

Full Length Research Paper

Designing an expert system model for pulmonary TB disease diagnose pre-screen laboratory process

Yonas Tesfaye^{1*} and Ketema Kifle²

¹Department of Information Technology, College of Engineering and Technology, Samara University Samara, Ethiopia.

²Department of Software Engineering, Institute of Technology, Debre Markos University, Debre Markos, Ethiopia.

Received 13 May, 2019; Accepted 2 August, 2019

Tuberculosis is a major public health problem in Ethiopia more than half a century ago. If the disease is not treated well and on time, it can lead to severe health problems like lungs, but can also affect other organs, including the central nervous system. Therefore, detecting those diseases at early stages enable us to overcome and treat them appropriately. However, among people in the developing countries like Ethiopia, permanent diseases are growing to be causes of death. These problems are becoming worse due to the scarcity of specialists, practitioners and health facilities. In Ethiopia, there has been observed a threat of increased prevalence of tuberculosis and the number of death rates imputed to tuberculosis reached above 29,000 deaths (35 per 100,000) due to TB, excluding HIV related deaths, during the same period, 2014. In an effort to address such problem, this study attempts to design an expert system model for tuberculosis disease diagnose pre-screen laboratory process that can provide advice for physicians and patients to facilitate the diagnosis before laboratory testing. To this end, knowledge is acquired using both structured and unstructured interviews from domain experts which are selected using purposive sampling technique from health agents. Relevant documents analysis method is also followed to capture explicit knowledge. Thereafter, the acquired knowledge is modeled using decision tree that represents concepts, procedures involved in diagnosis and treatment of tuberculosis and production rules are used to represent the domain knowledge, and knowledge-based system is developed using SWI Prolog editor tool. It uses backward chaining which begins with possible solutions or goals and tries to gather information that verifies the solution.

Key words: Model, expert system, knowledge representation, knowledge acquisition, tuberculosis disease.

INTRODUCTION

Tuberculosis (TB) is one of those infectious diseases caused by bacteria whose scientific name is mycobacterium. It was first isolated in 1882 by a German physician named Robert Koch. Tuberculosis has been

recognized as major public health problem in Ethiopia more than half a century ago. Many years ago, this disease was referred to us "consumption" because without effective diagnosis and treatment, these patients

*Corresponding author. E-mail: yonitesfu@gmail.com.

often would waste away. Today, of course, tuberculosis usually can be treated successfully with antibiotics but there is a problem with diagnosis of the disease (Abdi et al., 2010). Ethiopia's health care system is among the least developed in Sub-Saharan Africa and is not, at present, able to effectively cope with the significant health problems facing the country. Communicable diseases are the primary illnesses. Acute respiratory infections such as tuberculosis, upper respiratory infections, and malaria are the Ministry of Health's priority health problems. Poor sanitation, malnutrition, and a shortage of health facilities are some of the causes of spread of communicable diseases (Asia, 2012). The 30 high TB burden countries are given particular attention in the main body of WHO report, which includes Ethiopia (WHO, 2018). According to the WHO Global TB Report (2011), there were an estimated 220,000 (261 per 100,000) incident cases of TB in Ethiopia in 2010. Also, the prevalence of TB was estimated to be 330,000 (394 per 100,000). Particularly in developing countries, community health centers usually offer a broad range of services, including prenatal care, immunizations, treatment of childhood illnesses, treatment of malaria, tuberculosis and other common infectious diseases, and other basic medical care. Ethiopia's main health problems are said to be communicable diseases caused by poor sanitation and malnutrition (Richard, 2004). TB is a leading cause of morbidity and mortality in adults' worldwide, killing more than 1.5 million people every year (Elamy et al., 2010).

