

A Decision Support System for Tuberculosis Prevalence in South Africa

Mulalo Razwiedani[1] and Okuthe P. Kogeda[2(✉)]

[1] Tshwane University of Technology, Pretoria 0001, South Africa
razwiedanims@tut.ac.za
[2] University of the Free State, Bloemfontein 9300, South Africa
kogedapo@ufs.ac.za

Abstract. Tuberculosis is one of the most prevalent diseases, which is a threat to the lives of many South Africans. The disease has been spreading at a high rate in the past years. World health organization reported that South Africa is amongst 22 countries most burdened by the disease that has around 80% of the total global Tuberculosis cases. Despite the fact that the South African government is undertaking Tuberculosis campaigns, there is still a challenge in the control of the spread of Tuberculosis. The lack of awareness and access to information further contributes to the high rate at which Tuberculosis is spreading. Therefore, we developed a decision support system for Tuberculosis prevalence in South Africa. We used Bayesian network to aggregate, analyse and mine data received from health department of Mpumalanga province. The data is stored in MySQL database. We used WampServer environment, which enabled us to develop a web application. The proposed model educates, informs, and prescribes measures to take when visiting a high prevalence location. We then tested the system developed with data from Mpumalanga provincial health department showing that males are more at risk with prevalence rate of 53%, 0–35 age group are the most affected at 50%, Ehlanzeni location in Mpumalanga province is the most affected at 47%, pulmonary Tuberculosis is the most prevalence at 90% and the survival rate is at 87%. We evaluated the model achieving a utility of 88.2%. We believe this utility can be improved with more training and accurate data.

Keywords: Bayesian network · Decision support system · Tuberculosis · Data mining · ICT4D · Probability · Mobile health

1 Introduction

Tuberculosis (TB) is an airborne disease caused by a bacterium that mostly affects the lungs. It is highly infectious and can affect anybody exposed to the bacteria. The spread of the disease has been growing at an alarming rate. Globally about 9 million people got infected with TB in 2013 and about 16% of them died WHO [1]. South Africa is amongst the countries with a high prevalence rate and recorded approximately 320 000 infected people in 2013.

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12952, pp. 270–284, 2021.
https://doi.org/10.1007/978-3-030-86973-1_19

In South Africa, TB kills people who are mostly in the productive age of their employment and parenting lives, therefore, the country is losing members of the workforce, which is impacting negatively on the economy of the country. Children grow up without parents increasing the levels of the needy people without income, and the lack of income means that education is no longer a priority for them. The elderly are also affected as they lose those who support and care for them. Further, the productivity of the country drops resulting in a slowdown of development and growth, and the spread of the disease costs the country as more money has to be allocated to the health budget. Finally, the factors mentioned above discourage investors thereby affecting business and government revenues. This means that South Africa is in dire need of an effective TB epidemic control programme. This programme should have accurate information on prevalence of and the measures necessary in dealing with the epidemic. This can be achieved by creating a TB epidemic model.

In this study, we focused on predicting the chances of someone being infected if they visit an area with a high prevalence of TB. This prediction is done based on demographic data. Ultimately the study assists people to make a decision on what precautionary measures to take based on the prediction obtained. In this study, we used Bayesian networks to compute the probability of getting infected with TB.

This paper is organized as follows. In Sect. 2, we present related works. In Sect. 3, we present Bayesian network including reasons why we chose it. In Sect. 4, we provide system design and architecture. In Sect. 5, we provide implementation. Testing and results are presented in Sect. 6 with conclusion being presented in Sect. 7.

2 Related Works

There are a number of researchers who have done some work in this area. They are summarised as follows. Kumari, Rajan & Anshul [2] presented a prediction of diabetes using Bayesian network. The main objective of their work was to help in predicting diabetes by applying data mining. The discovery of knowledge from the medical data set was done in order to make effective medical diagnosis. They proposed the Bayesian network methodology to predict whether the person has diabetes or not. In their results they found that Bayesian network gave the prediction accuracy of 99.5%. This study is similar to ours as it uses the same methodology that we are using in our study, but they used it to predict whether the person has diabetes or not.

Thakkar *et al.* [3] presents a health care decision support system for swine flu prediction using Naive Bayes classifier. In their study, Naïve Bayes classifier was used to classify the patient of swine flu in three categories. Their research focused on medical diagnosis by learning pattern through the collected data for Swine Flu. They claimed that Naïve Bayes could identify all the significant medical predictors and had the accuracy rate of 63.33%. Though the approach of this study is similar to ours as we also used collected historical data to implement our system, however, they used Naïve Bayes classifier in predicting swine flu.

