DHEMPSTER-SHAFER IMPLEMENTATION IN OVERCOMING UNCERTAINTY IN THE INFERENCE ENGINE FOR DIAGNOSING ORAL CAVITY CANCER

Riduwan Napianto[1], Yuri Rahmanto[2], Rohmat Indra Borman*[3], Ova Lestari [4], Nurhasan Nugroho[5]
[1,2,3,4] Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Lampung
[5] Faculty of Teacher Training and Education, Universitas Bina Bangsa, Banten
e-mail: [1]riduwannapianto@teknokrat.ac.id, [2]yurirahmanto@teknokrat.ac.id,
*[3]rohmat_indra@teknokrat.ac.id, [4]ova_leatari@gmail.com, [5]nurhasan.nugroho@gmail.com

Abstract

The word uncertainty in an expert system is related to working with wrong data, wrong information, handling identical situations, the reliability of results, etc. Sources of uncertainty can come from unreliable information. This is usually caused by unclear domain concepts or for inaccurate data. One method for overcoming uncertainty is Dhempster-Shafer's theory. Dempster-shafers come up with approaches to calculate probabilities to look for evidence based on trust functions. In general the Dempster-Shafer theory is written at an interval [Confidence, Reasonable]. Belief (Bel) is a measure of the strength of evidence in support of a series of propositions. In this study an expert system will be developed to diagnose oral cancer that can recognize oral cancer based on the symptoms felt by the user. The results showed the Dempster-shafer was able to overcome the uncertainties in the construction of the inference engine, this is because the accuracy of the test results showed an accuracy of 86.6% Dempster-shafer.

Keywords— dempster-shafer, expert system, oral cavity cancer

1. INTRODUCTION

Expert systems or knowledge-based systems are part of artificial intelligence that allows computers to process and derive conclusions from a set of rules [1]. Expert system also known as knowledge based expert system is a computer program that information and experience in a specific area for decision making [2]. Expert system is not to replace the role of humans, but to substitute human knowledge in the form of a system, so that it can be used by many people. The purpose of expert system development is to produce a system that can assist human work, especially in relation to the utilization of expertise and experience in a particular field [3]. The expert system asks for facts that will show symptoms of certain diseases and can provide an explanation of the results of the consultations that have been carried out [4]. In the expert system the word uncertainty is related to working with incorrect data, incorrect information, handling identical situations, reliability of results etc. Sources of uncertainty come from information that can be unreliable, this is usually caused by unclear domain concepts or inaccurate data [5]. In statistical methods based on the assumption that uncertainty is the probability of an event / fact true or false [6]. One method in overcoming uncertainty is Dhempster-shafer theory.For this reason, the Dempster-shafer appears with an approach to calculating probabilities to look for evidence based on trust functions and reasonable reasoning, which is used to combine information (evidence) [7].

In this study Dempster-shafers are implemented on an inference engine for diagnosing oral cavity cancer. Oral cavity cancer, or just oral cancer, is cancer that starts in the mouth (also called the oral cavity) [8]. Globocan data states that in 2018 there were 18.1

million new cases with a mortality rate of 9.6 million deaths, where 1 in 5 men and 1 in 6 women in the world experience cancer. [9]. Therefore it is very important to be able to provide knowledge assistance to Indonesian people who are still low in knowledge about cancer, while cancer patients are always increasing every year. One way to help diagnose the disorder is to use an expert system. In this research, an expert system will be developed to diagnose oral cancer which can recognize oral cancer based on the symptoms felt by the user. The application also includes prevention solutions that can be done by the user against oral cancer that is experienced so as to help the community in preventing oral cancer.

2. RESEARCH METHODOLOGY

This research consists of several stages, namely problem identification, variable determination and analysis, system design, implementation and testing. After all the stages have been completed then documentation, reporting and publication are carried out. The stages of the research can be seen in the following figure 1.

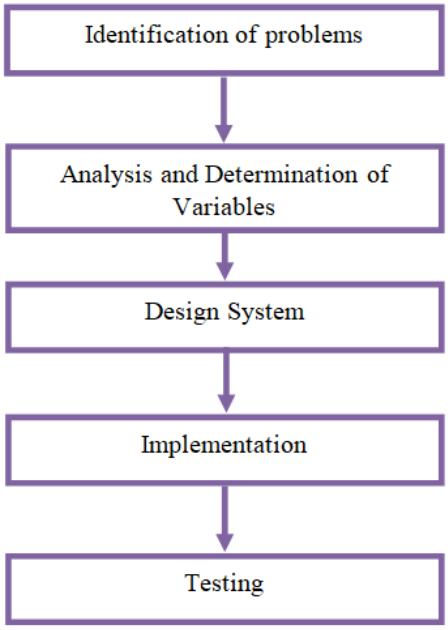

Figure 1. Research Stages

2.1 Identification of Problem

An initial stage of mastery of a problem where a particular object can be identified with a problem. At this stage the problem identification is carried out based on interviews and observations on diagnoses of oral cancer.

