Accuracy of Kidney Disease Expert System Based On Certainty Factor and Dempster Shafer Algorithm

Agussalim
Information Systems
Universitas Pembangunan Nasional
Veteran Jawa Timur
Surabaya, Indonesia
agussalim.si@upnjatim.ac.id

Nani Astuti Triana
Informatics
STMIK Handayani
Makassar, Indonesia
njcl481@gmail.com

Eristya Maya Safitri
Information Systems
Universitas Pembangunan Nasional
Veteran Jawa Timur
Surabaya, Indonesia
maya.si@upnjatim.ac.id

Anita Wulansari
Information Systems
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
Anita.wulansari.sisfo@upnjatim.ac.id

Seftin Fitri Ana Wati
Information Systems
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
seftin.fitri.si@upnjatim.ac.id

***Abstract*— The information technology developments today resulted in several innovations, including the existence of an Expert System. The systems help diagnose diseases without directly meeting with experts/doctors. Many researchers have proposed algorithms to improve the accuracy of the expert system to approach the diagnosis by experts/doctors, including certainty factor and dempster shader. This study compares the algorithm's accuracy with the results of expert diagnosis of kidney disease. The expert system was developed using UML and a web-based version. From the comparison results, the dempster shaver algorithm has an accuracy rate of 80%, while the certainty factor is 60% compared to expert diagnoses.**

Keywords—Certainty-Factor ; Dempster Shafer; Expert System; Kidney Disease; Accuracy

I. INTRODUCTION

An expert system is an associate of artificial intelligence used to solve some human problems. Expert systems adopt the expertise of an expert in a particular field into a computer system or program with a user-friendly user interface so that the users who are not experts can make decisions like an expert [1]. In developing an expert system, many methods can be used to make it easier to solve the existing problems, such as in the expert system for diagnosing kidney disease. Several methods could be applied to an expert system, including the Certainty Factor method and the Dempster Shafer method.

The certainty factor method is widely utilized in expert systems. A lot of studies have been done with this method. Such as research on expert systems for diagnosing skin diseases in humans. Using the certainty factor method, the accuracy value of the method is 73.15% same as the results given by expert doctors [2]. In similar research, an expert system for diagnosing dengue fever with the certainty factor method was conducted. The objective is to show the calculation of how certain the patient is exposed to dengue fever [3]. And further research is an expert system for detecting Osteoporosis and Osteoarthritis using the certainty factor method, with the result accuracy value reaching 80% [4].

Researcher in [5] has conducted research on an expert system to diagnose kidney disease using the dempster-Shafer method, the three values generated from the system are the same as the results of manual calculations on the dempster-Shafer method, so that the results are in accordance with the expected calculations. However, this study does not provide the percentage of the accuracy of the result generated.

Kidney disease is a disorder that occurs in the kidneys. It affects the performance in filtering and removing metabolic waste substances in the body. The main factors that cause kidney disease are lack of fluids or dehydration and an unhealthy lifestyle. The second factor is a medical risk, such as hypertension and diabetes mellitus, and heredity [6]. During this decade, a lot of research has been done to implement some methods in an expert system. This study [7] tested the accuracy of the Cat Disease Diagnosis Expert System Using the Naive Bayes and Certainty Factor Method on Android. The test achieved accuracy is 80% due to the 5 test data from 25 test data do not match to the expert results caused by similar symptoms.

Another study conducted by researcher in [8], this research compared Certainty Factor and Dempster-Shafer for Diagnosing Oral and Dental Diseases in Humans. The Dempster-Shafer was able to produce correct diagnosis results in 45 cases out of 50 cases studied and for the Certainty Factor as many as 41 cases out of 50 cases. This is because drawing conclusions with the Certainty Factor is only able to compare two symptoms in producing a belief value. At the same time, the Dempster-Shafer theory involves the probability of all possible diseases from each symptom.

Researcher in [9], Implemented Fuzzy and Dempster-Shafer on an Expert System for Diagnosing Chili Diseases. The fuzzy and dempster-shafer methods can be applied to an expert system for diagnosing chili plant diseases. Fuzzy is used to obtain belief values in the dempster shafer method and then calculate the combination of belief values from the symptoms suffered using the dempster-shafer formula where the combination of the largest belief values will be taken as a result of the diagnosis of chili plant disease. A similar study conducted by researcher in

https://doi.org/10/33005/ijconsist.v3i2.66

[10], building an expert system for early diagnosis of kidney disease using the fuzzy logic method for determining certainty factors, in this study states that the Fuzzy Logic and Certainty Factor methods can provide an initial diagnosis of kidney disease based on the symptoms experienced Patients with diagnostic validation results.

