International Journal of Informatics and Computer Science (The IJICS)
Vol 2 No 1, March 2018
ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

The Application of The Nearest Neighbor Algorithm in Detecting Appendicitis Disease

Sari Handayani, Nelly Astuti Hasibuan, Putri Ramadhani

Department of Computer Science STMIK Budi Darma, North Sumatra, Indonesia

Abstract – The expert system is a branch of artificial intelligence that facilitates the knowledge of an expert into an application that can be used to help the work of experts. The knowledge base is formed by an inference engine that translates an expert's knowledge. The result or output of the expert system can not guarantee its certainty. Therefore, it takes several methods that can be used to solve it. In this study, expert systems are applied to the medical or medical world. The following are related to the medical appendicitis disease or better known as appendicitis.Method used to improve the accuracy of the decision of the nearest neighbor algorithm. The nearest neighbor algorithm is an approach to finding a case with proximity or similarity between a new case and an old case. The results of this study is an engineering inference expertise with the aim to obtain clinical decisions of appendicitis diseases on the user (user) as medical treatment as early as.

Keywords – Experts System, Nearest Neighbor Algoritm, Appendicitis Desease

1 INTRODUCTION

The development of medical science is now progressing very rapidly. Characterized by the many findings of treatment of various diseases. Health is a very valuable thing for humans because anyone can experience health problems. Not a few ordinary people who do not know even seem to ignore the health disorders that often occur. The reason behind the author to take the problem of this disease is because appendicitis disease including the disease most often ignored, considering some of the community considers that this disease is a disease that does not need special treatment and serious so that people will see a doctor if the disease has been spelled out acute. This resulted in the people affected by the disease had to undergo surgery and had to spend a very expensive cost for his recovery. Note that the function of the appendix itself is as an immunologic organ that actively plays a role in the secretion of immunoglobulin (an immune system) in which the appendix has lymphoid glands. Appendicitis is the most common cause of acute abdomen.

Discuss the diseases that often occur in the human body may be endless especially if the human body itself does not get food intake with proper nutrition, unhealthy lifestyle, and other factors that can hinder the body's performance system. Problems faced by each patient will certainly experience different symptoms, such as problems that often occur in appendicitis disease, there maybe some different diseases perform similar symptoms. Therefore, it is very important to build a computerized system that can help solve the problem. The computerized system built must be able to present the statement and the appropriate solutions, accurate, and efficient about the disease appendicitis.

With the limitations of experts, the system can function as an expert (expert) that is efficient for humans. One application of the expert system is to provide information about appendicitis disease, based on it then the expert system can detect appendicitis disease. In connection with the description above, the author intends to design the expert system as a tool for detecting appendicitis disease.

One of the methods that can be used in the expert system is the nearest neighbor algorithm. The nearest Neighbor algorithm is very easy and efficient in solving the problem of appendicitis disease. The Nearest Neighbor (k-nearest neighbor or k-NN) algorithm is an algorithm for classifying objects based on learning data closest to the object [1]. The result of this research is in the form of engineering inference of expertise with the aim to obtain a clinical decision on appendicitis disease. From the explanation then the existence of this system is expected to help the community in early diagnosis of appendicitis disease that is suffered by humans.

2 THEORY

2.1 Apendicitis Disease

Appendicitis is a disease commonly known by the general public as the appendix. Appendix in Latin is called Appendix Vermiformis, This organ is found in humans, mammals, birds, and some types of reptiles. At first, this organ is regarded as a supplemental organ that has no function, but it is now known that the function of the appendix is as an immunologic organ and actively plays a role in the secretion of immunoglobulin (an immune)

ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

that has or contains the lymphoid gland. Like other organs of the body, this appendix or appendix can be damaged or disrupted by disease. It is well known as appendicitis.

2.2 Nearest Algorithm

The nearest neighbor algorithm (sometimes called k-Nearest Neighbor) is an algorithm that classifies the proximity of the location (distance) of a data with other data. The simple principle adopted by the K-NN algorithm is "if an animal is like a duck, voices quack-duck like a duck, and looks like a duck, the animal might duck" [4].

