Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 22, No. 1, November 2022, pp. 215∼228
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v22i1.2280

Comparison of Naive Bayes and Dempster Shafer Methods in Expert System for Early Diagnosis of COVID-19

Nurdin[1], Erni Susanti[2], Hafizh Al-Kautsar Aidilof[3], Dadang Priyanto[4]
[1,2,3]Universitas Malikussaleh, Aceh, Indonesia
[4]Universitas Bumigora, Mataram, Indonesia

Article Info

Article history:

Received Augustus 12, 2022
Revised November 17, 2022
Accepted November 26, 2022

Keywords:

COVID-19
Dempster Shafer
Diagnosis
Expert system
Nave Bayes

ABSTRACT

COVID-19 is a respiratory infection disease caused by the corona virus. Transmission of this virus can spread very quickly so that the number of cases of the corona virus continues to grow and becomes an epidemic that spreads not only in Indonesia but also in other countries in the world. The purpose of this study is to build an expert system that is able to diagnose Covid-19 early by using a comparison of the Nave Bayes method and the Dempster Shafer method. The amount of data used in this study is 550 data, consisting of 500 training data and 50 testing data. While the variables used are symptoms related to COVID-19 as many as 17 symptoms consisting of G01, G02, G03, G04, G05, G06, G07, G08, G09, G10, G11, G12, G13, G14, G15, G16, G17. The diagnostic data consists of Suspected (PDP), Non-Suspected, and Close Contact (ODP). The results of the percentage test by comparing system diagnoses with expert diagnoses, for the nave Bayes method it has an accuracy of 96% with 48 diagnoses according to expert diagnoses from 50 tested data. Meanwhile, the Dempster Shafer method has an accuracy of 40% with 20 diagnoses according to expert diagnoses from 50 tested data. Based on the results of this study, the Naive Bayes and Dempster Shafer methods can be applied to an expert system for early diagnosis of COVID-19, from the results of the system testing the Naive Bayes method has better accuracy than the Dempster Shafer method.

Corresponding Author:

Nurdin,
Program Studi Magister Teknologi Informasi,
Universitas Malikussaleh, Aceh, Indonesia
Email: nurdin@unimal.ac.id

How to Cite: N. Nurdin, E. Susanti, H. Aidilof, and D. Priyanto, Comparison of Naive Bayes and Dempster Shafer Methods in Expert System for Early Diagnosis of COVID-19, MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 22, no. 1, pp. 217-230, Nov. 2022.

1. INTRODUCTION

As a result of the COVID-19 pandemic that has hit the world, it has had a major impact on various sectors of human life, especially in the economic sector [1, 2], environmental sector [3] and the education and teaching sector [4–6]. One of the efforts to prevent the increase and spread of COVID-19 requires a computer system based on an expert system that can diagnose COVID-19 early. COVID-19 is a new disease that must be watched out for because of its very fast transmission and has a mortality rate that should not be ignored [6]. COVID-19 is a respiratory infection disease caused by the Severe Acute Respiratory Syndrome virus (SARS-CoV-2) which is often referred to as the corona virus [7].

COVID-19 is a disease caused by a type of corona virus that mutates, can spread from person to person, making this virus spread quickly and widely in various countries and the number of infections is increasing, making COVID-19 designated as an epidemic that hit not only in Indonesia but also in Indonesia. also other countries in the world [8]. Symptoms caused by corona virus infection such as general symptoms, namely fever and cough, as well as special symptoms such as taste disorders, olfactory disorders, make people panic if they experience similar symptoms. One application that can help diagnose COVID-19 early is an expert system application [9]. Expert systems solve a problem according to their expertise by imitating or simulating the knowledge and skills of an expert in a particular area. An expert system has two main components, namely a knowledge base and an inference engine. The knowledge base is a place for storing knowledge in a computer, where this knowledge is taken from expert knowledge.

The Nave Bayes method is a simple probability classification method by calculating a set of possibilities that is calculated by adding up the frequency and value of the training data that becomes the information guide. [10]. The nave Bayes method is the basis of machine learning and a basic method in data mining to assist decision making [11]. The Nave Bayes method uses previous events to predict future opportunities by using previous cases to solve problems. The Dempster Shafer method is a method that uses a belief function from experts, this method is used by combining pieces of information and calculating the probability of an event [12]. The Dempster Shafer method uses the trust value of the expert as the weight of each event in the form of possible events. Dempster Shafer theory is a mathematical calculation theory that is strong enough to prove based on belief functions with plausible reasoning, both of which are used to balance and combine several separate information (evidence) to calculate the probability of an event [13].

