Mining Medicinal Information using Flowers

S.Vigneswari
M.Phil Research Scholar
Alagappa University
Karaikudi,India

S.S.Dhenakaran
Faculty ,Comp.sci & Engg.
Alagappa University
Karaikudi,India

ABSTRACT

Mining is the art of finding unknown patterns from voluminous data. Medical data mining becomes major concern now a day. The objective of the proposed work is the association of data and prediction of treatment to cure diseases using flowers. The data source used here is the collection of flowers, which have separate medicinal values. Based upon the medicinal worthiness, the vitamin contents, in the flowers, prediction is made for the treatment of diseases. For this purpose, the association rule mining technique is taken into account. Association rules represent knowledge embedded in data sets as probabilistic insinuation and are closely related to computation of frequent item sets that occurs on the given data set.

Keywords

Data Mining, Association Rule Mining, Medicinal flowers, prediction.

1. INTRODUCTION

Data Mining is the process of extracting information from large data sets through Statistics, Data Base Management Systems and algorithms [6]. Data mining technology provides a user oriented approach to trace hidden patterns in the data. Data mining is a crucial step in discovery of knowledge from large datasets [14]. In recent years, Data mining has found its significant hold in every field including health care [15].

Mining process is more than the data analysis which also spans other disciplines like Data Warehousing, Statistics, Machine learning and Artificial Intelligence [12]. Significantly useful patterns or meaning in raw data has previously been called KDD (knowledge discovery in databases), data mining, and knowledge discovery [5]

Knowledge discovery is the "non-trivial process of identifying ultimately understandable patterns in data." Data mining "consists of applying data

Analysis and discovery algorithms that produce particular patterns over the data". Data mining is typically a bottom-up knowledge engineering strategy, Knowledge discovery involves the additional steps of target data set selection, data pre processing, and data reduction, which occur prior to data mining.

In the following, the performance perspective about data mining proposed by agarwal et.al and the association rules classification predicted for medical data mining given by deepika et.al, along with uniqueness of data mining is being presented. Agarwal et al. [2] proposed three basic classes of data mining problems which are association rules ,sequence number and clustering. Association rules are used to identify relationships among a set of items in databases. These relationships are not based on inherent properties of the data themselves, but rather based on co-occurrence of the data items. Every data mining technique serves a different purpose depending on the modeling objective. They are divided into two types named as classification and prediction. Classification models predict categorical labels, while prediction models predict continuous-valued functions. Decision Trees and Neural Net works use classification algorithms while Regression, Association Rules and Clustering use prediction algorithms. [4].

Medical Data mining help to provide appropriate treatment to the users. Discovery of hidden patterns and relationships often goes unexploited [4].Due to the complexities to use for data mining implementation because of its multi dimensional attributes. This type of data can include text, images, and videos also. Therefore, observation-using text, numerical, images and videos type data provide the complete judgment about the human's disease [11].

The paper is constructed as follows. Section II briefly elaborates the literature review based upon the usage of association rule mining in medical data mining. Section III dealt with the Association Rule Mining Algorithm for detailed view. Section IV comprises of proposed floral data mining and Section V concludes the paper.

2. REVIEW OF LITERATURE

This section briefly explains about various medical data mining applications, their algorithms, outcomes and their usage related to our association rule mining algorithm.

N.Deepika, K.Chandra Shekar, D.Sujatha [4] worked upon the Association Rules for classification of Heart–Attack Patients which presents about the various effective heart attack prediction system using Pruning-Classification Association Rule (PCAR), an Efficient Approach for mining Association Rules. A proficient methodology for the generation of association rules from the heart disease.

MA.Jabbar, Dr.pritichandra, BL.Deekshatulu, [10], proposes cluster based association rule mining for heart attack prediction. They suggest a new approach for association rule mining based on sequence number and clustering the transactional data base for heart attack prediction. They proposed a new algorithm which combines the concept of sequence numbers and Clustering.