Palaniappan & Rafiah [4] presents intelligent heart disease prediction system using data mining techniques. Their research developed a prototype intelligent heart disease prediction system using data mining techniques, namely, Decision Trees, Naïve Bayes

and Neural Network. They used medical profiles such as age, sex, blood pressure and blood sugar to predict the likelihood of patients getting a heart disease. Their results showed that the most effective model to predict patients with heart disease appears to be Naïve Bayes followed by Neural Network and Decision Trees. Their study is similar to ours as it focuses on prediction. Our study focuses on the chances of contracting Tuberculosis but theirs was mainly focused on predicting the chances of having a heart disease.

Sharareh, Niakan & Xiao-Jun [5] present evaluation and comparison of different machine learning methods to predict outcome of Tuberculosis treatment course with the aim to provide prediction on the outcome of Tuberculosis treatment to the World Health Organisation. Their methodology involves the collection of Tuberculosis data from Tuberculosis control centres throughout Iran in 2005. They developed and evaluated models by six algorithms including decision tree, artificial neural network, logical regression, radial basis function, Bayesian networks and support vector machine. They used a dataset of 6450 patients with 4515 patients used for training and 1935 patients used for testing the models. Decision tree was found to be the best algorithm with 74.21% prediction accuracy. Their study is similar and related to the one we are doing as it uses the same methodology that we followed. The methodology that they used is similar to the one we are using of collecting historical data and constructing a model for prediction, however, their study is different to our study as they predict the outcome of the Tuberculosis treatment course.

Kogeda, Agbinya & Omlin [6] present a probabilistic approach to faults prediction in cellular networks. They derived probabilistic models of cellular network systems in which the independence between the relations between the variables of interest are represented explicitly. The methodology they used involved constructing a structure of a Bayesian network. They further evaluated the Bayesian network showing the causal relationship between variables with a belief factor. Their experimental results showed a fault prediction model. While they used Bayesian networks and looks at cellular network systems, in our study we are looking at Tuberculosis prevalence.

Ogutu, Kogeda & Lall [10] present a probabilistic assessment of location dependent failure trends in South Africa water distribution networks. Their work evaluated failure trends of pipelines in a region characterised by dolomitic grounds, outlining how the different soil tyes affect water pipeline failures. They used data from a selected metropolitan region in South Africa and applied Bayesian networks to accommodate inherent data uncertainty while modelling failure causal properties. While they used Bayesian networks and historical data to evaluate water pipeline failure trends, in our study we used the same method to model Tuberculosis prevalence in South Africa.

Kgoete & Kogeda [11] present a prediction model for mitigating Tuberculosis infection for HIV patients at greater Tubatse local municipality in South Africa. They developed a windows mobile application system using Visual Basic.NET to predict Tuberculosis prevalence in Tubatse area using Bayesian network model. They collected, aggregated and analysed data and developed a system infusing Bayesian network model. Their system showed a prediction rate of 98% utility. Though their work used the same methodology as ours, their work majored on HIV patients and their system was targeted towards the HIV patients.

Kogeda & Dladlu [12] present a decision support system for Cancer prevalence in South Africa. They collected data of Cancer patients from major hospitals in Eastern Cape Province of South Africa, then aggregated, classified and analysed using Bayesian network model. They attained accuracy of 97%. While they also used Bayesian network and historical data, their study majored on Cancer patients.

3 Bayesian Networks

A Bayesian Network also known as Bayes network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a Directed Acyclic Graph (DAG) [7]. A Bayesian Network is defined by a directed network structure in the form of a DAG and a set of the local probability distributions (local pdfs), one for each node. A Bayesian Network shows conditional probability and causality relationships between variables. Conditional probability is the probability of an event happening given that another event has already happened. Probability (or likelihood) is a measure or estimation of how likely it is that something occurs or that a statement is true. The Bayesian network efficiently expresses the Joint Probability Distribution (JPD) of a set of variables. A Bayesian Network is made up of nodes that represent random variables and events, while arcs or arrows represent the probabilistic relationships between these variables. Each arrow starts at the parent node and ends at the child node, with the parent influencing the child in one way or another. Each node in the network has its conditional probability table computed from the dataset.

The joint probability distribution of random variable $S = \{X_i \ldots, X\}$ in a Bayesian network is calculated by the product of the local conditional probabilities of all the nodes. Let a node X_i in S denote the random variable X_i, and let $Pa(X_i)$ denote the parent nodes of X_i. Then, the joint probability distribution of $S = \{X_i \ldots, X\}$ is given by Eq. (1):

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=0}^{n} P(X_i|Pa(X_i)) \tag{1}$$

Bayesian Networks have been used in the areas including: Medical diagnosis, Decision-making, Adaptive testing, Decision centric troubleshooting, Data mining, Pattern recognition, Bioinformatics/computational biology, Weather forecasting, etc.