The nave bayes method and the dempster shafer method have been widely used in research related to expert systems, including the application of the nave bayes method in expert system modeling for disease diagnosis in shallots with an accuracy of 93.54% testing results [14] and the application of the nave Bayes method on an expert system to diagnose diseases in corn with an accuracy rate of 92% [15], The nave Bayes method can also be used in expert systems for determining women's facial skin types with a validation level of 90% [16]. The Dempster Shafer method can also be implemented in an expert system for the diagnosis of thyroid disease [17], and an expert system to diagnose stroke with an accuracy rate of 80% [18] and an expert system for the detection of rice plant diseases with an accuracy of 91% [19]. Other research on the implementation of an expert system for diagnosing skin diseases using the Dempster Shafer method was successfully carried out, with an accuracy of 68.8% of the experimental sample data based on the symptoms experienced by the patient [20]. From the aspect of computational complexity, the Dempster Shafer method is more complicated than the Bayesian network. the bayesian network method produces 6 probabilities higher than 10 correct diagnoses, while the dempster shafer method produces the highest 4 possible correct correct diagnoses so that the bayesian network method is more effectively used according to research [21]. Meanwhile, according to research conducted by [22] about the comparative analysis of the expert system for diagnosing cattle disease using the bayesian network and the dempster shafer method with the results of the percentage of diagnosing the dempster shafer method being better, the result is 87.2% compared to the bayesian network method with the result of 75.3%.

Previous research that has been done by others related to both methods and related to COVID-19, such as research conducted by [23] implementation of an expert system using the Certanty Factor method for early diagnosis of Corona Virus Disease (COVID-19), the output generated from this study is in the form of grouping based on the symptoms experienced by the patient. Patients who experience serious symptoms will be grouped into PDP status, patients who experience milder symptoms will be grouped into ODP status, while patients who experience symptoms outside of the main symptoms will be grouped into NON ODP. Research conducted by [24] decision support system to diagnose coronavirus disease (COVID-19) Using Dempster-Shafer. This study uses three disease criteria that allow patients to suffer, namely COVD-19, influenza and the common cold with 14 symptoms with a weight value each of which has been given by the relevant expert, each weight selected is then calculated the weight value using the Dempster Shafer method and will produce the conclusion is the greatest value of the three criteria for the disease that the patient may suffer, namely, COVID-19, Influenza or the common cold.

Research conducted by [25] by comparing the nave Bayes method and the certainty factor on an expert system to detect the COVID-19 virus, it shows that the nave Bayes method requires data from previous cases as a reference in diagnosing future cases, while the certainty factor requires a weighted value given by the expert. The test results from this study get the percentage value of the confidence level with the certainty factor method of 86%. Other related research is an expert system model for diagnosing COVID-19 using the nave Bayes classifier using three categories, namely PDP, ODP and OTG. The test results of experimental cases calculated using Nave Bayes resulted in a PDP diagnosis with a value of 0.3400 or 34%, the nave Bayes classifier method was successfully implemented for the PDP, ODP and OTG categories [26]. Other research related to the author's research conducted by [27] The expert system for early diagnosis of COVID-19 which was built using the Fuzzy Expert System was successfully used in early diagnosis of COVID-19 with positive and non-positive diagnostic results, making it easier for doctors to diagnose the symptoms experienced.

Based on the results of previous studies, no one has investigated the problem of comparing the Naive Bayes and Dempster Shafer methods in an expert system for early diagnosis of COVID-19. Research conducted by [28] using the Fuzzy Expert System with positive and non-positive diagnostic results. Meanwhile, the research that the author conducted used as many as 17 data on COVID-19 symptoms with the results of the diagnosis of Suspect, Non-Suspect and Close Contact, so that the differences with previous studies were visible. The purpose of this research is to build an expert system that is able to diagnose COVID-19 early, so that it will help and make it easier for the community to diagnose COVID-19 early.

As for the problem in this study, because the COVID-19 outbreak is very troubling to the public, the symptoms experienced by COVID-19 sufferers are fairly common symptoms that are commonly experienced such as coughs and fever making it difficult for ordinary people to realize the presence of disease in people around or even themselves. . Therefore, researchers are interested in building an expert system that is able to diagnose systematically by comparing the performance of the Nave Bayes and Damster Shafer methods, so as to determine the accuracy of each method and the differences between the two methods to produce a better expert system.

The outputs of the expert system for early diagnosis of COVID-19 are Suspects, Non-Suspects, and Close Contacts. The variables used in this study were symptoms related to COVID-19 based on the COVID-19 SOP Management Manual Edition 5 of 2020 and the results of interviews with doctors as follows: fever / history of fever, dry cough, weakness, dizziness, headache, pain throat, runny nose/stuffy nose, shortness of breath, nausea, chest pain, vomiting, diarrhea, impaired smell, taste disturbance, history of close contact with confirmed COVID-19 patients, history of travel to cities infected with COVID-19 and medical treatment of COVID-19 patients 19. The purpose of this study is to build an expert system that is able to diagnose COVID-19 early by using the Nave Bayes and Dempster Shafer methods. With this expert system, it can be more efficient in conducting consultations, the medical party can diagnose patients and carry out treatment as needed effectively and efficiently and to find out the results of the comparison of the Nave Bayes and Dempster Shafer methods in the expert system for diagnosing COVID-19. In addition to this research, there are other studies that the author has done related to COVID-19, including clustering the spread of COVID-19 in Aceh Province using the Fuzzy C-Means algorithm [28].

