

Expert System for Diagnosing Diseases Tuberculosis Using Bayes Method In Puskesmas Petumbukan

Muhammad Ridho Syahwana¹, R. Mahdalena Simanjan²

^{1,2} Informatics Engineering Study Program, STMIK Pelita Nusantara, Jl. St. Iskandar Muda
No. 1 Medan, Nort Sumatera, Indonesia, 20154

E-mail:syahridho2@gmail.com

Abstract-Tuberculosis is an infection caused by acid-fast bacilli (AFB). Tuberculosis is an infectious disease that can affect anyone over the air. Tuberculosis is an infectious disease that can strike between the age of 15-35 years. How to diagnose tuberculosis is by way of the expert of experts interviewed then test sputum samples using a lab to determine the positive or negative tuberculosis disease so it takes a long time. Therefore it takes an expert system with Bayesian methods to facilitate the diagnosis of tuberculosis. This expert system developed using Microsoft Visual Studio 2010 as well as by using Microsoft Access 2010 database.

Keywords : Diagnosis of tuberculosis, expert systems, Bayesian method, Microsoft Visual Studio 2010, Microsoft Access 2010.

1. Introduction

The development of science and technology has progressed so rapidly. One is the increasing advances in technology to be used in the medical world. The need for a computer in the medical world that is as a supporting performance in order to speed up diagnosis of the disease in a swift, responsive, and accurate. Diagnosis of the disease in the medical world can sometimes take a long time, especially of rare diseases in the know by masyarakat. Ada steps to diagnose a disease such as interviewing patients about symptoms complained of the test samples in the laboratory to determine the diagnosis of the disease, but it is not effective and require considerable time to determine the results of laboratory samples. Examples of such cases occur when handling the diagnosing of the disease *Tuberculosis (TB)* which takes a long time to obtain results. Researchers using an expert system to facilitate later in diagnosing tuberculosis disease by entering some data like the signs of symptoms of the disease to the herd data by research the expert in helping Tuberculosis. Untuk disease progressed Expert System applications, researchers are also using Bayes Method. Bayes method is used because it can calculate the probabilities on all hypotheses, so that if the value of the probability of high value then it will know how many percent positivity rate of tuberculosis disease patients.

2. Theory

2.1 Expert system

According to Professor Edward Feigenbaum in Muhammad Arhami (2004: 2) Expert System adalah an intelligent computer program that uses *knowledge* (Knowledge) and inference procedures to solve the problem is quite difficult so it takes an expert to complete.

2.2 Metode Bayes

Bayes adalah method is one method to overcome the uncertainty of the data. In this method, the use of the conditional probability is as the base, where the statistical science of probability of X in Y is the probability of intersection X and Y of probability Y.

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

Where:

P(H | E): the probability of the hypothesis H if given evidence E

P(E | H): probability evidence E if known hypothesis H

P (H): the probability of the hypothesis H regardless evidence any

P (E): probability evidence E

2.3 Tuberculosis (TB)

According to Shinta Sunaryati Septi (2011: 81) tuberculosis is an infection which is caused by acid-fast bacilli (AFB). Tuberculosis is an infectious disease that can affect anyone over the air. The main organ in attacking this disease are the lungs, or often known as pulmonary tuberculosis, there remains a wide variety of more diseases like tuberculosis that TB skin, glands TB, and TB Brain

3. Research methods

The method in this study was developed through a framework terlebih first. The description of the framework in the methods of this study are as follows:

3.1 Identification of problems

Identify problems in the study as less efesiennya to diagnose tuberculosis for a long time, the lack of computerization in the data store.

3.2 analyze Problems

Analysis of the problem that is studying the issues that will be discussed in the study by obtaining data that will be used. In this research is the problem that caused the duration of a process results in the diagnosis of tuberculosis, so it must need a system to speed up the handling in diagnosing tuberculosis disease.

3.3 formulating Problems

At this stage of formulation of the problem that will be created is to create a system to facilitate an pendiagnosannya. The expert system is a system that is in use in this study.

3.4 Implementation of Algorithms Theorem Bayes To Diagnose Disease Tuberculosis

The application of the algorithm method *Bayes* selected for diagnosing tuberculosis because it can menghitung probability that arise from the patient complained of symptoms, so they will find a positive or negative outcome of patients affected by tuberculosis.

