Implementation of Certainty Factor Method for Diagnose Tuberculosis

Wilsen[1], Moh. Iwan Wahyuddin[2], Ratih Titi Komalasari[3]

Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional, Jakarta 12520, Indonesia

Email: wilsnaseng@gmail.com, iwan_wyd@yahoo.com, ukhuwahku01@gmail.com

ARTICLE INFO	ABSTRACT
	Tuberculosis is an infectious disease caused by the bacteria mycobacterium tuberculosis. In 2017 Indonesia entered into the third largest number of TB cases in the world. Lack of public knowledge of the dangers of tuberculosis makes this disease is growing rapidly. This is the main reason why it is necessary to create a system that can diagnose the early symptoms of the disease so that it can assist in tackling tuberculosis early. An expert system is one of the techniques in the diagnosis of disease. This research aims to develop applications early diagnosis of tuberculosis disease system is expected to facilitate the public in the early diagnosis of tuberculosis. The system uses the calculation of symptoms / complaints using CF (certainty factor). Results of testing performed by the system and test method validation black box shows that each feature can work with both the application and the content therein can be trusted. In addition, this system is quite powerful in handling more users are accessing the system simultaneously.

1. Introduction

Tuberculosis (TB) is an infectious disease caused by the bacteria mycobacterium tuberculosis. This disease usually attacks the lungs but the disease can also affect other body parts[1], Based on data from the WHO in 2017 of tuberculosis to the death of the 4th largest in Indonesia, as many as 116 thousand people die of TB disease in Indonesia.

The expert system is a system that mimics the expert knowledge into a computer, in order to solve problems like that done by experts. The system is designed to mimic the expertise of a specialist in answering questions and solve a problem either in health or medicine, business, economics, etc. An important role of an expert may be replaced by a computer program that principle works to provide solutions such as that done by experts. An expert system is a branch of Artificial Intelligence (AI)[2],

Methods Certainty Factor (CF) is a method that defines the size of the capacity of a fact or rule, in expressing a confidence level of an expert to the problem at hand[3], This method also has a pretty good system performance and run as needed and have a high percentage of accuracy.

In this study the authors implement the method of Certainty Factor (CF) in an expert system for diagnosing tuberculosis. Web-based system is built so that users can access the system from anywhere and with different platforms.

2. Study of Literature

Research on the theme of tuberculosis disease diagnosis expert system also has been studied in previous research[1] - [5], One was the Tuberculosis Detection Expert System Based Android[1], Expert systems are designed using the android platform. Any diagnosis process chainning forward method, in the process of diagnosis results obtained are presented in the classical probability. From the test results to validate the application between experts and the system, the system is able to deliver the output as expected

In a study conducted expert system Trio Alfianto[3]expert system made early diagnosis of tuberculosis with certainty factor method. The application is built with the android operating system, the process of testing the accuracy of applications reached 80%, an error occurs due to input symptoms are entered when testing is still less specific symptoms. From the test results it can be concluded that the certainty factor method has a low error in the calculation. Then Nita Novianti[4]build an expert system to diagnose tuberculosis with fuzzy logic fuzzy logic method is
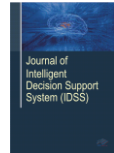
usually used for problems which contain elements of uncertainty and imprecision. From the results obtained degree of diagnostic accuracy using fuzzy logic by 70.33% the result is quite accurate but employ other methods need to be developed for a more accurate system. Then Ricky Surya[5]build an expert system for the detection of lung tuberculosis this system is built web-based calculations do that with certainty factor, the test results an expert system has an accuracy of up to 81.25% the figure can be said to have a high accuracy so that the certainty factor suitable for use in an expert system , Then Nur Aini[2] build an expert system diagnosis of tuberculosis it is built with web-based, and the methods used namely certainty factor, the results of system implementation expert system diagnosis of tuberculosis has a high degree of accuracy up to 85% of the figure accuracy of testing certainty factor is suitable for use in an expert system because it can generate high output accuracy level.

3. Research Methods

3.1. Flow Research

Broadly speaking, the stages of this research can be seen in Figure 1.