2. RESEARCH METHOD

2.1. Data Sources and Research Variables

This research was conducted at the Langsa City Hospital which is a COVID-19 referral hospital as the object of data collection. The amount of data taken and used in this study was 550 data, consisting of 500 training data and 50 testing data. Meanwhile, the variables used in this study are symptoms related to COVID-19 based on the COVID-19 SOP Management Manual Edition 5 of 2020 and the results of interviews with doctors, with symptom data of 17 symptoms as follows: G01 (fever / history of fever) , G02 (dry cough), G03 (weakness), G04 (dizziness), G05 (headache), G06 (throat pain), G07 (runny nose/stuffy nose), G08 (shortness of breath), G09 (nausea), G10 (chest pain), G11 (vomiting), G12 (diarrhea), G13 (olfactory disturbances), G14 (taste disorders), G15 (history of close contact with confirmed Covid-19 patients), G16 (travel history to cities affected by Covid-19) , G17 (carrying out medical care for Covid-19 patients). While the diagnostic data consists of Suspect, Non-Suspected, and Close Contact.

2.2. Naive Bayes and Dempster Shafer method

The following are the steps for completing the nave Bayes method: The first step is to find the Prior Probability value, the Prior Probability calculation is a calculation to find the probability of the previous occurrence of a case against the entire case used, namely the case on the test data, using equation (1) - (5):

$$P(h) = \frac{\Sigma \text{ incident of a case}}{Allcases} \tag{1}$$

The second step, for a case of symptoms that are indeed one of the symptoms in a case, with the equation:

$$P(e \mid h) = \frac{P(e \cap h)}{P(h)} \tag{2}$$

The third step is to find the Posterior value. Posterior value is the final probability, how to calculate it by multiplying the Prior Probability value of a case with the Likelihood value of each symptom in each case experienced by the patient, with the equation:

$$P(h \mid e) = P(h) * P(e_1, e_2, e_3, \ldots \mid h) \tag{3}$$

While the Dempster Shafer method is a mathematical theory to combine facts. Plausability can be denoted as follows [14, 19]:

$$PI(X) = 1 - Bel(X) \tag{4}$$

Where Bel (X) is a measure of the strength of evidence (evidence), if it is worth 0 then it identifies that there is no evidence and if it is worth 1 it indicates certainty. Similar to Belief, Plausability Pl (X) is worth 0 to 1. Plausabilty will reduce the level of trust and evidence. If there are 2 or more elements, a combination function of M1 and M2 will be formed as M3 with the following formula:

$$M3(Z) = \frac{\Sigma_{X \cap Y = Z} M1(X).M2(Y)}{1 - \Sigma_{X \cap Y = \phi} M1(X).M2(Y)} \tag{5}$$

2.3. System Schematic

The comparison system scheme for the Naive Bayes and Dempster Shafer methods on an expert system for early diagnosis of COVID-19 is shown in Figure 1. The system schema in Figure 1 is a schematic that describes the overall system design process as follows: the first step is initialization to start the application. After starting the program, the admin/user will be asked to input data from the patient to be diagnosed, then the admin/user will choose what symptoms the user is experiencing, possible symptoms related to covid-19 are available and admin/user just choose the symptoms experienced, the next step there is a process symbol after inputting data the user will choose the calculation method used, namely nave bayes and dempster shafer then the system will start processing according to the selected method, after doing the calculation, the system will display the output in the form of results calculation in the form of P01 (Suspected), P02 (Non Suspected) and P03 (Close Contact).

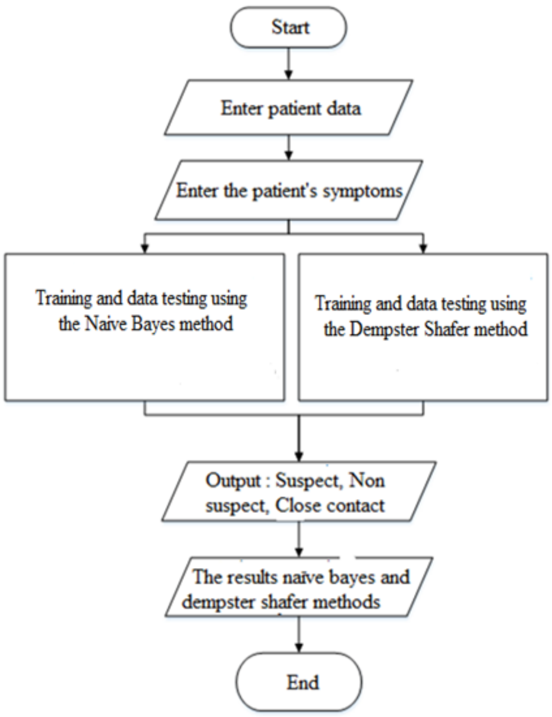

Figure 1. System Schematic

3. RESULT AND ANALYSIS

3.1. Data pre-processing and data processing with Naive Bayes method

The initial stage of the process of applying the nave Bayes method is the retrieval of training data, the data used is data from previous events, namely data on patients experiencing COVID-19 symptoms obtained at the COVID-19 referral General Hospital in Langsa City. The amount of data taken is 500 training data (training) and 50 testing data (testing) with 17 symptoms and 3 disease diagnosis categories, namely: Suspected, Non-Suspected and Close Contacts who act as training data. Diagnostic data are listed in Table 1.

