PREDICTION SYSTEM TO SUPPORT MEDICAL INFORMATION SYSTEM USING DATA MINING APPROACH

Mr. Dhiraj Pandey[1], Mr. Santosh kumar[2]
1(Department of Computer Science, JSSATE, Noida , U.P. India)
2 (Department of Computer Science, JSSATE, Noida, U.P. India)

ABSTRACT

Different health care societies & hospitals have different kind of system which has been used for data storage of each recapitulated patient's disease. However patient's disease is increasing, day by day. So there is a need of an application which can provide information for decision makers, on patient diseases collected data. Computational intelligence methods open up new prospects for diseases diagnostic criteria. Data mining is an approach which can help in decision making. Hybrid-dimension association rules based, data mining technique, is based on methodologies for analyzing the relationship between diseases with patient characteristics. Apriori algorithm based data mining can support for development of this type of methodologies. Hybrid dimension association rule is a multidimensional association rule that allows the repetition of the predicate. Each rule can be used to describe the rules of the relationship with the patient's disease and patient's characteristics. In the proposed approach an extendable and improved item set generation has been constructed, and developed, for mining the relationships of the symptoms and disorder in the medical databases. It will produce hybrid dimension association rules and the rules have been displayed in form of tables and graphs.

Keywords **-** Apriori algorithm; Model _Multi model; Multi-dimensional association rule mining model; Supporting System of Medical Decision

1 INTRODUCTION

Most of the information regarding medical care is paper-based and not easily accessible at the point of patient care. Even when the information is present in the form of electronic documents on the web or in clinical information systems, doctors are not realistically given adequate time to search for the information specific to a particular patient sitting in front of them. This environment is especially prone to an error in many hospitals for ordering drug prescriptions.

Data mining is an approach which can help in decision making. Hybrid-dimension association rules based, data mining technique, is based on methodologies for analyzing the relationship between diseases with patient characteristics. Apriori algorithm based data mining can support for development of this type of methodologies. Hybrid dimension association rule is a multidimensional association rule that allows the repetition of the predicate. Each rule can be used to describe the rules of the relationship with the patient's disease and patient's characteristics.

In the proposed approach an extendable and improved item set generation has been constructed, and developed, for mining the relationships of the symptoms and disorder in the medical databases. It will produce hybrid dimension association rules and the rules have been displayed in form of tables and graphs.

The proposed approach can be used for large medical and health databases for constructing association rules. For disorders frequently seen in the patient and determining the correlation of the health disorders and symptoms observed simultaneously. This will provide significant knowledge, e.g. patterns, relationships between medical factors related to diseas

2.1 DIABETES

Diabetes mellitus [9], often simply referred to as it, is a group of metabolic diseases in which a person has high blood sugar, either because the body does not produce enough insulin, or because cells do not respond to the insulin that is produced. This high blood sugar produces the classical symptoms of polyuria (frequent urination), polydipsia (increased thirst) andpolypha gia (increased hunger).

2.1.1 Diagnosis and analysis of diabetes

Diabetes mellitus, often simply referred to as diabetes [7], is a group of metabolic diseases in which a person has high blood sugar.

2.1.2 Types of diabetes

- **Type 1 diabetes:** This results from the body's failure to produce insulin, and presently requires the person to inject insulin [7]. (Also referred to as insulin-dependent diabetes mellitus, IDDM for short, and juvenile diabetes.)

- **Type 2diabetes**: These results from insulin resistance [12], a condition in which cells fail to use insulin properly, sometimes combined with an absolute insulin deficiency. (Formerly referred to as n insulindependent diabetesmellitus, NIDDM for short, and adult-onset diabetes.)

- **Gestational diabetes:** This is when pregnant women, who have never had diabetes before, have a high blood glucose level during pregnancy.

2.1. 3 Signs and symptoms

The classical symptoms of diabetes are [6] polyuria (frequent urination) polydipsia (increased thirst) and polyphagia (increased hunger). Symptoms may develop rapidly (weeks or months) in type 1 diabetes while in type 2 diabetes they usually develop much more slowly and may be subtle or absent.

Prolonged high blood glucose causes glucose absorption, which leads to changes in the shape of the lenses of the eyes, resulting in vision changes; sustained sensible glucose control usually returns the lens to its original shape. Blurred vision is a common complaint leading to a diabetes diagnosis; type 1 should always be suspected in cases of rapid vision change, whereas with type 2 changes are generally more gradual, but should still be suspected.

