International Journal of Engineering & Technology, 7 (4.38) (2018) 924-927

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Comparative Analysis of Veneral Diseases Diagnosed Using Tree and Certainty Factor Method

Elyza Gustri Wahyuni1 and Ricky Kurniawan2

1Department of Informatics, FTI
Indonesian Islamic University
Yogyakarta, Indonesia
2Department of Informatics, FTI
Indonesian Islamic University
Yogyakarta, Indonesia
**Corresponding author E-mail: elyza@uii.ac.id*

Abstract

The danger that can be generated from venereal diseases is, even infertility can lead to death. However, if the disease can be detected as early as possible by doing intense treatment of this disease can be more controlled and even cured. there have been several similar studies that examine the diagnosis of venereal disease, but no studies have tried to examine how effective the methods are applied to diagnose venereal diseases. So in this study wanted to compare the level of effectiveness of previous research with Decision Tree and Certainty Factor method. It was found that by comparative analysis Decision Tree and Certainty Factor methods show more effective system output if the system is made using Decision tree and Certainty Factor compared using Decision Tree only. It can be seen through the calculation using a Likert scale that shows the percentage of 92% (Strongly Agree).

Keywords: *Certainty Factor, Decision Tree, Venereal Diseases, Comparative Analysis.*

1. Introduction

The danger that can be generated from venereal disease, namely infertility, can even lead to death. However, if the disease can be detected as early as possible with the intense treatment, this disease can be more controlled and even cured. One way to detect the level of risk of genital disease is by creating an expert system that can be implemented to diagnose a venereal disease. Some researchers have already made a similar application, but none has tried to analyze the results of applications that have been made with aim to identify the accuracy of the diagnosis.

The previous research conducted by [1] was on the expert system to diagnose skin and venereal diseases in humans using Certainty Factor method. The study examined five types of skin and venereal diseases such as: Condyloma Acuminiata, Molluscum Contagiosum, Primary Syphilis, Chancroid and Genital Herpes by 40 symptoms, according to expert sources of skin and venereal diseases specialist Hj. Nafiah Chusniayati, Sp.KK., M.Sc as a doctor at the hospital PKU Muhammadiyah Yogyakarta. However, the drawback of the study lies on its lack of explanation about the general handling of the patient's diagnosed illness and lack of detailed explanation of the types of diseases that may be suffered by patients.

Another research conducted by [2] concerned on expert system diagnosis of skin and venereal diseases with CF and Fuzzy Logic method. The study combined two methods, CF and Fuzzy logic, but did not shed light on how many symptoms required to diagnose 10 different types of skin and venereal diseases for the 73% accuracy of the described test results with expert knowledge. The other drawback of the study was nearly the same with the previous study in that it did not thoroughly explained the handling of the patient's diagnosed illness and with lack of detailed explanation of the types of diseases that may be suffered by patients.

A subsequent research was conducted on the diagnosis of skin and genital diseases with the method of Decision Tree [3]. The test results have been described as 100% accuracy of results with expert knowledge, it was just because using the Decision Tree method and there was no measure of the certainty value of a person affected by venereal disease. For this research suggests to be developed by combining other methods that have the value of uncertainty.

Subsequent research on skin and genital diseases by combining 2 methods of Decision Tree with CF [4]. Continuing the previous research, this research attempts to combine 2 tree methods to determine the diagnosis result and CF to obtain the value of uncertainty. Hence, the result of this research will have a better conclusion compared to the previous research. Similarly, the response / expert test has obtained the same result with the diagnosis of specialist skin and genitalia.

Referring to some of these studies, this study would like to analyze in the perspective of two existing studies to observe the effectiveness of the method used first with the Decision Tree method and the second with the method of Decision Tree with Certainty Factor with the same case that is for Diagnosis of Sex Disease.

2. Theoretical Framework

Expert System

As one of the branches of artificial intelligence, expert system is built to solve the existing problems like an expert. The system

Copyright © 2016 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

works based on the provided knowledge basis stored in the system to solve the problems using inference reasoning. The knowledge basis is drawn from the discipline of knowledge or from expert scientist to be referred to as to resolve issues in some particular problems. Expert systems can help the people in general to solve complex problems like an expert. Interestingly, the system can be applied to assist an expert in resolving problems quickly and accurately. However, this expert system function is not designed to fully replace the role of an expert.

The inside of the expert system comprises of two main components, namely the knowledge basis and the inference engine [5]. There are two forms of commonly used knowledge basis, namely [5]:

1. Rule-based Reasoning.
2. The case-based reasoning.

Meanwhile, there are two methods to do the inference, namely [5]:

1. Forward Chaining method which is a reasoning derived from facts to test the truth of the hypothesis.
2. Backward Chaining method which is a reasoning derived from a hypothesis. In order to test the truth of the hypothesis it will look for existing facts in the knowledge basis.