Artificial intelligence has emerged after the introduction of the first computers. Currently, the concept of the artificial intelligence is understood as a branch of computer science, in which the computer programs have capability to simulate human behaviors. Computer programs are used both for experimental and practical purposes, such as formulation of medical diagnoses (Mohammed and Beshah, 2018). One of the earliest applications of expert systems in a biological discipline was in the medical area. As they are very rich in domain knowledge, knowledge-based systems (KBS) quickly became an essential tool for diagnostics and treatments in the medical area. Knowledge based systems are widely used in domains where knowledge is more common than data and that require heuristics and reasoning logic to derive new knowledge. Knowledge-based intelligent systems can be employed to reduce the number of deaths and the waiting time to see the specialist. The systems developed by imitating human intelligence could be employed to help the medical doctors in making decision without consulting the specialists. The system is not meant to replace the specialist or medical doctor, but it is developed to help general physicians and specialists in diagnosing, predicting condition of patient, and providing treatments from certain grounds or experiences. The expert's knowledge is available when the human expert might not be and so that the knowledge can be available at all

times and in many places, as necessary. Expert systems derive their input for decision making from prompts at the user interface, or from data files stored on the computer.

STATEMENT OF THE PROBLEM

The underlying research problem that necessitated this research is the existence of high death rate of TB at a national level. TB is a great problem in most low-income countries. It is the single most frequent cause of death in individuals aged fifteen to forty-nine years (Temurtas et al., 2008). Pulmonary TB is one of the types of TB that continues to be a major public health concern. The devastating impact of TB on vulnerable populations is also driven by its deadly synergy with HIV. HIV infection compounds the problems of accurate diagnosis as well as adequate treatment. TB causes more rapid deterioration of the immune systems of people with HIV or AIDS, and they are 100 times more likely to have active TB during their lifetime than people who are HIV-negative (Elamy et al., 2010). People with advanced HIV infection can have active TB that is smear-negative or without typical chest radiography features, which means that co-existent TB infections remain untreated. Nevertheless, a good treatment of the patients is periodically carried out in a challenge on the diagnosis of the disease and serious data quality issues with limited use of TB information for local decision. Reports of varying disease border across different localities with same-fits-all TB control efforts. Hence, by assuming the patients' disease can be TB, they started the therapy for the disease. After some weeks it may be discovered that it is wrongly diagnosed. This leads to delay in the control program of the disease and because of such kind of problems lots of patients die; also, in the rural area there is lack of doctors in clinical or lack of proper knowledge by community health workers to be the major challenges of TB controlling program in many developing countries. This proposed system is applicable and promising for assisting development agents who are working in rural areas where skilled health experts are unavailable for an early treatment to the infected patient before the condition get worse, and also used for urban which helps to manage their health problem from TB. In Ethiopia, most of these are in rural areas where skilled health experts or doctors are unavailable.

Objective

The general objective of this research was to design an expert system model for the diagnosis of tuberculosis disease in order to provide better information for physicians and patients.

The specific objectives of the study were:

- i) To understand the concept that can be helpful in

designing expert systems model for pulmonary TB disease diagnosis of the pre-laboratory screening process.

- ii) To extract knowledge from domain experts and manuals in the area of pulmonary tuberculosis disease diagnosis.
- iii) To model and represent the acquired knowledge using appropriate knowledge representation technique.

METHODOLOGY

Source of data

Both unstructured and structured interview were used to collect tacit knowledge from domain experts. The acquired knowledge is refined with the consultation of the expert. Moreover, secondary sources of knowledge are gathered from the internet, tuberculosis diseases diagnosis and treatment guidelines, research papers and articles by using document analysis technique.

Data collection method

Purposive sampling technique would be used to select domain experts for knowledge requirements acquisition. The selection criteria of domain experts for the study are based on the profession, educational qualification level, year of experience and their immediate position in the tuberculosis diseases diagnosis.

Knowledge requirement model

After the knowledge would be acquired, it was represented using rule-based knowledge requirement representation method. For this research, the rule based knowledge requirement representation method was chosen because it clearly demonstrates the domain knowledge. In rule-based system, much of the knowledge was represented as a rule that operates as conditional sentences (IF THEN) relating statements of facts with one another. As a result, rule-based representation method is more appropriate to represent and demonstrate the real domain knowledge in diagnosing tuberculosis diseases. Additionally, rule-based systems are the most commonly used knowledge representation language in health sector.

Implementation tool

The programming languages and tools would be used to develop an expert system. In this study, SWI-prolog programming language would be used to build the knowledge base of the system. The reason for the selection of this programming language is the features and abilities of the language that incorporates it. Prolog is a declarative language (where the problem intended to be solved is specified rather than how to solve it) and has the capacity to describe the real world. It has flexible and fast interface.