2.2 ASTHMA

Asthma is a chronic (long-lasting) inflammatory disease of the airways. In those susceptible to asthma, this inflammation causes the airways to spasm and swell periodically so that the airways narrow. The individual then must wheeze or gasp for air. Obstruction to air flow either resolves spontaneously or responds to a wide range of treatments, but continuing inflammation makes the airways hyper-responsive to stimuli such as cold air, exercise, and dust mites, pollutants in the air, and even stress and anxiety.

2.2.1 Signs and symptoms

Common symptoms of asthma include wheezing, shortness of breath [7], chest tightness and coughing, and use of accessory muscle. Symptoms are often worse at night or in the early morning, or in response to exercise or cold air. Some people with asthma only rarely experience symptoms, usually in response to triggers, whereas other may have marked persistent airflow obstruction.

2.2.2 Causes

Asthma is caused by environmental and genetic factors [7]. These factors influence how severe asthma is and how well it responds to medication. The interaction is complex and not fully understood. Studying the prevalence of asthma and related diseases such as eczema and hay fever have yielded important clues about some key risk factors. The strongest risk factor for developing asthma is a history of atopic disease; this increases one's risk of hay fever by up to 5x and the risk of asthma by 3-4x. In children between the ages of 3-14, a positive skin test for allergies and an increase in immunoglobulin E increases the chance of having asthma. In adults, the more allergens one reacts positively to in a skin test, the higher the odds of having asthma.

Because much allergic asthma is associated with sensitivity to indoor allergens and because Western styles of housing favour greater exposure to indoor allergens, much attention has focused on increased exposure to these allergens in infancy and early childhood as a primary cause of the rise in asthma. Primary prevention studies aimed at the aggressive reduction of airborne allergens in a home with infants have shown mixed findings. Strict reduction of dust mite allergens, for example, reduces the risk of allergic sensitization to dust mites, and modestly reduces the risk of developing asthma up until the age of 8 years old. However, studies also showed that the effects of exposure to cat and dog allergens worked in the converse fashion; exposure during the first year of life was found to reduce the risk of allergic sensitization and of developing asthma later in life.

2.2.3 Environmental

Many environmental risk factors have been associated with asthma development and morbidity in children [7, 21]. Recent studies show a relationship between exposure to air pollutants (e.g. from traffic) and childhood asthma.This research finds that both the occurrence of the disease and exacerbation of childhood asthma are affected by outdoor air pollutants. High levels of endotoxin exposure may contribute to asthma risk. Viral respiratory infections are not only one of the leading triggers of an exacerbation but may increase one's risk of developing asthma especially in young children. Respiratory infections such as rhinovirus, Chlamydia pneumoniae and Bordetella pertussis are correlated with asthma exacerbations. Psychological stress has long been suspected of being an asthma trigger, but only in recent decades has convincing scientific evidence substantiated this hypothesis. Rather than stress directly causing the asthma symptoms, it is thought that stress modulates the immune system to increase the magnitude of the airway inflammatory response to allergens and irritants.

2.3 DATA MINING

Data mining is the process of extracting patterns from data[3]. Data mining is seen as an increasingly important tool by modern business to transform data into an informational advantage. It is currently used in a wide range of profiling practices, such as marketing, surveillance, fraud detection, and scientific discovery. Observational studies, such as statistical learning and data mining, can establish the association of the variables to get the related & required result.

The related terms data dredging,[3] data fishing and data snooping refer to the use of data mining techniques to sample portions of the larger population data set that are (or may be) too small for reliable statistical inferences to be made about the validity of any patterns discovered. These techniques can however, be used in the creation of new hypothesizes to test against the larger data populations. Data Mining & Statistics Analysis is the search for valuable information in large volumes of data. It is now widely used in health care industry. When certain Data Mining & Statistics methods used and valuable information extracted from huge data, it can assist clinicians to make informed decision and improve health service. Especially breast cancer is the second most cause of cancer and the second most dangerous cancer. The best way to improve a breast cancer victim's chance of long-term survival is to detect it as early as possible with the help data mining techniques.

2.3.1 Different Methods of Data Mining

- Association Rule Learning
- Cluster Analysis
- Structured Data Analysis (statistics)
- Java Data Mining
- Data Analysis
- Predictive Analysis
- Knowledge Discovery

3.1 APRIORI BASED FRAMEWORK & ANALYSIS

In this apriori based framework, there are different steps related to whole methodology to find the frequent item-sets,