2.2 Analysis and Determination of Variables

At this stage data analysis is performed on data obtained from experts (primary data) and from literature (secondary data) as well as the determination of variables that have two main variables, namely negative and positive, which are useful for knowing problems and how to solve them so that applications are run will be in accordance with existing data. At this stage the knowledge acquisition process is implemented into a representation of knowledge that will be used as an inference engine

2.3 Design System

This stage performs modelling or design using a flowchart and User Interface which can later make it easier to conduct research so that a description of the system to be made clear what the intent and purpose. The sequence of the problem solving process in the expert system program outlines starting from the program displaying the symptom variables that exist then filling in the values or weights of the symptoms and then checking the values entered, whether filled in all or not. If all is filled, then go to the next process which is to calculate the value of the variable entered by the Dempster-shafer method, then produce output in the form of the calculation results along with the diagnosis of the disease. Whereas if it is not filled in all, then the process is immediately finished and the output cannot be displayed. To make it easier to describe the flow of the system described through a flowchart. Flowchart is an approach used to determine, build and visualize systems graphically and is practically useful on a very broad basis but still lacks formal and precise understanding [10]. Flowchart system of the Implementation of the Dempster-shafer Method in diagnosing oral disease can be seen in Figure 2 below.

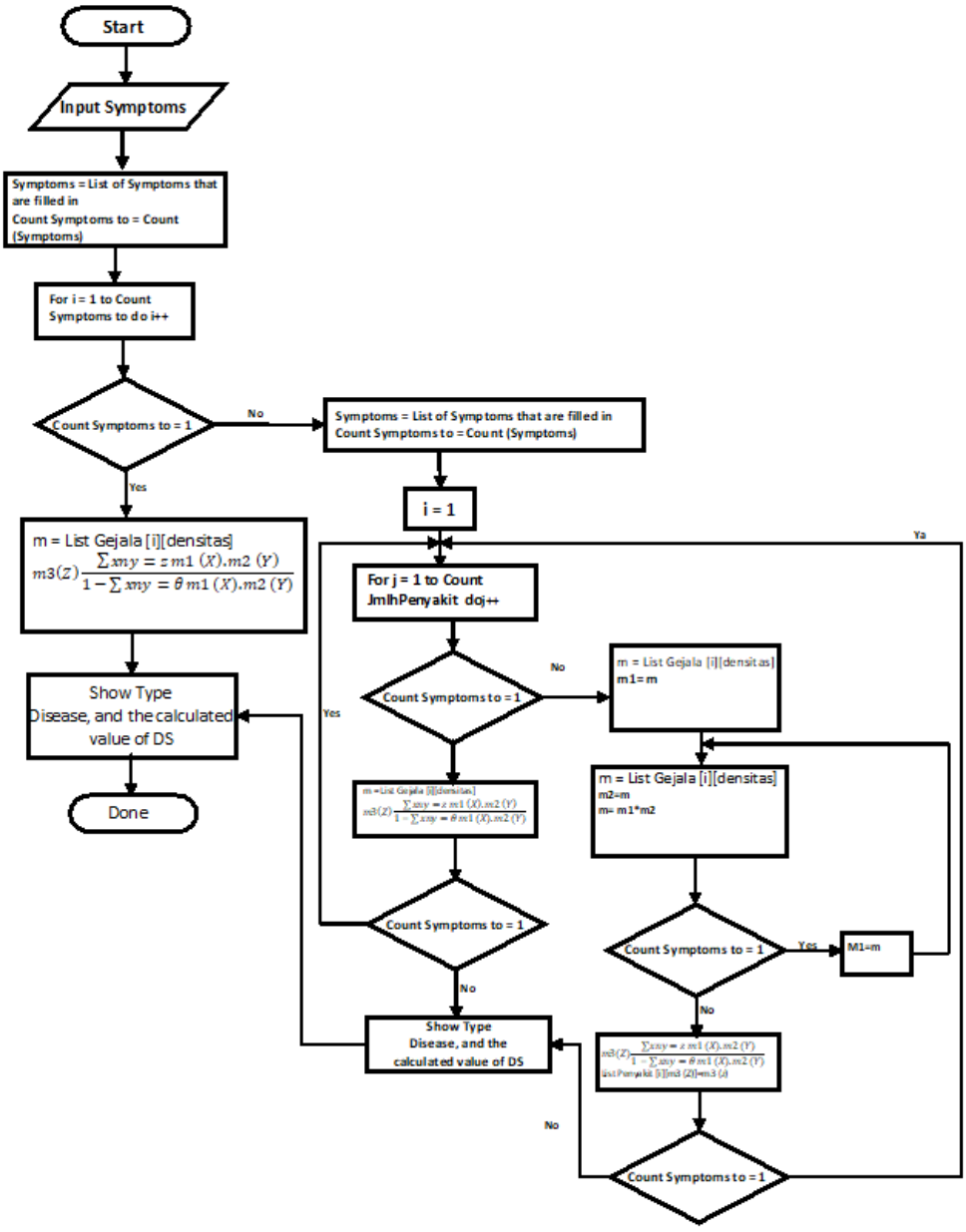

Figure 2. Flowchart Expert System for Diagnosing Oral Cancer

2.4 Implementation

Implementation is the stage where the designs that have been made previously are coded with a certain programming language to become an application [11]. At this stage the implementation of the rules into the inference engine is implemented using the Dhempster-Shafer algorithm into the PHP programming language with the Adobe Dreamweaver compiler so that the application can be implemented on the website platform.

2.5 Testing

Testing of systems that have been built by conducting trials with several parameters. The system that has been built must first be made in order to find errors [12]. In the testing stage, testing is done on the accuracy of the Dhempster-shafer method on the expert system of diagnosing oral cancer. The validation process of the system is done by entering test data into the system. This is done with the aim to find out to what extent the system has a success rate based on the test data that has been entered. In this study the success rate of the system is determined based on the accuracy of the diagnosis. Calculation of accuracy of diagnosis is obtained from a comparison between the results of the system diagnosis with a doctor's diagnosis.