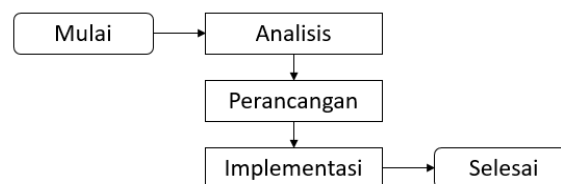**Fig 1.** Flowchart penelian

In the present study, there are 3 steps that should be based on the methodology used, namely:

a) analysis

In the analysis stage researchers to analyze the needs of the system and method of calculation

b) Design

At this stage, researchers studied the data and then do the design for the system to be implemented. In this design, researchers will model it with UML (Unified Modeling Language) diagrams, namely: Flowchart Diagram, Use Case Diagram

c) Implementation

At this stage, researchers will implement a design that has been created previously. Namely to build a system based on engineering calculations of Certain Factor and made it to the web platform

3.2. Method of Collecting Data

A. Study Documents

Studies document is data collection method is not directed to the subject of research. Studies document is a type of data collection that examines a wide range of useful documents for materials analysis

B. Interview

Interviews were conducted directly to the relevant parties in the research. In the present study interviews were conducted to a doctor in order to obtain precise data

3.3. System design

The design stage is the stage of the analysis phase wherein further at this stage of the design presented design applications such as interface design, and the design of a data base that will be applied to the system

a. diagram Flowchart

Flowchart diagram is a diagram that represents the type of algorithm, workflow or process, which displays the steps in the form of graphic symbols, and the order is connected with arrows. Expert System workflow Tuberculosis in the diagnosis menu can be seen in Figure 2. Menu diagnoses used by user to select the symptoms are felt, if the user has chosen symptoms are felt then the system will display the diagnostic results. The diagnosis shows the possibility of disease experienced by the user
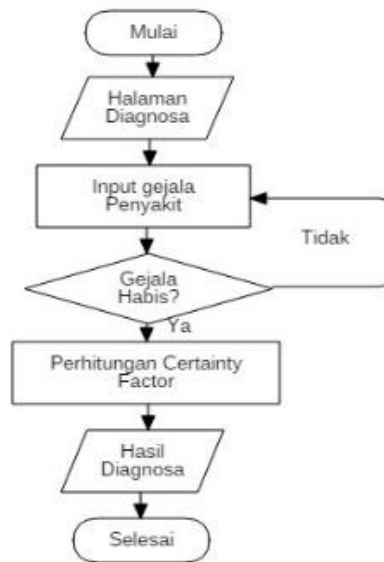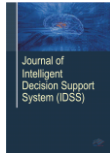
Fig 2, Flowchart Draft User Diagram

Workflow for admins can be seen in Figure 3. On the admin page, there are several menus the menu to set the symptoms experienced, manage disease, and also regulate the relationship between the symptoms of the disease

Fig 3. Flowchart Admin

b. Use Case Diagram
   Use Case diagram describing an interaction between one or more actors with the system to be created, Use Case Diagram for admin and user on the system can be seen in Figure 4. In general manual labor process is divided into three main processes data input process, the process of consultation and reporting process. As for the admin can process data of disease, symptom data and manage the relationships between the symptoms of the disease
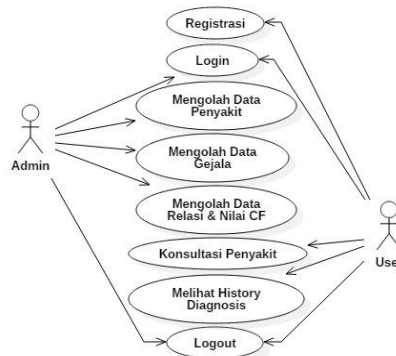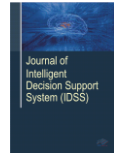
Fig 4. Use Case Diagram Draft User

3.4. Certainty Factor

Certainty factor is a method to manage uncertainty in a system based on rules. In mid-1975 Shortliffe and Buchanan to develop a model for MYCIN CF (type of expert system for diagnosis of the disease.