Gang Fang et.al, proposed an algorithm for improved association rule mining. They used the method of binary Boolean calculation to generate candidate frequent item sets by computing sequence number degree, which is gained through computing these sequence number of all these items obtained by candidate frequent item sets [7].

Tipawan Silwattananusarn, KulthidaTuamsuk, shows survey on data mining and its application for knowledge management. The discussion on the findings is divided into 4 topics: (i) knowledge resource; (ii) knowledge types and/or knowledge datasets; (iii) data mining tasks; and (iv) data mining techniques and applications used in knowledge management. The techniques in the process of knowledge management are summarized and discussed [13].

3. ASSOCIATION RULE MINING ALGORITHM

Association rule mining one of the most important and well researched techniques in data mining was introduced by R.Agarwal[2].The problem of association rule mining can be decomposed into two sub problems.

- Discovering frequent item sets
- Generating rules from frequent item sets.

Association rules are nothing different from classification rules except that does not predict only class labels but also predict any other attribute. It has freedom to produce a combination of attributes. Different association rules convey different regularities. The association rule is the number of instances for which it predicts correctly this is often called its support.

Its accuracy often called confidence is the number of instances that it predicts correctly, expressed as proportion of all instances to which it applies. The user has to specify the minimum coverage and accuracy values and look for only those rules whose values are at least of the specified minimum value[16].

An association rule (AR) is a pair (X, Y) of sets of attributes, denoted by $X \rightarrow Y$. X is the antecedent and Y is the consequent of the rule $X \rightarrow Y$. The simplest parameters associated to an AR are its support and confidence. The support of a rule $X \rightarrow Y$ is the number of records that contain all items of X. Clearly, the confidence of $X \rightarrow Y$ is an estimation of the probability that a record that contains the items of X, chosen at random, will contain the items of Y [3].

Association rules can be classified based on the type of vales, dimensions of data, and levels of abstractions involved in the rule. If a rule concerns associations between the presence or absence of items, it is called Boolean association rule and the dataset consisting of attributes which can assume only binary (0-absent, 1-present) values is called Boolean database[17].

4. PROPOSED METHODOLOGY

The extraction of significant patterns from the given data set is presented in this section. The data warehouse consists of the information about the flowers and their medicinal worthiness. Initially, the data warehouse is preprocessed to make mining process more efficient. In the First stage of our proposed study, we used the association rule mining algorithm for the categorization of data.

4.1 Data Set

The data set contains the attributes such as the name of the flower, medicinal value it holds, name of the disease that it can able to cure, the vitamin extraction it contains etc.,

In the dataset, the attributes shows the medicinal value of the flowers, the vitamins they holds etc. The dataset was a sample 20 flowers with their name, medicinal value it holds, the

disease that if cures,, the vitamin content that the flower contains.

The Begonia, flower consists of vitamin C, which able to cure headache and rid of the body toxins. The Bellis Perinneis , contains the vitamin C, that heals the wounds, arthritis and rheumatism. The third is the Black Cohosh with the vitamin A. Black Cohosh can be used as an emmenagogue, which means that it stimulates the uterus. Women with menstrual problems can effectively use low doses of this flower to help regulate their cycles and relieve pain. It can also cures the snake bites.

The Butterfly Weed also contains vitamin A, has ability to cure respiratory, lung related issues. The calendula consists of the vitamin A, have the power to heal the wounds, cuts and burns. The California Poopy with the vitamin C mineral, have capacity to reduce anxiety and insomnia, as well as bladder problems in children and adults. The carnation flower with the vitamin D, to reduce anxiety, agitation, and stress and fatigue healing effect on the skin and can bring down swelling.

The Chrysanthemum consists of the vitamin A and C, has many medicinal uses, including an aid in recovery from influenza. Due to the longevity of their blooms, chrysanthemum flowers become a favorite of florists for arrangements. The well known flower Lotus that is enriched with vitamin C shows its effectiveness against fever, diarrhea and also more serious illnesses such as cholera and bronchitis.