Decision Tree

Decision tree is a flowchart structure that resembles a tree where each internal node denotes a test on an attribute, each branch represents the results of the test, and the leaf node represents a class or class distribution. The flow of the decision tree starts from the pan of the root node to the leaf node that holds the class prediction of the example. Decision tree is easy to convert to the rules of classification [6].

Decision tree is one technique to be used to classify a set of objects or records. This technique consists of a set of decision nodes, connected by branches, moving downward from the root node until it ends at the leaf node [7].

Certainty Factor

Certainty Factor is one of the methods used for non monotonic reasoning. The *certainty factor* expresses trust in an event or fact (hypothesis) based on evidence or expert judgment [8]. *Certainty factor* uses a value to assume an expert's degree of confidence in a data. The *Certainty factor* was introduced by Shortliffe Buchanan in the making of MYCIN [9]. Notation of certainty:

CF [h, e] = MB [h, e] – MD [h,e] (1)

With:

CF [h, e] = certainty factor hypothesis influenced by the evidence e is known with certainty.

MB [h, e] = measure of belief, a measure of trust in hypothesis h, if influenced by evidence e (between 0 and 1).

MD [h, e] = measure of disbelief, size of distrust of hypothesis h, if influenced by evidence e (between 0 & 1).

The combination of uncertainty rules is 3 things, among others [9]:

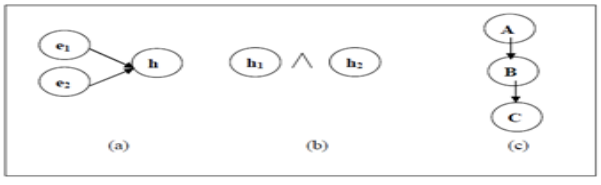


Figure 1. Combination of Uncertain Rules

1. Several combined evidences to determine the CF of a hypothesis (figure 1.a). If e1 and e2 are observations, then:

$$MB[h, e_1 \wedge e_2] = \left\{ \begin{array}{c} 0 \\ MB[h, e_1] + MB[h, e_2].(1 - MB[h, e_1]) \end{array} \right\} \quad (2)$$

$$MD[h, e_1 \wedge e_2] = \left\{ \begin{array}{c} 0 \\ MD[h, e_1] + MD[h, e_2].(1 - MD[h, e_1]) \end{array} \right\} \quad (3)$$

2. *Certainty Factor* is calculated from a combination of several hypotheses (Figure 1.b). If h1 and h2 are hypotheses, then:

$$MB[h_1 \wedge h_2, e] = \min(MB[h_1, e], MB[h_2, e]) \quad (4)$$

$$MB[h_1 \vee h_2, e] = \max(MB[h_1, e], MB[h_2, e]) \quad (5)$$

$$MD[h_1 \wedge h_2, e] = \min(MD[h_1, e], MD[h_2, e]) \quad (6)$$

$$MD[h_1 \vee h_2, e] = \max(MD[h_1, e], MD[h_2, e]) \quad (7)$$

3. Some rules go hand in hand, the uncertainty of a rule becomes the input for another rule (figure 1.c), then:

MB [h,s] = MB'[h,s] * max (0,CF [s,e])(8)

MB' [h,s] = measure of trust h based on full belief in validity s.

Likert Scale

The Likert scale is a bipolar scale method that measures either positive or negative responses to a statement. Four preferred scales are also sometimes used for Likert scale questionnaires that have made people to choose one of the poles because the "neutral" option is not available [10] the procedure steps that should be done by researchers in making Likert scale, including:

1. Collecting sufficient items, having relevance to the problem under study, and consisting of items that are quite obviously liked and disliked.
2. Then the items are tried to a group of respondents who are quite representative of the population to be studied.
3. The above respondent is asked to check each item whether he likes (+) or dislikes it (-). The responses were collected and the answers with a favored indication were given the highest score. There is no problem to give a 5 for the highest and score 1 for the lowest or vice versa. What is important is the consistency of the direction of the attitude shown. Likewise whether the "agree" or "disagree" answer is called the favor, depending on the content of the question and the contents of the items being prepared. While responding to questions on a Likert scale, respondents determine their level of approval of a statement by selecting one of the available options. Usually provided five scaled options with formats such as:

   ♣ Positive Questions (+)
   Score 1. Very (disagree / bad / less once)
   Score 2. No (agree / good /) or less
   Score 3. Neutral / Enough
   Score 4. (Agree / Good / like)
   Score 5. Very (agree / Good / Like)
   ♣ Negative Questions (-)
   Score 1. Very (agree / Good / Like)
   Score 2. (Agree / Good / like)
   Score 3. Neutral / Enough

Score 4. No (agree / good /) or less
Score 5. Very (disagree / bad / less once)

4. The total score of each individual is the sum of the scores of each individual item.
5. Responses are analyzed to determine which items are very real boundaries between high scores and low scores in total scale. For example, respondents in upper 25% and lower 25% are analyzed to see how far each item in this group is different. Items that do not show any real high score discarded to maintain the internal consistency of the question.