In theory certainty, qualitative data represented as the degree of confidence (degree of belief). In expressing some degree of confidence, certainty factor (CF) to assume a degree of belief an expert to the data. Certainty factor introduces the concept of belief / belief and disbelief / incertitude. This concept was formulated in the following basic formula:

CF [H, E] = MB [H, E] - MD [H, E]   .... (1)

Information:

CF (H, E)   : Certainty Factor (certainty factor) in the hypothesis H which is influenced by the fact E

MB (H, E) : Measure of Belief (confidence level), an increase in the size of the trust hypothesis H is affected by the fact E.

MD (H, E) : Measure of disbelief (the level of distrust) represents an increase of unbelief hypothesis H is affected by the fact E.

E   : Evidence (events or facts)

The next calculation is some combination of certainty factor to a certain premise

a)   Certainty Factor with a premise / single symptom:
CF_{gejala} = CF $[user]$ × $CF[pakar]$ .... (2)

b)   Certainty factor with more than one premise:
$CF_{kombinasi}$ $[CF1, CF2]$ = $CF1$ + $CF2$ * $(1 - CF1)$ .... (3)

c)   As for the possibility of disease calculates the percentage is calculated by the formula:
$CF_{persentase}$ = $CF_{kombinasi}$× 100% .... (4)

Certainty factor used to represent confidence value an expert to the data. The method used to obtain the value of confidence (CF) that interview, the values obtained from the interpretation of the "term" of the expert converted into CF value in accordance with Table 1.

Table 1.
Value Interpretation of experts

Uncertain Term (Condition Uncertain)	CF
definitely not	-1.0
Almost uncertain	-0.8
Unlikely	-0.6
Probably not	-0.4
Do not know	-0.2 to 0.2
Maybe	0.4
Most likely	0.6
almost certainly	0.8
Certainly	1.0

4.   Results and Discussion

4.1. Knowledge Base

The knowledge base is a representation of the knowledge of an expert. The knowledge base in this study obtained from interviews with experts and supplemented with references from books and journals related research. Symptoms of tuberculosis can be seen in Table 2.
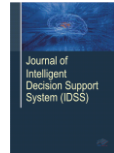
Table 2.
List of Symptoms

Code	Symptom
G01	Persistent cough with sputum for three weeks or more
G02	Productive cough out bloody
G03	Hard to breathe
G04	Pain in the chest
G05	Limp Agency
G06	Decreased appetite
G07	Weight loss
G08	Night sweats without any physical activity
G09	Fever fever (mild fever) more than a month
G10	Bumps on the emergence of problematic parts such as the neck, armpits, or groin
G11	Lumps can be shrunk first then enlarged and softened
G12	Enlargement lump gland and accompanied by fever

Application of the method requires a CF expert system rules (rule) as a benchmark in the calculation. This rule shows the relationship between the symptoms and the disease suffered by a person. Rule (rule) can be seen in Table 3.

Table 3.
Rule (Rule)

No.	rule
1	IF G01 AND G02 AND G03 AND G 04 AND G05 AND G06 AND G07 AND G08 AND G09 THEN Tuberculosis
2	IF G10 AND G11 AND G12 THEN Tuberculosis Glands

4.2. Manual calculation of Certainty Factor

The following example of a manual calculation methods certainty factor of lung tuberculosis
Users feel some of the symptoms and give the CF value of the perceived symptoms:
1. Bloody cough (0.4)
2. Cough for more than 2 weeks (0.6)
3. Weight Decrease (0.8)

Unknown value of CF that has been given the previous experts as follows:
1. Bloody cough (0.8)
2. Cough for more than 2 weeks (1.0)
3. Weight Decrease (1.0)

Then calculate the perceived value of CF symptoms user and the value of CF specialists provide
1. $CF_{symptom}$ 1 = CF_{user}(0.4) * $CF_{Specialists}$(0,8) = 0:32
2. $CF_{symptom}$ 2 = CF_{user}(0.6) * $CF_{Specialists}$(1.0) = 0.60
3. $CF_{symptom}$ 3 = CF_{user}(0.8) * $CF_{Specialists}$(1.0) = 0.80

After that calculates the value of CF disease by combining the values of CF symptoms
1. $CF_{Combination}$ (CF1, CF2) = 0.32 + 0.60 * (1 - 0:32) = 0728
2. $CF_{Combination}$ (CF1, CF2, CF3) = 0728 + 0.80 * (1-0728) = 0.9456

The last step that calculates the percentage chance of disease
1. 0.9456 * 100% = 94.56%

4.2. System Implementation

The following stages of the implementation of a system to the user
1. User Login page

It is the user login page for the application of expert systems.

