsets discovery and don't depend on target attribute for diagnosis. In this paper, an effective algorithm is proposed to mine two types of CARs: Precise Class Association Rules (PCARs) and Approximate Class Association Rules (ACARs), and to generate ECAR for each ACAR from medical datasets. The experimental results shows the effectiveness of the proposed algorithm.***

Keywords— Association Rule Mining; Knowledge Discovery; Medical Data Mining; Class Association Rule; Exception Class Association Rule.

I. INTRODUCTION

Information is gathered almost everywhere in our everyday life and the amount of this information stored in modern databases makes manual analysis intractable [1]. Discovering knowledge in medical systems and health care industry considered as crucial task that helps in decision making [2]. There are an urgent need in the health care industry to store and organize huge clinical data, and to develop data mining techniques to mine hidden patterns and discover valuable knowledge from clinical data [3, 4].

Association rule mining has been, mainly, developed to identify the relationships strongly associated among itemsets that have high frequency and strong correlation [5]. Association rule is of the form $X \rightarrow Y$, where X (antecedent part) and Y (consequent part) are disjoint itemsets. The support (*sup*) of the association rule $X \rightarrow Y$ is defined as:

$$sup(X \rightarrow Y) = \frac{|X \cap Y|}{|D|} \quad (1)$$

where $|X \cap Y|$ is the total number of transactions containing both X and Y and $|D|$ is the total number of transactions in a dataset.. The confidence (*conf*) can be defined as:

$$conf(X \rightarrow Y) = \frac{|X \cap Y|}{|X|} \quad (2)$$

where |X| is the number of transactions containing X. An association rule is considered interesting if its *sup* and *conf* are larger than some thresholds, usually denoted by *minsup* (for support) and *minconf* (for confidence). The thresholds are usually defined by the user [5]. Association rule mining process basically consists of two steps [6]: (i) finding all the frequent itemsets that satisfy *minsup* threshold and (ii) generating strong association rules from the derived frequent itemsets by applying *minconf* threshold.

Classification rule mining aims to discover set of rules in the dataset leading to a specific class label [3]. Three main standards of classification techniques include, statistical based classification, machine learning based classification and neural network based classification, which play an important role in decision support systems [7]. For association rule mining, the target attribute (or class attribute) is not pre-determined. However, the target attribute must be defined in classification problems.

Class Association Rule (CAR) integrates the techniques of association rule mining and classification rule mining [8]. CAR is a small subset of association rules whose right hand sides are restricted to the class label. CAR has high *sup* and high *conf*. In the proposed work, the CARs are classified into two types of rules: Precise CAR (PCAR) and Appropriate CAR (ACAR). CAR is said to be an PCAR if it has high *sup* but the items contained in it are highly correlated or the *conf* of the rule 100%. CAR is said to be ACAR if it has high *sup* and confidence value less than 100% and greater than or equal to *minconf*.

This sort of knowledge discovery is very useful in many real time applications including disease detection, diagnosis, disease classification, predicting breast cancer survivability, finding whether back surgery fails or succeed and decision support systems in clinical applications [9]. Although, CAR mining is able to find a lot of rules, some important rules are still not considered. These rules called Exception Class Association Rules (ECARs), which have high *conf* and low *sup* [10]. Exception rules deal with a different type of knowledge, provide a more comprehensible understanding of the information provided by a dataset and can improve the quality of decisions [11]. Exception rules are always interesting to discover, as they challenged the existing knowledge, often lead to the growth of knowledge in new directions, deserves more attention and still remains a great challenge [12, 13].

The proposed algorithm is developed to find CAR, to get PCAR and ACAR, and to generate ECAR for each ACAR from medical datasets. The rest of this paper is organized as follows: Section II presents problem statement, In Section III overview of ECAR Mining is presented. Related works is presented in Section IV. Section V describes the proposed algorithm. Computational results and comparative study are given in Section VI and Section VII respectively. Conclusion and future work are obtained in Section VIII.

II. PROBLEM STATEMENT

Discovering association rules is the most general mechanism in the field of data mining and it can be used also for classification [14]. The current algorithms focus in frequent itemsets discovery, ignoring rare itemsets and don't depend on target attribute for diagnosis, all of that are considered very important in medical domain. Those rare instances and infrequent itemsets related to exceptions are considered very important by domain's expert in medical field, which lead to make perfect diagnosis.

