

Automated Healthcare System

Rohan Panchal¹, Ashwin Nikam¹, Abhishek Patil¹, Ajinkya Ghodekar¹

Student, Dept of Computer Engineering, NBN Sinhgad School of Engineering, Pune, India¹

Abstract: Statistics from the World Health Organization (WHO) have recently shown that the disease and mortality rates of the entire world population greatly depend on the quality of healthcare access. Access to proper healthcare in each and every part of the world is a major hurdle we have yet to overcome. The attention and focus towards providing reliable healthcare access to even the most remote areas of the world is increasing day by day and has become a major issue. This paper presents a feasible solution in the form of automated medical diagnosis and treatment of diseases for those masses that are deprived the access to experienced professional healthcare. It focuses more on the diagnosis part and the challenges which should be overcome while achieving this feat.

Keywords: Decision making, Pattern recognition, Association rule.

I. INTRODUCTION

The problem associated with medical diagnosis has been attracting a lot of attention since the past few decades but there has been no significant change in providing a solution to such problem. Even today, a large section of the world population is still deprived from proper medical healthcare access. Proper medical healthcare may refer to the prevention and cure of diseases and medical condition. The cure part of a disease first consists of diagnosing the exact disease which the patient is suffering from and then undergoing the proper treatment required. This paper focuses more on the diagnosis part of medical healthcare. The ultimate aim of this paper is to develop the software to provide reliable and proper medical healthcare access to even the most underprivileged people who haven't got any access to experienced physicians. There have already been some developments in this direction which include MYCIN. It was developed at Stanford in the 1970s which was an expert system used for diagnosing blood clotting diseases although it had around 600 rules. The solution discussed in this paper will make the use of a large database consisting of diseases and their associated symptoms. The proposed system will ask the patient for confirmation of a particular symptom. Based on the patient's responses, association rules will help guide the system into predicting the next most probable symptom that the patient might have. Next, the system identifies a disease corresponding to the pattern of symptoms entered by the patient. This paper extends the ideas of Decision Making, Pattern Recognition and Association Rules for solving a critical problem impacting a major chunk of population – that of medical diagnosis.

A) Contributions of the paper

The main contribution of the paper is to provide a solution to the problem of reliable medical diagnosis by using the concepts of Decision Making, Pattern Recognition and Association.

The combination of the above concepts is unique and provides a feasible solution which is a) based on previous diagnosis b) general and adequate enough to be used by masses c) improves accuracy on each use.

B) Paper organization

The paper is organized categorically for better understanding of the entire concept. Section 2 contains the literature survey consisting of related work done in the past. Section 3 discusses about the methodology and implementation of the proposed system. Section 4 explains about the technologies that can be used to realise the suggested methodology. Section 5 explains all the possible hurdles and challenges which we need to overcome. It also suggests a few improvements to the existing project. Section 6 concludes the paper.

II. LITERATURE SURVEY

Significant amount of research and work has been dedicated to developing projects similar to the one being discussed in this paper. Some projects which can be related to our vision and contribute in realizing our objectives have been cited in this literature survey. Scheuermann et al. present a framework in [1] to represent ontologies of the diseases and diagnoses and relationships between them. They use the ontology approach to help with medical diagnosis. The other category of papers consists of work related to the machine learning realm. One such is [2] where three different Machine Learning algorithms are used to predict the onset of type II diabetes and compare the results. In [3], the authors apply machine learning (SVM Adaboost) to predict addiction to smoking based on brain features as an alternative to the traditional techniques in addiction related neurobiology. Authors of [4] show how linear machine learning algorithms such as Logistic Regression takes just minutes to do predictions, while other powerful methods like Support Vector Machines (SVM) and Random Forests take hours or even days. They use it to predict mortality of men diagnosed with prostate cancer.

Healthcare delivery via teleconsulting, though requiring manual expertise for diagnosis is an important step in realizing our vision. In [5], the authors describe their working model of providing teleconsulting to the underprivileged. The authors of [6] use a Naive-Bayes

based machine learning system to associate phrases in clinical notes with medical concepts. Annotation techniques such as this can help our discussion because our text mining approach relies on clinical notes made by different physicians who can possibly use varying phrases to describe the same concept. So, a mapping between the phrases in the notes and medical concepts can result in better accuracy in mining the notes.

III. METHODOLOGY

The methodology discussed in this section uses the concepts of Decision Making, Pattern Recognition and Association rules.