Each method has various ways of diagnosing the disease. These dissimilarities will make the presentation of trust in the information result also different. It is important to analyze the comparison of the two algorithms in the expert system to determine which method is better for diagnosing kidney disease. This study aims to design an expert system with the Dempster-Shafer method and the Certainty Factor method in kidney disease. Also, to analyze the trust of information results the Dempster-Shafer Method and the Certainty Factor Method on the kidney disease expert system.

II. State of The Art

A. Expert System

The expert system is interpreted by referring to several opinions. An expert system adopts the expertise of an expert in a particular field into a computer system or program with a user-friendly GUI so that users who are not experts can use it easily [1]. Users can create a decision or determine a policy like an expert. It also has a knowledge base and employs inference reasoning to resemble an expert to find the solution to the problems. The expert system adopts the knowledge of an expert into computer software. Whose way of working is based on specific knowledge to solve particular problems.

The expert system is delivered in the development and consulting environments (runtime) [11]. The ES builder uses the development environment to build components and insert knowledge into the knowledge base, while non-experts use the consulting environment to obtain expert knowledge and advice. This environment separated after the system was complete.

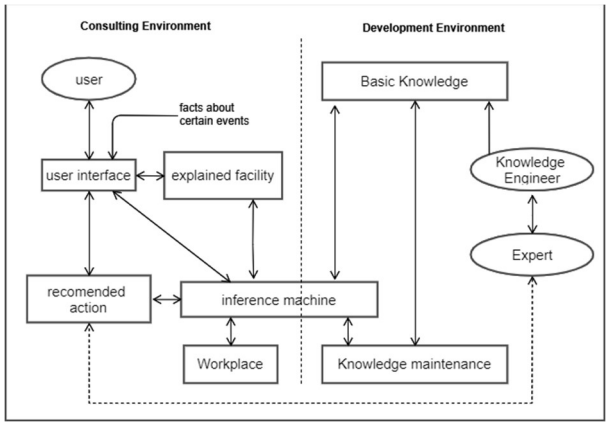

Figure 1. Expert System Structure

Figure 1 the expert system structure. The existence of an expert in a consulting environment can be replaced by someone who is not an expert with the benefit of an inference engine on an expert system. An expert system that already contains an expert knowledge base could help the limitation of availability of experts by interpreting their knowledge into the knowledge base for improvement of knowledge to be processed by the expert system inference engine.

B. Certainty Factor

A certainty factor is an algorithm to verify whether a fact is certain or uncertain based on a metric. It is usually employed in an expert system because it is appropriate for interpreting something that is not certain [12]. The certainty factor method is an alternative method to the Bayesian reasoning method or fuzzy method, or other similar methods in creating inferences in uncertain conditions (uncertainty).

Shortliffe and Buchanan proposed the certainty factor in 1975 to accommodate the unjustified sense of an expert. This theory developed simultaneously with the MYCIN expert system. The MYCIN Development team reports that doctors usually examine available information with terms: probably, most likely, almost certainly, etc. The MYCIN team employed the certainty factor (CF) to describe the expert confidence level in the problem [13].

C. Dempster-Shafer

Dempster introduced The Dempster-Shafer method, which sampled with uncertainty models with a degree of probabilities as a single probability. Shafer published Dempster's theory in a book entitled Mathematical Theory of Evident in 1976. According to the facts collected, the dempster-Shafer's shows a way to give weight to belief. This theory can indicate uncertainty and ignorance [14].

The Dempster-Shafer approach is implemented in an interval: [Belief, Plausibility]. Belief (Bel) estimates the strength of evidence in assisting a set of propositions. If it is 0, it implies no proof, and if it is 1, it indicates certainty. Plausibility (Pls) will reduce the level of confidence in the evidence. The probability is 0 to 1. For example, if you are confident about X', then it can be said that Bel(X') = 1, so the formula above has the value of Pls(X) = 0.