Exception rules are a suitable tool for representing infrequent itemsets that can be present in the data [12]. This is often true in areas such as medical diagnosis and treatment where many deviations/symptom combinations will only manifest in a small number of patient cases [15]. Experts in medical domain agreed that mining exception rule that can change the behavior of the diagnosis are very useful and further research in this field is needed. Nowadays, searching for specific kind of knowledge that deviates from the usual standards is very useful in medical diagnosis [12]. The result of medical data mining can help doctors to make the correct diagnosis and treatment, and this is very significant for human health. Also, another problem is found in the increasing availability of the biomedical literature and datasets which makes it difficult to discover valuable knowledge without help of automatic knowledge discovery techniques [16]. One of the major challenges in the medical domain today is how to exploit the huge amount of data that this field generates and to extract the novel and usable information or knowledge is very complicated and time consuming task [3, 17]. To do this, an efficient and accurate automation of such systems that are capable of discovering knowledge that is useful for decision making in medical field would be extremely advantageous.

In this context, an effective algorithm for mining PCARs, ACARs and to discover ECAR for each ACAR from medical datasets is presented. The proposed algorithm express a kind of knowledge which is very useful in medical domain. With this kind of discovered rules, the usual patterns (PCAR and ACAR), and the exception rule (ECAR) representing a change in the common patterns are obtained.

III. EXCEPTION CLASS ASSOCIATION RULE MINING

CAR mining is an approach that applies association rule mining to build classifier [8]. Rules found by CAR mining are of high *sup* and *conf from* the given dataset. These rules represent the common trends in the databases

and are valuable for the common states. The rules with low *sup* are just as valuable as they may contain unusual, and unexpected knowledge about database. These rules are referred to as ECAR, which have high *conf* but low *sup*. They are important because they are contrary to the CAR mining and may give us some extra knowledge. ECAR can be defined as [10]:-

Let $X \rightarrow y_1$ be a ACAR, an ECAR has a form of $\{X, Z \rightarrow y_2\}$, $X, Z \in I$ where I represents set of frequent itemsets , $X \cap Z = \emptyset$, $y_1, y_2 \in Y$, $y_1 \neq y_2$. The *conf* of ECAR satisfies the minconf requirement, but does not satisfy the *minsup*. The attribute X interacting with another attribute Z may change the consequent (target class) of ACAR from y_1 to y_2. For example [13], if X represents antibiotics, Y recovery, Z staphylococci, and E death, then the following rule might be discovered:

If the patient takes antibiotic (X) **Then** it tends to recover, (Y) **Unless** staphylococcus appears, (Z)

If the patient takes antibiotic (X) $\wedge$ staphylococcus appears (Z) **Then** death, (E)

These exception rules indicate that there is some kind of interaction between two factors, X and Y, so that the presence of Z alters the usual behavior (Y). This is a very interesting kind of knowledge which cannot be detected by traditional association rules because the exceptions are hidden by a dominant rule [12].

IV. RELATED WORKS

Association rules are widely used in various areas such as telecommunication networks, risk and market management, medical diagnosis, medical research and health care industry [18]. A methodology to discover locally frequent diseases with the help of Apriori data mining techniques is presented in [9]. They used visualization techniques to present the trends graphically and built a prototype application that demonstrates the efficiency of the method. A boosted Apriori as an effective data mining association rules for heart disease prediction system is proposed in [2]. In this system hidden patterns from a data collected together by an hospital are extracted. Apriori algorithm for rules extraction from Breast Cancer dataset is presented in [14]. The discovered association rules are used for classification of Breast Cancer patients. The study was carried out by using the Weka data mining tool. An intelligent predictive system using classification techniques for heart disease diagnosis is introduced in [19]. Their research objective is to study some of classification techniques to predict the heart disease and find the best technique of prediction. A statistical prototype is developed in [20] to perform clinical diagnosis of malaria using rule-based classification and statistical models. An associative based classification algorithm for diabetes disease prediction is introduced in [21]. Experimental results proved that the proposed model has high true positive rate and precision compared to traditional classification methods. A heart disease prediction system using associative classification and genetic algorithm is proposed in [22]. This method performs high level rules that are accurate and comprehensible, and contains high interesting value. A class based approach for medical classification of chest pain is suggested in [23]. The class label is used in classification to minimize the searching space. A novel approach for mining relevant CARs that consider constraints on the rule consequent is proposed in [24]. Rules found by CAR mining are of high *sup* and *conf* from the given data, some important ECARs are still not discovered. An effective method to find CAR, to get Useful Class Association Rules (UCAR) by removing the Spurious Class Association Rules (SCAR), and to generate ECAR for each UCAR, is presented in [10]. A method for computing both normal and abnormal classification models in one phase is presented in [25] .This method shows the important complementary role of abnormal models with respect to normal models in classification through experimentation on UCI datasets. Approaches for discovering exception and anomalous rules in association rules are proposed in [12].

