

DATA MINING TECHNIQUES FOR VERIFICATION OF MEDICINE CONTENTS RELATED TO CARDIAC PROBLEMS

Shaikh Abdul Hannan ¹, Pravin Yannawar², Dr. R.R. Manza ², Dr. R. J. Ramteke ³,

¹Vivekanand College, Aurangabad, India (hannan_7us@yahoo.com)

² Department of Computer Science and Information Technology, Dr. B.A.M.U., Aurangabad

³ Department of Computer Science and IT, Reader, NMU, Jalgaon (rakeshramteke@gmail.com)

ABSTRACT: To handle application of multidisciplinary areas, the knowledge discovery and data mining (KDD) field has made significant progress in the past decade. The field is vibrant with significant impact on a wide variety of science, business and technology areas where it has become necessary to deal with exponentially increasing volume of data. In this paper the association rules as a means of identifying relationships among sets of symptoms, side effects and medicine, which can be used to evaluate trends and classify groups. This concept is applied on medicine database to identify diseases on the basis of symptoms and by avoiding medicine side effects prepare prescription for patients. In this paper more than 100 patients' data have been collected and output is verified by expert Doctor. As per Doctors opinion this output is nearby their outputs. The outcome of the result shown here will be useful for Doctors to give appropriate medicine to patient on the basis of symptoms [1][2][3]

KEYWORDS: Data Mining, Medicine, Symptoms, Disease, Side effects.

Introduction: The scalability of mining algorithms has become a major research topic. One approach to the scalability problem is to run data mining algorithm on a small subset of the data. This strategy can yield useful approximate results in a fraction of the time required to compute the exact solution, thereby speeding up the mining process by orders of magnitude[4][5][6][7][8][9][10]. Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information. Information that can be used to increase revenue, cuts costs, or both[11][12][13][14]. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified [15][16][17][18]. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. One of the reasons behind maintaining any database is to enable the user to find interesting patterns and trends in the data.[16][17][18][19][20].

Heart disease symptoms and medicine :

Coronary artery disease, heart attack -- each type of heart disease has different symptoms, although many heart problems have similar symptoms The

symptoms you experience depend on the type and severity of your heart condition. Learn to recognize your symptoms and the situations that cause them. Call your doctor if you begin to have new symptoms or if they become more frequent or severe. The most common symptom is angina. Angina can be described as a discomfort, heaviness, pressure, aching, burning, fullness, squeezing or painful feeling in your chest. It can be mistaken for indigestion or heartburn.

MID	Medicine Name	MSE
22	Corticosteriod	35,36,37,38,39,40,41
23	Aspirin	1,2,3,4,42,43,44
24	Ticlopidine	1,2,6,26,46

Table 1: Medicine ID, Medicine name and Medicine side Effect

SEID	Side Effect Description
40	Myopathy
41	Alopecia
42	Intolerance

Table 2 : Side effect ID and its description

DID	Disease Name	SID1
1	Heart attack with chest pain	1,2,3,12,13
2	Heart attack with breathlessness	1,3,12,13,16
3	Heart attack with palpitation	1,12,13,16,20

Table 3 Disease ID, Disease Name and Symptoms ID

SID	Symptom Description
11	weakness
12	sweating
13	perspiration

Table 4 Symptoms ID and Symptoms Description

DID	MEDID
1	4, 10,56,60,63,71,72,73
2	4,10,56,60,71,72
3	4,10,56,60,63,71,72,73

Table 5 Diseases ID and Medicine ID

Experimental Analysis : In experiment samples from 100 patients has been collected from Sahara Hospital, under the guidance of Dr. Abdul Jabbar (MD Medicine). According to observation of

samples the information has been arranged in formats like, Medicine which consists of Medicine Id, Medicine Name and Medicine Side effect as shown in table 1. Where as table 2 shows the side effect ID and description about side effect. The corresponding ID Number of side effect of table 2 has written in appropriate row of medicine ID and Medicine name in table 1. Table 4 has given the information about symptoms description and its ID number. The Symptoms ID number has been listed out in the corresponding rows of disease ID and disease description in table 3. The last table 5 as given the relationship among disease ID with the suitable medicine by giving medicine ID number along with disease ID number. In this way the samples information has been arranged in various tables. Each table consists of around 150 entries out of them few important has given in tables 1 to 5.