Bayes theorem was originally developed by Thomas Bayes, a theologian and mathematician in the 18th century. Bayes is the theory that gave rise to Bayesian networks. The theorem was first published in 1763 and is stated as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

Where:

- P(A) is the prior probability of the event A, without any information about the event B,
- P(A/B) is the conditional probability of the event A, given the event B,
- P(B/A) is the conditional probability of the event B, given the event A,
- P(B) is the marginal probability of the event B, acting like a normalizing constant.

Bayesian network is defined by:

- a directed cyclic graph in which each node denotes a random variable, either discrete or continuous values
- a set of conditional probability tables (CPT) for each node [5].

3.1 Why Bayesian Networks?

Classical statistics do not handle uncertainty well. For this reason, some of the grounds for using a Bayesian Network to tackle this problem included:

- They are graphical models and can represent relationships clearly and easily.
- They clearly depict the direction, thereby showing the cause-effect relationship which other Artificial Intelligence domains (decision trees, artificial neural networks, mixtures of basic functions, Markov networks, etc.) are not able to represent.
- They are able to handle uncertainty.
- They handle uncertainty using the theory of probability [8].

A Bayesian Network was the best algorithm to use as it relies on a well-established theoretical or mathematical base and it is very simple to understand and use.

4 System Design and Architecture

The main system requirement of this study is to create a prototype that computes the probability of getting infected with Tuberculosis using Bayesian network. The Mobile Health Decision Support System (MHDSS) was constructed according to the following functional requirements:

- Allow user to query for the probability of getting infected with Tuberculosis based on their age group, gender, location and type of Tuberculosis.
- Allow user to view the precautionary measures to avoid being infected with Tuberculosis.
- Allow health practitioner to add more patient records on the system.

4.1 Object-Oriented Requirements Analysis

An object-oriented approach with Unified Modelling Language (UML) has been used in the design of the MHDSS. UML is the standard language for modelling software systems [9]. The UML has several types of diagrams in order to model the static and dynamic behaviour of a system. The use cases of the MHDSS are presented. Figure 1, shows user use case diagram. The diagram illustrates that the user is able to request for the probability of getting infected with Tuberculosis in the location of interest by selecting and submitting information about their location, age group, gender and type of Tuberculosis. The system employs the Bayesian Network to compute the probability of getting infected with Tuberculosis. The system then presents the user with the probability of getting infected with Tuberculosis. The user is able to view precautionary measures to take when avoiding being infected with Tuberculosis.

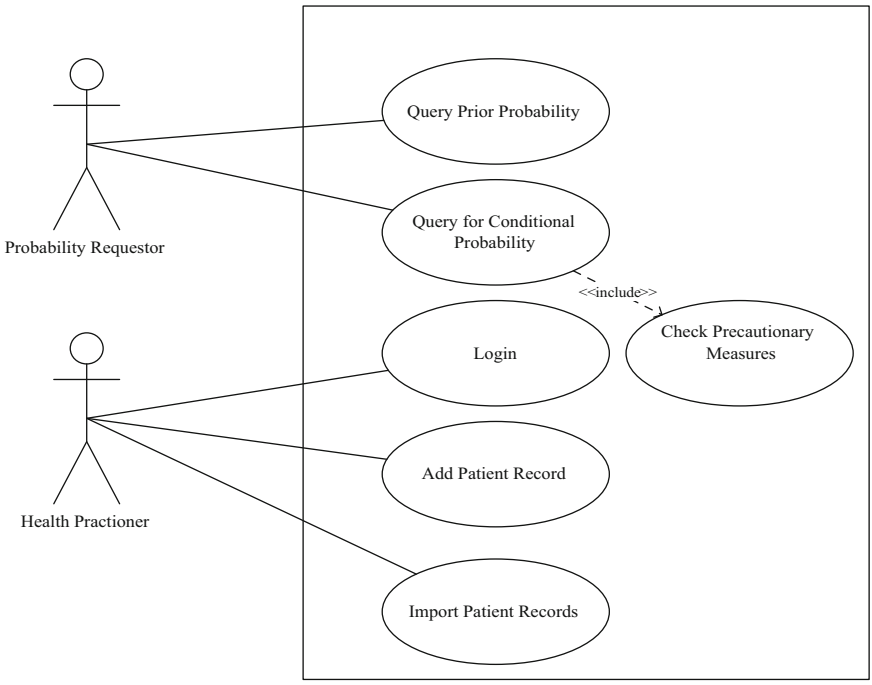


Fig. 1. Top level use case diagram

4.2 System Architecture

The MHDSS architecture is based on a three-tier architecture (see Fig. 2). This system architecture approach was chosen as it brought with it system robustness, flexibility and resistance to potential change. A three-tier architecture provides the following benefits: scalability, maintainability, reliability, availability, extensibility, performance, manageability, and security. The chosen architecture is composed of three layers: the user interface layer, the application logic layer and the data layer.