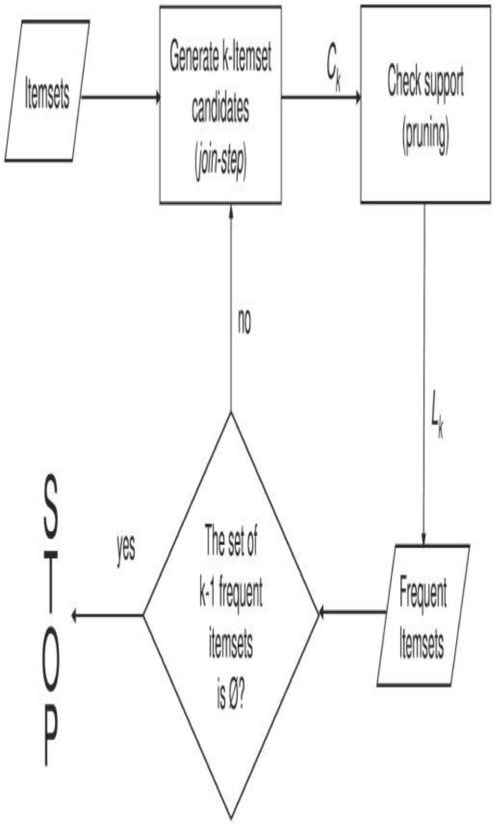

Itemsets
Generate k-Itemset candidates (join-step)
C_k
Check support (pruning)
no
L_k
STOP
yes
The set of k-1 frequent itemsets is Ø?
Frequent Itemsets

FIGURE 3.1 APRIORI BASED FRAMEWORK & ANALYSIS

Apriori algorithm uses a search order of the cycle-level approach (also known as the layers of iterative search method) to complete the set of frequent excavation work. The fig. 3.3 shows framework of Apriori Algorithm.

The characteristics of Apriori algorithm:
(1) Apriori optimizes the use of the method; the so-called Apriori is to optimize the use of Apriori nature of the optimization.
(2) Service database for the mining association rules.
(3) For sparse data sets, based on previous studies, this

method can only sparse data sets for the mining association rules, which is frequently set the length of the project smaller data sets.

Apriori based algorithm is given as follows:

```
Pass 1
    1. Generate the candidate itemsets in $C_1$
    2. Save the frequent itemsets in $L_1$
Pass k
1. Generate the candidate itemsets in $C_k$ from the frequent itemsets in $L_{k-1}$
    1.1 Join $L_{k-1}$ p with $L_{k-1}$q, as follows:
            insert into $C_k$
            select p.item$_1$, p.item$_2$, . . . , p.item$_{k-1}$, q.item$_{k-1}$
            from $L_{k-1}$ p, $L_{k-1}$q
            where p.item$_1$ = q.item$_1$, . . . p.item$_{k-2}$ = q.item$_{k-2}$, p.item$_{k-1}$ < q.item$_{k-1}$
    1.2 Generate all (k-1)-subsets from the candidate itemsets in $C_k$
    1.3 Prune all candidate itemsets from $C_k$ where some (k-1)-subset of
        the candidate itemset is not in the frequent itemset $L_{k-1}$
2. Scan the transaction database to determine the support for each
    Candidate itemset in $C_k$
3. Save the frequent item sets in $E_{lk}$
```

4.1 FIGURES AND TABLES

The system architecture has been depicted in figure 3.1. There are two modules in the proposed system. In the first module, following steps has been done.

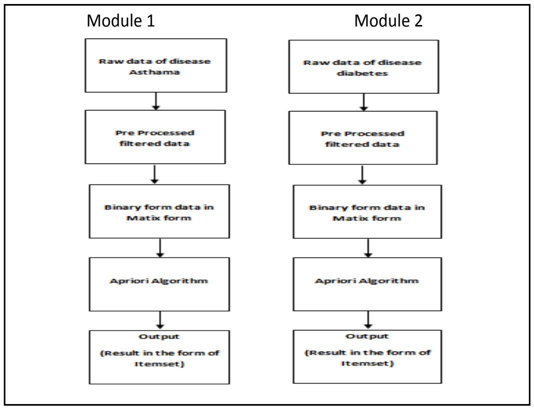

Module 1
Module 2
Raw data of disease Asthama
Raw data of disease diabetes
Pre Processed filtered data
Pre Processed filtered data
Binary form data in Matix form
Binary form data in Matix form
Apriori Algorithm
Apriori Algorithm
Output (Result in the form of Itemset)
Output (Result in the form of Itemset)

Figure 4.1 System Architecture Diagram

1. Raw data of asthma is collected.
2. Collected data of asthma has been filtered and preprocessed in the binary
   Form according to the data mining methodology.
3. Data of asthma will be changed into binary matrix to apply it into application.
4. Application software based on apriori algorithm.
5. Finally analysis of asthma is done by the developed application software.

In the second module, following steps has been done.

1. Raw data of diabetes is collected.
2. Collected data of diabetes has been filtered and preprocessed in the binary
   form according to the data mining methodology.
3. Data of diabetes will be changed into binary matrix to apply it into application.
4. Application software based on apriori algorithm.
5. Finally analysis of diabetes is done by the developed application software.