III. Methodology

To perform accuracy analysis of the Certainty Factor and Dempster-Shafer methods in cases of kidney disease diagnosis. A web-based kidney disease expert system application was developed. The system includes an online kidney disease initial examination process and provides some information about kidney disease. This expert system is expected to assist the work of an expert or doctor in conducting patient examinations. In addition, this system also allows a patient to carry out an early examination independently without having to meet face-to-face with a doctor. Patients will get the results of the examination as well as information about the possible disease experienced. Users will also be given advice and information about other kidney diseases..

A. Use Case

In this study, the system modeling uses the Unified Modeling Language (UML) to describe the proposed system process

according to the procedure using a use case diagram as shown in Figure 2 and Figure 3.

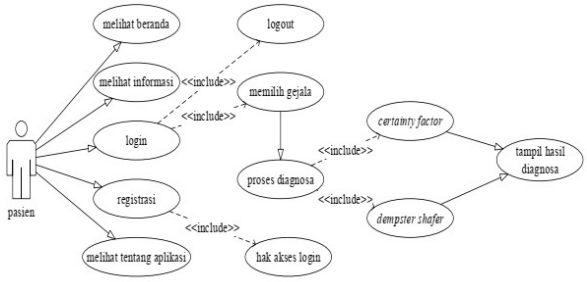

Figure 2. Use case of patient and expert system

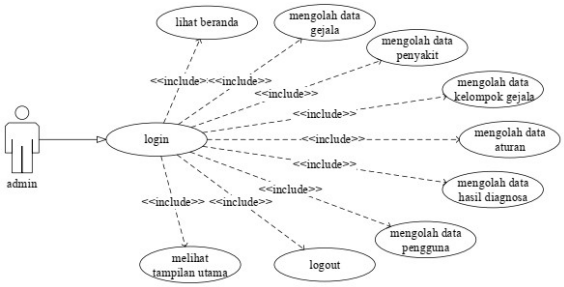

Figure 3. Use case of admin and expert system

As shown in Figure 2 and figure 3, administrators/experts have to log in to start using the system and manage symptom data, disease data, and management of kidney disease according to the rules. Before using the system, the patient/user has to register. After registering, the patient will be given access to the system. The patient logs into the system using the username and password given in the registration step. After that, the patient could use the system and enter the symptoms complained of. Then the system will diagnose the symptoms by using one of the dempster-shafer methods or the certainty factor method. Then, the system will publish a diagnosis result and a solution if there are symptoms of kidney disease in the patient.

B. Class Diagram

Class diagrams show classes and packages in the system. It delivers a static graph of the system and the associations between classes. In this system, there are nine classes that process and relate to each other, as shown in the following Figure 4.

C. Sequence Diagram

Figure 5 shows the sequence diagram Admin. First, the Admin opens the application to log in using the registered username and password. After successfully logging in, the Admin selects a menu to process symptom data, symptom group data, disease data, rules data for processing weight data, and diagnostic data and selects the user menu to process user data stored in the database. Then can Admin logout to exit the system when the proses are finished.

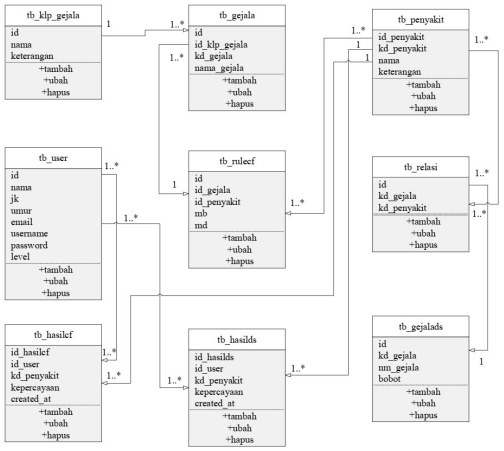

Figure 4. Clas diagram of kidney disease expert systems

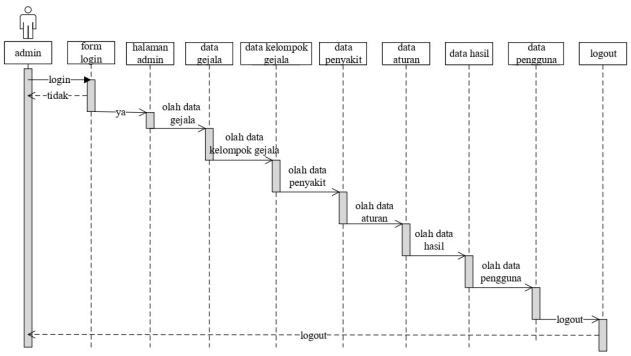

Figure 5. sequence diagram of kidney disease expert systems

D. Algorithm Implementation

An algorithm is a step-by-step procedure for computation. Algorithms are used for calculation, data processing, and automatic reasoning. Certainty Factor and Dempster Shafer are the algorithms used in this system. There are main step to implement the algorithm, first is Knowledge Base Development and determine the rules.