Table 1. COVID-19 diagnostic data

No.	Code	Name
1	P01	suspect
2	P02	Non - Suspect
3	P03	Close Contact

While the symptom data for early diagnosis of COVID-19 is in Table 2.

Table 2. Data on COVID-19 symptoms

No .	Code	Name
1	G01	Fever $>$ 38 / History of fever $>$14 days
2	G02	Dry Cough
3	G03	Weak
4	G04	Dizzy
5	G05	Headache
6	G06	Sore Throat
7	G07	Runny nose / stuffy nose
8	G08	Out of breath
9	G09	Nauseous
10	G10	Chest pain
11	G11	Vomit
12	G12	Diarrhea
13	G13	Olfactory Disorder
14	G14	Taste Disorder
15	G15	History of Close Contact with Patients Confirmed
16	G16	Travel History to the Infected City
17	G17	Performing Medical Treatment with COVID-19 Confirmed Patients

The following 500 training data (training) that will be used as a reference for calculating the nave Bayes method are listed in Table 3.

Table 3. Naive Bayes method training data

No.	Symptom Code																	Diagnosis
	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	G13	G14	G15	G16	G17	
1				√	√											√		√
2				√	√		√		√									√
3			√	√			√				√					√		√
4															√	√		√
5	√	√		√	√		√										√	√
6			√		√											√		√
7	√	√	√		√			√										√
8	√		√		√		√											√
9	√	√					√									√		√
10	√										√				√			√
11	√											√			√			√
12	√														√			√
13	√				√		√											√
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
500		√														√		√

Patient 1 experienced symptoms of G01, G08 and G13, the solution using the nave Bayes method was as follows:

a. Finding the Prior Probability value using the equation (1)

$$P(P01) = \frac{220}{500} = 0,440$$

$$P(P02) = \frac{218}{500} = 0,436$$

$$P(P03) = \frac{62}{500} = 0,124$$

b. Finding the Likelihood value using the equation (2)

$$P(G01 \mid P01) = \frac{182}{220} = 0,8273$$
$$P(G01 \mid P02) = \frac{76}{218} = 0,3486$$
$$P(G01 \mid P03) = \frac{19}{62} = 0,3065$$

$$
\begin{aligned}
P(G08 \mid P01) &= \frac{109}{220} = 0,4955 \\
P(G08 \mid P02) &= \frac{58}{218} = 0,2661 \\
P(G08 \mid P03) &= \frac{1}{62} = 0,0161 \\
P(G013 \mid P01) &= \frac{41}{220} = 0,1864 \\
P(G013 \mid P02) &= \frac{0}{218} = 0 \\
P(G013 \mid P03) &= \frac{0}{62} = 0
\end{aligned}
$$

c. Finding the Posterior value using the equation (3)

$$
\begin{aligned}
P(P01 \mid e) &= P(P01)^{*}P(G01 \mid P01)^{*}P(G08 \mid P01)^{*}P(G013 \mid P01) \\
&= 0,440^{*}0,8273^{*}0,4955^{*}0,1864 = 0,033610 \\
P(P02 \mid e) &= P(P02)^{*}P(G01 \mid P02)^{*}P(G08 \mid P02)^{*}P(G013 \mid P02) \\
&= 0,436^{*}0,3486^{*}0,2661^{*}0 = 0 \\
P(P03 \mid e) &= P(P03)^{*}P(G01 \mid P03)^{*}P(G08 \mid P03)^{*}P(G013 \mid P03) \\
&= 0,124^{*}0,3065^{*}0,0161^{*}0 = 0
\end{aligned}
$$

So the case of patient 1, who experienced symptoms of fever (G01), shortness of breath (G08) and impaired smell (G13) was included in the suspect patient (P01) with the highest probability value of 0.03361.

3.2. Data pre-processing and data processing with the Dempster Shafer Method

The Dempster Shafer method is a method that uses the belief function (belief function), Belief (Bel) is a measure of the strength of evidence (evidence) in supporting a set of propositions, if it is 0 then it identifies that there is no evidence and if it is 1, it indicates certainty. Trust value is the value that the expert believes in a symptom, in the form of the weight given by the expert to a symptom. The first step is to interview experts to get the confidence value of each symptom. The following is the data obtained in the form of trust values from experts in Table 4.