3. RESULTS AND DISCUSSION

In this study the method used is Dempster-shafer, because it has a certainty of decision making, the conclusion drawn is obtained from data provided by the user through various symptoms that occur. Dempster-Shafer theory is a representation, combination and propogation of uncertainty, where this theory has several characteristics that are institutively in accordance with the way of thinking of an expert [13]. Dempster-shafer algorithm has the value of belief that serves to determine the influence between symptoms obtained from an expert [14]. The following furniture is a decision table that has been obtained from an expert, and the expert has given the required density value.

Table 1. Decision Table

No	Symptom Code	Disease Code					Density
		P1	P2	P3	P4	P5	
1	G01	*					0,6
2	G02	*					0,8
3	G03	*					0,7
4	G04	*	*		*		0,9
5	G05	*					0,8
6	G06	*	*	*	*	*	0,8
7	G07		*				0,7
8	G08		*				0,8
9	G09		*				0,6
10	G10			*			0,9
11	G11			*			0,7
12	G12			*			0,7
13	G13			*			0,7

No	Symptom Code	Disease Code					Density
		P1	P2	P3	P4	P5	
14	G14				*		0,7
15	G15				*		0,8
16	G16				*		0,7
17	G17					*	0,8
18	G18					*	0,6
19	G19					*	0,8
20	G20					*	0,7

Example calculation using the Dempster-shafer method applied to cancer with selected symptoms are symptoms that are in the oral cavity or experienced by the user:

1. Lip color does not appear pink (G01)
2. Wounds on the mouth that are difficult to heal (G04)
3. Often experience numbness in the oral cavity (G06)
4. Jaw is experiencing a freezing (G19)

Symptom-1: Lip color does not appear pink; Based on table I of the decision table, values are obtained:
m1 { P1 } = 0,6
m1 { θ } = 1-0,6 = 0,4

Symptom-2: Wounds on the mouth that are difficult to heal; based on table I of the decision table, values are obtained:
m2 {P1,P2,P4} = 0,9
m2 { θ } = 1-0,9 = 0,1

In general the Dempster-Shafer theory is written in an interval [Belief, Plausibility]. Belief (Bel) is a measure of the strength of evidence in supporting a set of propositions. If it has a value of 0 (zero) then it indicates that there is no evidence, and if it has a value of 1 indicates there is certainty. Based on the initial density determination in symptoms 1 and 2, a new density value can be obtained by making a combination rule table in advance using the formula:

$$m3(Z)\frac{\sum xny=z\ m1\ (X).m2\ (Y)}{1-\sum xny=\theta\ m1\ (X).m2\ (Y)} \qquad (1)$$

From the results of this combination will be used to calculate the value of new symptoms.