In the current work, the proposed algorithm discovers exception rules that are related to CARs representing the usual and normal patterns. According to this, it can discover from scratch the usual patterns in terms of PCARs and ACARs and also the different exceptions associated with ACARs, resulting in ECARs. Therefore, the proposed algorithm differs from these techniques that only search for the exceptions in the whole dataset. Also, the proposed algorithm is developed to deal with the problem of imbalanced datasets effectively. The class with the lowest number of instances is usually the class of interest from the view point of the learning task and has many applications especially in medical diagnosis [26].

V. THE PROPOSED ALGORITHM

In this section, the proposed algorithm for mining CARs and ECARs from medical datasets is introduced. In the proposed algorithm, all the possible attribute values in the dataset are mapped to the same target class. In Step 1, the set of all supported itemsets with respect to the same target-class are generated. Traditional association rule mining algorithms use only a single *minsup* in rule generation, which is inadequate for imbalanced class distribution. The proposed algorithm suggests a new minimum support (*minsup-new*) for each class *Ci*. A dataset

is class imbalanced if the classification categories are not approximately equal represented [26]. Suppose a dataset has 2 classes, C_1 and C_2 with freqDistr (C_1) =98% and freqDistr (C_2) =2%. Consider we set *minsup*=3%, we will not find any rule of class C_2. To solve this problem, the multiple minimum class supports is suggested. For each class attribute C_i, a different minimum class support is assigned (i.e., one *minsup* value for each target class). This value is called *minsup-new*. The user only gives a total *minsup*, which is distributed to each class according to their class distributions as follows:

$$minsup\text{-}new = (minsup/100)*\text{number of instances covered by the target class } (C_i) \quad (3)$$

This formula gives frequent classes higher *minsup* and infrequent classes lower *minsup*. This ensures that sufficient rules for frequent and infrequent classes will be generated. The set of all supported itemsets (satisfy the *minsup* requirement) in L_k only are used to generate CARs. But Apriori algorithm uses the set of all supported itemsets in $L_1 \cup L_2 \cup \ldots \cup L_k$ to generate association rules. In this case, the proposed algorithm reduces the number of dataset scans, facilities support counting of candidates, and then generates less number of rules comparing to Apriori algorithm. The generated rules called CARs with respect to the target class under consideration. These rules satisfy the *minsup-new* and *minconf* requirements as shown in Step 2.

In Step 3, the proposed algorithm discovers PCARs and ACARs as follows: for every rule (Z) found in CARs in the form of CAR(x → y). If *conf* (Z) =1, then this rule has no exception and called PCAR. Otherwise this rule has exception/exceptions and called ACAR. In Step 4, the proposed algorithm finds ECAR for each ACAR(x → y) as follows: for every instance that satisfy the condition (=x and !=y) , this instance is considered as an exception candidate. For all the exception candidates, the proposed algorithm scans the whole data once to check their *conf* requirements. Those satisfying the *minconf* are considered ECARs. The proposed algorithm for mining CARs (PCARs, ACARs) and ECARs is shown in Figure 1:

Input: A dataset D, minsup, minconf and target class.
Output: CARs and ECARs
Step 1*: //To find frequent itemsets*

Create C_1 = set of itemsets of cardinality one only in instances containing the target class where class not be include in C_1 *set*.
minsup-new =((*minsup* /100) * number of instances covered by the target class))
Create L_1 = set of supported itemsets of cardinality one only in the target class
Set k to 2
While (L_{k-1} *!* = ø) {
Create C_k from L_{k-1}
For all candidates c ∈ C_k compute *sup*(c) for itemsets covered by the target class
If (*sup*(c) >= *minsup-new*){
Add c set to frequent in L_k
Else
Prune all the itemsets in C_k *that* are not supported} *//end if*
Increase *k* by 1 //k++ } *end For* //
The set of all supported itemsets in L_k*}* // end While