4.1.1 Acquiring data sets

In data mining, collection of right data and use of this data in proper way is most important task. To perform data mining, Different patient's data is collected with their different symptoms, sign and their corresponding treatments. Before applying the acquired data of diabetes and asthma, this is highly required that it must be pre-processed according to the requirement of association rule based algorithm for analysis. This data is changed into a matrix form. This matrix keeps data in binary form which is actually used, in which '1' is representing the existence of particular sign & symptoms and '0' is representing the nonexistence of particular sign, symptoms or corresponding treatments. In this data set, following form of data is used for asthma.

1	2	3	4	5	6	7	8	9	10
1	0	1	0	1	1	0	1	0	1
1	1	0	0	1	0	1	1	0	0
1	1	0	1	1	0	1	0	0	1
1	0	0	1	0	1	0	1	1	1
1	1	0	1	0	0	1	1	0	1

Table4.1.1 Binary representation of Sign, Symptoms & Treatment

There are ten columns, Its four columns representing symptoms and next three representing heredity, environment, sex respectively and last three are representing medications (1**:** wheezing & shortness of breath, 2: cough & cold, 3: any worse diseases, 4: exercise induced shortness of breath 5:heredity, 6:environment, 7:sex 8:only medication 9: steroids based medication 10: inject able steroids based medication)

Table 3.1 shows a binary form data in which every column representing a particular symptom or treatment. “1” is representing the existence of the symptom or treatment and “0” is representing absence of the symptom or treatment. In this data set, following form of data is used for diabetes.

4.2 Flow diagram & implementation of apriori algorithm

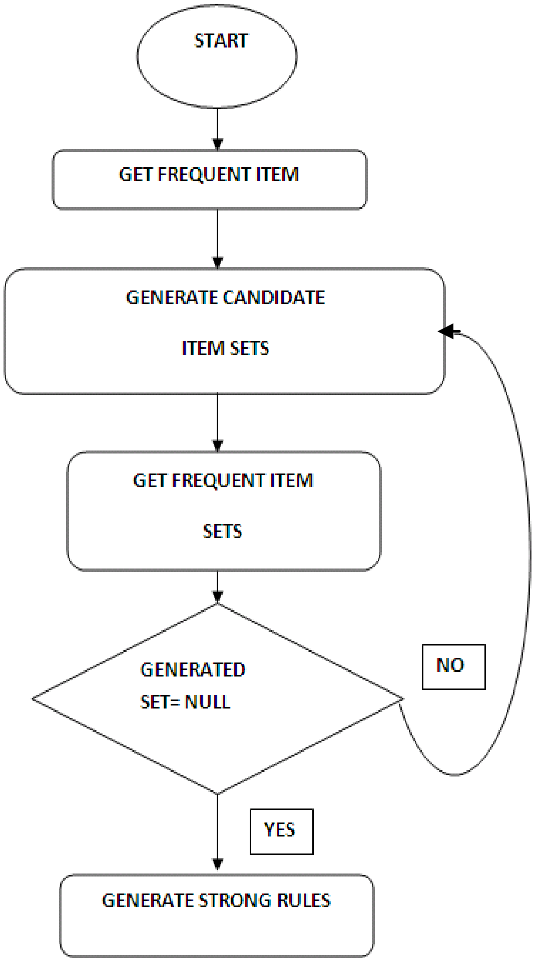


Figure 4.2 : Flow diagram & Implementation of Apriori algorithm

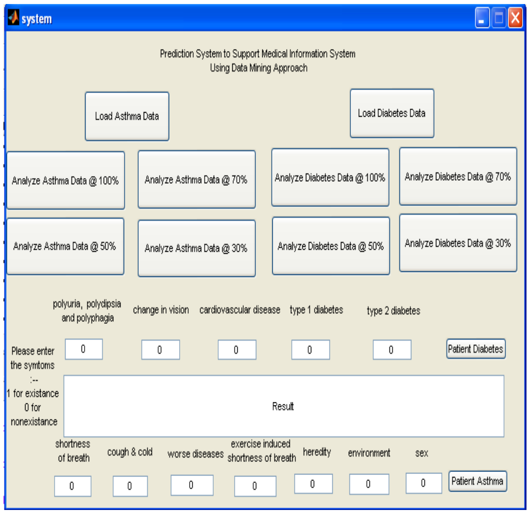
Previous given apriori based algorithm is used for analysis. it is applied on 2000 data sets (two data set each contain 1000 data of each disease). It gave different item sets of different symptoms and treatment according to different minimum support (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or100).