After determining the knowledge base, then proceed with the process of making rules. The formation of rules is done by combining symptom data and combined with disease data. From discussions that have been carried out with experts, it can also be obtained a weight value based on symptoms of the disease with a value scale of 0-100%. To determine the value of the rule-based the value scale of the expert is converted to a 0-1 value scale. By calculating the conversion of the percent number divided by 100, for example, 10% is converted to a number. The conversion process is 10/100 = 0.1, and the conversion result of 10% is equal to 0.1.

IV. RESULT AND DISCUSSION

A. Certainty Factor Implementation

The following is the implementation of the certainty factor method in the PHP programming language:

Displays the Symptom Checkbox:

```
if (!$this->input->post('gejala')) {
        $data['view'] = 'user/diagnosa/form';
        $data['listKelompok'] = $this->Kelompok_model->get_list_data();
        $this->load->view('user/diagnosa/index', $data);
}else{
        $data["view"]="user/diagnosa/hasil_diagnosa";
        $gejala = implode(",", $this->input->post("gejala"));
        $data["listGejala"]=$this->Gejala_model->get_list_by_id($gejala);
```

This step is to display symptom data and symptom groups. In this procedure, each symptom in the symptom table and symptom group will be called and displayed on the diagnosis page.

Figure 6. The Symptoms Checkbox

B. *Certainty Factor Calculation*

This step is the process of calculating the Certainty Factor method from the symptoms entered by the user, with the following formula:

$$CF = MB - MB \quad (1)$$

$$CF_{combinate} = MB_{old} + \left(MB_{new} * (1 - MB_{old})\right) \quad (2)$$

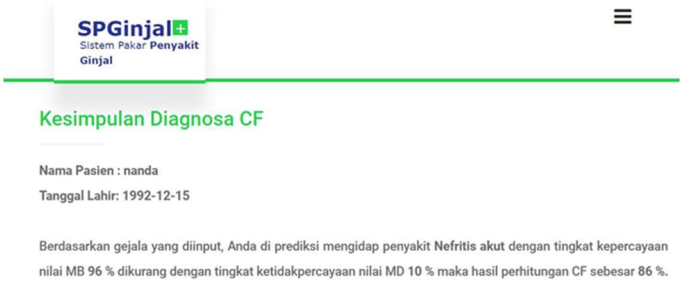

Figure 7. The Symptoms Checkbox

C. *Manual Calculation of Certainty Factor*

The manual calculation of the certainty factor method for an expert system for diagnosing kidney disease can be seen in the table described below:

TABLE I. SYMTOMS SELECTED BY USER

Symptoms Selected by User	
Symptoms Code	**Symptoms Name**
G003	Fever
G008	Abnormal urine colour
G024	Low back pain
G011	Swelling in the body part

TABLE II. MANUAL CALCULATIONS OF ACUTE NEPHRITIS

No	Symptoms	MB Calculation	MD Calculation
1	Fever^ Swelling in the body part	0.8+0.8 x (1-0.8) = 0.96	0.2+0 x (1-0.2) = 0.2
2	Fever^ Swelling in the body part^ Abnormal urine color	0.96+0.7 x (1-0.96) = 0.988	0.2+0.2 x (1-0.2) = 0.36
3	Fever^ Swelling in the body part ^ Abnormal urine color ^ Low back pain	0.988+0.8 x (1-0.988) = 0.9976	0.36+0.2 x (1-0.36) = 0.074

Nilai CF = 0.9976 - 0.488

= 0.5096

After manual calculations on the Certainty Factor method, the results obtained are that the patient is likely to be affected by acute nephritis with a confidence level of 0.9976 and a distrust of 0.074, then the CF result is 0.9236. On the system, the level of distrust is 0.1, with a difference of 0.0636. Then these results are the same as the calculations in the expert system for diagnosing kidney disease.