The Corn Flower is highly cerium with vitamin A and C, which acts as a laxative and also as a marts cleanse. Jasmine flower with high mineral content used as an antiseptic, anti depressant, analgesic Jasmine Tea relieves stress and anxiety. The Tulips are used for gaining strength best for poultice for insect bites, bee sting, and ban run rashes with quick relief.

The Iris consists of the Vitamin C, cure wounds and excellent remedy for bites. Another highly available sunflower consists of calcium with the vitamin E used for Lung ailments & malaria, reduces high fever. The Peruvian Lity heals the visible inflammations and ulcers, and applied to hard ture to stop them equally.

The Dandelian with Vitamin A, should be used to cleans the blood & recycle nutrients use pancreas, stomach. The Orchid can be used to cure headache, cough, fever, aches, wounds, cure lung. The well known rose is highly enriched with vitamin C that increases blood circulation, relive depression. Improve skin condition cure stomas and for Controlling Cancer growth.

The Rosy Periwinkle can stub disease like diabetes, blood pressure, cancer & hodgkines's diseases. The Fox glove is used to treat heart failure disease. The valerian is one of the powerful vervine stimulant, and antiseptic. It cures cholera.

4.2 Association Rule Generation

The proposed method consists of the followed sub steps. As a first stage, the if …Then rule is applied.

IF conditions THEN conclusion

This kind of rule consists of two parts. The rule antecedent (the IF part) contains one or more conditions about value of predictor attributes where as the rule consequent (THEN part) contains a prediction about the value of a goal attribute. An accurate prediction of the value of a goal attribute will improve decision-making process. IF-THEN prediction rules are very popular in data mining.

They represent discovered with the knowledge at a high level of abstraction. In the health care system it can be applied as follows:

Symptoms -> Flower with Vitamin Detail -> the Disease can be cured.

Rule 1:

IF (Flower = Lupinus AND Medicine = Lubriderm SF (Cream)) THEN Diseases to be cured = helps relieves from Itchy Dry Skin.

Rule 2:

IF (Flower = Dandelion AND Medicine = Vitamin A Dry) Diseases to be cured = Support Eye Health and Immunity Function.

Rule 3:

IF (Flower = Carnation n AND Medicine = KAL Vitamin D) Diseases can be cured = Healthy skeletal growth, Strong bones ion.

Rule 4:

IF (Flower = Lotus n AND Medicine = Acroplex Forte) Diseases that can be cured = scurvy.

Rule 5:

IF (Flower = Bergonia AND Medicine = Celin)) THEN Diseases to be cured = cure headache and rid of the body toxins)

Rule 6:

IF (Flower = Bellis Perinnis AND Medicine = Cobadex Forte) Diseases to be cured = heals the wounds, arthritis and rheumatism

Rule 7:

IF (Flower = Black Cohosh AND Medicine = emmenagogue) Diseases can be cured = stimulates the uterus.

Rule 8:

IF (Flower = Butterfly Weed AND Medicine = Glacex) Diseases that can be cured = Ability to cure respiratory, lung related issues.

Rule 9:

IF (Flower = California Poppy AND Medicine = Zevit) THEN Diseases to be cured = Heal the wounds, cuts and burns.

Rule 10:

IF (Flower = Chrysanthemum AND Medicine = Retinol) Diseases to be cured = Recovery from Influenza.

Rule 11:

IF (Flower = Corn Flower AND Medicine = Carotnoids) Diseases can be cured = Acts as a laxative.

Rule 12

IF (Flower = Jasmine AND Medicine = mineral content) Diseases that can be cured = used as an antiseptic, anti depressant, analgesic.

Rule 13:

IF (Flower = Iris AND Medicine = cobadex Forte)) THEN Diseases to be cured = cure wounds, remedy for bytes.