Significance of the study

The general benefits and beneficiaries of the research will be of two types, viz; primary and secondary. The primary beneficiaries of this research output are those experts who are involved in pulmonary tuberculosis diagnosis activities. The expert system can help experts in controlling and managing consistently during tuberculosis diagnosis, to manage wrong tuberculosis diagnosis. Those

inconsistencies created by human experts during tuberculosis diagnosis will be solved if this system is applied. The focus of all medical systems that are developed is to build better health care facility in order to reduce time, cost and medical error (Margret et al, 2013).

The secondary beneficiaries of this research output are those subjects who receive tuberculosis diagnosis. Those patients who need a tuberculosis diagnosis can get the correct tuberculosis type for the right patient. An expert cannot have any doubts about diagnosis and it is also advantageous for remote and rural areas that have scarcity of medical professionals and medication facilities. It also serves as temporary assistance to those who are in need of instant help when expert consultant is not readily available due to time and distance.

IMPLEMENTATION AND EXPERIMENTATIONS

Model

The knowledge modeling phase, the specialist's knowledge (elicited by various techniques) is represented in a knowledge model. A knowledge model is a structured representation of knowledge using symbols to represent pieces of knowledge and the relationships between them (Aebissa, 2012; Buchanan and Shortliffe, 1984). Based on Richard et al. (1999), one of the most extensively applied methods for conceptual modeling is decision tree. Decision tree commonly acts a key role in a knowledge modeling process. Figure 1 shows the decision tree structure that has the flow of knowledge in the diagnosis and treatment of TB disease.

Decision tree for modeling pulmonary TB diagnosis

Decision trees are one of the modeling tools used in variety of settings to organize and break down cluster of data. It is a graphical representation of the information in the factor table to determine a course of actions and models the possible consequences of a series of decisions in some situations (Bethune et al., 2007). Decision analysis is explicit because it forces the decision maker to separate the problem into its component parts without losing the context of the big picture. A decision tree is formalism for expressing and mapping attributes linked to two or more sub-trees and leaves or decision nodes (Podgorelec et al., 2002).

Knowledge representation in decision tree for pulmonary TB and prototype system architecture

Knowledge representation in decision tree for pulmonary TB along with architecture of the prototype system is as shown in Figures 1 and 2.

Expert system prototype

The user interface of the prototype system is allowing non-expert and physicians' users to query (question) an expert system, and to see delivery of advices. In this prototype, users can interact with the system through yes or no responses. Based on the user response, the system draws a conclusion and displays it in the user interface window. Figure 3 displays the home window that has different options for users to interact with the system.

After the prototype displays the greeting page, a user can interact directly with the system by answering the question "yes or no" followed by dot. Figure 4 shows the sample dialogue windows between the user and the system to identify the infections on tuberculosis disease.

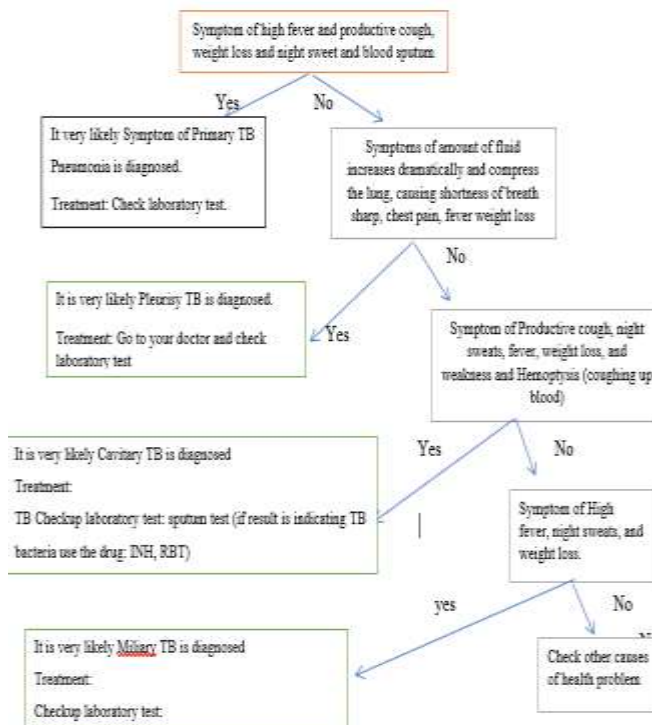


Figure 1. Decision trees for diagnosis and treatment for pulmonary TB.