The database required for the system should primarily consist of a 'symptoms' table which stores the diseases along with its corresponding symptoms. The first column in this table should be the 'id' column which corresponds to each disease in the database. The remaining subsequent columns should be labelled with the symptom names. For a particular row (which corresponds to a disease), put the value '1' (for a particular column) if that particular symptom is related to that disease. It is to be noted that the data relating to the diseases must be collected from reliable sources.

Another important table required for the diagnosing process would be the 'Association Table' as shown in Figure 1. This table should consist of the association of one particular symptom to all other symptoms. In a more practical sense, if you have 100 symptoms in your database, create a table of the size 100x101. The first column should have the symptom names. For a particular row (symptom) give values for every column that are equal to the number of times that particular column label (symptom) occurs with the symptom mentioned in the first column of the considered row. The significance of this table will be elaborated in the next paragraph.

symptoms	High_fever	Loss_of_apetite	Body_pain	Drowsiness	Stomach_pain	Sore_throat	Cough
Abdominal_pain	2	0	1	0	0	0	0
Abnormal_sounds_while_breathing	1	0	0	0	0	0	1
Back_becoming_arched	1	0	0	0	1	0	0
Back_pain	0	1	1	0	0	0	0
Babies_by_an_animal(dog)	0	1	0	0	0	0	1
Bleeding_in_the_eye	0	1	0	0	0	0	0
Blisters	0	1	0	0	0	0	0
Blocked_nose	1	0	1	0	0	0	0
Blood_in_stools	0	0	0	0	0	0	0
Blood_in_urine	0	0	0	0	0	0	0
Blurred_vision	0	0	0	0	0	0	0
Body_jerks	0	0	0	0	0	0	0
Body_pain	4	3	0	1	1	2	0
Body_paresis	0	0	1	0	0	0	0
Burning_sensation_in_chest	1	0	1	0	0	0	0
Changes_in_behaviour/personality	0	0	0	0	0	0	0
Chest_pain	3	1	0	0	0	0	0
Chills	5	1	3	0	0	0	3

Figure 1: Snapshot of an association table

The general working of the proposed expert system would be as follows: Firstly, it wants you to create an account if you are not an existing user

When a patient starts a new diagnosis, the system should first ask the symptoms that were present for that particular patient in all previous diagnosis. The system keeps track

of all the previous diagnosis of all the patients for this purpose. If a user is diagnosing for the first time, then the highest occurring symptom in the database should be asked to that user.

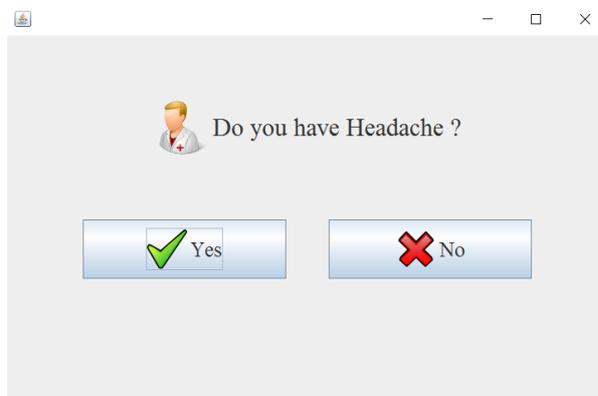


Figure 2: System diagnosing the user

For every 'No' given by a patient for a particular symptom, display the next highest occurring symptom to be asked to that user. For every 'Yes' or a confirmation of a particular symptom, query the 'symptoms' table to get the rows (diseases) when the value of that particular symptom is one. If the result of the query shows more than one row, find the symptom that has maximum association to the previously agreed symptom and ask the patient for his confirmation. The 'Association Table' should be used to achieve this. In this way we can make the diagnosing process faster by asking the symptoms that are related to the symptoms, previously agreed by the patient.

At this stage, for every subsequent 'No' (for a symptom), the system should enquire about the next highest associated symptom. For all subsequent 'Yes' (for a symptom), query the 'symptoms' table and find the resultant table where values of all agreed symptoms is one. Evidently this process of asking symptoms to the user continues until only one row (disease) exists in the query result. This will be the effective diagnosed disease. Hence, by recognizing a unique pattern of symptoms, the system would be able to perform decision making in order to output the diagnosed disease.