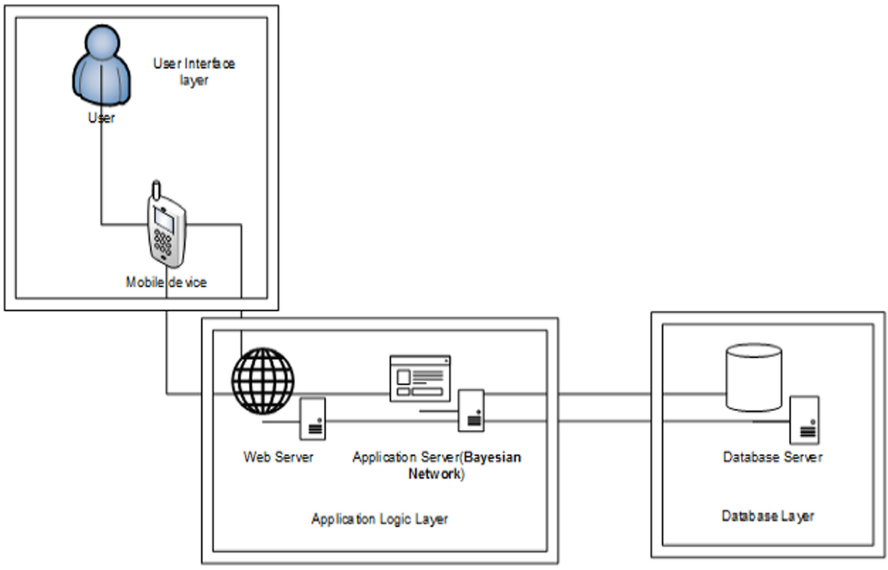


Fig. 2. System architecture

5 Implementation

The research implementation involved developing the mobile health decision support system for Tuberculosis prevalence. The data collected were stored in a database and the model designed during the modelling stage was then implemented. The implementation involved designing the software system, reusing software components e.g., MySQL, and choosing appropriate programming language. The mobile graphical user interface was then developed for the system. After testing, a system evaluation was conducted to ensure that the system implemented was as designed. The evaluation process also assisted to make sure that all the system requirements were achieved.

From literature review, we have learned that Tuberculosis is caused by breathing in an air droplet that contains a Tuberculosis bacterium. The bacterium spreads when an infected person coughs, spits, sneezes or shouts. This means that being in an area where there is a high prevalence of Tuberculosis, your chance of contracting the disease increases. There are a number of stages that we are going to take in order to build a Bayesian Network.

5.1 The Database

The available dataset has been collected from gathered records by health practitioners, nurses, and doctors at local Clinics and hospitals throughout Mpumalanga in 2012. Table 1 presents a sample from the dataset from Mpumalanga health department.

Table 1. Sample of TB patient dataset

Location	Agegroup	Gender	Classification	Outcome
Ehlanzeni	35 to 65	Male	Pulmonary	Died
Ehlanzeni	0 to 35	Female	Extra-Pulmonary	Died
Gert Sibande	35 to 65	Female	Pulmonary	Died
Ehlanzeni	35 to 65	Male	Pulmonary	Died
Nkangala	35 to 65	Male	Pulmonary	Died
Nkangala	35 to 65	Male	Extra-Pulmonary	Alive
Nkangala	0 to 35	Male	Pulmonary	Alive
Ehlanzeni	0 to 35	Male	Pulmonary	Alive
Nkangala	35 to 65	Male	Pulmonary	Alive
Ehlanzeni	65+	Female	Pulmonary	Alive

The Tuberculosis patient dataset consist of Location (i.e., District in Mpumalanga shown in Fig. 3), Agegroup, Gender, Tuberculosis classification and Treatment outcome. Table 2 presents the variables and their possible states.

Table 2. Definition of the possible states of variables TB, L, A, G and TO

Types of TB (TB)	Location (L)	Age group (A)	Gender (G)	Treatment outcome (TO)
Pulmonary = p	Nkangala = n	0 to 35	Male = m	Alive = a
Extra-Pulmonary = xp	Gert Sibande = g	35 to 65	Female = f	Died = d
	Ehlanzeni = e	65+		

5.2 Network Structure

The structure of the network captures the relationships between the variables. In our problem the factors that affect a user's chance of having Tuberculosis are Age group, Location and Gender. Then we add arcs from Age group, Location and Gender to Tuberculosis (TB) to show the cause effect relationship of TB. Similarly, having a particular type of TB affects the Treatment outcome of TB. So we then add arcs from TB to Treatment outcome. The resultant structure is shown in Fig. 4.