4.3 IMPLEMENTATION RESULTS

A computer aided model has been developed for the analysis of diseases. Two different concept have been dealt in this research. In this thesis, a broaden and advance itemset generation methodology has been assembled and developed for excavating the connections of the symptoms and disorders in the asthma and diabetes so that will produce hybrid dimension association rules. The implementation of the algorithm is using matlab 7 software. The methodology of the developed software finds the frequent illnesses with medication and generates association rules using Apriori algorithm. The developed software can be employable for huge medical and health data sets for building association rules for confusions frequently seen in the patient and concluding the relationship of the health disorders and symptoms viewed simultaneously. It facilitates important facts like correlations between medical issues related to disease.

4.3.1 Graphical user interface

A graphical user interface has been made in Matlab 7 to do different tasks. In this system many options are provided to users. The options are:
1. Load Asthma Data (This Option load the data set of asthma)
2. Load Diabetes Data (This Option load the data set of diabetes)
3. Analyze Asthma Data @ 30% (This Option analyze the data set of asthma at 30% minimum support


system
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
Patient Diabetes
Result
Patient Asthma

Prediction System to Support Medical Information System
Using Data Mining Approach

Please enter the symtoms :-- 1 for existance 0 for nonexistance

polyuria, polydipsia and polyphagia	change in vision	cardiovascular disease	type 1 diabetes	type 2 diabetes
0	0	0	0	0

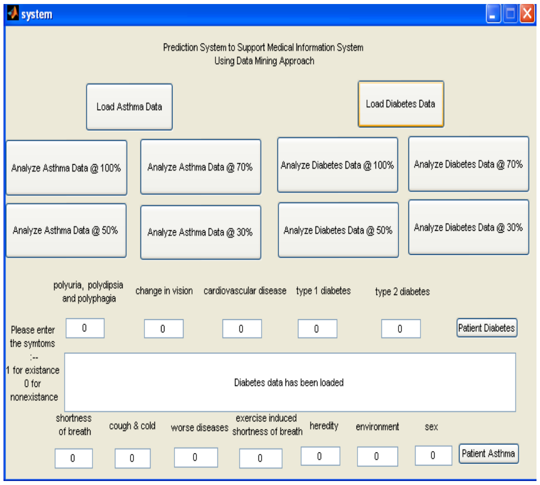
shortness of breath	cough & cold	worse diseases	exercise induced shortness of breath	heredity	environment	sex
0	0	0	0	0	0	0

Figure 4.3.1 Graphical User Interface

4. Analyze Asthma Data @ 50% (This Option analyze the data set of asthma at 50% minimum support)
5. Analyze Asthma Data @ 70% (This Option analyze the data set of asthma at 70% minimum support)
6. Analyze Asthma Data @ 100% (This Option analyze the data set of asthma at 100% minimum support)
7. Analyze diabetes Data @ 30% (This Option analyze the data set of diabetes at 30% minimum support)
8. Analyze diabetes Data @ 50% (This Option analyze the data set of diabetes at 50% minimum support)
9. Analyze diabetes Data @ 70% (This Option analyze the data set of diabetes at 70% minimum support)
10. Analyze diabetes Data @ 100% (This Option analyze the data set of diabetes at 100% minimum support)
11 Last blank box is used to display concluded result.

4.3.2 Implementation result loading asthma data

In this part Load asthma data option has been opted to load the data for further processing. Result box shows the status about data loading.


system
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

Prediction System to Support Medical Information System
Using Data Mining Approach

Please enter the symtoms :-- 1 for existance 0 for nonexistance

0 0 0 0 0

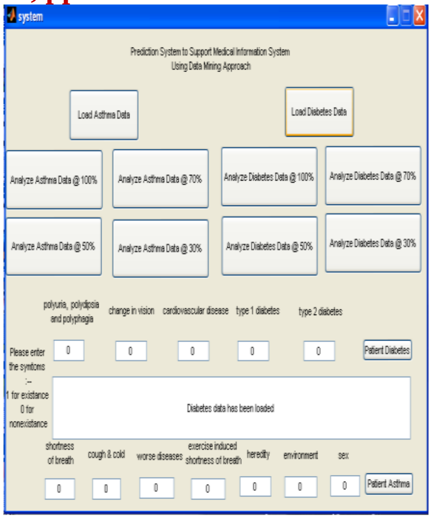
Diabetes data has been loaded

0 0 0 0 0 0 0

Figure 4.3.2 Implementation result loading asthma data

4.3.3 Implementation result loading diabetes data

In this part Load diabetes data option has been opted to load the data for further processing. Result box shows the status about data loading.


