

Implementation of the Bayes method for diagnosing tuberculosis

Nina Sari¹⁾*, Volvo Sihombing²⁾, Deci Irmayani³⁾

^{1) 2) 3)} Labuhan Batu University, Rantauprapat, Indonesia

¹⁾ninasari@gmail.com, ²⁾volvolumbantoruan@gmail.com, ³⁾deacyirmayani@gmail.com

Submitted: Mar 4, 2021 | **Accepted:** Apr 26, 2021 | **Published:** May 4, 2021

Abstract: Tuberculosis is an infection caused by acid-resistant bacilli or an infectious disease that can attack anyone through the air is also a dangerous infectious disease besides it is also a chronic or chronic disease that can strike between the ages of 15-35 years. The purpose of this study is to help prevent tuberculosis by implementing an expert system using the Bayes method. The method used in this research includes identifying problems faced in the medical world for treating tuberculosis, analyzing the problem, then formulating the problem and applying an expert system with the Bayes method to solve the problems that are obtained, the next stage is designing an application as needed, testing the application with the aim of knowing the success rate of the system. The implementation of the Bayes method in diagnosing tuberculosis is found. The result is that the calculation process using the Bayes method is based on the symptoms experienced by the patient. It can be seen that the patient is "most likely" to have pulmonary tuberculosis with a confidence value of 0.64 or 64%. From the results of the research conducted, it can be concluded that in diagnosing Tuberculosis by using the Bayes method expert system, it can help medical parties handle cases more quickly in terms of recognizing the symptoms of Tuberculosis so that people quickly know the disease they are experiencing.

Keywords: Bayes Method, Expert System, Microsoft Visual Studio 2010, Microsoft Access 2010, Tuberculosis Diagnosis.

INTRODUCTION

The development of science and technology has progressed so rapidly. One of them is the increasing technological advances for use in the medical world. The need for computers in the medical world, namely to support performance in order to accelerate disease diagnosis in an alert, responsive, and accurate manner (Aji et al., 2018)

Diagnosis of disease in the medical world (Aji et al., 2018) sometimes it takes a long time, especially diseases that are rarely known by the public. There are several steps to diagnosing a disease such as interviewing the patient about the symptoms that have been complained of and then testing the sample in the laboratory to find out the diagnosis of the disease, but this is ineffective and takes a long time to find out the results of the laboratory sample. Sample case (Cubfritua et al., 2018) like this happened when handling the diagnosis of tuberculosis (TB) sufferers who took a long time to find out the results.

Expert systems are used to facilitate later diagnosis of Tuberculosis by entering some data such as signs of symptoms of the disease by collecting data by conducting research on experts in Tuberculosis. (Putu & Dewi, 2016). The Bayes method is used because it can calculate the probability of all hypotheses (Saleh, 2015), so that if the probability value is large, it will know what percentage of the patient's level of positivity for tuberculosis (Gustina & Chandra, 2015).

The importance of this research is to quickly find out the complaints of the community who come to consult a doctor, which sometimes takes a long time, especially diseases that are rarely known by the public. There are several steps to diagnose a disease, such as interviewing a patient about the symptoms he has complained about and then testing the sample in the laboratory to find out the diagnosis of the disease, but by doing this research, patients no longer need to queue or wait a long time to find out what the disease is based on symptoms experienced by only matching the symptoms experienced to the questions displayed on the system being made.

The main objective of research carried out using the Bayes method to diagnose tuberculosis is to produce a system that is easy to visit and the accuracy of the system can be ascertained on the level of possibility to diagnose

* Corresponding Author



the disease based on the symptoms present in the disease. The Bayes method used in research can help medical problems in treating patients for the diagnosis of tuberculosis disease.

LITERATURE REVIEW

Expert system

According to (Syahputra & Halim, 2019) and the Expert System is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough that they require an expert to solve them. In the journal (Dewi et al., 2015) revealed that the expert system has advantages including:

- a. Making a layman can work like an expert.
- b. Can work with incomplete or uncertain information.
- c. Increase output and productivity. Expert systems can work faster than humans. This advantage means reducing the number of workers required, and will ultimately reduce costs.
- d. Improve quality.
- e. Expert systems provide consistent advice and reduce error rates.
- f. Makes complex equipment easier to operate because expert systems can train inexperienced workers.
- g. Reliable (Reliability).
- h. Expert systems can't get tired or bored. Also consistent in giving answers and always giving full attention.
- i. Have the ability to solve complex problems.

Expert system compensation according to (Resmiati & Supriatna, 2016):

1. User Interface
2. Knowledge Base
3. Knowledge Acquisition
4. Inference Engine
5. Workplace

Bayes Method

The Bayes method is a method for overcoming data uncertainty. In this method the use of conditional probability is the basis, where in statistics the probability of X in Y is the probability of infection X and Y of probability Y (Dewi et al., 2015).