A special case where classification is predicted based on the closest learning data (in other words, k = 1) is called the nearest neighbor algorithm. The nearest neighbor algorithm is based on the learning process using analogies or learning by analogy. The sample training is described in terms of numerical n-dimensional attributes. Each sample represents a point in the dimension space. In this way, all sample training is stored in the n-dimensional space pattern. When given the "unknown" sample, k nearest neighbor classifier searches for the sample K training space closest to the "unknown" sample. K training of this sample is k nearest neighbor of unknown samples. Unknown samples are set with the most common class among k nearest neighbors. When k = 1, the unknown sample is specified with the class of the sample training closest to the spatial pattern. The formula for computing similarity weights with nearest neighbor retrieval is as follows [1]:

$$\textit{Similarity (problem,case)} = \frac{s1*w1 + s2*w2 + ... + sn*wn}{w1 + w2 + ... + wn} \quad (1)$$

Information:
S = similarity (similarity value if symptom is equal then value 1 and if different then value 0)
W = weight (weight given)

Nearest Neighbor is a simple classification technique but has a pretty good work. However, K-NN also has advantages and disadvantages. Some of the characteristics of K-NN are as follows [4]:

1. K-NN (Nearest Neighbor Algorithm) is an algorithm that uses all training data to perform the process of classification (complete storage). This results in a very long prediction process for very large amounts of data. Another approach is to use the mean data of each class, then calculate the closest distance of the test data to the mean data of each class. This gives a faster job advantage, but the result is less satisfactory because the model only forms a linear hyperplane right in the middle between the 2 classes separating the 2 classes (for case 2 classes). The more train data, the smoother the hyperplane is made. There is a trade-off relation between the amount of trainee data on computational costs and the resulting decision boundary quality.
2. The K-NN algorithm does not distinguish each feature with a weight as in the Artificial Neural Network (ANN) which attempts to suppress features that do not contribute to the classification to 0 in the weights. K-NN (Nearest Neighbor Algorithm) has no weight for each feature.
3. Since K-NN belongs to the lazy learning category that stores some or all of the data and almost no training process, K-NN (Almost Nearest Neighbor Algorithm) is very fast in the training process (because it does not exist), but very slow in the prediction process.
4. The most complicated thing of the nearest neighbor algorithm (K-NN) is to determine the most appropriate value of K.
5. Since K-NN (Nearest Neighbor Algorithm) in principle is choosing the nearest neighbor, distance parameter is also very important to be considered in accordance with the case data. Euclidean is perfect for using the shortest distance (straight) between two data, but Manhattan is very robust to detect outliers in the data.

3 RESULT AND DISCUSSION

Appendicitis disease begins with symptoms that are often found in diseases associated with other abdominal problems. Many of the general population are not aware of the same symptoms so that a person suffering from the disease will only know after the existence of sustained symptoms which can then cause the disease becomes very severe. In this stage is done search and data collection and knowledge obtained from an expert, so that in the end analysis should be obtained a system with a good and clear. Symptoms of appendicitis disease obtained based on the results of data collection and search conducted by researchers, then confirmed to experts that the symptoms are truly experienced by people with appendicitis disease. Therefore, designed an expert system as a tool for early diagnosis of appendicitis disease based.

ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

Table 1. Rules of Apendicitis Disease Symptoms

Code	Symptoms of Disease
G001	Fever
G002	The pulse rate increases
G003	Feels cramps in the abdomen
G004	Nausea and vomiting
G005	Body temperature above 37^0C
G006	Pale face
G007	The stomach often feels bloated
G008	Lasseque sign, pain or tenderness in the waist area due to nerve strain when lasseque test is performed
G009	High Blood Endarm (LED) is characterized by pain in the joints and the presence of infections in the body such as influenza
G010	Leukocytes, the rising number of white blood cells in the blood is characterized by fever, pain, and heat during BAK, pale face, and often tired
G011	Muscle pain, pain in the right and lower left abdomen when in the knock
G012	Tenderness, painful when the right lower quadrant is suppressed
G013	Obturator sign, a pain that occurs when the pelvis and knee are reflected
G014	Psoas sign, pain in the lower right abdomen
G015	Rovsing sign, pain in the lower right abdomen when doing emphasis on the left lower abdomen

Source: Results Consultation with dr. Robert F. Siregar, S.Pd

With the table above it can be presented as the rule below. This representation is used to obtain diagnostic results, here is the discussion:

Rule 1:
IF fever
AND pulse rate increases
AND cramps in the abdomen
AND lasseque sign, pain or pain in the waist area due to nerve strain when lasseque test is done
AND leukositisis, the increase in the number of white blood cells in the blood is characterized by the presence of fever, pain, and heat when the BAK, pale face
AND excessive nausea and vomiting
AND pain knock, pain on the right and left underbelly when in the knock
AND is tender, painful when the right lower quadrant is suppressed
AND obturatur sign, the pain that occurs when the pelvis and knee are reflected
AND psoas sign, pain in the lower right abdomen
AND rovsing sign, pain in the lower right abdomen when doing the emphasis on the left lower abdomen
AND body temperature above 370C
AND faces often become pale
THEN name of acute appendicitis disease