Table 4. Dempster Shafer data

No	Symptom Name	Diagnosis			bilief value	Uncertainty Value
		P01	P02	P03		
1	Fever > 38 / History of fever > 14 days	√	√		0,8	0,2
2	Dry Cough	√	√		0,8	0,2
3	Weak	√	√	√	0,6	0,4
4	Dizzy	√	√	√	0,6	0,4
5	Headache	√	√	√	0,6	0,4
6	Sore Throat	√	√	√	0,7	0,3
7	Runny nose / stuffy nose	√	√	√	0,7	0,3
8	Out of breath	√	√		0,8	0,2
9	Nauseous	√	√	√	0,6	0,4
10	Chest pain	√	√	√	0,5	0,5
11	Vomit	√	√	√	0,6	0,4
12	Diarrhea	√	√	√	0,7	0,3
13	Olfactory Disorder	√			1,0	0,0
14	Taste Disorder	√			1,0	0,0
15	History of Close Contact with Patients Confirmed	√		√	1,0	0,0
16	Travel History to the Infected City	√		√	1,0	0,0
17	Performing Medical Treatment with COVID-19 Confirmed Patients	√		√	1,0	0,0

Sample Data Patient 1 experienced symptoms of G01, G08 and G13. The solution using the Dempster Shafer method is as follows:

a. G01 (fever)
   Belief Value M1 $(P01, P02)$ = $0,8$
   Plausibility Value M1 (θ) = $0,2$
b. G08 (out of breath)
   Belief Value M2 $(P01, P02)$ = $0,8$
   Nilai Plausibility M2 (θ) = $0,2$

Then find the new Density M3 using the equation (5)

Table 5. Density M3

	M2 (P01, P02) 0,8	M2 (θ) 0,2
M1 (P01,P02) 0,8	(P01, P02) 0,64	(P01,P02) 0,16
M1 (θ) 0,2	(P01, P02) 0,16	(θ) 0,04

Then the density of M3 is:

$$M3(P01, P02) = \frac{0,64 + 0,16 + 0,16}{1 - 0} = 0,96$$
$$M3\theta = \frac{0,04}{1 - 0} = 0,04$$

c. G13 (olfactory disorder)
   Belief Value M4 $(P01)$ = 1
   Nilai Plausibility M4 (θ) = 0

Then find the new density M5 using equation (5)

Table 6. Density M5

	M4 (P01) 1	M4 (θ) 0
M3 (P01,P02) 0,96	(P01) 0,96	(P01,P02) 0
M3 (θ) 0,04	(P01) 0,04	(θ) 0

Then the density of M5 is:

$$M5(P01) = \frac{0,96 + 0,04}{1 - 0} = 1$$
$$M5(P01, P02) = \frac{0}{1 - 0} = 0$$
$$M5(\theta) = \frac{0}{1 - 0} = 0$$

Then the calculation using the Dempster Shafer method on patient data 1 who has symptoms of fever (G01), shortness of breath (G08) and impaired smell (G13) is included in the suspect patient (P01) with the highest value of 1.

3.3. Testing the Naive Bayes and Dempster Shafer Methods

This study uses 50 test data, then 50 test data will be processed using the Nave Bayes method and the Dempster Shafer method. The test results of the two methods will be compared with the answers of the experts. The results of testing 50 test data are in Table 7.

Table 7. Data on the results of testing the Nave Bayes method and the Dempster Shafer method

No	Symptom	Dempster Shafer	Nave Bayes	Expert Diagnosis
1	Fever, dry cough, weakness, headache, sore throat, impaired smell, impaired taste.	P01	P01	P01
2	Dry cough, headache, shortness of breath, history of close contact with confirmed cases of COVID-19.	P01	P01	P01
3	Headache and dizziness	P01, P02, P03	P02	P02
4	Nausea, chest pain, vomiting.	P01, P02, P03	P02	P02
5	History of close contact with confirmed cases of COVID-19	P01, P03	P03	P03
6	Performing medical care for COVID-19 positive patients	P01, P03	P01	P02
7	Fever, dry cough, weakness, dizziness, nausea, vomit	P01, P02	P01	P01
8	Weakness, headache	P01, P02, P03	P02	P02
9	Fever, dry cough, weakness, sore throat, impaired olfaction	P01	P01	P01
10	Dry cough, weakness, colds, shortness of breath, history of close contact with confirmed COVID-19 patients.	P01	P01	P01
11	Fever, shortness of breath, impaired smell	P01	P01	P01
12	Weakness, dizziness, vomiting.	P01, P02, P03	P02	P02
13	Nausea, vomiting, diarrhea.	P01, P02, P03	P02	P02
14	Dry cough, weakness, sore throat, diarrhea.	P01, P02	P01	P02
15	Weakness, nausea, vomiting, diarrhea.	P01, P02, P03	P02	P02
16	Olfactory disorders, taste disorders, performing medical treatment for covid positive patients.	P01	P01	P01
17	Weakness, colds, diarrhea	P01, P02, P03	P02	P02
18	Cough, shortness of breath, travel history to infected cities.	P01	P01	P01
19	Dry cough, cold, sore throat, history close contact with a confirmed COVID-19 patient.	P01	P01	P01
20	Weakness, dizziness, headache	P01, P02, P03	P02	P02
21	Headache, nausea, chest pain, vomiting.	P01, P02, P03	P02	P02
22	Fever, shortness of breath, chest pain.	P01, P02	P01	P01
23	Dry cough, cold, olfactory disorders.	P01	P01	P01
24	Fever, shortness of breath, history of close contact with confirmed COVID-19 patients	P01	P01	P01
25	Fever, dry cough, shortness of breath	P01, P02	P01	P01
26	Fever, dry cough, sore throat, cold, history of close contact with confirmed patient.	P01	P01	P01
27	Dizziness, headache, chest pain	P01, P02, P03	P02	P02
28	Dizziness, nausea, chest pain, vomiting.	P01, P02, P03	P02	P02
29	Fever, dry cough, cold, shortness of breath	P01, P02	P01	P01
30	Weakness, nausea, chest pain.	P01, P02, P03	P02	P02
31	Fever, dry cough, performing medical treatment for COVID-19 positive patients.	P01	P01	P01
32	Chest pain, vomiting, diarrhea	P01,P02,P03	P02	P02
33	Fever, dry cough, weakness, shortness of breath.	P01, P02	P01	P01
34	Fever, dry cough, cold, impaired olfaction	P01	P01	P01
35	Weakness, diarrhea.	P01, P02, P03	P02	P02
36	Olfactory disorders, taste disorders.	P01	P01	P01
37	Weakness, dizziness, nausea, vomiting.	P01, P02, P03	P02	P02
38	Fever, dry cough, sore throat, cold.	P01, P02	P01	P01
39	Fever, dry cough, runny nose, history of close contact with confirmed COVID-19 patients.	P01	P01	P01
40	Colds, nausea, chest pain.	P01, P02, P03	P02	P02
41	Weakness, colds, nausea, vomiting.	P01, P02, P03	P02	P02
42	Fever, dry cough, weakness, dizziness, shortness of breath	P01, P02	P01	P01
43	Weakness, dizziness, sore throat, nausea.	P01, P02 ,P03	P02	P02
44	Weakness, impaired olfaction, impaired taste.	P01	P01	P01
45	Weakness, nausea, diarrhea.	P01,P02,P03	P02	P02