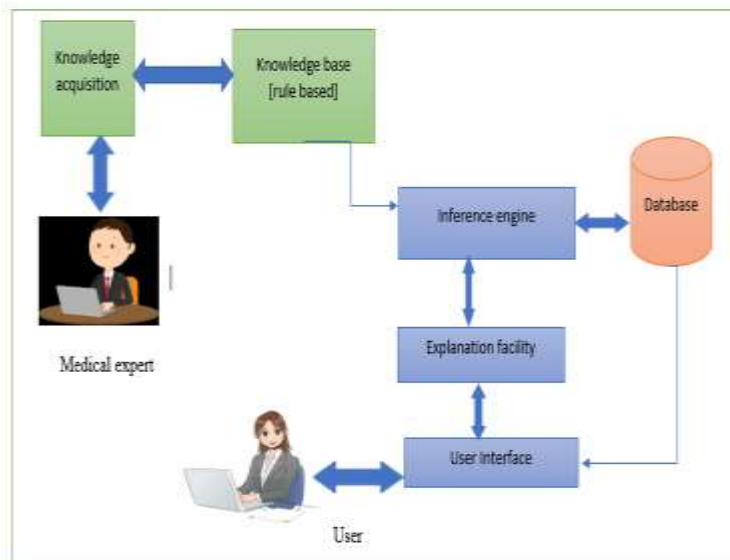


Figure 2. Architecture of the prototype system.

Conclusions

Tuberculosis has been a major public health problem in Ethiopia since more than half a century ago. If the disease is not treated well and on time, it can lead to

severe health problems like lung disease and any extra body can be easily affected. In the developing countries like Ethiopia, TB as well as permanent diseases are growing to be causes of death. Hence, in this study an effort has been made to design an expert system that

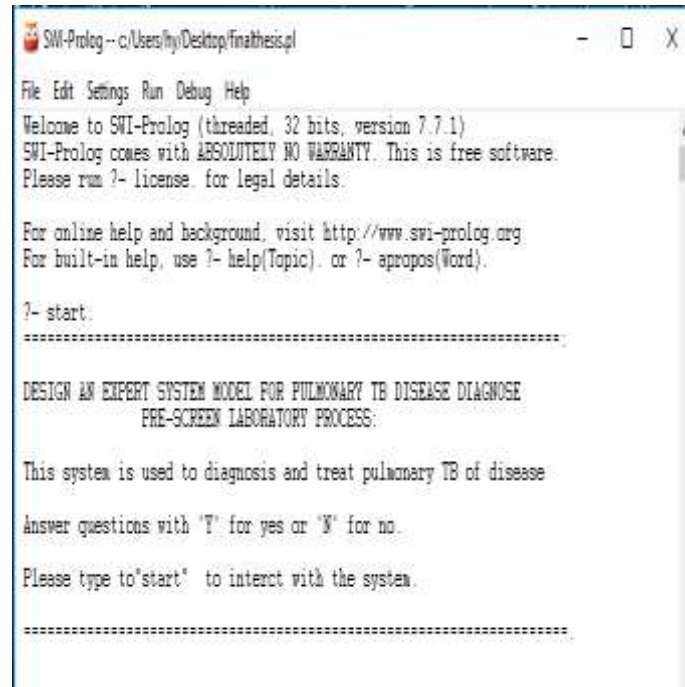


Figure 3. Welcoming windows of user interface.