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

Where:

- P (H | E): probability of hypothesis H if provided with evidence E
- P (E | H): probability of the emergence of evidence E if the hypothesis is known H
- P (H) : probability of hypothesis H regardless of any evidence
- P (E) : probability of evidence E (Manalu et al., 2017)

Tuberculosis (TB)

According to (Amin & Bahar, 2014) *Tuberculosis* is an infection caused by acid-resistant bacilli (AFB). Tuberculosis is a contagious disease that can attack anyone through the air (Nugroho, 2018). The main organ that is attacked by this disease is the lungs, or what is often known as pulmonary tuberculosis, there are still various kinds of tuberculosis, namely skin tuberculosis, glandular tuberculosis, and brain tuberculosis.

METHOD

The method in this research is structured through a framework that is in advance. The description of the framework in this research method is as follows:

1. Identification of problems
Identifying problems in research such as inefficiency in diagnosing tuberculosis because of a long time, there is no computerization in storing data.
2. Analyzing Problems
Problem analysis is studying the problems that will be discussed in the research by getting the data that will be used later. In research the problem that arises is the length of a process result in diagnosing Tuberculosis, so a system must be needed to speed up handling in diagnosing Tuberculosis (Hidayat et al., 2019) (Murni & Riandari, 2018).
3. Formulating Problems
At this stage the problem formulation that will be made is by creating a system to facilitate a diagnosis. The expert system is a system used in this study.



4. Data collection

The data obtained were obtained from the Cinta Mama clinic by directly interviewing the experts to obtain data on the symptoms and types of tuberculosis, from the research results found several symptoms and diseases including 16 symptoms and 5 types of tuberculosis disease (Aziz Amrullah & Ekojono, 2017).

5. Application of Bayes Theorem Algorithm to Diagnose Tuberculosis

The application of the algorithm using the Bayes method was chosen to diagnose tuberculosis because it can calculate the probability that the symptoms that patients complain about, so that later they will find positive or negative results for patients with tuberculosis.

6. Application Design

The application design stage was carried out in Visual Basic 2010 by submitting to the needs of users in the Cinta Mama Clinic medical midwife and the database used was MySQL.

7. Evaluation Testing

The testing phase is carried out directly with experts to diagnose the disease of patients whose data is consulted at the Cinta Mama clinic at the same time to find out the accuracy of the system and the results obtained. Evaluation is carried out when the results of the system are not in accordance with the needs of the clinic in terms of diagnosing Tuberculosis

8. Implementation

After testing and the results obtained are in accordance with the accuracy of the expert, the application built is implemented directly on the patient who comes and immediately knows the results of the consultation carried out on the system (Informatics et al., 2012) (- AMIK BSI Purwokerto & - AMIK BSI Pontianak, 2018).

RESULTS

Questionnaire Results

The results of research that have been carried out, obtained data on Tuberculosis disease as follows:

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

Table 1.
List of Tuberculosis Diseases

Code	Disease	Information	Weight
P1	Pulmonary TB	Tuberculosis disease that attacks and is infected in the lungs	0.3
P2	Skin TB	Tuberculosis that attacks the skin, usually like a chronic ulcer that doesn't heal.	0.5
P3	Glandular TB	Tuberculosis disease that attacks the lymph, usually lymph enlargement in the neck.	0.6
P4	Bone TB	Tuberculosis attacks the spine in the thoracic area (the back of the chest).	0.5
P5	Brain TB	<i>Tuberculosis</i> This usually affects children by attacking the lining of the brain or commonly called meningitis.	0.4

Table 2.
List of Tuberculosis Symptoms

CODE	SYMPTOMS	WEIGHT
G01	Cough continuously with phlegm for three weeks or more.	0.5
G02	Sputum mixed with blood	0.7
G03	Shortness of breath and pain in the chest	0.4
G04	Weakness, decreased appetite, and decreased body weight	0.6
G05	Sweating at night even without activity	0.2
G06	Chills (low-grade fever) for more than a month	0.6
G07	The appearance of a lump in the neck, armpits and groin area	0.6
G08	Recurrent diarrhea that does not resolve with ordinary diarrhea medicine	0.2
G09	If there is partial blockage of the bronchi (tubes leading to the lungs) due to pressure on the enlarged lymph nodes, it will cause a wheezing sound (weak breath sound) accompanied by shortness of breath.	0.4