Step 2: *//To add target class for all supported itemsets in* L_k *to generate CARs.*

For each subset x (set of frequent itemsets) ∈ L_k {
Add target class to each subset x (in consequent part)
Extract CARs that *conf* (CAR) ≥ *minconf*} // *end For*

Step 3: *// PCARs and ACARs discovery*

CARs= {Z: z ∈CARs} // Z: sets of frequent itemsets covered by the target class in the form of x → y where x (set of frequent itemsets) and y(target class).

For each subset Z in CARs
{Compute *conf* (Z)
If (*conf* (Z) ≥ *minconf* && *conf* (Z) < 1.00)
Add Z to ACARs
Else If (*conf* (Z) =1)
Add Z to PCARs // end For

Step 4*: //ECARs discovery*

For each subset x in ACARs (x → y)
Find every instance that satisfies the condition (= x and ! =y (another target class))
Add the extracted rules to ECAR set. //the end

Figure 1. The proposed algorithm.

VI. COMPUTAIONAL RESULTS

The performance of the proposed algorithm is validated on real-world medical datasets obtained from UCI Machine Learning Respiratory, which is a collection of widely used benchmark and real-world datasets for data mining and KDD community [27]. The performance of the proposed algorithm is evaluated and compared with well-known Apriori algorithm which is implemented in a public domain tool called WEKA. WEKA is a collection of machine learning algorithms for data mining task [28].

All experiments are performed on a laptop with c# 2012, windows 7.0, 4 GB RAM and 2.4 GHZ core i5 processor. In the following experiments, each class in a dataset is dealt with separately. During each run, the same class under consideration is assigned to all the attributes values containing for this class in the dataset. The set of discovered rules are divided into CARs (PCARs, ACARs) and ECARs. The performance of the proposed algorithm on different dataset is demonstrated below:

A. EXPERIMENT 1

The Hepatitis dataset contains 155 data instances and 20 attributes. The last attribute is the class attribute which can take two values (die or live). The proposed algorithm discovered two ACARs (the best rule for each class). The *conf* of those two ACARs are not equal to one, which means some exceptions could be discovered for each ACAR. So, the proposed algorithm, would discover those exceptions for each ACAR resulting in ECARs (R2 and R4). Those exceptions are in bold italic font. The proposed algorithm would not discover any PCAR from this dataset. The discovered rules are shown in Table I.

TABLE I. RESULT FOR THE HEPATITIS DATASET.

No.	Mined rules	*minsup*	*minsup-new*	*minconf*	Conf. of Mined Rules		
					PCAR	ACAR	ECAR
R1	**If** Anorexia=yes ,Asictes=yes,spiders=yes ,spleenpalpable=yes,varies=yes **Then** class=live.	0.6	73.8	0.7	--	0.96	--
R2	**If** Anorexia=yes ,Asictes=yes, spiders=yes, spleenpalpable=yes,varies=yes and ***hisology=no,antiviral=no*** **Then** Class=die.	0.6	------	0.7	--	--	1.00
R3	**If** antiviral=yes,liver_big=yes,protime=low **Then** Class=die	0.4	12.8	0.7	--	0.92	--
R4	**If** antiviral=yes, liver_big=yes, protime=low, ***albumin=low,ascites=no*** **Then** Class=live	0.4	-------	0.7	--	--	1.00

B. EXPERIMENT 2

The Wisconsin Breast Cancer dataset was used in this experiment. The total number of instances are 699 and the number of attributes is 10 plus the class attribute. Each instance has one of 2 possible classes: benign or malignant tumor. The proposed algorithm would discover one PCAR for class=2, which means there is no exception for this class. Also, the proposed algorithm would discover one ACAR for class=2 and its ECAR is discovered as shown in R3. Accordingly, the proposed algorithm would discover one ACAR (R4) for class =4 and its ECAR is discovered as shown in R5. The exceptions for every ECAR are shown in bold italic font. Table II shows the mined rules.

TABLE II. RESULT FOR THE BREAST CANCER DATASET.