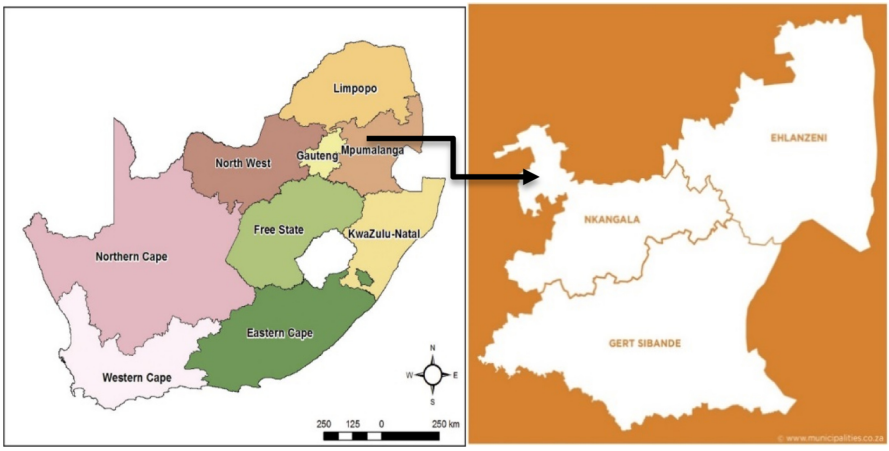


Fig. 3. Mpumalanga districts

5.3 Joint Probabilities Distribution

Figure 4 shows the construction of a Bayesian network that predicts the Tuberculosis prevalence. The joint probability distribution for a Bayesian network is obtained by the product of the local probabilities distributions for each node. Then the joint probability of our Bayesian network may be calculated as given by Eq. (3):

$$P(A, L, G, TB, TO) = P(A) \times P(L) \times P(G) \times P(TB|A, L, G) \times P(TO|TB) \quad (3)$$

For example, we can calculate joint probability using Eq. (3):
P(L = Nkangala, A = 0 to 35, G = Male, TB = Pulmonary, TO = Alive)
= P(L = Nkangala) × P(A = 0 to 35) × P(G = Male)
× P(TB = Pulmonary| L = Nkangala, A = 0 to 35, G = Male)
× P(TO = Alive| TB = Pulmonary)
= 0,53685 * 0,50279 * 0,28141 * 0,87179 * 0,86811 = 0.05749.

5.4 Conditional Probabilities

Given the prior probability, the conditional probability distribution for each types of TB given the state of Location, Gender and Age group can be computed. Then having TB causes the patient to go through the TB treatment thereby affecting the treatment outcome.

Given variable A and variable B are independent on each other. Then conditional probability can be computed using Eq. (2). The probability of A occurring given that B has already occurred is given by Eq. (2).

Suppose we want to calculate the probability of a female of age 0 to 35 living in Nkangala district of getting infected with Pulmonary TB. The conditional probability is calculated below:

P(TB = p| A = 0–35, G = f, L = n) = P(TB = p, A = 0–35, G = f, L = n)/P(A = 0–35, G = f, A = n).