system
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

Prediction System to Support Medical Information System
Using Data Mining Approach

Please enter the symtoms :-- 1 for existance 0 for nonexistance

0 0 0 0 0

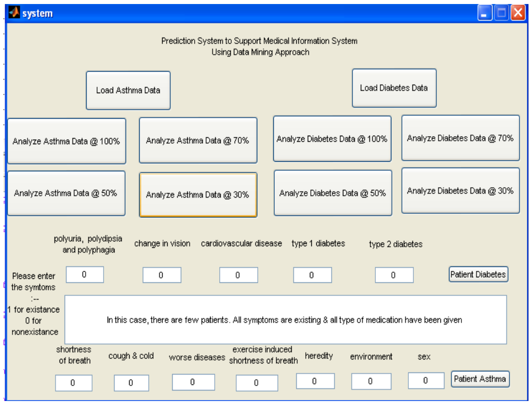
Diabetes data has been loaded

0 0 0 0 0 0 0

Figure4.3.3 Implementation result loading diabetes data

4.3.4 Implementation result after analyzing asthma data at 30% minimum support

In this part Analyze Asthma Data @ 30% option has been opted to analyze Data when minimum support is 30%. Result box shows the output (In this case there are few patients, all symptoms are existing and all type of medication have been given). This result emphasizes that there are 30 % cases showing all symptoms and getting all possible treatments.


system
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

Prediction System to Support Medical Information System
Using Data Mining Approach

Please enter the symtoms :-- 1 for existance 0 for nonexistance

0 0 0 0 0

In this case, there are few patients. All symptoms are existing & all type of medication have been given

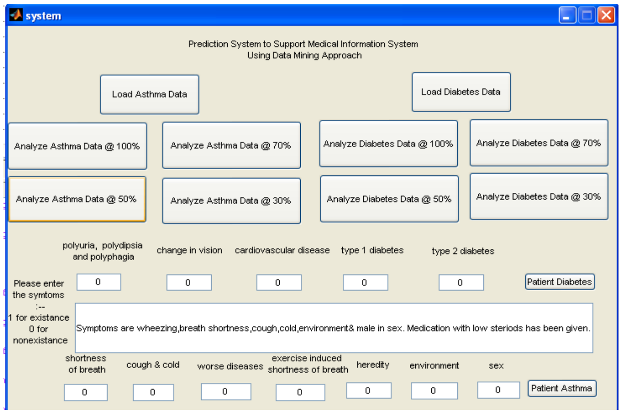
0 0 0 0 0 0 0

Figure 4.3.4 Implementation Result after analyzing asthma data at 30% minimum support

4.3.5 Implementation Result after analyzing asthma data at 50% minimum support

In this part Analyze Asthma Data @ 50% option has been opted to analyze Data when minimum support is

50%. Result box shows that there are 50 % cases showing only wheezing & shortness of breath, cough & cold, bad environment and male sex as positive symptoms and getting only medication with low steroids based treatments.


system
Prediction System to Support Medical Information System
Using Data Mining Approach
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
Please enter the symtoms :-- 1 for existance 0 for nonexistance
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

0 0 0 0 0

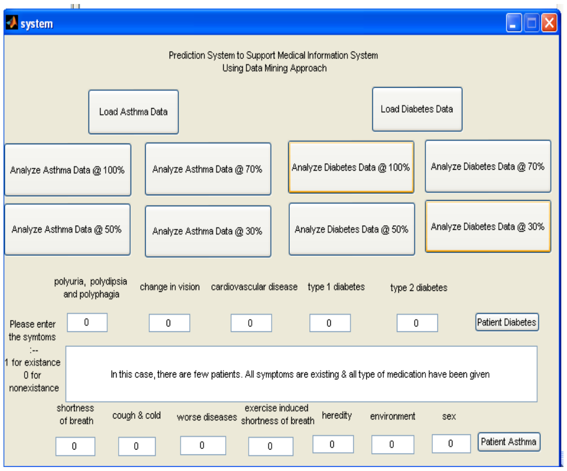
Symptoms are wheezing,breath shortness,cough,cold,environment& male in sex. Medication with low steriods has been given.