Fig 5. Login page

2. Main page

Here the main page after logging in, in the main page of the menu, there are several consultations, history, about disease, find doctors and about us.
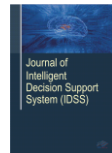
Fig 6. Main page

3. Weather Consulting
   Following consultation page. In this page is used to input the perceived symptoms user.

Fig 7. Weather Consulting

4. Consultation Results Page
   Following consultations result pages the user.

Fig 8. Diagnosis Results Page

5. Form a history of consultation
   The following pages consultations user history.

Fig 9. Consultation History page

The following stages of implementation of the system for admin

1. Admin Login Page
   It is for the admin login page of the application.

Fig 10. Login Page admin

2. Home Applications
   Here are the main admin page views, pages there are several menus namely: Data disease, symptom data and data relationships.
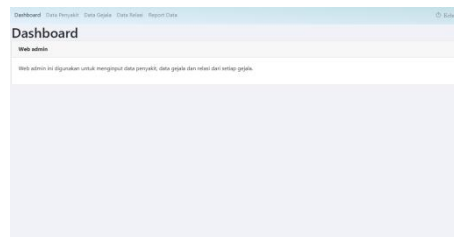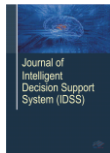
Fig 11. Home

3. Weather Data Disease
   The following data pages in the application of existing disease.
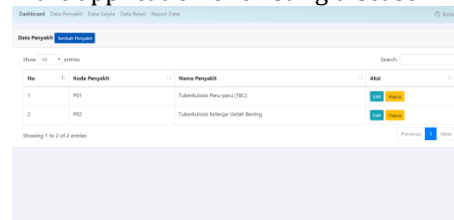

Fig 12. Page Data Disease

4. Data page Symptoms
   The following data page symptom of the disease
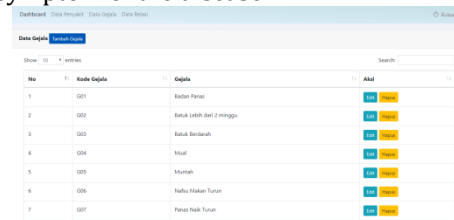
Fig 13. Page Data symptom

5. Maps Relationships
   The following relations page showing the relationship between the symptoms of the disease.
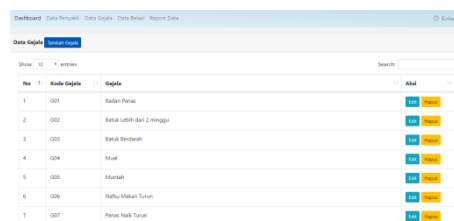
Fig 14. Page Relationships

6. Weather Report Data
   The following report page data indicating a history of diagnosis performed by users.

Fig 15. Weather Report Data

4.3. Testing Systems

Tests conducted to determine the function of every menu and button in the application, testing is done by black box method

Table 4.
Testing plan

Testing item	Testing Details	Types of Tests
Function Home User	Testing function "consultation"	Black Box
	Testing the function of "history"	Black Box
	Testing function "of the disease"	Black Box
	Testing function "find a doctor"	Black Box

Testing item	Testing Details	Types of Tests
	Testing function "on application"	Black Box
Function Main Page Admin	Testing function "Data disease" Testing function "symptom data" Testing the function of "data relationships" Testing function "Data report"	Black Box Black Box Black Box Black Box
Function Consultation page	Tests enter the "select box" Tests enter the "process"	Black Box Black Box
Consultation Results Page Functions	Testing the function of the results of consultations Testing function "finished"	Black Box Black Box
Search page function Doctors	Function test of "make an appointment" Testing function "practice schedule"	Black Box Black Box