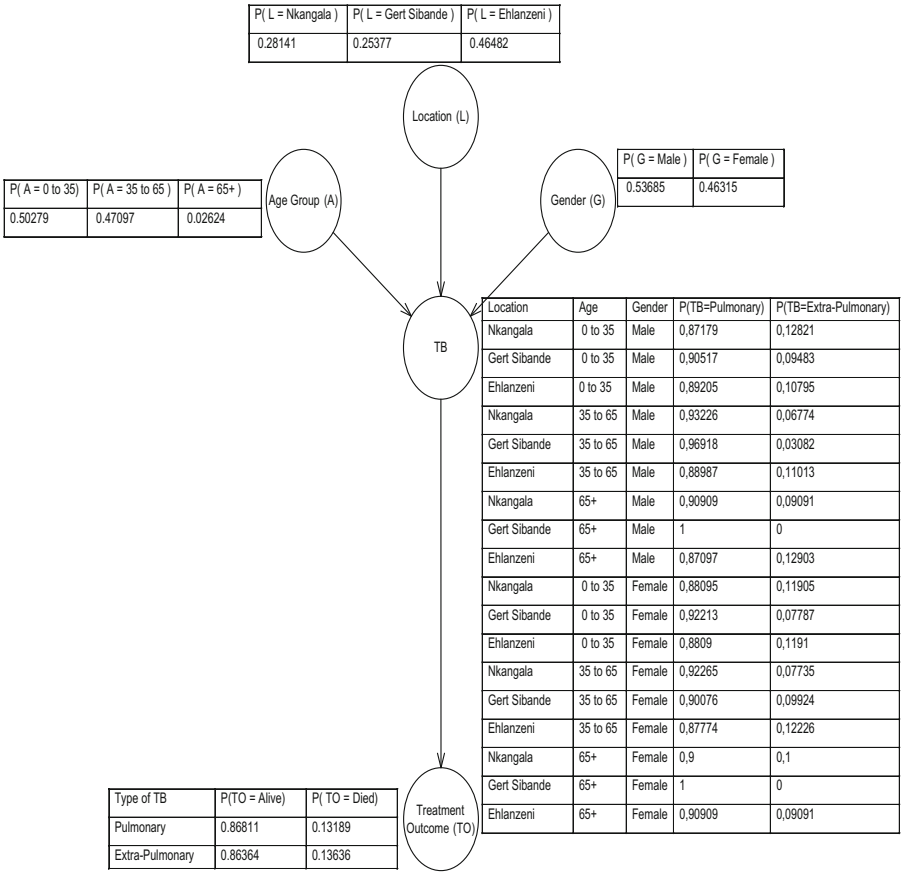


Fig. 4. Bayesian network for the tuberculosis example.

$P(TB = p\dashv \mid A = 0–35, G = f, L = n) = 0.05773/0.065531 = 0.88095$.

Therefore, the probability of a female of age group 0 to 35 in Nkangala district getting infected with Pulmonary TB is 88.095%.

6 Testing and Results

The results show that Pulmonary TB is the most prevalent and make more than 90% of the TB cases in Mpumalanga Province. Figure 5 shows the prior probability distribution of the types of TB.

Ehlanzeni district has the most cases of TB as shown in Fig. 6. Figure 7 shows that males are the most infected with TB in Mpumalanga as compared to females. The prevalence of TB across age group showed that age group 0 to 35 and 35 to 65 are the most vulnerable groups. The older groups of 65+ are less vulnerable as shown in Fig. 8. The rate of patients being cured and alive is high according to our results. However, more than 10% of the patient's treatment outcomes show that they have died. This is

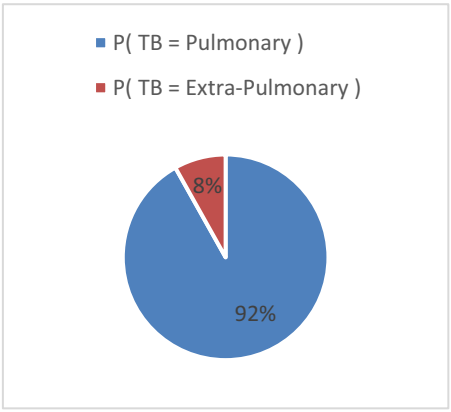


Fig. 5. Prior probability distribution of TB in terms of type of TB

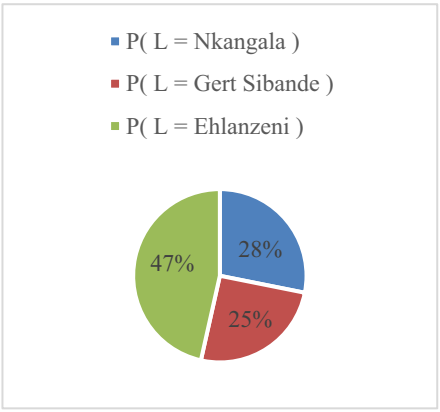


Fig. 6. Prior probability distribution of TB in terms of location

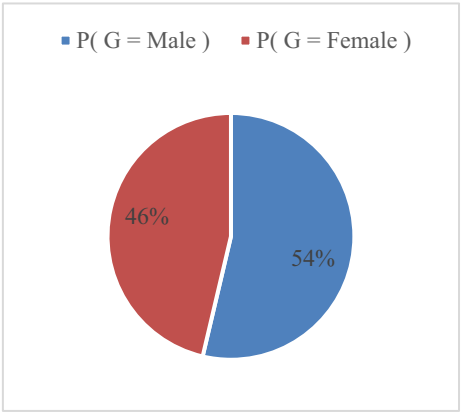


Fig. 7. Prior probability distribution of TB in terms of gender

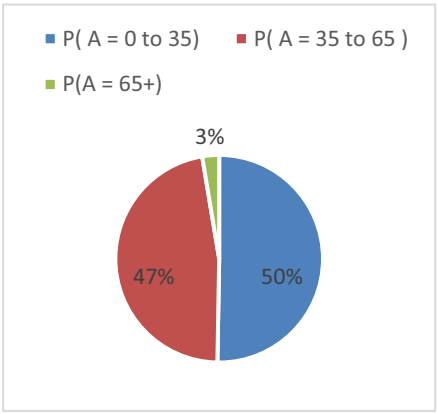


Fig. 8. Prior probability distribution of TB in terms of type of TB

shown in Fig. 9. The conditional probability distribution results for Treatment outcome, TO shown in Table 3.

Testing was done to validate and verify that the system is implemented correctly and according to the functional requirements. The following scenarios were used during testing.

