Scientific Journal of Informatics
Vol. 4, No. 2, November 2017
p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Expert System Diagnosis of Bowel Disease Using Case Based Reasoning with Nearest Neighbor Algorithm

Lucky Gagah Vedayoko[1], Endang Sugiharti[2], Much Aziz Muslim[3]

[1,2,3] Computer Science Departement, FMIPA, Universitas Negeri Semarang, Indonesia
Email: [1]stevano.lucky@gmail.com, [2]endangsugiharti@mail.unnes.ac.id, [3]a212muslim@yahoo.com

Abstract

Expert System is a computer system that has been entered the base of knowledge and set of rules to solve problems like an expert. One method in the expert system is Case Based Reasoning. To strengthen the retrieve stage of this method, the Nearest Neighbor algorithm is used. Bowel is one of the digestive organs susceptible to disease. The purpose of this study is to implement expert systems using Case Based Reasoning with Nearest Neighbor algorithm in diagnosing bowel disease and determine the accuracy of the system. Data used in this research are 60 data, obtained from medical record RSUD dr. Soetrasno Rembang. Variables used are general symptoms and types of diseases. The level of system accuracy resulting from scenario are 40 data as source case, and 20 data as target case that is equal to 95%.

Keywords: Expert System, Bowel Disease, Case Based Reasoning, Nearest Neighbor.

1. INTRODUCTION

Expert System is part of high-level specialized software or high-level programming language, which is trying to duplicate the functionality of an expert in a particular field of expertise [1]. Expert systems can be used to overcome multiple problems by reasoning using knowledge like an expert [2]. On implementation of expert system are not only the problems that rely on algorithm, but also sometimes also the problem that is difficult to comprehend [3]. Within the expert system, there is a knowledge base and an inference engine (set of rules) [4]. In other words, the expert system is a software-based system that makes or evaluates decisions based on rules, defined in the software [5]. The purpose of the expert system is not to replace human role, but to present human knowledge into a system form, so it can be used by many people [6]. From some previous studies, expert systems provide good results for solving cases that use complex data in cases of prediction or diagnosis, such as diagnosis of lung disease, diagnosis of matic motorcycle damage, treatment of cases based on symptoms, and diagnosis of digestion disease in humans [1, 7-10]

Case Based Reasoning is one part of Artificial Intelligence, which can solve problems in search of solutions of a new case, the system will search for solutions from old cases that have similar problems [11]. In looking for a case resemblance, all feature weights

are considered equal, meaning no feature is considered more important, so an algorithm needs to be added on it. Nearest Neighbor is an approach method to finding cases by calculating the proximity between new cases and old cases, which based on matching the weights of a number of existing features [12]. With this algorithm, not only all features have the same weight but also, there are features that have a higher priority.

Bowel is one of the digestive organs that are susceptible to disease. One of the bowel diseases, diarrhea is the second leading cause of child death in the world with 15 million children dying every year [13]. Typhoid fever is also an infectious disease that is easily contagious and can attack many people, so it can cause an outbreak [14]. Usually a doctor can diagnosing the disease by analyzing the symptoms that felt by the patient. In the current technological advances, a disease will be detected more quickly through perceived symptoms using a system [15].

Based on the above description, the purpose of this research is to implement expert system Case Based Reasoning with Nearest Neighbor algorithm and know its accuracy in diagnosing bowel disease.

2. METHODS

2.1. Case Based Reasoning

Case-based reasoning is based on the hypothesis, that the solutions of past problems can be helpful for solving a current problem, whenever there is any resemblance between them [16]. There are four stages in Case Based Reasoning [17]:

1) Retrieve, looking for similarities between old cases with new cases and taking the most similar
2) Reuse, reusing the solution of the old cases
3) Revise, repaired the failed solution
4) Retain, save sections of issues that may be useful for the future.

2.2. Nearest Neighbor

Nearest Neighbor algorithm is implemented in retrieve Case Based Reasoning stage. The algorithm is to find the closest match of the case that has been stored in the database to the new case by using distance calculation, which determines how similar the two cases by comparing the feature [18]. The function of Nearest Neighbor similarities can be formulated as follows [12].

$$Similarity(T,S) = \frac{\sum_{i=1}^{n} f(T_i, S_i) * w_i}{w_i} \quad (1)$$

Information:

Q: new case

S: existing cases in storage

n: number of attributes in each case

i: an individual attribute between 1 s / d n
f: similarity function between case T and case S
wi: the weight assigned to the i-th attribute

2.3. Process Design

The steps of making an expert system for the diagnosis of bowel disease are shown in Figure 1.