Rule 2:
IF fever
AND cramps in the abdomen
AND lasseque sign, pain or pain in the waist area due to nerve strain when lasseque test is done
AND leukositisis, the increase in the number of white blood cells in the blood is characterized by the presence of fever, pain, and heat when the BAK, pale face
AND excessive nausea and vomiting
AND pain knock, pain on the right and left underbelly when in the knock
AND is tender, painful when the right lower quadrant is suppressed

ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

AND obturator sign, the pain that occurs when the pelvis and knee are reflected
AND psoas sign, pain in the lower right abdomen
AND roving sign, pain in the lower right abdomen when doing the emphasis on the left lower abdomen
AND stomach often feels bloated
AND LED (High Blood Depth Rate) is characterized by the pain in the joints and the presence of infections in the body such as influenza
THEN name of chronic appendicitis disease

The analysis of the expert system that is built is a rule or rules that apply the method nearest neighbor algorithm. The method of the nearest neighbor algorithm or K-NN is an algorithm that classifies by proximity of location or distance of data with other data in overcoming difficulties at the time of diagnosing symptoms of appendicitis disease. In expressing the proximity of location or distance between the data one with other data, the nearest neighbor algorithm or K-NN generate calculations based on the weighting process that has been given. Then calculate the level of resemblance to the new case entered by the user shown in the following example. As the completion stage, then first done weighting to determine the value of each symptom. Based on the similarity level obtained, the system will produce the diagnosis of disease suffered by the patient.

Case in point: A patient came to the surgical poly with a major complaint of heartburn and spread to the lower right abdomen of a continuous nature. Patients also complain of fever and no appetite that causes the stomach to become bloated. On physical examination, the patient is diagnosed with an elevated ESR. Local examination of the abdomen is found at the point Mc. Burney tenderness is also sore and the presence of obturator sign.

Table 2. Weighted Value Symptoms

No.	Disease Case / Symptoms	1	2	3	4	5	6	Value
1.	Fever	√	√	√	√	√	√	$\frac{6}{6} = 1$
2.	The pulse rate increases	-	-	-	√	-	√	$\frac{2}{6} = 0{,}33$
3.	Feels cramps in the abdomen	√	-	-	√	-	-	$\frac{2}{6} = 0{,}33$
4.	Nausea and vomiting	-	√	√	-	√	√	$\frac{4}{6} = 0{,}66$
5.	Body temperature above 37^0C	-	-	-	-	-	-	-
6.	Pale face	-	-	-	-	√	-	$\frac{1}{6} = 0{,}16$
7.	The stomach often feels bloated	-	√	-	-	-	-	$\frac{1}{6} = 0{,}16$
8.	*Lasseque sign*	-	-	-	√	-	-	$\frac{1}{6} = 0{,}16$
9.	LED is high	-	√	-	-	-	√	$\frac{2}{6} = 0{,}33$
10.	Leukocytes	-	-	-	-	-	-	-
11.	Pain Knock	√	√	-	√	√	-	$\frac{4}{6} = 0{,}66$
12.	Tenderness	√	-	√	√	√	√	$\frac{5}{6} = 0{,}88$
13.	*Obturatur sign*	-	√	-	-	√	-	$\frac{2}{6} = 0{,}33$
14.	*Psoas sign*	√	-	-	-	-	-	$\frac{1}{6} = 0{,}16$
15.	*Rovsing sign*	-	-	-	-	-	-	-

Table 3. Similarity with Illness 1

Case of Illness 1	New Disease Cases	Similarity value
- Fever	- Fever	1
- Tenderness	- Tenderness	1
- Tenderness	- Tenderness	1
- Stomach cramps	- Bloated	0
- Psoas sign	- LED increases	0
	- Obturator sign	0

ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}88) + (1 * 0{,}66)}{1 + 0{,}88 + 0{,}66 + 0{,}33 + 0{,}16 + 0{,}16 + 0{,}33 + 0{,}33} = \frac{2{,}54}{3{,}85} = 0{,}659$$

Table 4. Similarity with Illness 2

Case of Illness 2	New Disease Cases	Similarity value
- Fever	- Fever	1
- Nausea and vomiting	- Tenderness	0
- Tenderness	- Tenderness	1
- Bloated	- Bloated	1
- LED increases	- LED increases	1
- Obturatur sign	- Obturator sign	1