No.	Mined rules	*minsup*	*minsup- new*	*minconf*	Conf. of Mined rules		
					PCAR	ACAR	ECAR
R1	**If** Bare=1,Clump=1,Normal=1 , Uniformitycellsize=1 **Then** Class=2	0.6	275	0.7	1.00	--	--
R2	**If** Bare=1,Clump=1, Normal=1, Single=1 **Then** Class=2	0.2	91	0.7	--	0.99	--
R3	**If** Bare=1,Clump=1, Normal=1, Single=1, ***Uniformitycellsize=4 , Uiformitycellshape=1*** **Then** Class=4	0.6	--	0.7	--	--	1.00
R4	**If** Clump=1,Marginal=10 **Then** Class=4	0.2	48	0.7	--	0.97	--
R5	**If** Clump=1,Marginal=10,***Mitoses=10 ,normal=1*** **Then** Class=2	0.6	--	0.7	--	--	1.00

C. EXPERIMENT 3

Dermatology dataset was used in this experiment. This dataset contains 366 examples and 34 non-categorical attributes (33 of which are linear valued and one of them is nominal. It has six classes (1- psoriasis, 2- seboreic dermatitis, 3- lichen planus, 4- pityriasis rosea, 5- cronic dermatitis, and 6- pityriasis rubra pilaris). In Dermatology dataset, fourteen rules (four PCARs, five ACARs, and five ECARs) are generated as shown in Table III.

TABLE III. RESULT FOR THE DERMATOLOGY DATASET

No.	Mined rules	*minsup*	*minsup- new*	*minconf*	Conf. of Mined rules		
					PCAR	ACAR	ECAR
R1	**If** eosinophils =0, Exocytosis=0, follicular =0 , inflammatory =2, melanin =0,perifollicular=0, spongiosis=0,vacuolization=0 **Then** class=1	0.5	56	0.7	--	0.92	--
R2	**If** acanthosis=2 ,eosinophils =0 , Exocytosis=0 , follicular =0 and melanin =0, perifollicular=0, spongiosis=0 ,vacuolization=0 **Then** class=1	0.5	56	0.7	--	0.85	--
R3	**If** eosinophils =0, Exocytosis=0, follicular =0 ,inflammatory=2, melanin=0 perifollicular=0, spongiosis=0 ,vacuolization=0 , ***koebner_phenomenon=0,acanthosis=2,Erythema=1*** **Then** class = 5	0.5	--	0.7	--	--	1.00
R4	**If** acanthosis=2 ,eosinophils =0 , Exocytosis=0 , follicular =0 . melanin =0, perifollicular=0, spongiosis=0 , vacuolization=0) , (***koebner_phenomenon=0, parakeratosis=0, Erythema=1*** **The**n class= 5	0.5	--	0.7	--	--	1.00
R5	**If** acanthosis=2 ,eosinophils =0 , melanin=0, parakeratosis=0, spongiosis=0 , vacuolization=0, perifollicular=0 **Then** class= 5	0.3	15	0.7	1.00	--	--
R6	**If** koebner_phenomenon=0,,eosinophils =0 , melanin=0, parakeratosis=0, spongiform=0 , follicular =0 vacuolization=0, perifollicular=0 **Then** class= 5	0.3	15	0.7	--	0.94	--
R7	**If** koebner_phenomenon=0,,eosinophils =0 , melanin=0, parakeratosis=0, spongiform=0 , follicular =0 vacuolization=0, perifollicular=0 , ***inflammatory =0, Erythema=2*** **Then** class= 2	0.3	--	0.7	--	--	1.00
R8	**If** elongation=0, koebner_phenomenon=0, melanin=0, perifollicular=0, spongiform=0, vacuolization=0,scaling=2 **Then** class= 2	0.4	24	0.7	--	0.82	--
R9	**If** elongation=0, koebner_phenomenon=0, melanin=0, perifollicular=0, spongiform=0, vacuolization=0,scaling=2 , ***inflammatory =2, acanthosis=2, Erythema=2*** **Then** class= 4	0.4	--	0.7	--	--	1.00
R10	**If** elongation=0, spongiform=0, perifollicular=0, vacuolization=2, follicular =0 **Then** class= 3	0.5	36	0.7	1.00	--	--
R11	**If** elongation=0, spongiform=0, follicular =0, melanin=2, perifollicular=0 **Then** class= 3	0.5	36	0.7	1.00	--	--
R12	**If** elongation=0, eosinophils =0, follicular =0, melanin=0, perifollicular=0, vacuolization=0, spongiform=0 **Then** class= 4	0.5	24	0.7	--	0.94	--
R13	**If** elongation=0, eosinophils =0, follicular =0, melanin=0, perifollicular=0, vacuolization=0, spongiform=0 , ***inflammatory =3, koebner_phenomenon=0*** **Then** class= 2	0.5	--	0.7	--	--	1.00
R14	**If** elongation=0, eosinophils =0, inflammatory =2, koebner_phenomenon=0, perifollicular=0, vacuolization=0 **Then** class= 6	0.5	10	0.7	1.00	--	--