3. System Analysis

There have been several studies conducted on expert systems for the diagnosis of venereal disease with several different methods and case bases. Several existing studies have concluded that with the uncertainty method suitable for diagnosing venereal disease there has been no research that tries to examine how effective the methods are applied to diagnose venereal diseases. Thus, with reference to previous research on the same case with its implementation using the Decision Tree method, this study aims to compare the level of effectiveness with previous research with the method of Decision Tree and Certainty Factor.
By analyzing the two methods, it is expected that the results of this research will get the conclusion of the application of which method is more accurate, effective and can be measured the level of certainty

4. Experimental Result

Testing System

System testing aims to measure the results of the application, so that it can be compared with expert diagnosis. Display the results of the application using the tree shown in Figure 2, while the application view using the tree and Certainty factor can be seen drawn 3. which distinguishes the figure 2 and figure 3 of the Certainty factor calculation, if using the tree can not know the value of uncertainty and not explained treatment advice.



Figure 2. Diagnostic Results for Tree Method



Figure 3. Diagnostic Results for Tree and CF Method

B. User Acceptance Test

User acceptance test is a testing process conducted by the user intended to produce a document that is used as evidence that the software has been accepted by the user. In this report the testing process is done with Likert scale.
Likert scale is a scale used to measure perceptions, attitudes or opinions of a person about a social event or phenomenon and most commonly used in the questionnaire. In this testing process, the authors have given 2 pieces of questionnaires to experts namely dr. Rosmelia, M.kes. Sp.KK. each of which consists of 4 questions. For more details can be seen in Table 1.

Table 1. Questionnaire for Experts

No	Question
1 1	The existence of this expert system makes your work more effective, and efficient
2 2	The features available in this system are interesting
3 3	This system is easy to use
4 4	The look of this expert system is interesting
5 5	Expert system results built using the decision tree method in accordance with the realization on the direct practice

In each question a number of answer choices have been available: very agree, agree, quite agree, disagree, and strongly disagree. This answer will then be converted into the form of assessment scale which will be the weight of each answer to facilitate in doing calculations on the Likert scale. The result of convert the scoring scale of the answer and the percentage of values can be seen in table 2 and table 3.

Table 2. The results convert to the rating scale

Answer	Weight
Very Agree	5
Agree	4
Enough	3
Disagree	2
Very disagree	1

Table 3. Percentage value

Information	Percentage
Very Agree	80 – 100%
Agree	60 – 79,99%
Enough	40 – 59,99%
Disagree	20 – 39,99%
Very disagree	0 – 19,99%

The next step, we need to collect data from our sources that is dr. Rosmelia, M.kes. Sp.KK. From the data obtained, then processed by summing the weight of the answers obtained from each question, then divided by Y which is the number of questions * the highest answer weight. Index score% = Total weight / Y * 100. Then the results of the calculation of each questionnaire is as follows:

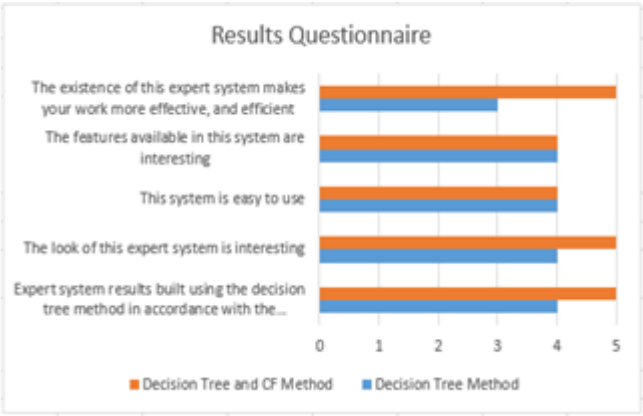


Figure 4. Results Questionnaire

1. Expert System Questionnaire with Decision Tree Meth-

od
• The first question answered agree to get a weight of 4
• The second question answered agree to get a weight of 4
• The third question answered agree to get a weight of 4
• The fourth question answered agree to get a weight of 4
• The fifth question answered quite agree to get weight 3