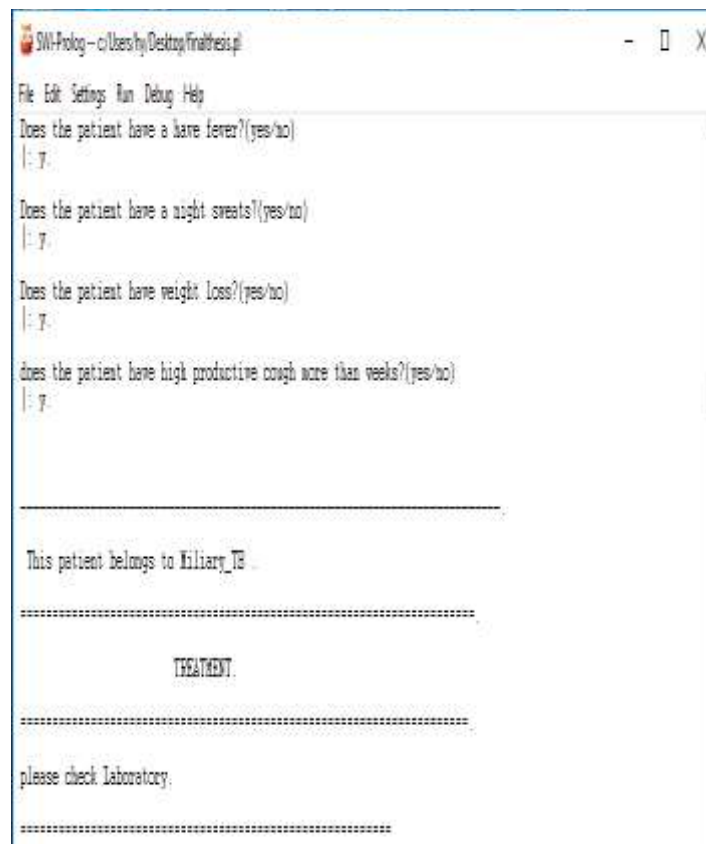


Figure 4. Sample dialogue windows between the user and the system to identify TB infections.

models a prototype of knowledge based system that can provide advice before the patient go to check laboratory test for physicians and patients to facilitate the diagnosis and treatment of patients affected by TB. The acquired knowledge is modeled using decision tree that represents concepts and procedures involved in diagnosis and treatment of TB. Thereafter, the validated knowledge is represented using rule-based representation technique and codified using SWI-Prolog editor tool for building the knowledge-based system to provide advice for patients and experts.

RECOMMENDATION

This rule-based system is not self-learning. In the future, learning component should be integrated to reason and remember when new circumstances and unknown facts are asked by users to suggest solutions. To enhance the performance of the prototype expert system, the hybrid strategy approaches should be investigated which combines case-based reasoning. The inclusion of case-based reasoning helps the system to learn from documented experiences. In its present version, this system is a standalone one. For the future, it will be more attractive and effective if a Web-based version of it is developed that would make the diagnostic system accessible to anyone with a computer and an Internet connection.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

- Aebissa B (2012). Developing a knowledge-based system for. Addis Ababa University, Ethiopia.
- Asia (2012). Mining patients' data for effective tuberculosis diagnosis: The case of Minilik II Hospital. Addis Ababa, AAU.
- Bonner R, Galant V, Owoc M (1999). Features of Decision Trees as a Technique of Knowledge Modelling. In The Workshop on Computer Science and Information Technologies (CSIT) pp. 135-137.
- Buchanan BG, Shortliffe EH (1984). "Rule-based expert systems: the mycin experiments of the stanford heuristic programming project. I." In g.b buchanan and h.e. shortliffe, eds. pp. 2-67.
- Elamy AH, Mandal M, Behrouz F, Anup B, Irene C, Richard L (2010). "An Intelligent Cad System for Automated Detection of Pulmonary Tuberculosis on Chest Radiograph and Ct Thorax: "A Road Map".
- Margret S, Clara MLJ, Jeevitha P, Nandhini RT (2013). "Design of a Diabetic Diagnosis System Using Rough Sets." *Cybernetics and Information Technologies* 13(3):124-139.
- Mohammed S, Beshah T (2018). Amharic based Knowledge-Based System for Diagnosis and Treatment of Chronic Kidney Disease using Machine Learning. *International Journal of Advanced Computer Science and Applications* 9(11):252-260.
- Podgorelec V, Kokol P, Stiglic B, Rozman I (2002). Decision trees: an overview and their use in medicine. *Journal of medical systems*, 26(5):445-463..
- Richard BF, Violetta G, Mieczyslaw OL (1999). "On Features of Decision Trees as a Technique of Knowledge Modelling , Proceedings of the Workshop on Computer Science and Information Technologies CSIT'99." *Moscow, Russia*.
- Richard GW (2004) "Reviewing Ethiopia's Health System Development'." *JMAJ, International Medical Community* 52(4):279-286.