G10	If there is fluid in the pleural cavity (covering the lungs), it can be accompanied by complaints of chest pain.	0.4
G11	When it hits the bone, there will be symptoms such as bone infection which at one time can form a channel and lead to the skin above it, at this estuary pus will come out.	0.8
G12	In children it can hit the brain (the lining of the brain) and is called meningitis (inflammation of the lining of the brain), the symptoms are high fever, decreased consciousness and seizures.	0.8
G13	The presence of scrophuloderma or skin tuberculosis (such as chronic, persistent sores).	0.8
G14	The presence of phlycternular conjunctivitis (sometimes there is red in the eye, then there is a white spot)	0.2
G15	Presence of specific lymphadenopathy (enlarged lymph nodes in the neck)	0.8
G16	In tuberculosis, usually enlarged glands will be lined up or more than one	0.4

Bayes Method Analysis

Bayes method calculation in diagnosing Tuberculosis disease in a system designed based on the Bayes algorithm which is discussed in the calculation according to the symptoms of a person's illness. The Bayes Method Equation formula is as follows:

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

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

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

Table 3.
Bayes Value Table

Bayes' Theorem	Bayes Probability Value
Nothing	0 - 0.2
Maybe	0.3 - 0.4
Most likely	0.5 - 0.6
Almost certain	0.7 - 0.8
Certainly	0.9 - 1.0

A case of Tuberculosis, where the symptoms and answers are as follows:

1. Sputum mixed with blood = 0.7 = P (E | H1)
2. Continuous coughing 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. Chills (mild fever) more than a month = 0.6 = P (E | H6)

Calculation Using the Bayes Method

1. Determining Probability Value

First define the probability value of each evidence for each hypothesis based on the available sample data using the Bayes probability formula.

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

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

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

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

2. Determining the Value of the Universe

Find the value of the universe by adding from the hypothesis

$$\sum_{Gn}^n G1 + \dots + Gn$$

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

$$G01 = P(E | H2) = 0.6$$

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



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

$$\sum_{Gn}^n 0.7 + 0.5 + 0.4 + 0.6 = 2.2$$

After the sum is known, the formula is obtained to calculate the value of the universe:

$$P(Hi) = \frac{P(Hi)}{\sum_{Gn}^n}$$

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

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

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

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

3. Determine the Probability Value P (Hi)

After the P value (Hi) is known, the probability value of the hypothesis H is without looking at any evidence.

$$\sum_{Gn}^n (0.31 * 0.7) + (0.22 * 0.5) + (0.18 * 0.4) + (0.27 * 0.6) = 0.561$$

4. Determine the P Value (Hi | E)

Looking for the P value (Hi | E) or the probability that the hypothesis Hi is true if evidence E.

$$P(Hi) = \frac{P(Hi) * P(E|Hi)}{\sum_{Gn}^n}$$

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

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

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

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

5. Determining Bayes Values

After all P values (Hi | E) are known, then add up all Bayes values with the following formula:

$$\sum_{Gn}^n = (P(E|H1) * P(H1|E1)) + ... + (P(E|Hi) * P(Hi|Ei))$$

a. Pulmonary Tuberculosis

$$\sum_{Gn}^n = (0.7 * 0.38) + (0.5 * 0.19) + (0.4 * 0.12) + (0.6 * 0.28)$$

$$= 0.64 * 100\%$$

$$= 64\%$$

b. Tuberculosis of the skin

$$\sum_{Gn}^n = (0.6 * 0.28)$$

$$= 0.168 * 100\%$$

$$= 16.8\%$$

c. Glandular Tuberculosis

$$\sum_{Gn}^n = (0.6 * 0.28)$$

$$= 0.168 * 100\%$$

$$= 16.8\%$$

d. Bone Tuberculosis



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

e. Tuberculosis of the brain

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

From the calculation process using the Bayes method above, it can be seen that the patient is "most likely" to experience pulmonary tuberculosis.ru with confidence value 0.64 or 64%.

From the results of tests carried out on patients with the Bayes method can be seen in table 4.

Table 4.
Test result

No.	Name	Belief Value	Information
1	Rusdiana said	64%	Most likely
2	Arisma Hanafi	66%	Most likely
3	Romianti Sitepu	85%	Certainly

DISCUSSIONS

The application of an expert system in the software to diagnose Tuberculosis is one way that can be done to get the optimal value, from the various symptoms that appear, a more accurate disease will be found in determining the diagnosis of Tuberculosis. This is done in many ways and must use appropriate methods and should not be done in a manner. arranged at will. The final results obtained from the analysis illustrate that the results of the calculation of the Bayes method for tuberculosis diagnosis obtained a success rate of 85%. The certainty obtained was only 64%, but in this study the certainty score was higher because the symptoms were more than in previous studies.