Table 5.
Test result

Name Testing	Testing Forms	Expected Results	Test Result
Test the functionality of the consultation page	Click the menu button "consultation" on the main page	Displays menu consultation	succeed
Function test history page	Click the menu button "history"	Displays menu history	succeed
Test function on disease pages	Click the menu button "on disease"	Displays about the disease	succeed
Test the functionality of the page for a doctor	Click the menu button "find a doctor"	Displays the name of the doctor	succeed
Test the functionality of the app page	Click the menu button "about us"	Displays about the application	succeed
Test the functionality of the data page of disease	Click the menu button "Data disease"	Displaying data existing disease	succeed
Function test data pages symptoms	Click the menu button "symptom data"	Displaying data existing symptoms	succeed
Test the functionality of the data page of the relation	Click the menu button "data relationships"	Displaying data relationships	succeed
Test history page function diagnostics	Click the menu button on the "report a data"	Featuring user diagnosis history	succeed
Test the menu functions consultation	Give input in the "select box"	Displays the value of CF	succeed
Test the menu functions consultation	Click the "process"	Showing diagnosis	succeed
Test the menu functions consultation results	Click the "finished"	Done and displays the main page	succeed
Test the menu functions for a doctor	Click the "make an appointment"	Displays information about doctors	succeed
Test the menu functions for a doctor	Click the "schedule of practice"	Displays information about the doctor's schedule	succeed
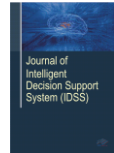
4.4. Validity testing

Validation testing conducted to determine the accuracy of the data between the symptoms and impact of disease

Table 6.
Validity Testing Data

Disease	symptom	System	book I	book II	book III
tuberculosis	Persistent cough with sputum for three weeks or more	✓	✓	✓	✓
	Productive cough out bloody	✓	✓	-	✓
	Hard to breathe	✓	✓	✓	✓
	Pain in the chest	✓	✓	✓	-
	weakness	✓	✓	✓	✓
	Appetite Down	✓	✓	-	✓
	Weight loss	✓	✓	✓	✓
	Berkeringan hours a day without any physical activity	✓	✓	✓	✓
	Fever fever (mild fever) more than a month	✓	✓	-	✓
tuberculosis glands	Bumps on the emergence of problematic parts such as the neck, armpits, or groin	✓	-	✓	-
	Lumps can be shrunk first then enlarged and softened	✓	-	✓	-
	Enlargement lump gland and accompanied by fever	✓	-	✓	-

From the table, testing the validity of the data, use the three books as reference. Among them is the "Public Health TB. Disease & Prevention Method "by dr. Yoanes Y. Laban[6], "Infectious Disease Extra pulmonary TB and Lung" by Misnadiarly[7]"Textbook of Diagnostic Microbiological Tuberculosis" by Ni Made Mertaniasih[8], Textbooks are written on the chart testing the validity of the data sequentially "Book I", "Book II", "Book III".

4.5. Testing Reliability Systems

System reliability testing was conducted to determine the performance of the system. Tests carried out using Jmeter Apache applications, this application is open source software made in Java designed to load test functional behavior and measure the performance. Testing is done with the user virtual simulations of 200 to 3000 users and the waiting time of 60 seconds. Here is a view after the test run.
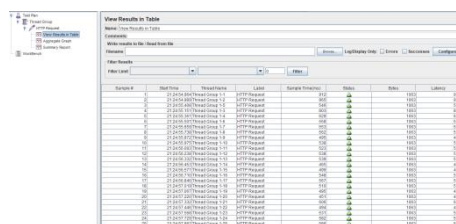

Fig 16. Test result

Figure 16 is a display using Apache JMeter test results using the laptop with Intel Core i3 2.0GHz processor and 4GB of RAM.