In training the model, five different cases from doctors were injected to test the quality of the MHDSS developed. Each time a case was injected, a calculation was conducted to verify if the MHDSS system has predicted correctly. Let us examine one of the cases injected to the system. To calculate, we used probabilities from Table 3.

To Predict Treatment Outcome of new sample X:

L = Nkangala, A = 0 to 35, G = male, TB = Pulmonary.

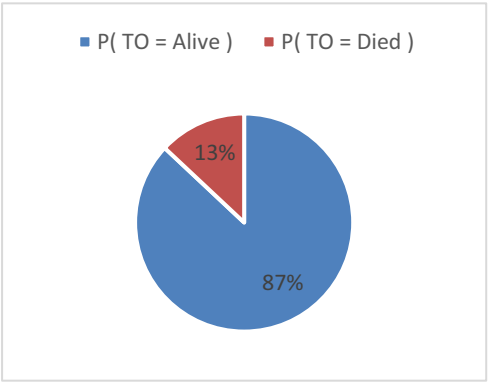


Fig. 9. Prior probability distribution of TB in terms of type of TB

Table 3. Probability distribution for node TO

P(TO = a)	0.86998	P(TO = d)	0.13002
P(L = n | TO = a)	0.28498	P(L = n | TO = d)	0.25751
P(L = g | TO = a)	0.2516	P(L = g | TO = d)	0.26824
P(L = e | TO = a)	0.46341	P(L = e | TO = d)	0.47425
P(A = 0 to 35 | TO = a)	0.52985	P(A = 0 to 35 | TO = d)	0.32189
P(A = 35 to 65 | TO = a)	0.44961	P(A = 35 to 65 | TO = d)	0.61373
P(A = 65 + | TO = a)	0.02054	P(A = 65 + | TO = d)	0.06438
P(G = m | TO = a)	0.53273	P(G = m | TO = d)	0.56438
P(G = f | TO = a)	0.46727	P(G = f | TO = d)	0.43562
P(TB = p | TO = a)	0.89987	P(TB = p | TO = d)	0.91416
P(TB = xp | TO = a)	0.07927	P(TB = xp | TO = d)	0.08369

P(TO = a | X) = P(TO = a) * P(L = n | TO = a) * P(A = 0 to 35 | TO = a) * P(G = m | TO = a) * P(TB = p | TO = a) = 0.86998 * 0.28498 * 0.52985 * 0.532738 * 0.89987 = 0.062975.

P(TO-d | X) = P(TO = d) * P(L = n | TO = d) * P(A = 0 to 35 | TO = d) * P(G = m | TO = d) * P(TB = p | TO = d) = 0.13002 * 0.25751 * 0.32189 * 0.56438 * 0.91416 = 0.00556.

The highest probability is for treatment outcome being alive. To evaluate if the system has predicted correctly, we compute the utility of the system. Table 4 shows that the system was able to correctly predict 3 out the five cases. So it showed a utility of 60%. However, we performed further training with data we got from the hospitals.

The system showed good performance as it could execute a probability results in about 0.15 μs. Further training of the proposed system was done to get the utility of the model. New hourly data was further collected from the department of health. The

Table 4. Results after five tests

Location	Agegroup	Gender	Classification	Outcome	Prediction
Nkangala	0 to 35	Male	Pulmonary	Alive	Alive
Gert Sibande	0 to 35	Male	Pulmonary	Died	Alive
Nkangala	35 to 65	Female	Pulmonary	Alive	Alive
Ehlanzeni	0 to 35	Female	Pulmonary	Alive	Died
Gert Sibande	0 to 35	Male	Pulmonary	Alive	Alive

first hour, new data was imported to update the MHDSS system and new 'belief' would be computed for each and every variable. This further trained the Bayesian Network, resulting in improved accuracy of the model. After each and every hour data was updated, the conditional probability of a person contracting Tuberculosis was computed and prediction calculation similar to the one shown above was done for each case in the new hourly dataset batch. We then compared the prediction with the data that we received from the department to confirm if the system has been predicted correctly or incorrectly. This was then done 63 times (runs) by feeding the system with hourly updated data and each time a percentage of correctly and incorrectly predicted cases were recorded to produce the utility graph of the MHDSS shown in Fig. 10. It shows utility of 88.2% which can be improved by further training of the MHDSS with new data when it becomes available.