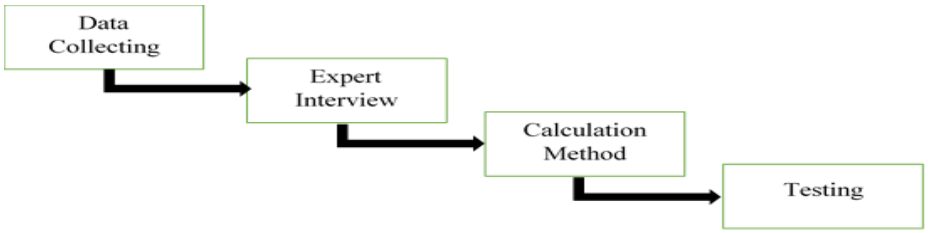

Figure 1. The Steps of Making the Expert System

The program flowchart is shown in Figure 2.

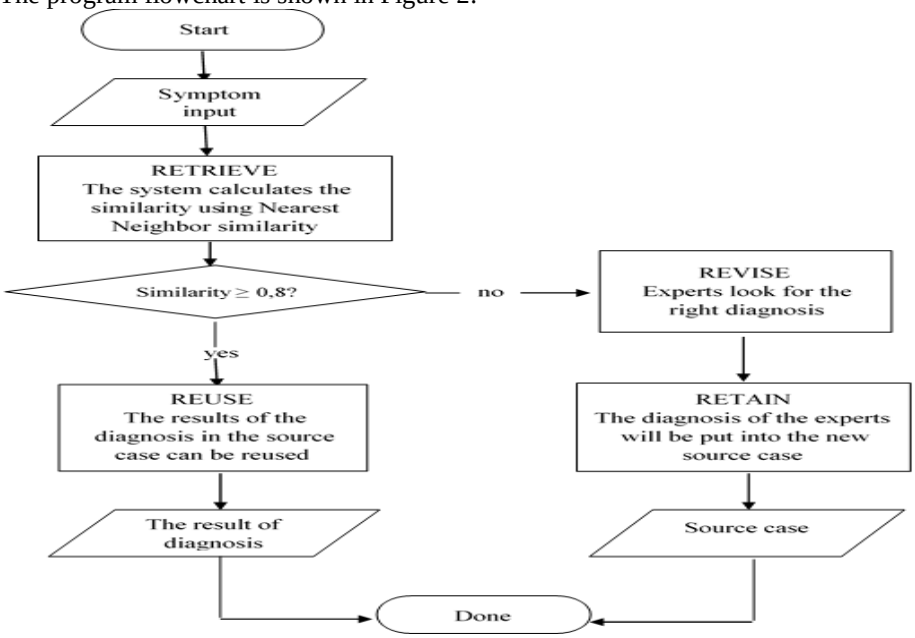

Figure 2. The Program Flowchart

2.4. System Development

The design of this expert system are using waterfall model. Waterfall is an approach method based on the assumption that major decisions must be made before encoding begins [19]. This model is often used by systems analysts in general [20]. There are

four stages in the waterfall model. The stages are, needs analysis, design, implementation and testing [21].

1) The requirement analysis stage is defining the entire software format, identifying all the needs, and outlines of the created system [22].
2) The design stage is to design applications including interface design, and database structure design [23]. To create an interesting web-based application program (website) it must be designed beforehand, so the achieved results are suitable with the predetermined objectives [24].
3) The implementation stage is designing software which realized as a series of program or program unit [25].
4) The testing stage is to test whether the system is ready and feasible to use. The tester can define the set of input conditions and perform testing on functional specifications of the program [26].

3. RESULTS AND DISCUSSION

3.1. Data Collecting

The data used in this research is the result of medical records of 60 patients on 2016, who suffer from bowel diseases such as Acute Diarrhea, Tipoid Fever, Appendicitis, Gastroentritis and Colitis taken at RSUD dr. Soetrasno Rembang Regency. The data that used on this system are symptoms experienced by patients in RSUD dr. Soetrasno and the diagnosis result from the responsible doctor.

3.2. Expert Interview

The interview was conducted with a specialist in internal disease who is a doctor in RSUD dr. Soetrasno. From the interview, the result is weight value of each symptom in bowel disease shown in Table 1.