Rule 14:

IF (Flower = Tulips AND Medicine = Glacex) Diseases to be cured = used for insect bites, bee sting, and run rashes with quick relif.

Rule 15

IF (Flower = Sunflower AND Medicine = calcium) Diseases can be cured = Used for Lung ailments and malaria.

Rule 16:

IF (Flower = Peruvian Lilly AND Medicine = Retinyl Palmitate) Diseases that can be cured = heals the visible inflammations and ulcers,

Rule 17:

IF (Flower = Orchid AND Medicine = Acroplex Forte) Diseases that can be cured = used to cure headache, cough, fever, aches, wounds, cure lung

Rule 18:

IF (Flower = Rose AND Medicine = Cobadex Forte)) THEN Diseases to be cured = increases blood circulation, relive depression)

Rule 19:

IF (Flower = Rosy Periwinkle AND Medicine = Beta Carotine) Diseases to be cured = can stub disease like diabetes, blood pressure, cancer & hodgkines's diseases.

The sample data set is given below:

Table 1. Sample Data Set

Flower	Medicine	Diseases that can be cured	Vitamin
Lupinus	Lubriderm SF(Cream)	Helps to relieve Itchy Dry Skin.	Vitamin E
Dandelion	Vitamin A Dry	Support Eye Health and Immune Function	Vitamin A
Carnation	KAL Vitamin D	healthy skeletal growth, strong bones	Vitamin D
Lotus	Acroplex Forte	scurvy	Vitamin C
Bergonia	Celin	cure headache and rid of the body toxins	Vitamin C
Bellis Perinneis	Cobadex Forte	heals the wounds, arthritis and rheumatism	vitamin C,
Black Cohosh	emmenagogue	stimulates the uterus	vitamin A.
Butterfly Weed	Glacex	ability to cure respiratory, lung related issues	vitamin A.
California Poopy	Zevit	heal the wounds, cuts and burns	vitamin C,
Chrysanthem um	Retinol	recovery from influenza	vitamin A
Corn Flower	Carotenoids	acts as a laxative	vitamin A
Jasmine	mineral content	used as an antiseptic, anti depressant, analgesic	vitamin A
Iris	Cobadex Forte	cure wounds and excellent remedy for bites	vitamin C
Tulips	Glacex	used for insect bites, bee sting, and ban run rashes with quick relief.	vitamin A
Dandelian	Vitamin A Dry	used to cleans the blood & recycle nutrients use	vitamin A

		pancreas	
sunflower	calcium	used for Lung ailments & malaria, reduces high fever	vitamin E
Peruvian Lity	Retinyl Palmitate	heals the visible inflammations and ulcers,	vitamin A
Orchid	Acroplex Forte	used to cure headache, cough, fever, aches, wounds, cure lung	vitamin c
Rose	Cobadex Forte	increases blood circulation, relive depression	vitamin C
Rosy Periwinkle	Beta Carotine	can stub disease like diabetes, blood pressure, cancer	vitamin A

5. RESULT AND DISCUSSION

The proposed method is to find the chances of can stub disease like diabetes, blood pressure, cancer Diseases based on input data. The application was built in Microsoft platform using Visual Studio 2010.

Fig. 1 Prediction window based on input.

An output window predicting the chances of can stub disease like diabetes, blood pressure, cancer diseases based on input data is represented in Fig I. Also the chances of occurrence various medicine, flower. Vitamin details and diseases to be cured are contained in our database.

6. CONCLUSION

The proposed implementation work is based upon the If... Then rule generation. It mainly concern for the decision making. The proposed work nearly consists of 400 flowers along with their attributes mentioned as in the table given above. It is better from other in the way that instead of using the production of drugs or pharmaceutical materials, consequent flowers can be used as an alternative. By this, considerable cost can be reduced for the production of drug manufacturing.The proposed article is aimed to present medical information with flowers, which are seen in the world. It brings useful information to the user to heal diseases with early available flowers. The dataset used is only a sample, which can be increased to include the medical value with small plants and its stems. It can be referred for drug manufacturing also.