VII. COMPARATIVE STUDY

Experiments were then made to compare the execution time in seconds(s) between the proposed algorithm and Apriori algorithm which is implemented in WEKA. Table VI shows the execution time comparison of the proposed algorithm and Apriori algorithm with different minimum support values when applied to the Hepatitis, Breast Cancer and Dermatology datasets.

TABLE VI. TIME COMPARISON OF THE PROPOSED ALGORITHM AND APRIORI ALGORITHM.

Dataset	*minsup*	Running time of Apriori algorithm(s)	Running time of the proposed algorithm(s)
Hepatitis	0.5	4	2
	0.4	6	2
	0.3	7	3
	0.2	7	3
	0.1	8	4
Breast cancer	0.5	5	2
	0.4	6	2
	0.3	7	3
	0.2	8	4
	0.1	7	4
Dermatology	0.5	8	3
	0.4	9	3
	0.3	9	4
	0.2	9	4
	0.1	10	5

It is expected that the mining time keeps on increasing when the value of *minsup* decreasing. In the proposed algorithm, the frequent itemsets with class label are generated. So, we only need to join the frequent k-itemsets within the same class label to generate the k+1 frequent itemsets. Additionally, searching exception rules in these transactions satisfying the antecedent of the rule prunes the search space, feature that is very important when we handle with very large datasets. Comparing to Apriori algorithm, which scans the whole dataset and generates all the frequent itemsets, a lot of time can be saved. The experimental results proved that the proposed algorithm is much more efficient than Apriori algorithm in terms of mining time.

VIII. CONCLUSION AND FUTURE WORK

This paper proposes a new effective algorithm to find CARs and ECARs from medical datasets. The proposed algorithm generates two types of CAR: PCAR and ACAR. Consequently, ECAR for each ACAR is discovered. This algorithm is developed to deal with class imbalanced problem effectively. The proposed algorithm was compared with respect to the mining time with Apriori algorithm. The results proved that the proposed algorithm can attain considerable performance improvement comparing to the Apriori algorithm. One of the most important future directions would be the discovery of ECARs from large datasets using the multi-objective evolutionary algorithm.

REFERENCES

[1] M. Shrider, and M. Parmar, "Survey on association rule mining and its approaches," International Journal of Computer Sciences and Engineering (IJCSE), vol.5, Issue 3, pp. 129-135, 2017.

[2] R. Thanigaivel, and K. R. Kumar, " Boosted Apriori: - An effective data mining association rules for heart disease prediction system," Midde-East Journal of Scientific Research, vol. 24, no. 1, pp. 192-200, 2016.

[3] R. J. Oskouei, N. M. Kor, and S.A. Maleki, " Data mining and medical world: breast cancers, diagnosis, treatment, prognosis and challenges, " American Journal of Cancer Research (AJCR), vol.7, no. 3, pp.610-627, 2017.

[4] S. Sharath, M. N. Rao, and H. G. Chetan," A survey on the principles of mining clinical datasets by utilizing data mining techniques," International Journal of Innovative Research in Computer and Communication Engineering, vol. 2, Issue 4, 2014.

[5] R. Agrawal, and R. Srikant, "Fast algorithms for mining association rules in large databases," 20th International Conference on Very Large Data Bases, 1994.

[6] D. J. Prajapati, S. Garg and N. C. Chauhan," Interesting association rule mining with consistent and inconsistent rule detection from big sales data in distributed environment, " Future Computing and Informatics Journal, pp. 1-12, 2017.