The researcher conducted the test 1. From the data entered, the researcher had 4 symptoms that were found to be tested in knowing the disease that was obtained. From the data shared, tested using the Bayes method. The results of manual calculations obtained a higher confidence percentage level of 64% with the information that the possibility of having pulmonary tuberculosis disease was high. The researcher then carried out the second test. Just like the first test, there are 5 symptoms that are used for testing.

The results of manual testing obtained a percentage level of confidence with the Bayes method of 66% most likely to experience pulmonary tuberculosis. The test results using the Bayes method indicate that all the symptoms chosen by the patient have a major effect on the level of confidence of the disease experienced by the patient. The researcher also tested the application that was built, namely the visual studio to find out the similarity of manual calculations carried out by testing the application that was built. With the tests carried out, the results are similar to manual calculations, so it can be concluded that the synchronization between manual calculations and applications with the same symptoms is the same. Thus the Bayes method is precisely used to diagnose pulmonary tuberculosis.

CONCLUSION

Based on the research that has been done, it can be concluded that the results of the analysis carried out can diagnose the disease with the bayaes method, and can design a system to diagnose tuberculosis using the Bayes Method Expert System into the Microsoft Access database. Also can apply an expert system program to diagnose tuberculosis using the Bayes method.

REFERENCES

- Aji, A. H., Furqon, M. T., & Widodo, A. W. (2018). Sistem Pakar Diagnosa Penyakit Ibu Hamil Menggunakan Metode Certainty Factor (CF). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(5), 2127–2134. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1556>
- Amin, Z., & Bahar, A. (2014). Tuberkulosis Paru. In *Buku Ajar Ilmu Penyakit Dalam*.
- Aziz Amrullah, A. J., & Ekojono, E. (2017). Sistem pakar diagnosa penyakit paru-paru dengan metode forward chaining. *Jurnal Informatika Polinema*. <https://doi.org/10.33795/jip.v2i1.48>
- Cubfritua, C., Sianturi, F. A., & Gea, A. (2018). Sistem Pakar Untuk Mendiagnosa Pertumbuhan Gigi Balita



- Dengan Menggunakan Metode Dempster Shaper. Jurnal Armada Informatika. <https://doi.org/10.36520/jai.v2i2.28>
- Dewi, P. S., Lestari, R. D., & Lestari, R. T. (2015). Sistem Pakar Diagnosis Penyakit Ikan Koi Dengan Metode Bayes. Komputa : Jurnal Ilmiah Komputer Dan Informatika. <https://doi.org/10.34010/komputa.v4i1.2404>
- Gustina, D., & Chandra, Y. I. (2015). Aplikasi Sistem Pakar Untuk Mendiagnosa Penyakit Paru Pada Anak Menggunakan Metode Rapid Application Development (RAD). Jurnal UMJ.
- Hidayat, T., Nasution, H. N., Nasution, S. W. R., & Fauzi, R. (2019). Sistem pakar untuk mendiagnosa penyakit lupus dengan menggunakan metode certainty factor. Jurnal education and development. <https://doi.org/10.37081/ed.v7i3.1201>
- Informatika, J. T., Informatika, T., Industri, S. I., Informatika, T., & Riau, P. C. (2012). Sistem Pakar Diagnosa Penyakit Ginjal Menggunakan Metode. Sistem Pakar Diagnosa Penyakit Ginjal Menggunakan Metode Hill Climbing.
- Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv . Papa dan Mama Pastries. Mantik Penusa, 1(2), 16–21. <http://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/257>
- Murni, S., & Riandari, F. (2018). Penerapan Metode Teorema Bayes Pada Sistem Pakar Untuk Mendiagnosa Penyakit Lambung. Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP). <https://doi.org/10.34012/jutikomp.v1i2.226>
- Nugroho, F. A. (2018). Perancangan Sistem Pakar Diagnosa Penyakit Jantung dengan Metode Forward Chaining. Jurnal Informatika Universitas Pamulang. <https://doi.org/10.32493/informatika.v3i2.1431>
- Putu, D., & Dewi, S. (2016). Sistem Pakar Diagnosa Penyakit Jantung dan Paru dengan Fuzzy Logic dan Certainty Factor. Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi).
- Resmiati, R., & Supriatna, A. D. (2016). Pengembangan Sistem Pakar Diagnosis Penyakit Cabai Paprika Berbasis Android. Jurnal Algoritma. <https://doi.org/10.33364/algoritma/v.13-1.191>
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. Creative Information Technology Journal.
- Syahputra, T., & Halim, J. (2019). Sistem Pakar Untuk Mendiagnosa Penyakit Menular Seksual (HIV / AIDS) Dengan Menggunakan Metode Case Based Reasoning (CBR). Jurnal Sains Dan Komputer (SAINTIKOM).