Table 1. The Value of Symptom Weights

No	Symptom	P1	P2	P3	P4	P5	Weight
1	liquid faeces	√	√	√	√		5
2	faeces are black/ there is blood					√	5
3	Cough	√	√				1
4	weight loss	√			√		1
5	Dehydration	√	√				3
6	Fever	√	√	√			5
7	fever for days		√				5
8	sometimes fainted		√				1
9	cold feet				√		1
10	bloated		√	√			1
11	stomach cramps				√		3
12	limp	√	√		√	√	1
13	white/ dirty tongue		√				3

14	eyes limp	√					1
15	shiver		√				1
16	nausea	√	√	√	√	√	1
17	heartburn	√		√		√	1
18	gag	√	√	√	√	√	3
19	appetite down	√					1
20	abdominal pain		√				3
21	lower right abdominal pain			√			5
22	heart pain		√				3
23	cold		√				1
24	pale				√		1
25	dizzy	√	√				1
26	appendicitis history			√			3
27	often urinating	√	√				1
28	often defecate	√					3
29	often sweating				√		1
30	hard to defecate		√				1
31	Vertigo					√	1

Keterangan:
P1: Acute Diarrhea
P2: Tipoid Fever
P3: Appendicitis
P4: Gastroentritis
P5: Colitis

3.3. Calculation Methods

1). Retrieve Stage

This stage looks for similarities between the target case (new case) with the source case (old case). The example of target case 1 is in Table 2.

Table 2. The Example of Target Case 1

Name	Symptom
P013	Liquid faeces, gag

Target case 1 will be searched for resemblance between source case 1 and source case 2. Example source case 1 is in Table 3.

Table 3. The Example of Source Case 1

Name	Symptom	Disease
P004	Liquid faeces, gag, limp	Acute Darrhea

a. Value of proximity attribute liquid faeces symptom: 1
b. Weight of liquid faeces symptom: 5
c. Value of proximityt attribute limp symptoms: 0
d. Weight of limp symptoms: 1

e. Value of proximity attribute gag symptoms: 1
f. Weight of gag symptoms: 3

$$\text{Similarity} = \frac{[(a*b) + (c*d) + (e*f)]}{b + d + f}$$

$$= \frac{[(1*5) + (0*1) + (1*3)]}{5 + 1 + 3} = 0{,}89$$

The example of source case 2 in Table 4.

Table 4. The Example of Source Case 2

Name	Symptom	Disease
P008	Liquid faeces, gag, fever	Acute Diarrhea

a. Value of proximity attribute liquid faeces symptom: 1
b. Weight of liquid faeces symptoms: 5
c. Value of proximity attribute fever symptoms: 0
d. Weight of fever symptoms: 5
e. Value of proximity attribute gag symptoms: 1
f. Weight of gag symptoms: 3

$$\text{Similarity} = \frac{[(a*b) + (c*d) + (e*f)]}{b + d + f}$$

$$= \frac{[(1*5) + (0*5) + (1*3)]}{5 + 5 + 3} = 0{,}61$$

2). Reuse Stage
This stage reuses the diagnosis of the old case. If the minimum similarity score is in the range 0.8 - 1, then the diagnosis of the old case will be reused for the new case. The similarity value of target case 1 with source case 1 is 0.89, so the diagnosis from source case 1 that is acute diarrhea, can be reused in target case 1.

3). Revise Stage
This stage stores a case that has a similarity score below 0.8. So a target case that has a low resemblance is not immediately discarded but stored for future study materials.

4.) Retain Stage
A case that the expert has reviewed and can be found to be diagnosed at the revision stage will be saved to the database into a new source case for future learning.

3.4. Testing

After testing on 60 medical records data of patients with bowel disease in RSUD dr. Soetrano Rembang with 40 data scenario as source case and 20 data as target case, the result of accuracy equal to 95% as seen in Table 5.

Table 5. Results of Data Testing

Number of test data	Data matching	Data not matching	Accuracy
20	19	1	95%

From the result of medical record test data using the system that has been made, the result are 19 accurate test data with doctor's diagnosis and 1 data which is less accurate because the system uses strong calculation base, so it is not easily influenced by other factors. While the specialist in internal disease as an expert, has other considerations in determining the disease suffered by the patient.

3.5. Implementation

The consultation page is a page that users can use to consult. Without the need to login, the user can consult with the system. In consulting, users fill in the name, gender, age and any symptoms experienced. Figure 3 shows the consultation page.