Table 7.
System Reliability Test Results

Thread (user)	Average	Std. dev	Throughput (/ second)
200	541	263.3246	3.321707
500	546	272.7109	8.267332
1000	553	276.1158	16.47827
1500	575	281.2749	24.73411
2000	591	285.7071	26.74691
2500	634	393.3974	32.18497
3000	655	400.1971	37 103

In Table 7 Thread (user) merupaan number of simulated virtual users who make requests to the system. Average is the average time it takes to respond to requests. Standard Deviation is a request data that deviates from the average sample. Throughput is the number of requests processed per time by the server.
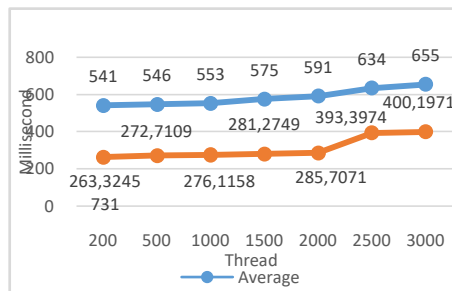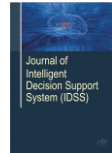
Fig 17, Graphs Average & Deviation Testing

In the figure 17 it can be seen that each thread (user), the greater the waiting time to respond to the request and the greater the value of the Standard Deviation. The limits of the standard deviation should be smaller or equal to the average value of the above if it then the test is said to be invalid[9],
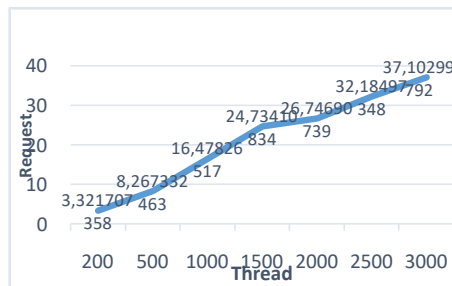

Fig 18. Graphs Throughput Testing

In Figure 18 it can be seen that the more thread then the value will be greater throughput.

5. Conclusion

Based on the research that is done it can be concluded that the symptoms of tuberculosis disease if not detected early can cause harm to the cause of death. To detect the disease early made expert system for diagnosing tuberculosis.

This research method is suitable certainty factor for the resolution of a problem because every problem there must be some factors that uncertain outcome. Based on black box testing and validation test can be concluded that the system has reliable information data. The researcher also testing the reliability of the system and the result of this system is able to work well in dealing with the many users up to 3000 users. By means of this expert system can function properly and accurately.

6. References

[1] E. D. Widianto, Y. W. Zaituun, and I. P. Windasari, "Aplikasi Sistem Pakar Pendeteksi Penyakit Tuberkulosis Berbasis Android," Khazanah Inform. J. Ilmu Komput. dan Inform., vol. 4, no. 1, p. 47, 2018.
[2] N. Aini, R. Ramadiani, and H. R. Hatta, "Sistem Pakar Pendiagnosa Penyakit Tuberkulosis," Inform. Mulawarman J. Ilm. Ilmu Komput., vol. 12, no. 1, p. 56, 2017.
[3] T. Alfianto and Benisius, "Aplikasi Diagnosa Dini Penyakit Tuberculosis Dengan Menggunakan Metode Certainty Factor," pp. 121–127, 2018.
[4] N. Novianti, D. Pribadi, and R. A. Saputra, "Sistem Pakar Diagnosa Pulmonary TB Menggunakan Metode Fuzzy Logic," vol. 5, no. 2, pp. 228–236, 2018.
[5] R. Surya and D. Gunawan, "Situsparu: Sistem Pakar Untuk Deteksi Penyakit Tuberkulosis Paru," J. Ultim., vol. 10, no. 1, pp. 41–47, 2018.
[6] dr. Yoannes Y. Laban, Kesehatan Masyarakat TBC. Penyakit & Cara Pencegahan. Penerbit Kanisius, 2008.
[7] Misnadiarly, Penyakit Infeksi TB Paru dan Ekstra Paru: Mengenal, Mencegah, Menanggulangi TBC Paru, Ekstra Paru, Anak, Pada Kehamilan. Jakarta: Pustaka Populer Obor, 2006.
[8] N. M. Mertaniasih, Ed., Buku Ajar Tuberkulosis Diagnostik Mikrobiologis. Surabaya: Pusat Penerbitan dan Percetakan Universitas Airlangga, 2013.
[9] Rishill Bhatt, "Understand and Analyze Summary Report in Jmeter," 2017. [Online]. Available: http://www.testingjournals.com/understand-summary-report-jmeter/. [Accessed: 29-Dec-2019].