3.5. Data analysis

The data analysis is an attempt to process data into information so that the characteristics of the data can be understood and useful for the solution of problems, particularly issues related to research. Another definition of the data analysis is an activity undertaken to change the data results of research into information that can later be used to draw conclusions.

The results of research that has been done, the data obtained tuberculosis diseases as follows:

1. Pulmonary TB
2. TB glands
3. TB Skin
4. Bone TB
5. TB Brain

Table 1
List of tuberculosis disease

Code	Disease	Information	Weight
P1	Pulmonary TB	Tuberculosis disease that attacks and infected at the lung	0.3
P2	TB Skin	Tuberculosis disease that attacks on the skin, usually as a chronic sore that never healed.	0.5
P3	TB glands	Tuberculosis disease that attacks on the lymph, usually occurring enlarged lymph nodes in the neck.	0.6
P4	Bone TB	The tuberculosis disease attack on the spine in the thoracic area (back of the chest).	0.5
P5	TB Brain	<i>tuberculosis</i> This usually affects children with an attack on the lining of the brain or in commonly called meningitis.	0.4

Table 2
List of Symptoms Tuberculosis

CODE	SYMPTOMS	WEIGHT
G01	Persistent cough with sputum for three weeks or more.	0.5
G02	Sputum mixed with blood	0.7

CODE	SYMPTOMS	WEIGHT
G03	Shortness of breath and pain in the chest	0.4
G04	Weakness, loss of appetite, and weight loss also decreased	0.6
G05	Sweat at night even without activity	0.2
G06	Fever fever (mild fever) more than a month	0.6
G07	The emergence of a lump in the neck area, armpits, and groin	0.6
G08	Repeated diarrhea that does not heal with ordinary diarrhea drug	0.2
G09	In case of partial bronchial obstruction (channel into the lungs) due to compression of the enlarged lymph nodes will cause wheezing sound (the sound of breathing weakened) accompanied by shortness	0.4
G10	If there is fluid in the pleural cavity (wrapping lungs), can be accompanied by complaints of chest pain.	0.4
G11	When the bone, there will be symptoms such as bone infection at some point be able to form a channel and lead to the overlying skin, in this estuary will discharge pus.	0.8
G12	In children can about the brain (brain wrapper layer) and is referred to as meningitis (inflammation of the lining of the brain), the symptoms are high fever, the loss of consciousness and convulsions seizures	0.8
G13	Their scrophuloderma or TB skin (such as chronic ulceration and never healed).	0.8
G14	Phlycternular their conjungtivitis (sometimes in the eyes there is red, then there are white spots)	0.2
G15	The existence of specific lymphadenopathy (enlarged lymph nodes in the neck)	0.8
G16	In tuberculosis, enlarged glands normally would in a row or more than one	0.4

3.6. Analysis Bayes Method

Calculation of Bayesian methods in diagnosing tuberculosis disease the system is designed based on Bayes algorithm discussed in the calculation according to the symptoms suffered by a person. Persamaaa formula Bayes method is as follows:

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)} \quad (1)$$

$$P(H_i|E) = \frac{P(E|H_i) \cdot P(H_i)}{\sum_{k=1}^n P(E|H_k) \cdot P(H_k)} \quad (2)$$

$$P(H|E, e) = P(H|E) \cdot \frac{P(e|H)}{P(e)} \quad (3)$$

Table 3
Table Value Bayes

Bayes theorem	Value Probability Bayes
There is no	0 - 0.2
Maybe	0.3 - 0.4
Most likely	0.5 - 0.6
definitely Hampi	0.7 - 0.8
Certainly	0.9 - 1.0

3.7. Diagnose Disease Expert System Rule Tuberculosis

rule 1: IF Persistent cough with sputum for three weeks or more is *True*
AND Sputum mixed with blood is *True*
AND Shortness of breath and pain in the chest is *True*
AND Weakness, loss of appetite, and weight loss also decreased is *True*
AND Sweat at night even without activity is *True*
AND Fever fever (mild fever) more than a month is *True*
AND The emergence of a lump in the neck area, armpits, and groin

is True

AND In case of partial bronchial obstruction (channel into the lungs) due to compression of the enlarged lymph nodes will cause wheezing sound (the sound of breathing weakened) accompanied by shortness *is True*

AND If there is fluid in the pleural cavity (wrapping lungs), can be accompanied by complaints of chest pain *is True*

THEN Pulmonary TB

rule 2: IF Fever (mild fever) more than a month's chest *is True*

AND Their scrophuloderma or TB skin (such as chronic ulceration and never healed) *is True*