D. *Implementasi Algoritma Dempster Shafer*

At this stage, the Dempster Shafer method is calculated from the symptoms that have been inputted by the user, with the following formula:

$$Bel(X) = \sum m(y) \quad (3)$$

$$m3(Z) = \frac{\sum X \cap Y = m1(X).m2(Y)}{1 - \sum X \cap Y = m1(X).m2(Y)} \quad (4)$$

Figure 8. Dempster Shafer Calculation Result

E. *Accuracy of Certainty Factor and Dempster Shafer*

As shown in Table III, the comparative results of the certainty factor method and the Dempster shaper method was compared in ten cases. In this comparison, we compare the results of the expert system developed with the results of expert doctors' diagnoses. There were ten cases tested based on the existing symptoms to test the accuracy of the Certainty Factor algorithm and Dempster Shafer algorithm for diagnosing kidney disease. For comparative analysis, the method is calculated by the following formula:

$$level\ of\ accuracy = \frac{amount\ of\ correct\ data}{amount\ of\ data\ test} x100\% \quad (5)$$

TABLE III. COMPARISON CASE

No	Diagnostic Result		
	Expert Result (Doctor's verdict)	Expert System	
		Certainty*Factor*	Dempster*Shafer*
1	Chronic kidney failure	Chronic kidney failure	Chronic kidney failure
2	Chronic kidney failure	Nefritis acute	Chronic kidney failure
3	Kidney cyst	Glomerulonefritis	Chronic kidney failure
4	Kidney stones	Kidney stones	Kidney stones
5	Kidney cancer	Kidney stones	Nefritis acute
6	Nefritis acute	Kidney cyst	Nefritis acute
7	Nefritis acute	Nefritis akut	Nefritis acute
8	Kidney Cancer	Kidney cancer	Kidney cancer
9	Chronic kidney failure	Chronic kidney failure	Chronic kidney failure
10	Nefritis acute	Nefritis acute	Nefritis acute

Based on the comparative result shown in Table III, the certainty factor achieved six correct test data from ten cases of test data. The Certainty factor failed to diagnose four kidney diseases, i.e., chronic kidney failure, kidney cysts, kidney cancer, and acute nephritis. Especially for acute nephritis, the test was done twice. In the first test, the certainty factor resulted in an incorrect diagnosis, while in the second experiment, the diagnosis results are the same as the expert diagnosis results. From these results, the accuracy of the certainty factor is calculated using formula 5, and the result is the certainty factor is 60% in the case study to diagnose kidney disease.

Dempster Shafer has a higher level of accuracy than the Certainty factor, as shown in Table III. Out of 10 tests, Dempster-Safer only failed in the diagnostic test for acute nephritis at point 2 and kidney cancer at point 5. Then the value of accuracy is measured using equation 5. Based on the above calculation, the accuracy of the Dempster Shafer is 80% in the case study diagnosing kidney disease.

After the calculation process on the certainty factor method and the Dempster Shafer method was found in the case study of diagnosing kidney disease, the Dempster Shafer method had a higher accuracy rate of 80%, compared to the certainty factor method had an accuracy rate of 60%, but in different cases, it could be just given different results. This is because in diagnosing kidney disease, drawing conclusions using the certainty factor method is only able to compare two symptoms to produce a confidence value. At the same time, the Dempster Shafer method in its calculation involves the probability of all possible diseases from each symptom.

V. CONCLUSION AND FUTURE WORK

This research has developed an expert system to find the accuracy of two algorithms i.e., certainty factor and dempster-Shafer for diagnosing kidney disease. The system develops using UML and web-based to make it easier for users to make an initial diagnosis before meeting directly with expert doctors (experts). The two algorithms were compared to determine the level of accuracy in diagnosing kidney disease compared to the results of expert doctors' diagnoses. Testing is done by analyzing it ten times according to types of kidney diseases. The test results show that Dempster Shafer has an accuracy rate of 80% higher than the Certainty Factor, which only has an accuracy rate of 60%. For further research, it is necessary to test several expert system algorithms to determine the level of accuracy in diagnosing certain diseases. It is also essential to research whether an expert system is still crucial to prospective patients. On the other hand, several telemedicine applications have been introduced that allow patients to communicate directly with expert doctors through applications without going to the hospital.