0 0 0 0 0 0 0

Figure 4.3.5 Implementation Result after analyzing asthma data at 50% minimum support

4.3.6 Implementation result after analyzing diabetes data at 30% minimum support

In this part Analyze diabetes Data @ 30% option has been opted to analyze Data when minimum support is 30%. Result box shows that there are 30 % cases showing all symptoms and getting all possible treatments.


system
Prediction System to Support Medical Information System
Using Data Mining Approach
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
Please enter the symtoms :-- 1 for existance 0 for nonexistance
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

0 0 0 0 0

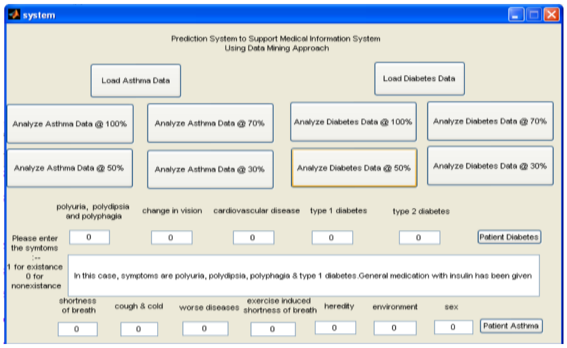
In this case, there are few patients. All symptoms are existing & all type of medication have been given

0 0 0 0 0 0 0

Figure 4.3.6 Implementation Result after analyzing

4.3.7 Implementation result after analyzing diabetes data at 50% minimum support

In this part Analyze diabetes Data @ 50% option has been opted to analyze Data when minimum support is 50%. Result box shows that there are 50 % cases showing only polyuria, polydipsia, polyphagia and type 1 diabetes as positive symptoms and getting only medication with insulin based treatments.


system
Prediction System to Support Medical Information System
Using Data Mining Approach
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
Please enter the symtoms :-- 1 for existance 0 for nonexistance
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

0 0 0 0 0

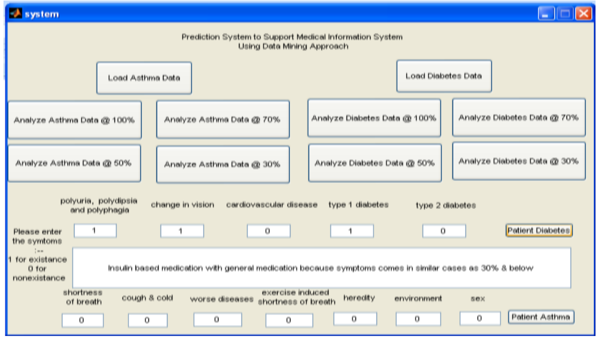
In this case, symptoms are polyuria, polydipsia, polyphagia & type 1 diabetes.General medication with insulin has been given

0 0 0 0 0 0 0

Figure 4.3.7 Implementation Result after analyzing diabetes data at 50% minimum support

4.3.8 Implementation result after enter new patient diabetes symptoms

In this part Analyze diabetes Data when we enter symptom of disease (polyuria, polydipsia and polyphagia, change in vision and type 1 diabetes) as positive. Result box shows the output to getting Insulin based medication with general medication because symptoms come in similar cases as 30% and below.


system
Prediction System to Support Medical Information System
Using Data Mining Approach
Load Asthma Data
Load Diabetes Data
Analyze Asthma Data @ 100%
Analyze Asthma Data @ 70%
Analyze Diabetes Data @ 100%
Analyze Diabetes Data @ 70%
Analyze Asthma Data @ 50%
Analyze Asthma Data @ 30%
Analyze Diabetes Data @ 50%
Analyze Diabetes Data @ 30%
polyuria, polydipsia and polyphagia
change in vision
cardiovascular disease
type 1 diabetes
type 2 diabetes
Patient Diabetes
Please enter the symtoms :-- 1 for existance 0 for nonexistance
shortness of breath
cough & cold
worse diseases
exercise induced shortness of breath
heredity
environment
sex
Patient Asthma

1 1 0 1 0

Insulin based medication with general medication because symptoms comes in similar cases as 30% & below

0 0 0 0 0 0 0

Figure 4.3.8 Implementation result after enter new patient diabetes symptoms

5.1 CONCLUSION

A computer aided model has been developed for the analysis of diseases. This research uses the methodology based on the Apriori algorithm, calculating multidimensional association rules by the frequent predicate verb set, has good scalability, it can deal with large data, Mining result is to display the correlation between data (association rules) along with information on support that can be analyzed. Information provides additional considerations for the user in decision making further. Applications can process data recapitulation of the patients at any hospital to find frequent item set that meets the minimum support and generate Dimension Hybrid Association Rules. Specifically, the whole application software finds the frequent illnesses with medication The developed software can be useable for huge medical and health data sets for building association rules for confusions frequently seen in the patient and concluding the relationship of the health disorders and symptoms viewed simultaneously. This research provides important facts like correlations between medical issues related to disease finally.

5.2 FUTURE DIRECTIONS

In this thesis, supervised learning techniques have been used. this thesis can be further extended by using various other supervised and unsupervised learning techniques like density based clustering, Naive Bayes based classification and frequent pattern matching based association. It is also possible to extend the proposed methodology on other type of disease and analysis can be carried out.