Table 2. Combination Rules for m3

		{P1,P2,P4}	(0,9)	θ	(0,1)
{P1}	(0,6)	{ P1}	(0,54)	{P1}	(0,06)
θ	(0,4)	{P1, P2, P4}	(0,36)	θ	(0,04)

From the Dempster-shafer formula, then θ m1 (X) m2 (Y) has a value of 1 (1-0), so the calculation can be made as follows:

m3 {P1,P2,P4} $= \frac{0,36}{1-0}$ = 0,36
m3 {P1} $= \frac{0,54+0,06}{1-0}$ = 0,6
m3 {θ} $= \frac{0,04}{1-0}$ = 0,04

Symptom-3: Often experience numbness in the oral cavity area; based on table I of the decision table, values are obtained:
m4 {P1,P2,P3,P4,P5} = 0,8
m4 { θ } = 1-0,8 = 0,2

Table 3. Combination Rules for m5

	{P1,P2,P3,P4,P5} (0,8)	ϴ (0,2)
{ P1,P2,P4} (0,36)	{P1,P2,P4} (0,288)	{ P1,P2,P4} (0,072)
{P1} (0,6)	{P1} (0,48)	{P1} (0,12)
ϴ (0,04)	{P1,P2,P3,P4,P5} (0,032)	ϴ (0,008)

So the calculation can be done as follows:
m5 { P1,P2,P4} $= \frac{0,288+0,072}{1-(0)}$ = 0,36
m5 {P1,P2,P3,P4,P5} $= \frac{0,032}{1-(0)}$ = 0,032
m5 {P1} $= \frac{0,48+0,12}{1-(0)}$ = 0,6
m5 {θ} $= \frac{0,008}{1-(0)}$ = 0,008

Symptom-4: The jaw is swollen; Based on table I of the decision table, values are obtained:
m6 { P5} = 0,8
m6 { θ } = 1-0,8 = 0,2

Table 3 Combination Rules for m7

	{ P5} (0,8)	ϴ (0,2)
{P1,P2,P4} (0,36)	{ Ø } (0,288)	{P1,P2,P4} (0,072)
{P1,P2,P3,P4,P5} (0,032)	{ P5} (0,026)	{P1,P2,P3,P4,P5} (0,006)
{P1} (0,6)	{-Ø} (0,48)	{P1} (0,12)
ϴ (0,008)	{P5} (0,006)	ϴ (0,007)

So the calculation can be done as follows:
m7 { P1,P2,P4} $= \frac{0,072}{1-(0,288+0,48)}$ = 0,310
m7 { P1,P2,P3,P4,P5} $= \frac{0,006}{1-(0,288+0,48)}$ = 0,026
m7 {P1} $= \frac{0,12}{1-(0,288+0,48)}$ = 0,517

m7 { P5 } $= \frac{0{,}026+0{,}006}{1-(0{,}288+0{,}48)} = 0{,}138$

m7 { θ } $= \frac{0{,}007}{1-(0{,}288+0{,}48)} = 0{,}030$

Because no new symptoms appear, it can be concluded that the highest density is the symptom {P1}, which is Lip Cancer with a value of 0.517 or if it is presented to be 51.7%.

After the inference engine has been designed, then the programming language implementation. Implementation is the application of how the system works based on the results of the analysis and also the design that has been made previously into a particular programming language. Application of expert system for diagnosing oral cancer has functional requirements including: the system can manage symptoms data, disease data, and representations of knowledge, diagnose diseases and see the results of diagnosing the disease and its solutions.

Figure 3. Main Menu Admin Expert System for Diagnosing Oral Cancer

In this application the results of the diagnosis of the disease and the solution will display the results of the diagnosis that has been selected by the user and displays the type of disease in accordance with the symptoms chosen by the user. The diagnosis results will be strengthened by the level of confidence percentage from the use of the Dempster-shafer method.