7. REFERENCES

[1] Abdullah Saad Almalaise Alghamdi, *"Efficient Implementation of FP Growth Algorithm-Data Mining on Medical Data"*, in International Journal of Computer Science and Network Security, VOL.11 No.12, December 2011, pp: 7 – 16.

[2] Agrawal, R., Imielinski, T. & Swami, A. (1993). *"Database Mining: A Performance Perspective"*. IEEE Transactions on Knowledge and Data Engineering5 (6): 914–925.

[3] Dan A. Simovici, *"Data Mining of Medical Data: Opportunities and Challenges in Mining Association Rules "*, pp: 1-25.

[4] N.Deepika, K.Chandra Shekar, D.Sujatha, *"Association Rules for classification of Heart–AttackPatients"*, in International Journal of Advanced engineering Sciences and Technologies, Vol 11, issue 2, pp:253 – 257.

[5] Fayyad, U. M., Piatetsky-Shapiro, G. & Smyth, P. (1996a). *"From Data Mining to Knowledge Discovery: An Overview"*. In Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. & Uthurusamy, R. (eds.)Advances in Knowledge Discovery and Data Mining, 1–36. AAAI Press/MIT Press.

[6] Feelders, A., Daniels, H. and Holsheimer,M. (2000) '*Methodological and Practical Aspects of Data Mining'*, Information and Management, pp.271-281.

[7] Gang Fang, Zu-Kuan Wei, Yu-Lu Liu," *"An algorithm of improved association rules mining"*, Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009.

[8] Hai Wang, Shouhong Wang, *"Medical Knowledge Acquisition through Data Mining",* Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education.

[9] Hian Chye Koh and Gerald Tan, *"The data mining applications in healthcare management"*, in Journal of Healthcare information Management — Vol. 19, No. 2, pp: 64 – 72.

[10] MA.Jabbar,Dr.pritichandra, L.Deekshatulu,, *"Cluster based association rule mining for heart attack prediction"*, in Journal of Theoretical and Applied Information Technology", vol 32, no 2, pp: 196 – 201.

[11] Krzysztof J. Cios and G. William Moore, "*Uniqueness of Medical Data Mining*", Artificial Intelligence in Medicine Journal, 2002.

[12] Larose T. Daniel (2005). *"Discovering Knowledge in Data: An Introduction to Data Mining"*, John Wiley & Sons, Inc., Hoboken, New Jersey

[13]Tipawan Silwattananusarn, KulthidaTuamsuk*," Data Mining and Its Applications for Knowledge Management* .

[14] Mrs S.M Uma and Dr.E.Kirubakaran,"*Intelligent Heart Diseases Prediction System Using A New Hybrid Metaheuristic Algorithm*",International Journal of Engineering Research & Technology (IJERT)Vol. 1 Issue 8, October – 2012..

[15] M. ANBARASI,E. ANUPRIYA and N.CH.S.N.IYENGAR, "*Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic*", International Journal of Engineering Science and Technology Vol. 2(10), 2010, 5370-5376 .

[16] Ramya Rathan,Shridhar and R Balasubramanian S ,"*Association Rule- Spatial Data Mining Approach for Exploration of Endometrial Cancer Data*", International Journal of Advanced Research in Computer Science and Software Engineering,Volume 3, Issue 10, October 2013 ISSN: 2277 128X.

[17] P. Sunil Kumar and Ashok Kumar Panda," *USE OF ASSOCIATION RULE MINING IN HIGHER SECONDARY EDUCATION IN ODISHA* ",International Journal on Advanced Computer Theory and Engineering (IJACTE),ISSN (Print) : 2319 – 2526, Volume-2, Issue-6, 2013