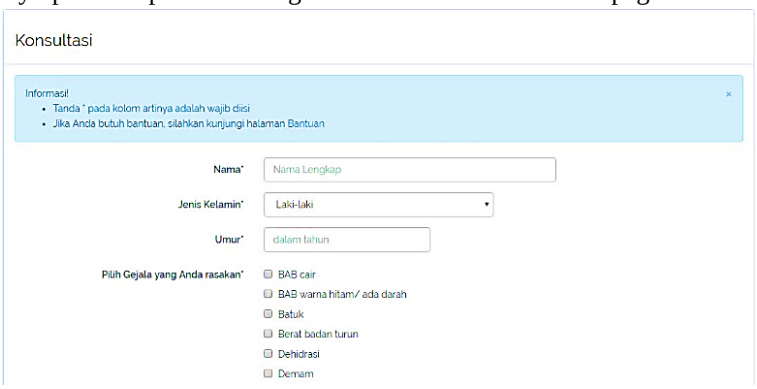

Figure 3. Consultation Page

The diagnosed page is a page that provides diagnosis information from the symptoms input. For the user page, it only displays the symptoms experienced and the results of his diagnosis. In contrast to admin and experts, in addition to displaying the symptoms and diagnosis results, it also displays the results of the calculation method, or level of similarity with the existing source case. Figure 4 show the admin and expert diagnosis page.

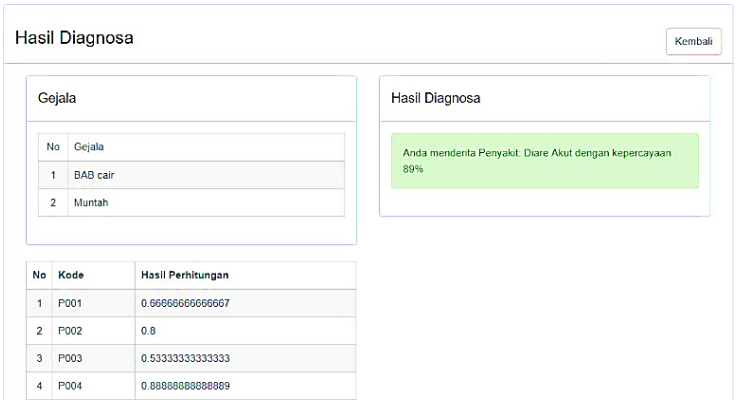

Figure 4. Result The Admin and Expert Diagnosis Page

4. CONCLUSION

From the results of the research related to implementation of Case Based Reasoning with Nearest Neighbor algorithm in diagnosing bowel disease in humans using medical record from RSUD dr. Soetrasno Rembang, it can be concluded that Case Based Reasoning has 4 stages. The retrieve stage, by finding the similarity between cases with the calculation of the Nearest Neighbor algorithm. Reuse stage, with a minimum similarity of 0.8 diagnostic results from source case can be reused. Revise phase, target case with a similarity value less than 0.8 will be reviewed by experts. Retain stage, the results of expert review at the revise stage will be stored into the database to serve as a new source case. After implementing 60 medical records data taken from RSUD dr. Soetrasno Rembang with scenarios of 40 source case and 20 target cases, it can be seen that the accuracy of the diagnosis of the system that is equal to 95%.

5. REFERENCES

[1] Hustinawaty, R. A. 2014. The Development of Web Based Expert System for Diagnosing Children Diseases Using PHP and MySQL. *International Journal of Computer Trends and Technology (IJCTT)*, *10*(4), 197-202.

[2] Naik, V. M., & Lokhanday, S. 2012. Building a Legal Expert System for Legal Reasoning in Specific Domain-A Survey. *International Journal of Computer Science & Information Technology*, *4*(5), 175.

[3] Setiabudi, W. U., Sugiharti, E., & Arini, F. Y. 2017. Expert System Diagnosis Dental Disease Using Certainty Factor Method. *Scientific Journal of Informatics*, *4*(1), 43-50.

[4] Pramesti, A. A., Arifudin, R., & Sugiharti, E. 2016. Expert System for Determination of Type Lenses Glasses Using Forward Chaining Method. *Scientific Journal of Informatics*, *3*(2), 177-188.

[5] Josephine, M. S., & Jeyabalaraja, V. 2012. Expert System and Knowledge Management for Software Developer in Software Companies. *Software Engineering and Technology*, *4*(4), 147-151.

[6] Muslim, M. A., Kurniawati, I. I. N., & Sugiharti, E. 2015. Expert system diagnosis chronic kidney disease based on mamdani fuzzy inference system. *Journal of Theoretical and Applied Information Technology*, *78*(1), 70.

[7] Kusuma, D. A., & Chairani, C. 2014. Rancang Bangun Sistem Pakar Pendiagnosa Penyakit Paru-Paru Menggunakan Metode Case Based Reasoning. *Jurnal Infotel*, *6*(2), 57-62.

[8] Kosasi, S. 2015. Pembuatan Aplikasi Diagnosa Kerusakan Mesin Sepeda Motor Matic dengan Case-Based Reasoning. *Creative Information Technology Journal*, *2*(3), 192-206.