REFERENCES

1) Yuguang Y., Chunyan W., Min L. "Application of the Model_Multi based on Apriori algorithm in Supporting System of Medical Decision" IEEE Third International Conference on Measuring Technology and Mechatronics Automation 2011.

2) Rostianingsih S., Budhi G. S., Dwijayanti N.W.Y. "Hybrid-Dimension Association Rules for Diseases Track Record Analysis at Dr. Soetomo General Hospital" IEEE International Conference on Uncertainty Reasoning and Knowledge Engineering 2011

3) Han, J., Kamber, M.: "Data Mining Concepts and Techniques", Morgan Kaufmann Publishers, 2006.
4) http://www.asthama histry.com

5) Kaur, H., Wasan, S. K.: "Empirical Study on Applications of Data Mining Techniques in Healthcare", Journal of Computer Science 2(2), 194-200, 2006.

6) Sellappan, P., Chua, S.L.: "Model-based Healthcare Decision Support System", Proc. Of Int. Conf. on Information Technology in Asia CITA'05, 45-50, Kuching, Sarawak, Malaysia, 2005
7) Blake, C. L., Mertz, C.J.: "UCI Machine Learning Databases", http://mlearn.ics.uci.edu/databases/heart-disease/, 2004.

8) Chapman, P., Clinton, J., Kerber, R. Khabeza, T., Reinartz, T., Shearer, C., Wirth, R.: "CRISP-DM 1.0: Step by step data mining guide", SPSS, 1-78, 2000.

9) http://www.diabetes .com

10) http://www. Asthma .com

11) Geodic, P.: "Applied Data Mining: Statistical Methods for Business and Industry", New York: John Wiley, 2003.

12) Mehmed, K.: "Data mining: Concepts, Models, Methods and Algorithms", New Jersey: John Wiley, 2003..

13) Mbenshain, M.K: "Application of Data Mining Techniques to Healthcare Data Infection Control and Hospital Epidemiology, 25(8), 690–695, 2004. 45-50, Kuching, Sarawak, Malaysia,

14) http://www.diabeteshistry .com/

15) Wu, R., Peters, W., Morgan, M.W.: "The Next Generation Clinical Decision Support: Linking Evidence to Best Practice", J Healthcare Information Managment. 16(4), 50-55, 2002.

16) Farley and Piatetsky-Shapiro, 1996. KnowledgeDiscovery in Databases: An Overview. The AAAI/MIT Press, Menlo Park, C.A.
17) Glamour, C., D. Madigan, D. Presidion and P.Smyth, 1996. Statistical inference and data mining. Communication of the ACM, pp: 35-41.

18) Shams, K. and M. Frashita, 2001. Data Warehousing Toward Knowledge Management.
19) Jones, A.W., 1990. Physiological Aspects of Breath-Alcohol Measurements. Alcohol Drugs Driving.
20) Han, J. and M. Kamber, 2001. Data Mining: Concepts and Techniques. San Francisco, Morgan Kauffmann Publishers.
21) Lu, H., R. Section and H. Liu, 1996. Effective data mining using neural networks. IEEE Trans. On Knowledge and Data Engineering.

22) Miller, A., B. Blot and T. Hams, 1992. Review of neural network applications in medical imaging and signal processing. Med. Biol. Engg. Comp.

23) Miller, A., 1993. The application of neural networks to imaging and signal processing in astronomy and medicine. Ph.D. Thesis, Faculty of

Science, Department of Physics, University of Southampton.

24) Weinstein, J., K. Kohn and M. Graver *et al*., 1992.Neural computing in cancer drug development: Predicting mechanism of action. Science, 258: 447-451.

25) Stanford, G.C., P.E. Kelley, J.E.P. Skye, W.E. Reynolds and J.F. Todd, 1984. Recent improvements in and analytical applications of advanced ion-trap technology. Intl. J. Mass Spectrometry Ion Processes, 60: 85-98.

26) Robinson, P.J., 1997. Radiology's Achilles' heel: Error and variation in the interpretation of the Roentgen image. Radiol. Brit. J.

27) Itchhaporia, D., P.B. Snow, R.J. Almassy and W.J. Oetgen, 1996. Artificial neural networks: Current status in cardiovascular medicine.

28) Schnorrenberg, F., C.S. Pastiches, C.N. Schemas, K. Kyriako and M. Vassiliou, 1996. Computer aided classification of breast cancer nuclei.

29) Choy, H.K., T. Jarring, E. Bentsen, J. Vashon, K. Western, P.U. Maelstrom and C. Bausch, 1997. Image analysis based carcinoma. Comparison of object