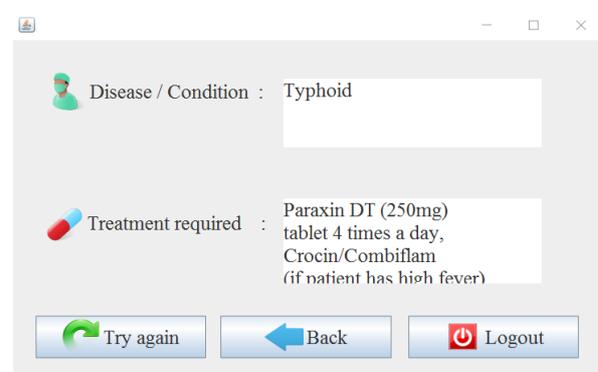


Figure 3: Diagnosis successfully performed

Here another table comes into picture. There should be a table in the database which stores the treatment for every disease. Therefore, after a patient is diagnosed with a

particular disease, its corresponding treatment should be displayed to the patient for cure as shown in Figure 3.

Two more tables should be included in the existing database. Firstly, a table to store the basic credentials of the user accounts and another table to store the medical diagnosis history of each patient/user.

The above suggested methodology could be deployed as a website which would be used by masses all over the world through internet. Another option for deployment could be by making an installer of the software along with all the required databases and making it available on the internet for download and use.

IV. FINDINGS

There are a number of observations and results that approve the above methodology. The suggested logic for implementation of the system can be realised by using MySQL and Java environment for the database and programming respectively.

Another approach which could be used is to develop a web application using the methodology discussed in the previous section. Firstly, a domain will be required to be purchased. Next, a Java Web Application will have to be developed that will run on the server machine of the domain.

V. DISCUSSIONS AND LIMITATIONS

There are a number of aspects of the software which could be improved. These aspects would help make a more reliable medical healthcare system. Let us look at the limitations and challenges involved while realising this project.

A. Questionable quality of data

The data regarding the diseases and related symptoms cannot be predicted unless it is from a very trusted source. This is the same for the treatment advised to patients.

B. Quantity of data

The amount of data (number of diseases) in the database determines the usability and accuracy of the system. For example, a database with information about two diseases will only have the capability to diagnose those two diseases.

C. Privacy

The user/patient credentials and the diagnosing data should be effectively maintained as private. The confidentiality of this data should not be hampered in any way.

Further, we have a few improvements to the existing design. With the current methodology, only those diseases that do not require a laboratory test can be diagnosed effectively. Hence, to accommodate the diagnosing capability for those diseases, sensors can be attached to the body of the patients that will capture the patient's medical data like heartbeat, blood pressure etc. to diagnose other diseases. But this would lead to increasing the cost of

using the system, which will be quite contradictory to our original motivation.

Another challenge would be to provide the treatment (tablets) based on their availability in a particular geographical area or region. For example, a particular tablet available in India might not be available in South Africa.

VI. CONCLUSION

We have presented a feasible solution to realise the vision of providing healthcare to people who do not have access to good physicians. The methodology discussed in the third section uses the concepts of Decision Making, Pattern Recognition and Association rules to generate an efficient diagnosis. The technologies suggested in section 4 can be effectively used to develop this project. We have also listed a few challenges and improvements in section 5 which will help in making this project a more complete one.

REFERENCES

- [1] cheuermann, R. H., Ceusters, W., & Smith, B. (2009). Toward an onto- logical treatment of disease and diagnosis. *Summit on translational bioinformatics*, 2009, 116. Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. *Intelligent Systems, IEEE*, 21(3), 96-101.
- [2] Sarwar, A., & Sharma, V. (2014). Comparative analysis of machine learning techniques in prognosis of type II diabetes. *AI & society*, 29(1), 123-129.
- [3] Pariyadath, V., Stein, E. A., & Ross, T. J. (2014). Machine learning classification of resting state functional connectivity predicts smoking status. *Frontiers in Human Neuroscience*, 8, 425.
- [4] Ngufor, C., Wojtusiak, J., Hooker, A., Oz, T., & Hadley, J. (2014, March). Extreme Logistic Regression: A Large Scale Learning Algorithm with Application to Prostate Cancer Mortality Prediction. In *The Twenty-Seventh International Flairs Conference*.
- [5] Sapounas, D., Jackson, K., & Ervin, D. (2011, October). International Consultants in Medicine: A Framework for Medical Expertise and Social Telemedicine Addressing Medical Disparities. In *Global Humanitarian Technology Conference (GHTC), 2011 IEEE* (pp. 208-211). IEEE.
- [6] Gobbel, G. T., Reeves, R., Jayaramaraja, S., Giuse, D., Speroff, T., Brown, S. H., ... & Matheny, M. E. (2014). Development and evaluation of RapTAT: a machine learning system for concept mapping of phrases from medical narratives. *Journal of biomedical informatics*, 48, 54-65.