[9] Risfianti, W.R., Pudjiantoro T.H., & Hadiana A.I. 2016. Penentuan Penanganan Kasus Terhadap Penyakit Berdasarkan Gejala Menggunakan Case Base Reasoning Dan Algoritma Nearest Neighbor (Studi Kasus: Klinik Citra Medika Cianjur). *Prosiding SNST ke-7 Tahun 2016*. Semarang: Universitas Wahid Hasyim Semarang.

[10] Situmorang, A. H., Hakim, I. N., & Shofyan, M. 2016. Aplikasi Sistem Pakar

Untuk Mendiagnosa Penyakit Pencernaan Pada Manusia Menggunakan Metode Forward Channing. *Semnasteknomedia Online*, *4*(1), 3-6.

[11] Atanassov, A., & Antonov, L. 2012. Comparative analysis of case based reasoning software frameworks jCOLIBRI and myCBR. *Journal of the University of Chemical Technology & Metallurgy*, *47*(1), 83-90.

[12] Kusrini. 2009. *Algoritma Data Mining*. Yogyakarta: Andi.

[13] Utami, R. S., & Wulandari, D. 2015. Studi Kasus: Asuhan Keperawatan pada Anak dengan Gastroenteritis Dehidrasi Sedang (Case Study: Nursing Care In Children With Gastroenteritis Moderate Dehydration). *IJMS-Indonesian Journal on Medical Science*, *2*(1).

[14] Nuruzzaman, H. & Syahrul, F. 2016. Analisis Risiko Kejadian Demam Tifoid Berdasarkan Kebersihan Diri dan Kebiasaan Jajan di Rumah. *Jurnal Berkala Epidemiologi*, *4*(1), 75.

[15] Handarko, J. L., & Alamsyah, A. 2015. Implementasi Fuzzy Decision Tree untuk Mendiagnosa Penyakit Hepatitis. *Unnes Journal of Mathematics*, *4*(2). .

[16] Guessoum, S., Laskri, M. T., & Lieber, J. 2014. RespiDiag: a case-based reasoning system for the diagnosis of chronic obstructive pulmonary disease. *Expert Systems with Applications*, *41*(2), 267-273.

[17] Aamodt, A. & Plaza E. 1994. Case Based Reasoning: Foundation Issues Methodological Variations, and Sistem Approaches. *AI Communication IOS Press*, *7*(1), 46-50.

[18] Salem, A. B. M., Roushdy, M., & HodHod, R. A. 2005. A case based expert system for supporting diagnosis of heart diseases. *AIML Journal*, *5*(1), 33-39.

[19] Patel, U. A., & Jain, N. K. 2013. New idea in waterfall model for real time software development. *International Journal of Engineering Research & Technology (IJERT)*, *2*(4), 2278-0181.

[20] Roviaji, R., & Muslim, M. A. 2017. Pembuatan Sistem Informasi Gardu Induk PT. PLN (Persero) App Semarang Se-Kota Semarang dengan Java Android. in *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi (SAKTI)* (Vol. 2, No. 1, pp. 182-185).

[21] Pressman, R.S. 2001. *Software Engineering*. Online. Tersedia di http://www.resource.mitfiles.com/ [diakses 18-2-2017].

[22] Nugroho, Z. A., & Arifudin, R. 2015. Sistem Informasi Tracer Study Alumni Universitas Negeri Semarang Dengan Aplikasi Digital Maps. *Scientific Journal of Informatics*, *1*(2), 153-160.

[23] Putra, A. T. 2015. Pengembangan E-Lecture menggunakan Web Service Sikadu untuk Mendukung Perkuliahan di Universitas Negeri Semarang. *Scientific Journal of Informatics*, *1*(2), 168-176.

[24] Muslim, M. A. 2012. Pengembangan Sistem Informasi Jurusan Berbasis Web Untuk Meningkatkan Pelayanan dan Akses Informasi. *Jurnal MIPA*, *35*(1).

[25] Purwinarko, A., & Sukestiyarno, Y. L. 2015. Model Expertise Management System di Universitas Negeri Semarang. *Scientific Journal of Informatics*, *1*(2), 177-184.

[26] Mustaqbal, M. S., Firdaus, R. F., & Rahmadi, H. 2016. Pengujian Aplikasi Menggunakan Black Box Testing Boundary Value Analysis (Studi Kasus: Aplikasi Prediksi Kelulusan SMNPTN). *Jurnal Ilmiah Teknologi Informasi Terapan*, *1*(3).