International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Differential Diagnosis in Medical Field Using SOA and Classifier Techniques: A Review

Rashmi M. Jadhav[1], Roshani Ade[2]

[1]D.Y. Patil School of Engineering and Technology, Lohgaon, SPPU, Pune, Maharashtra, India

[2]D.Y. Patil School of Engineering and Technology, Lohgaon, SPPU, Pune, Maharashtra, India

Abstract: ***Recent diagnosis in medical field reveals that the probable causes of failure in patients treatment is the misdiagnosis factor that leads to inadequate treatment of the patients. This paper focuses on the misdiagnosis attribute that if mined properly may produce accurate results. The data obtained in hospitals is vast and contains enormous amount of information. Findings for the patients symptoms is based on the current condition of the patient but in some cases the patient's medical history also plays a crucial role. The misdiagnosis attribute is taken into consideration in this paper. The system developed predicts the probability of a disease using an algorithm that combines the key features of neural networks, large memory storage and retrieval and k-NN classifier. Thus diagnosis of a patient's ailment can be made effectively using such algorithm.***

Keywords: Data Mining, Diagnosis, Service-oriented architecture (SOA), knowledge discovery and symptoms.

1. Introduction

Technology has made a rapid development in the field of diagnosis in medical sciences. This has also helped doctors in different ways from surgery images to X-ray photography. But technology is still lagging behind for diagnosis of sheer variables that constitute different factors. These factors involve medical history to geographical conditions. No medical model has been successful yet to analyse such variables [1].Medical decision support systems are coming into picture to help doctors assist in decision making.

Diseases showing similar symptoms are hard to predict. Factors leading to misdiagnosis may differ with inexperience of doctors. Doctors may also be in a restless state of mind which may affect their diagnosis. In some cases doctors may diagnose a particular symptom depending on their habitual and repeated diagnosis and also factors due to misinterpretations, ambiguities in symptoms, and inadequate information. Research scholars have been striving hard to diagnose this mislead factor so as to assist doctors in treatment of diseases.

Algorithms used sometimes superintend the variables involved in symptoms of a particular patient. These factors may include predominant conditions, medical history of the patient, family history. All the misdiagnosis is due to only unknown values of complete variables. Research depicts that such values can be made available if the medical data is mined properly and adequately. The paper deals with such mining in medical field.

Classification of diseases is based on differential diagnosis method [2].This is done by the doctors by slandering down the differential process in steps that range from root cause of the disease to its treatment. A list of similar symptoms is traversed through to get the exact matching disease to the input symptom. If only one symptom is input then the algorithm returns in lesser time giving output as per the input of symptom. But if the symptoms from the patient are large in number then the complexity of the algorithm increases. Experienced doctors use classifiers to reach to the ground level of the disease. This is accomplished by knowledge of doctors and their previous experience in curing the disease. However this needs skill of doctors to some extent. The problem gets intensified if the doctor is new and lacks training. The scenario comes into focus in developing countries [3].

This paper proposes an inventive methodology in assisting doctors and making treatment results better, thus making tougher task easier to a large extent. Smart pattern matching technique is used by including k-NN classifier [4] and the next probable diseases by performing differential diagnosis using the Hopfield neural networks theory and Large Memory Storage and Retrieval (LAMSTAR) networks. The algorithm outputs the disease from the symptoms entered and gives the next probability of disease so that the necessary treatment can be decided upon.

The system was built on service-oriented architecture (SOA) [8]and has been implemented on a web server to be accessible by anyone including doctors from rural and remote areas. After entering the symptoms, the system makes use of various techniques and in turn gives the results of most probable disease to be likely to be caused to the patient. This helps the doctors to recommend tests related to the disease thereby reducing the number of tests that the doctor is not sure of but had suggested the patient only for his confirmation. The system thus helps in saving time and money for both the patient and doctor.

2. Related Research

The search engine gives possible list of symptoms with the required medication for any disease, consisting of both accurate and inaccurate remedies. However, this non verified over information may also prove fatal. The authenticity of such results cannot be validated. Some sites give the feature to diagnose the disease based purely on the input of symptoms. However, these sites too cannot be entirely relied upon information as the symptom/diagnosis may not be

Licensed Under Creative Commons Attribution CC BY

relevant for the patient. The patient ultimately has to consult doctors for authentic and safe prescription.