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}66) + (1 * 0{,}16) + (1 * 0{,}33) + (1 * 0{,}33)}{1 + 0{,}66 + 0{,}16 + 0{,}33 + 0{,}33 + 0{,}66 + 0{,}88} = \frac{2{,}48}{4{,}02} = 0{,}6169$$

Table 5. Similarity with Illness 3

Case of Illness 3	New Disease Cases	Similarity value
- Fever	- Fever	1
- Tenderness	- Tenderness	1
- Nausea and vomiting	- Tenderness	0
	- Bloated	0
	- LED increases	0
	- Obturator sign	0

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}88)}{1 + 0{,}88 + 0{,}66 + 0{,}66 + 0{,}16 + 0{,}33 + 0{,}33} = \frac{1{,}88}{4{,}02} = 0{,}4676$$

Table 6. Similarity with Illness 4

Case of Illness 4	New Disease Cases	Similarity value
- Fever	- Fever	1
- Tenderness	- Tenderness	1
- Tenderness	- Tenderness	1
- Stomach cramps	- Bloated	0
- The pulse rate increases	- LED increases	0
	- Obturator sign	0

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}88) + (1 * 0{,}66)}{1 + 0{,}88 + 0{,}66 + 0{,}33 + 0{,}33 + 0{,}16 + 0{,}33 + 0{,}33} = \frac{2{,}54}{4{,}02} = 0{,}6318$$

Table 7. Similarity with Illness 5

Case of Illness 5	New Disease Cases	Similarity value
- Fever	- Fever	1
- Tenderness	- Tenderness	1
- Tenderness	- Tenderness	1
- Pale face	- Bloated	0

ISSN 2548-8384 (online)
ISSN 2548-8449 (print)
Page: 1-6

Case of Illness 5	New Disease Cases	Similarity value
- Lasseuque sign	- LED increases	0
- Obturatur sign	- Obturator sign	1
- Nausea and vomiting		0

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}88) + (1 * 0{,}66) + (1 * 0{,}33)}{1 + 0{,}88 + 0{,}66 + 0{,}33 + 0{,}16 + 0{,}16 + 0{,}16 + 0{,}33 + 0{,}66} = \frac{2{,}87}{4{,}34} = 0{,}6612$$

Table 8. Similarity with Illness 6

Case of Illness 6	New Disease Cases	Similarity value
- Fever	- Fever	1
- Tenderness	- Tenderness	1
- Nausea and vomiting	- Tenderness	0
- The pulse rises	- Bloated	0
- LED increases	- LED increases	1
	- Obturator sign	0

The Solution:

$$= \frac{(1 * 1) + (1 * 0{,}88) + (1 * 0{,}33)}{1 + 0{,}88 + 0{,}66 + 0{,}66 + 0{,}16 + 0{,}33 + 0{,}33 + 0{,}33} = \frac{2{,}21}{4{,}35} = 0{,}5080$$

Based on the calculation of 6 existing old cases, cases that have the lowest similarity weight is case number 3 that is equal to 0.4676. The case that has the highest weight of similarity is case number 5 of 0.6612. The results of the calculation using the nearest neighbor algorithm with the weights indicating the closest value of the distance is the case at number 5. So, the diagnosis that will be recommended to the user refers to case number 5 which is acute appendicitis disease.

4 CONCLUSION

The results of analysis of the previous chapters, it can be concluded, where the conclusions are likely to be useful for the readers. Thus, the writing of this article can be more useful. The conclusions are as follows:

1. Appendicitis disease can be diagnosed with an expert system, where the expert system can make it easier for experts and society in need to get the early information about the symptoms of this disease
2. The nearest neighbor algorithm can be applied to the expert system because basically this algorithm is only to get the information of the distance of proximity between new disease with pre-existing disease.
3. Diagnosis of appendicitis disease can be designed by using web-based programming language.

REFERENCES

[1] N. Mariana, "Penerapan Algoritma k-NN (Nearest Neighbor) Untuk Deteksi Penyakit (Kanker Serviks)," Dinamika Informatika, vol. 7, p. 27, 2015.
[2] Sutojo, Kecerdasan Buatan, Yogyakarta: ANDI, 2011.
[3] M. Arhami, Konsep Dasar Sistem Pakar, Yogyakarta: ANDI, 2005.
[4] E. Prasetyo, DATA MINING - Konsep dan Aplikasi Menggunakan MATLAB, Yogyakarta: ANDI, 2012.
[5] M. John H. Tilden, "Appendicitis: The Etiology, Hygenic, and Dietetic Treatment," The Project Gutenberg Literary Archive Foundation, US, 2003.