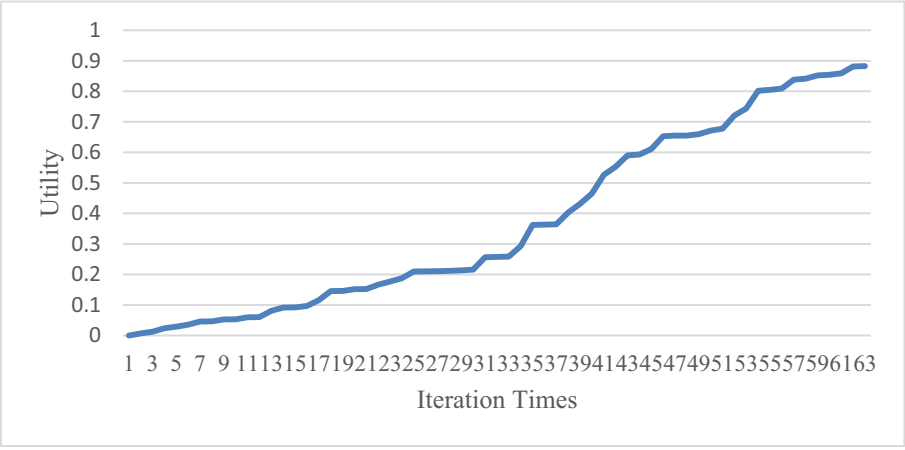


Fig. 10. MHDSS utility

7 Conclusion

Upon testing the model, it predicted the probability of getting infected with Tuberculosis given location, age group and gender. We then did system training, making 63 runs of tests, getting a utility of 88.2%. The MHDSS was scalable, robust and efficient. Bayesian Networks are suitable for use in problems that require reasoning and decision-making under uncertain conditions. Because of their strong mathematical base they handle uncertainty very well using probabilities, and are also simple to build and easy to use. The Bayesian networks thus provide a probability of a specific event occurring. Further research is required to enhance the response of the system. The scalability, security and interoperability of the system could be areas of further research. It can be thought provoking to investigate on how the performance can be improved perhaps by using other models (i.e., neural network, genetic algorithm, etc.) instead of Bayesian Network. Predict the prevalence of Tuberculosis given the rate at which the population is growing. This could be used by government health departments for the control of Tuberculosis. Adoption of the system by the users, government and health practitioners in decision making could be investigated further.

Acknowledgement. The authors would like to acknowledge Tshwane University of Technology for financial support, Mpumalanga provincial health department for data and Tshwane University of Technology ethics committee for ethical approval (Ref. REC2014/01/002) of this project.

References

1. WHO: Tuberculosis Prevalence Surveys. World Health Organisation, New York (2014)
2. Kumari, M., Rajan, V., Anshul, A.: Prediction of diabetes using Bayesian network. Int. J. Comput. Sci. Inf. Technol. **30**(4), 5174–5178 (2014)
3. Thakkar, B.A., Hasan, M.I., Desai, M.A.: Health care decision support system for swine flu prediction using Naives Bayes classifier, Spain (2010)
4. Palaniappan, S., Rafiah, A.: Intelligent heart disease prediction system using data mining techniques. In: IEEE/ACS International Conference, Doha, pp. 108–115 (2008)
5. Sharareh, R., Niakan, K., Xiao-Jun, Z.: Evaluation and comparison of different machine learning methods to predict outcome of tuberculosis treatment course. J. Intell. Learn. Syst. Appl. Suzhou **2013**, 184–193 (2013)
6. Kogeda, O.P., Agbinya, J.I., Omlin, C.W.: A probabilistic approach to faults prediction in cellular networks. In: International Conference on Systems and International Conference on Mobile Communications and Learning Technologies, Sydney (2006)
7. Horvatovich, P., Bischoff, R.: Comprehensive Biomarker Discovery and Validation for Clinical Application. Royal Society of Chemistry, Cambridge (2013)
8. Margaritis, D.: Learning Bayesian Network Model Structure from Data. Carnegie Mellon University, Pittsburgh (2003)
9. Booch, G., Jocobson, I., Rumbaugh, J.: The Unified Modeling Language User Guide. Addison-Wesley, Boston (1999)
10. Achieng, O.G., Kogeda, O.P., Lall, M.: A review of probabilistic modelling of pipeline leakage using Bayesian networks. J. Eng. Appl. Sci. **12**(12), 3163–3173 (2017)

11. Kgoete, S.F., Kogeda, O.P.: A prediction model for mitigating tuberculosis infection for HIV patients at greater Tubatse local municipality in South Africa. In: The Proceedings of Southern Africa Telecommunication Networks and Applications Conference (SATNAC), Fairmont, Zimbali Resort in Ballito, KwaZulu-Natal, South Africa, 1–4 September 2019, pp. 240–247 (2019)
12. Dladlu, N., Kogeda, O.P.: A decision support system for cancer prevalence in South Africa. In: The Proceedings of IEEE 8th International Conference on Computing Technology and Information Management (NCM & ICNIT), vol. 1: NCM Track, Grand Hilton Hotel, Seoul, Korea, 24–26 April 2012, pp. 532–537 (2012)