No	Symptom	Dempster Shafer	Nave Bayes	Expert Diagnosis
46	Weakness, olfactory disorders, taste disorders, performing medical treatment for covid positive patients.	P01	P01	P01
47	Fever, colds, shortness of breath, history of traveling to the city affected.	P01	P01	P01
48	Weakness, nausea, vomiting.	P01, P02, P03	P02	P02
49	Fever, colds, olfactory disorders, travel history.	P01	P01	P01
50	Cough, sore throat, cold, impaired olfaction.	P01	P01	P01

1. Naive Bayes Method Test Results

In testing using 50 test data in calculations using the nave Bayes method through the system and through manual calculations, it produces 1 Close Contact (P03), 20 Non-Suspected (P02) and 29 Suspected (P01). The following is a graph that visually displays the test results of 50 test data in Figure 2.

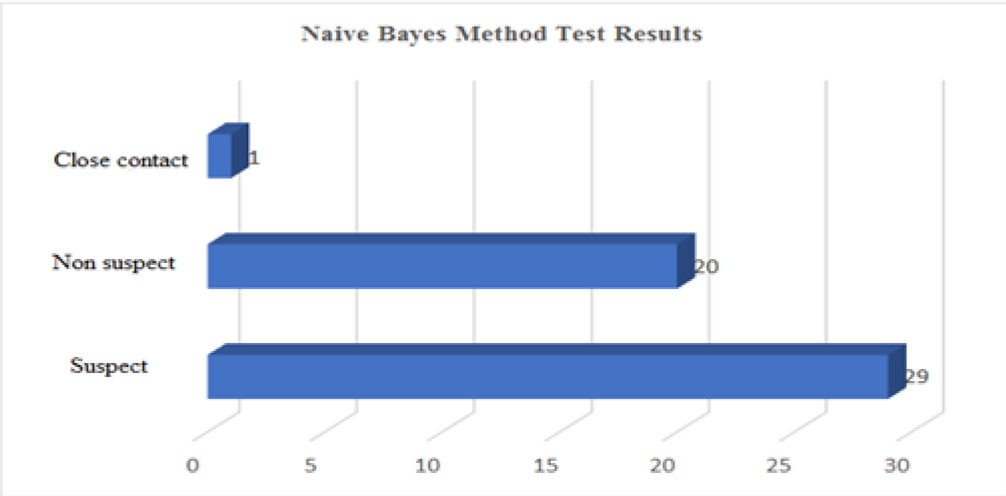

Figure 2. Graph of test calculations using the Naive Bayes method

The results of the test data are then tabulated in an equation, then calculate the accuracy of the nave Bayes method using equation (6)

$$\text{Accuracy value} = \frac{\Sigma \text{ of data matched}}{\Sigma \text{ of data}} \times 100\% \tag{6}$$

$$\text{Nave Bayes Method Accuracy} = \frac{48}{50} \times 100\% = 96\%$$

The following is a graph of the accuracy of the nave Bayes method of 96%, with 48 test data according to the calculation of the nave Bayes method of testing with expert diagnoses (experts) and 4%, with 2 test data that have different answers from expert diagnoses as in Figure 3. The results of this study indicate that there are differences in the results of the analysis with previous research conducted by [25] by comparing the nave Bayes method and certainty factor in an expert system to detect the COVID-19 virus. The test results from this study obtained a percentage value of 86% confidence level. Other related studies also show differences in the results of the analysis conducted by the [26] expert system model to diagnose COVID-19 using the nave Bayes classifier using three categories, namely PDP, ODP and OTG. The test results of experimental cases calculated using Nave Bayes resulted in a PDP diagnosis with a value of 0.3400 or 34%. The nave Bayes classifier method was successfully implemented for the PDP, ODP and OTG categories.