Iliad [13] is an expert system program using Bayesian reasoning to calculate the posterior probabilities of various diagnoses when input is given as the findings. Also, DXplain [14] is a decision support system which acts on a set of clinical findings to produce a ranked list of diagnoses which may give the clinical indicators. DXplain also provides justification for why each of these diseases might be considered, suggests what further clinical information would be useful to collect for each disease. DXplain formulates this list using a large database.

MediQuery[4] puts to use storage of information so that diagnosis based on this historical data can be made. There are systems to predict diseases of the heart, brain and lungs based on past data collected from the patients. The paper focuses on computing the probability of occurrence of a particular ailment from the medical data by mining it using a unique algorithm which increases accuracy of such diagnosis by combining Neural Networks, Bayesian Classification and Differential Diagnosis all integrated into one single approach.

The availability of huge amounts of medical data leads to the need for powerful data analysis tools to extract useful knowledge. Researchers have long been concerned with applying statistical and data mining tools to improve data analysis on large data sets. Heart disease is the leading cause of death all over the world in the past ten years [2]. Several researchers are using statistical and data mining tools to help health care professionals in the diagnosis of heart disease. This paper finds gaps in the research on heart disease diagnosis and treatment and proposes a model to systematically fill gaps.

Healthcare recommender systems for Chronic Disease Diagnosis (CDD) can play a major role in controlling the disease, through providing accurate and trustworthy disease risk diagnosis prediction and medical advices recommendation. It has 24/7 remote patient monitoring system and assist patients to have 24 hour access to the medical care. A CDD recommender system is based on a hybrid method using multiple classifications and unified Collaborative Filtering. Multiple decision tree classifiers have been applied to achieve high accuracy in predicting diseases.

3. System Workflow and Algorithm

The system developed for the differential analysis and prediction of probable disease diagnosis uses the workflow pattern as shown in Figure 1.1.

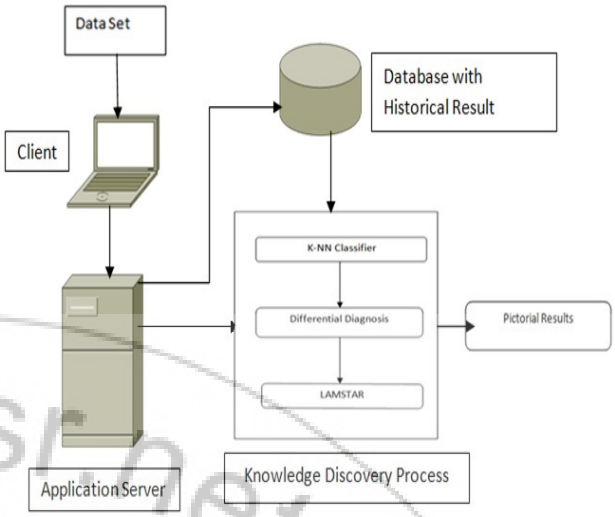

Figure 1.1: The workflow pattern

The workflow is as explained below:
Step 1: The doctor or the medical personnel may enter the symptom as seen in the patient.
Step 2: The application server uses the knowledge discovery module to reach to the correct diagnosis.
Step 3: The Knowledge discovery process consists of classification of the disease based on the symptom, refining that classification using differential diagnosis and finally storing the results in LAMSTAR networks.
Step 4: The output shown by the knowledge discovery process aids the doctor or the medical personnel in estimating the exact disease without having the patient undergo certain specific tests.

The entire process needs to refer to two databases namely Disease/symptom and Records database. The disease/ symptom database stores the list of existing known disease along with its corresponding symptoms. The records database maintains records of patients from all hospitals that are in the network. The databases are replicated across various servers to achieve fault tolerant nature.

The system has been tested for differential cases including environmental factors also. The authors call it as the recent trend service. The service uses the results from pattern matching and input from doctor for the symptom along with the geographical location of that particular area and the current climatic factors. Based on these inputs the service gives the probable disease for the symptom.