REFERENCES

[1] Andriani, Anik. 2017. Pemrograman Sistem Pakar : Konsep Dasar dan Aplikasinya Menggunakan Visual Basic 6. Mediakom, Yogyakarta.

[2] Yastita, Sri, Yohana Dewi Lulu, dan Rika Perdana Sari. 2012. Sistem Pakar Penyakit Kulit Pada Manusia Menggunakan Metode Certainty Factor Berbasis Web. Jurnal Seminar Nasional Teknologi Informasi Komunikasi dan Industri (SNTIKI) 4 Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau Pekanbaru, ISSN : 2085-9902.

[3] Sari, Nur Anjas. 2013. Sistem Pakar Mendiagnosa Penyakit Demam Berdarah Menggunakan Metode Certainty Factor. Jurnal Pelita Informatika Budi Darma, Volume : IV, ISSN : 2301-9425.

[4] Halim, Stephanie, dan Seng Hansun. 2015. Penerapan Metode Certainty Factor dalam Sistem Pakar Pendeteksi Resiko Osteoporosis dan Osteoarthritis. Jurnal ULTIMA Computing, Vol. VII, No. 2, ISSN 2355-3286.

[5] Sulistyohati, Aprilia, dan Taufiq Hidayat. 2008. Aplikasi Sistem Pakar Diagnosa Penyakit Ginjal dengan Metode Dempster-Shafer. Jurnal Seminar Nasional Aplikasi Teknologi Informasi, ISSN: 1907-5022.

[6] Marianti. 2017. Penyakit Ginjal. [online]: https://www.alodokter.com/penyakit-ginjal/

[7] Nugraha, Achmad Affan Suprayogi, Nurul Hidayat, dan Lutfi Fanani. 2018. Sistem Pakar Diagnosis Penyakit Kucing Menggunakan Metode Naive Bayes – Certainty Factor Berbasis Android. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, universitas Brawijaya. Vol. 2, No. 2, e-ISSN: 2548-964X.

[8] Darwanis, dan Yul Hendra. 2016. Perbandingan Efektifitas Metode Forward Chaining dan Dempster Shafer Mendiagnosa Penyakit Tanaman Kelapa Sawit. Jurnal Fakultas Ilmu Komputer (FIKOM) Universitas Almuslim.

[9] Husain, Agung Toto Wibowo, dan Endro Ariyanto. 2010. Analisis Perbandingan Sistem Pakar untuk Mendiagnosis Penyakit Mulut dan Gigi pada Manusia dengan Menggunakan Certainty Factor dan Teori Dempster-Shafer. Jurnal Universitas Telkom.

[10] Azka, Nurul, Andi Farmadi dan Dwi Kartini. 2017. Sistem Pakar Diagnosa Awal Penyakit Ginjal Dengan Metode Fuzzy Logic Untuk Penentuan Certainty Factor. Jurnal Elektronik Nasional Teknologi dan Ilmu Komputer (JENTIK).

[11] Prabowo, dkk. 2013. Sistem Pakar untuk Mendiagnosa dan Menanggulangi Penyakit pada Ikan Lele Dumbo(Clarias Gariepinus) Menggunakan Metode Backward Chaining. Jurnal Simposium Nasional Teknologi Terapan(SNTT). ISSN : 2339-028X.

[12] Wulandari, Fitri dan Ihsan Yuliandri. 2014. Diagnosa Gangguan Gizi Menggunakan Metode Certainty Factor. Jurnal Sains, Teknologi dan Industri, Vol. 11, No. 2 ISSN 1693-2390.

[13] Ikorasaki, Frans. 2015. Sistem Pakar Mendiagnosa Penyakit pada Tulang dengan Menggunakan Metode Certainty Factor. Jurnal Prosiding Seminar Nasional Informatika, Vol 1, No 1.

[14] Wahyuni , Elyza Gustri dan Widodo Prijodiprojo. 2013. Prototype Sistem Pakar untuk Mendeteksi Tingkat Resiko Penyakit Jantung Koroner dengan Metode Dempster-Shafer. Jurnal Teknik Informatika, FTI UII, Yogyakarta. IJCCS, Vol.7, No.2, ISSN: 1978-1520.