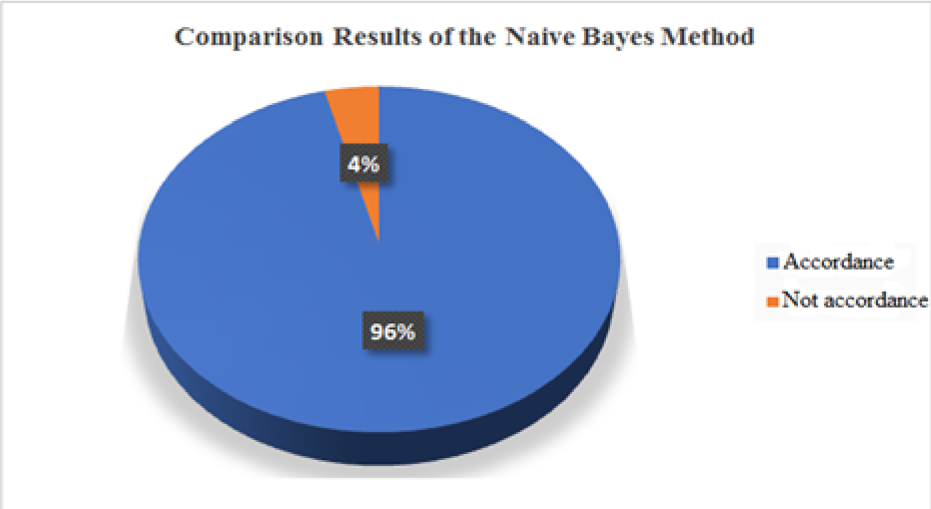

Figure 3. Graph of the Comparison of the Naive Bayes Method

2. Dempster Shafer Method Test Results

Of the 50 test data carried out using the system and using manual calculations with the dempster shafer method, 20 diagnoses were suspected (P01), 8 suspected/non-suspected (P01/P02), 2 suspected/close contact (P01/P03) and 20 suspected/ Non Suspect/Close Contact (P01/P02/P03). The following is a graph that visually displays the test results of 50 test data with dempster shafts in Figure 4.

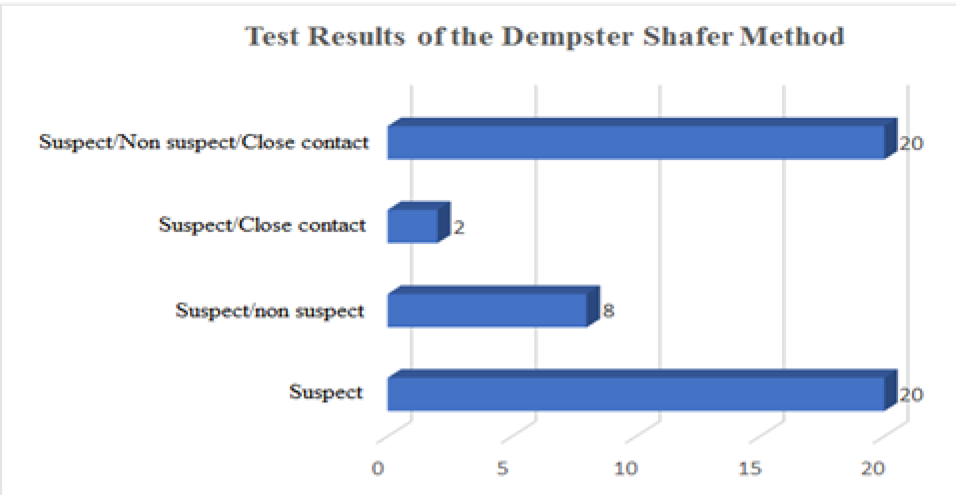

Figure 4. Graph of test calculations using the dempster shafer method

The results of the test data are then tabulated in an equation, then calculate the accuracy of the dempster shafer method using equation (7)

$$\text{Accuracy value } = \frac{\Sigma \text{ of data matched}}{\Sigma \text{ of data}} \times 100\% \tag{7}$$

$$\text{Dempster Shafer Method Accuracy} = \frac{20}{50} \times 100\% = 40\%$$

The following graph shows the accuracy of the Dempster Shafer method by 40%, with 20 test data that match the calculation of the Dempster Shafer method with expert diagnoses (experts) and 60%, with 30 test data that have different answers from expert diagnoses as shown in Figure 5.