In experimental work more than 100 patients information, symptoms of disease are taken as input by considering all physical examination were normal. Based on given inputs the appropriate medicines were suggested as prescription for patient in result. By applying set of association rules to avoid such medicine which are harmful to the patient the final prescription is generated as output. In the experimental work near about 100 side effects, 75 medicines, 10 different diseases few symptoms has stored in data base. The results produced are in various forms like individual patient, all patient, disease and medicine information, disease with symptoms report, specific disease wise symptom report, all patients, symptom and suggested medicine information, individual patient, symptom and suggested medicine information, specific disease and medicine information etc. One sample output of individual patient which consists of patient name, age, symptoms, disease and list of suggested medicine. Like this more than 5 different formats as used to produce patient reports as output of this research work. All results are satisfactory and verified by Dr. Abdul Jabbar (MD Medicine).

Individual report of the patient :

Jamkar Gangadhar Keshawrao Age 66
Bhim nagar Osmanpura, A'bad.

PE Y -->Y: YES, Physical Examination,
Aurangabad found Normal (* PE : Physical Examination)

Symptom Description

Chest pain, Breathlessness, Palpitation

List of Possible Disease

Heart attack with chest pain

List of Suggested Medicine

Atorvastatin, Alprazolam, Clopidogrel,
Trimetazidine, Ramipril, Iso sorbide dinitrate,
Metoprolol, Nikorandil

Specific Disease and Medicine Information

Heart attack with chest pain

Medicine Name

Iso sorbide dinitrate, Ramipril, nikorandil,
trimetazidine, metoprolol, clopidogrel, atorvastatin,
alprazolam

Conclusion

This work is based on association rules of data mining technique for verification of medicine contents related to the cardiac problems. In the experimental work it has been already considered in all the cases of patients that the physical examination of all the patients is found normal. In the result the information generated gives the relation about disease with medicine, disease with symptoms, patient details with medicine. The final results have been verified by Dr. Abdul Jabbar (MD Medicine). He is satisfactory about the work and all results. Further he has suggested that the work can extend in future by considering physical examination along with current inputs to generate good results. Therefore according to his suggestions this work in future may be extended to that direction

References:

1. Keon-Myung Lee, "Mining generalized fuzzy quantitative association rules with fuzzy generalization hierarchies", International Conference, IFSA World Congress and 20th NAFIPS, Page(s):2977 - 2982 vol.5, 25-28 July 2001.
2. De-Xing Wang; Xue-Gang Hu; Xiao-Ping Liu; Hao Wang; Jun Guo, "Research on model of association rules mining with added-newly measure criteria", Proceedings of 2004 International Conference on, Machine Learning and Cybernetics, Page(s):1187 - 1190 vol.2 , 2004.
3. Rastogi, R.; Kyuseok Shim, "Mining optimized association rules with categorical and numeric attributes", IEEE Transactions on, Knowledge and Data Engineering, Volume 14, Issue 1, Page(s) : 29 – 50, Jan.-Feb. 2002.
4. Yingjiu Li; Peng Ning; Wang, X.S.; Jajodia, S., "Discovering calendar-based temporal association rules", Proceedings. Eighth International Symposium on, Temporal Representation and Reasoning, Page(s):111 – 118, 14-16 June, 2001.
5. Xunwei Zhou; Hong Bao, "Mining Double-Connective Association Rules from Multiple Tables of Relational Databases", International Conference on, Computer Science and Software Engineering, Page(s):271 – 279, Volume 4, 12-14 Dec. 2008 .
6. Ishibuchi, H.; Yamamoto, T.; Nakashima, T., "Determination of rule weights of fuzzy association rules", The 10th IEEE International Conference on, Page(s):1555 – 1558, Volume 3, 2-5 Dec. 2001.