THEN TB Skin

rule 3: IF Persistent cough with sputum for three weeks or more *is True*

AND Fever (mild fever) more than a month *is True*

AND The emergence of a lump in the neck area, armpits, and groin *is True*

AND In case of partial bronchial obstruction (channel into the lungs) due to compression of the enlarged lymph nodes will cause wheezing sound (the sound of breathing weakened) accompanied by shortness *is True*

AND The existence of specific lymphadenopathy (enlarged lymph nodes in the neck) *is True*

AND In tuberculosis, enlarged glands normally would in a row or more than one *is True*

THEN TB glands

rule 4: IF Fever (mild fever) more than a month *is True*

AND When the bone, there will be symptoms such as bone infection at some point be able to form a channel and lead to the overlying skin, in this estuary will discharge pus. *is True*

THEN Bone TB

rule 5: IF Fever (mild fever) more than a month *is True*

AND In children can about the brain (brain wrapper layer) and is referred to as meningitis (inflammation of the lining of the brain), the symptoms are high fever, the loss of consciousness and convulsions seizures *is True*

AND Repeated diarrhea that does not heal with ordinary diarrhea drug *is True*

THEN TB Brain

3.8 Case study

Analysis of tuberculosis disease diagnosis with Bayesian methods can be loaded as in the following cases:

A case of tuberculosis disease, in which symptoms of the disease and answer as follows:

1. Sputum mixed with blood = $0.7 = P(E | H1)$
2. Persistent cough with sputum for three weeks or more = $0.5 = P(E | H2)$
3. Shortness of breath and pain in the chest = $0.4 = P(E | H5)$
4. Fever (mild fever) more than a month = $0.6 = P(E | H6)$

Calculation Method Using Bayes

a) Determining the value of Probability

Define in advance the value of the probability of any evidence for each hypothesis based on the available sample data using Bayes probability formula.

$$G02 = P(E | H1) = 0.7$$

$$G01 = P(E | H2) = 0.5$$

$$G03 = P(E | H5) = 0.4$$

$$G06 = P(E | H6) = 0.6$$

b) Determining the value of the Universe

Looking universe by summing the values of the hypothesis

n

$$G1 + \dots + Gn$$

Gn

$$\begin{aligned} G02 &= P(E | H1) = 0.7 \\ G01 &= P(E | H2) = 0.6 \\ G03 &= P(E | H5) = 0.4 \\ G06 &= P(E | H6) = 0.6 \end{aligned}$$

$$0.7 + 0.5 + 0.4 + 0.6 = 2.2$$

G_n

Once the sum is known, then obtained the formula for calculating the value of the universe:

$$P(H_i) = \frac{P(H_i)}{G_n}$$

$$G02 = P(H1) = \frac{0.7}{2.2} = 0.31$$

$$G01 = P(H2) = \frac{0.5}{2.2} = 0.22$$

$$G03 = P(H5) = \frac{0.4}{2.2} = 0.18$$

$$G06 = P(H6) = \frac{0.6}{2.2} = 0.27$$

c) Determining the value of probability P (Hi)

Once the value of P (Hi) is known, the value of the probability of the hypothesis H regardless of any evidence.

G_n

$$0.31 * 0.7 + 0.22 * 0.5 + 0.18 * 0.4 + 0.27 * 0.6 = 0.561$$

G_n

d) Determining the value of P (Hi | E)

Finding the value of P (Hi | E) or the probability of the hypothesis Hi is true if given evidence E.

$$P(H_i | E) = \frac{P(H_i) * P(E|H_i)}{G_n}$$

$$P(H1 | E) = \frac{0.31 * 0.7}{0.561} = 0.38$$

$$P(H2 | E) = \frac{0.22 * 0.5}{0.561} = 0.19$$

$$P(H5 | E) = \frac{0.18 * 0.4}{0.561} = 0.12$$

$$P(H6 | E) = \frac{0.27 * 0.6}{0.561} = 0.28$$

e) Determining the value of Bayes

After all values P (Hi | E) is known, then the total the entire value Bayes with the following formula:

G_n

$$= P(E|H1) * P(H1|E1) + ... + P(E|Hi) * P(Hi|Ei)$$

G_n

a. Pulmonary tuberculosis disease

G_n

$$= 0.7 * 0.38 + 0.5 * 0.19 + 0.4 * 0.12 + 0.6 * 0.28$$

G_n

$$= 0.64 * 100\%$$

$$= 64\%$$

b. Skin tuberculosis disease

G_n

$$= 0.6 * 0.28$$

G_n

$$= 0.168 * 100\%$$

$$= 16.8\%$$

c. Gland tuberculosis disease

$$\begin{aligned}n &= 0.6 * 0.28 \\Gn &= 0.168 * 100\% \\&= 16.8\%\end{aligned}$$

- d. Bone tuberculosis disease

$$\begin{aligned}n &= 0.6 * 0.28 \\Gn &= 0.168 * 100\% \\&= 16.8\%\end{aligned}$$

- e. Brain tuberculosis disease

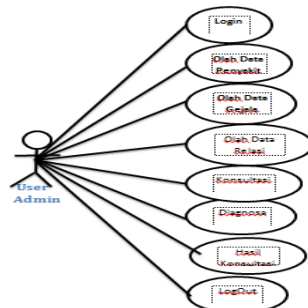
$$\begin{aligned}n &= 0.6 * 0.28 \\Gn &= 0.138 * 100\% \\&= 16.8\%\end{aligned}$$

Of the calculation process using Bayesian methods above, it can be seen that the patient "Great Possibilities" underwent pulmonary tuberculosis disease-paru with value of 0.64 or 64% confidence.

4. Results and Discussion

4.1 Use Case Diagram

Use case diagrams digunakan untuk menggambarkan dan menganalisis interaksi the user terjadinya antara sistem sehingga dapat dipahami dengan lebih mudah. Here is a use case diagram tuberculosis disease diagnosis expert system using Bayesian methods.



Picture 1. Use Case Diagram Diagnosis of tuberculosis disease

Explanation Use Case Diagram:

- 1) Users must login first to get into the system.
- 2) The user then perform data processing both the data relationships, symptom data, and data of diseases.
- 3) Users can also enter data consultation as symptoms in the patient's natural to eventually get the results of the consultation.

Implementation is the execution and at the same time for sistem yang testing has been approved and Bayes Method which is used as a draft in an expert system.

4.2 Form symptom

Figure 2. Interface Form Symptoms

4.3 Form Consultation

Figure 3. Interface Consultation Form

4.4 Form reports Results

Figure 4. Interface Form Reports

5. Conclusion

Based on the research and discussion above, it can be concluded as follows:

- 1) Can be analyzed in diagnosing tuberculosis.
- 2) Can design a system for diagnosing tuberculosis expert menggunakan Sistem with Bayes Method into Microsoft Access database.
- 3) Program can apply expert system to diagnose tuberculosis using Bayes methods.

6. Reference

- [1]. Erick Kurniawan, 2010. Visual Basic 2010. Yogyakarta: Andi
- [2]. Double Anggara, Gede Pramayu, Arif Wicaksana. Build Expert System Using Bayes Theorem To Diagnose Lung Disease. National Seminar on Information and Multimedia Technology, 2016, 81.
- [3]. Haer Talib, 2011. The Complete Guide to Microsoft Access 2010. Jakarta: PT Gramedia.
- [4]. Hengki Tamando Sihotang. Perancangan Expert Systems Applications. TIKomSIN, 2018, 6 (1), 37.
- [5]. Muhammad Arhami, 2004. Basic Concepts of Expert Systems. Yogyakarta: Andi.
- [6]. Johan Wahyudi Muhammad, Abdul Fadl. Sistem Pakar For Identifying Disease Using Galah Methods Bayes Theorem, Srijana Journal of Information Engineering, 2013, 1 (1), 15.
- [7]. Nur Aini, Ramadiani, Heliza Rahmania Hatta. Expert System Pendiagnosa tuberculosis disease. Mulawarman Informatics Journal, 2017, 12 (1), 57.
- [8]. Septi Shinta Sunaryati, 2011. 14 Most Frequently Diseases Attacking and Very Mematikan. Yogyakarta: Flash Books.
- [9]. Windah Supartini, Hindarto. *Web-Based Expert System Forward Chaining Method In Early Disease Diagnosing tuberculosis in East Java*. KINETIC, 2016, 1 (3), 149.
- [10]. Yossi Octavina, Abdul Fadlil. Expert System to Diagnose Respiratory Disease and Lung On Method Using Certainty Factor. Bachelor of Engineering Informatics Journal, 2016, 2 (2), 1126.