[7] L. T.T. Nguyen, B. Vo, T-p. Hong and H. C. Thanh, "CAR-Miner: An efficient algorithm for mining class-association rules," Expert Systems with Application, vol. 40, Issue 6 , pp. 2305-2311,2013.

[8] B. Liu , W. Hsu , and Y. Ma , "Integrating classification and association rule mining, " In: Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, New York, pp. 80–86, 1998.

[9] M. A Khaleel, S. K. Pradhan, and G. N. Dash, " Finding locally frequent diseases using modified Apriori algorithm," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, Issue 10, 2013.

[10] Z. Aoying , W. Li and Y. Fang, " Effective discovery of exception class association rules," Journal of Computer Science and Technology, vol. 17, Issue 3, pp. 304-313, 2002.

[11] E. Suzuki, and M. Zytkow, "Unified algorithm for undirected discovery of exception rules," International Journal of Intelligent Systems, vol. 20, pp. 673–691, 2005.

[12] M. D. Ruiz, D. Sanchez, M. Deldado, and M. J. Martin-Bautista, " Discovering Fuzzy Exception and Anomalous Rules," IEEE Transactions on Fuzzy Systems, vol.24, no. 4,pp . 930-944, 2016.

[13] E. Suzuki, "In pursuit of interesting patterns with undirected discovery of exception rules, " In Progress Discovery Science, vol. 2281, pp. 504–517, Lecture Notes in Artificial Intelligence, 2001.

[14] T. Pala, I. Ycedag , and H. Biberoglu, "Association rule for classification of breast cancer patients, " Sigma J. Eng & Nat Sci ,vol. 8, Issue 2, pp.155-160, 2017.

[15] L. Szathmary, P. Valtchev , and A. Napoli , “Finding minimal rare itemsets and rare association rules, " In Proceedings of the 4th International Conference on Knowledge Science, Engineering and Management (KSEM 2010), pp. 16-27, 2010.
[16] F. Shen and Y. lee, "Knowledge discovery from biomedical ontologies in cross domains," vol. 11. Issue 8, 2016.
[17] A. Anguera, J.M. Barreiro, and D. lizocano, " Applying data mining techniques to medical time series: An empirical case study in electroencephalography and stabilomerty, " Computational and Structural Biotechnology Journal, 2016.
[18] G. Kaur, “Association rule mining: A survey," International Journal of Computer Science and Information Technologies, (IJCSIT) , vol. 5 no. 2 , pp. 2320-2324, 2014.
[19] B. M. Al-Maqaleh, and A. M. G. Abdullah, "Intelligent predictive system using classification techniques for heart disease diagnosis," International Journal of Computer Science Engineering (IJCSE), vol. 6, no. 6, pp. 145-151, 2017.
[20] F. Bbosa, R. Wesonga , and P. Jehopio , " Clinical malaria diagnosis rule-based classification statistical prototype ," Springer-verlag , vol. 5 , Issue 1, 2016.
[21] N.G. Deepkia, Y. Surekha and G. L. Kumami, "Associative based classification algorithm for diabetes disease prediction," IJEET, vol.41, no. 3, pp. 159-164, 2016.
[22] M. A. Jabbar, P. Chandra, and B. L. Deekshatulu, " Heart disease prediction system using associative classification and genetic algorithm," International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies(ICECIT), 2012.
[23] D. Sasirekha and A. Punitha, "A comprehensive analysis on associative classification in medical datasets," Indian Journal of Science and Technology, vol 8(33), 2015.
[24] D. Nguyen, L.T.T. Nguyen, B. Vo, and T.-P. Hong," A novel method for constrained class association rule mining ," Journal of Information Sciences, vol. 320, pp. 107-125, 2015.
[25] V. Luong, “Mining normal and abnormal class-association rules”, IEEE 27th International Conference on Advanced Information Networking and Applications, 2013.
[26] A. M. Mahmood, “Class imbalance learning in data mining: A survey,” International Journal of Communication Technology for Social Networking Services, vol.3, no.2, pp.17-36, 2015.
[27] UCI Repository of Machine Learning Databases, Department of Information and Computer Science University of California, [http://www.ics.uci.edu/ ~mlearn/MLRepositry.html], 1994.
[28] WEKA: http://www.cs.Waikato.ac.nz/ ml/weka/ index.html.