2. Expert System Questionnaire with Decision Tree Method and Certainty Factor
• The first question answered strongly agree to get a weight of 5
• The second question answered agree to get a weight of 4
• The third question answered agree to get a weight of 4
• The fourth question answered agree to get a weight of 5
• The fifth question answered agree to get weight 5

From the points mentioned above, it can be seen in the first questionnaire weighs 19 and in the second questionnaire weighs 23. Then, we divide by Y which is the number of questions * the highest answer weight is 5 * 5 = 25. Then, stay inserted into the formula index% listed above, total weight / Y * 100.
The following results were obtained:
1. Expert System Questionnaire with Decision Tree Method 19/25 * 100 = 76 (Agree)
2. Expert System Questionnaire with Decision Tree Method and Certainty Factor 23/25 * 100 = 92 (Strongly Agree)
Through the above calculation results, it can be concluded that the Expert System Diagnosis of Sex Disease with Decision Tree Method and Certainty Factor shows the output of the system more effective than using the method of Decision Tree. It can be seen through the calculation using a Likert scale that shows the percentage of 92% (Strongly Agree).

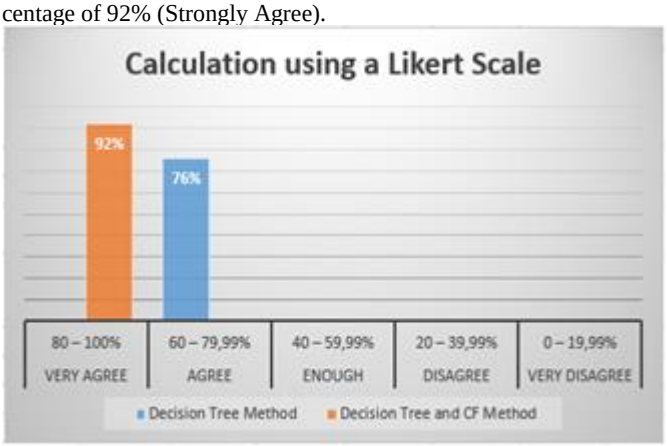


Figure 5. Calculation using a Likert Scale

Conclusions

After a comparison of two methods, Decision Tree with Decision Tree and Certainty Factor, the results showed that the Diagnostic Disease Expert System with the Decision Tree and Certainty Factor methods showed more effective system output than the Decision Tree method. It can be seen through the calculation using a Likert scale that shows the percentage of 92% (Strongly Agree).
ACKNOWLEDGMENT
Thank you to Department of Informatics, Faculty of Industrial Technology, Islamic University of Indonesia, the staff and college teacher.

References

[1] Maulana, A., 2016. Sistem Pakar Untuk Mendignosa Penyakit Kulit Pada Manusia dengan Metode Certainty Factor. Skripsi UII Yogyakarta.
[2] Krisnawan dkk, 2014. Sistem Pakar Diagnosa Penyakit Kulit dan Kelamin Dengan Metode Certainty Factor dan Fuzzy Logic. MERPATI. Vol 2- No 3.
[3] Wahyuni, E.G., and Al-amin, M.Z., 2017. "Expert System For Diagnosis Venereal Disease With Decision Tree". International Conference on Internet Services Technology and Information Engineering (ISTIE).
[4] Kurniawan & Wahyuni, 2017. "Aplikasi Untuk Mendiagnosis Penyakit Kelamin dengan Metode Decision Tree dan Certainty Factor". Seminar Nasional Informatika Medis (SNIMED) VIII.
[5] Kusumadewi, Sri. 2003. Artificial Intelligence (Teknik dan Aplikasinya). Yogyakarta: Graha Ilmu.
[6] Kusnawi, 2007. Pengantar Solusi Data Mining. Seminar Nasional Teknologi (SNT). STMIK Amikom Yogyakarta, D-1 - D-9.
[7] Yusuf, Y.W. 2007. Perbandingan Performansi Algoritma Decision Tree C5.0, CART, dan CHAID: Kasus Prediksi Status Resiko Kredit di Bank X. Seminar Nasional Aplikasi Teknologi Informasi (SNATi). B59 - B62.
[8] Turban, E. 2005. Decision Support System and Intelligent Systems. Yogyakarta: Penerbit Andi.
[9] Kusumadewi, Sri. 2003. Artificial Intelligence (teknik dan aplikasinya). Yogyakarta: Penerbit Graha Ilmu.
[10] Nazir, M., 2005. Metode Penelitian. Bogor: Ghalia Indonesia