4. Conclusion

This system proposed in the paper has been aimed to provide essential medical services with clinical precision which needs high accuracy. Even though the system is to be used by doctors only, and the doctors have the final decision to make, the accuracy of the system is promising and will help the practitioners in their verdict. To verify this, the results obtained by this system were compared with the differential diagnosis provided by various other medical systems, including the information that is available at various online medical portals, and these were also verified by a panel of experts, consisting five reputed doctors at local level. The

Licensed Under Creative Commons Attribution CC BY

results obtained matched up to the doctors' expectations, and since the system is self-learning, with time, as the database grows, the accuracy of the system improves.

Different data mining algorithms can be used for disease diagnosis. The algorithms applicable can be summarized as in Table 1.1 shown below.

Table 1.1: Algorithms and their utilities as applicable to the diseases

Techniques	*Utility*	*Disease*
Apiori and FP growth	Association rule miming for finding frequent item sets	Hereditary diseases, diabetes
Genetic algorithm	Extracting patterns, detecting trends	Diabetic diseases
Bayesian Classification	Classification	Liver diseases
Bayesian network algorithm	Modelling and analysis of medical data	Coronary heart diseases
Support vector machine	Diseases classification	Diabetes

On initial comparisons with existing sources providing differential diagnosis, this system gave similar results, with the added information of the list of other probable diseases. This list of diseases obtained by performing differential diagnosis was compared with the existing differential diagnosis relationships already established, and the list was also found to be accurate to its expectations.

References

[1] Shouman, Mai, Tim Turner and Rob Stocker. "Using data mining techniques in heart disease diagnosis and treatment" *Electronics, Communications and Computers (JEC-ECC), 2012 Japan-Egypt Conference on*IEEE, 2012.
[2] El-Bathy, Naser, et al. "Intelligent optimization models for disease diagnosis using a service-oriented architecture and management science"*Electro/Information Technology (EIT), 2012 IEEE International Conference on* IEEE, 2012.
[3] Hussein, Asmaa S., et al. "Efficient Chronic Disease Diagnosis prediction and recommendation system" *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*IEEE, 2012.
[4] R. Carvalho, R. Isola, and A. Tripathy, "MediQuery—An automated decisionsupport system," in *Proc. 24th Int. Symp. Comput.-Based Med. Syst.*,Jun. 27–30, 2011, pp. 1–6.
[5] W. Siegenthaler, *Differential Diagnosis in Internal Medicine: From Symptomto Diagnosis*. New York: Thieme Medical Publishers, 2011.
[6] J. Han and M. Kamber, *Data Mining Concepts and Techniques*.SanMateo, CA: Morgan Kaufmann, 2011.
[7] D. Graupe, H. Kordylewsky, and N. Schneider, *Principles of ArtificialNeural Networks*, 2nd ed. Singapore: World Scientific, 2010.
[8] M.Bell, *SOAModeling Patterns for Service-Oriented Discovery and Analysis*.New York: Wiley, 2010, p. 390.
[9] (Jun. 10, 2012). WebMD: Better Information, Better Health.[Online]http://symptoms.webmd.com/symptomchecker
[10] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, "Improvingclinical practice using clinical decision support systems: A systematicreviewof trials to identify features critical to success,"*Br.Med. J.*, vol. 330,p. 765, 2005.
[11] Department of Medicine Massachusetts Hospital, Boston, DXplainSystem. (2011). Available at http://dxplain.org/dxpdemopp/dxpdemobrief_files/frame.htm
[12] K. A. Mohan. (Jun. 2012). National Electronic Health Record Models.Tata Consultancy Services (TCS) Whitepaper.[Online]. http://www.tcs.com/resources/white_papers/Pages/NationalElectronic/HealthRecordModels.aspx
[13] D. S. Rizzuto and M. J. Kahana, "An autoassociative neural networkmodel of paired-associate learning," *Neural Comput.*, vol. 13, pp. 2075–2092, 2001.
[14] Department of Medicine Massachusetts Hospital, Boston, DXplain System. (2011). http://dxplain.org/dxpdemopp/dxpdemobrief_files/frame.htm

Author Profile

Rashmi M. Jadhav received the B.E degree in Computer Science Engineering from Rajarambapu Institute of Technology in 2008. She is currently pursuing her master's degree in computer engineering from D. Y. Patil School of Engineering and technology, Lohgaon Pune.

Licensed Under Creative Commons Attribution CC BY