Figure 4. Form the Results of the Diagnosis System Expert Diagnosis of Oral Cancer

Next after the application has been implemented then testing the application is carried out. The test carried out is testing accuracy, by comparing the results of expert system diagnoses with the results of an expert's analysis. From 30 test cases conducted by random

sampling for symptoms chosen by the expert system user, he was able to diagnose according to the results of an expert's analysis of 26 cases. This shows the accuracy rate of the expert system is 86.6%

4. CONCLUSION

Based on the research that has been done, the conclusions that can be drawn are as follows:

1. The Dhempster-shafer algorithm overcomes uncertainty by combining evidence from several sources and incorporating or providing a level of confidence from the available evidence
2. Based on accuracy testing by comparing the results of expert system diagnostics with expert analysis, it shows an accuracy of 86.6%. The accuracy value is influenced by the accuracy of the weight previously determined, so that it affects the plausability value and its combination

5. RECOMMENDATION

The results of the accuracy of the research that have been carried out are still not optimal, it is necessary to conduct further studies related to other uncertainty methods for cases of oral cancer.

REFERENCES

[1] J. Durkin, Expert Systems: Design and Development, New York: Maxwell Macmillan International, 1994.

[2] M. Akram, I. A. Rahman and I. Memon, "A Review on Expert System and its Applications in Civil Engineering," *International Journal of Civil Engineering and Built Environment,* pp. 24-29, 2014.

[3] G. C. Prasetyadi and M. Mahfudin, "Web-Based Expert System Application to Recommend Computer Specifications for Gaming Using Backward Chaining Inference Method," *Journal of Information System,* pp. 110-117, 2017.

[4] A. Sucipto, Y. Fernando, R. I. Borman and N. Mahmuda, "Penerapan Metode Certainty Factor Pada Diagnosa Penyakit Saraf Tulang Belakang," *JURNAL ILMIAH FIFO,* vol. X, no. 2, 2018.

[5] S. Dubey, R. K. Pandey and S. S. Gautam, "Dealing with Uncertainty in Expert Systems," *International Journal of Soft Computing and Engineering (IJSCE),* vol. 4, no. 3, 2014.

[6] W. U. Setiabudi, E. Sugiharti and F. Y. Arini, "Expert System Diagnosis Dental Disease Using Certainty Factor Method," *Scientific Journal of Informatics,* vol. 4, no. 1, 2017.

[7] A. M. Alfatah, R. Arifudin and M. A. Muslim, "Implementation of Decision Tree and Dempster Shafer on Expert System for Lung Disease Diagnosis," *Scientific Journal of Informatics,* vol. 5, no. 1, 2018.

[8] S.-H. Huang and B. O´Sullivan, "Oral cancer: Current role of radiotherapy and chemotherapy," *Journal section: Oral Medicine and Pathology*, pp. 233-240, 2013.

[9] Depkes, "Hari Kanker Sedunia 2019," January 2019. [Online]. Available: http://www.depkes.go.id/article/view/19020100003/hari-kanker-sedunia-2019.html. [Accessed June 2019].

[10] N. Tiwari and L. Prasad, "A Comparative Study: Reverse Engineering Flowcharting Tools," *INTERNATIONAL JOURNAL OF INNOVATIVE TRENDS IN ENGINEERING (IJITE)*, vol. 7, no. 1, 2015.

[11] T. Monica and R. I. Borman, "Implementasi Konsep Media Sosial Dalam Sistem Informasi Kegiatan Kesiswaan (Studi Kasus : SMK XYZ)," *Jurnal TEKNO KOMPAK*, vol. 11, no. 2, 2017.

[12] A. E. Kumala, R. I. Borman and P. Prasetyawan, "Sistem Informasi Monitoring Perkembangan Sapi Di Lokasi Uji Performance (Studi Kasus : Dinas Peternakan Dan Kesehatan Hewan Provinsi Lampung)," *Jurnal TEKNOKOMPAK*, vol. 12, no. 1, 2018.

[13] N. S. B. Sembiring and M. D. Sinaga, "Penerapan Metode Dempster Shafer Untuk Mendiagnosa Penyakit Dari Akibat Bakteri Treponema Pallidum," *CSRID Journal*, vol. 9, no. 3, 2017.

[14] M. G. Nababan, R. R. M. Putri and I. Indriati, "Pembangkitan Nilai Belief Pada Dempster-Shafer Dengan Particle Swarm Optimization (PSO) Untuk Penentuan Pasal Kasus Penganiayaan," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 10, 2017.