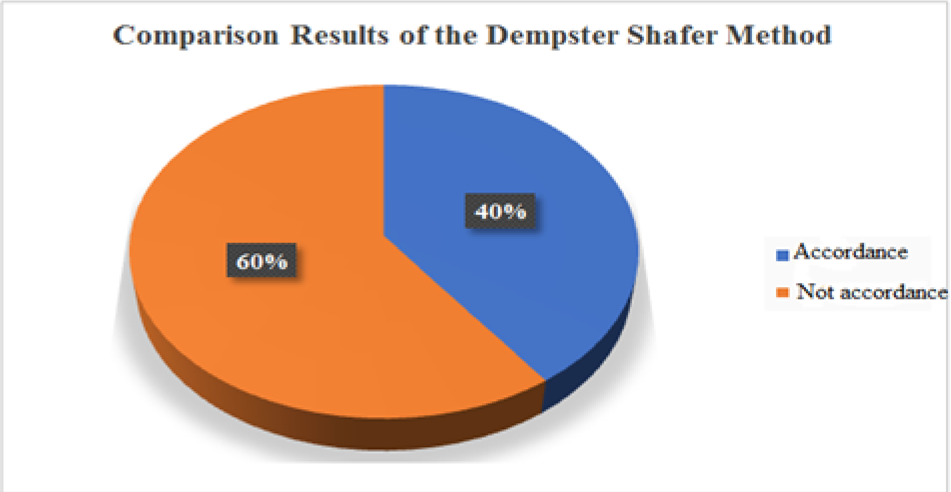

Figure 5. Graph of Comparative Results of the Dempster Shafer Method

Based on the results of this study, the application of the nave bayes method and the dempster shafer method to an expert system for early diagnosis of COVID-19 uses 50 test data. The results of the analysis show that the nave Bayes method has an accuracy of 96%, which is 48 data according to expert diagnoses. While the dempster shafer method has an accuracy of 40%, because as many as 20 data are in accordance with expert diagnoses.

4. CONCLUSION

Based on the results of the research, the Nave Bayes method and the Dempster Shafer method can be used and applied properly to an expert system for early diagnosis of COVID-19. The amount of data used in this study was 550 data, consisting of 500 training data and 50 testing data. While the variables used are symptoms related to COVID-19 as many as 17 symptoms consisting of G01 (fever / history of fever), G02 (dry cough), G03 (weakness), G04 (dizziness), G05 (headache), G06 (sore throat), G07 (runny nose/stuffy nose), G08 (shortness of breath), G09 (nausea), G10 (chest pain), G11 (vomiting), G12 (diarrhea), G13 (impaired smell), G14 (impaired taste) , G15 (history of close contact with confirmed Covid-19 patients), G16 (history of travel to cities affected by Covid-19), G17 (carrying out medical care for Covid-19 patients). The diagnostic data consists of suspects, non-suspects, and close contacts. The results of the test using 50 test data using the nave Bayes method through the application system and through manual calculation testing resulted in 1 Close Contact (P03), 20 Non-Suspected (P02) and 29 Suspected (P01). While the test results using the application system and using manual calculation testing with the dempster shafer method resulted in 20 diagnoses of suspect (P01), 8 suspected/non-suspected (P01/P02), 2 suspected/close contact (P01/P03) and 20 suspected/ Non Suspect/Close Contact (P01/P02/P03). From the results of the percentage test by comparing system diagnoses with expert diagnoses, the nave Bayes method has an accuracy of 96% with 48 diagnoses according to expert diagnoses from 50 tested data. As for the Dempster Shafer method, it has an accuracy of 40% with 20 diagnoses according to expert diagnoses from 50 tested data.

Further research is recommended to use other methods to find out the most effective method and the level of accuracy in this expert system. This system can be further developed in the form of an Android application, making it easier for the general public to diagnose COVID-19 early.

5. ACKNOWLEDGEMENTS

The Acknowledgments section is optional. Research sources can be included in this section.

6. DECLARATIONS

AUTHOR CONTIBUTION

Author 1: Constructing an idea or hypothesis for research and manuscrip, create and compile literature reviews, planning methodology to reach the conclusion and wrote the paper. Author 2: colected the data, performed the analysis and making expert system software and coding. Author 3: make a system schematic design, implementation and testing the system. Author 4: Reviewing the article before submission not only for spelling and grammar but also for its intellectual content. All authors read and approved the final manuscript.

FUNDING STATEMENT

This research is independent research with funding from the researchers themselves.

COMPETING INTEREST

The author declares that it has no competing interests.

REFERENCES

[1] N. Fernandes, "Economic Effects of Coronavirus Outbreak (COVID-19) on The World Economy," Tech. Rep., 2020.

[2] P. K. Ozili and T. Arun, "Spillover of COVID-19: Impact on The Global Economy," in *Managing Inflation and Supply Chain Disruptions in the Global Economy*. IGI GLOBAL, 2022, pp. 1–29.

[3] M. Vadiati, A. Beynaghi, P. Bhattacharya, E. R. Bandala, and M. Mozafari, "Indirect Effects of COVID-19 on The Environment : How Deep and How Long ?" *Science of the Total Environment*, vol. 810, no. March, pp. 1–5, 2022.

[4] M. Akat and K. Karatas, "Psychological Effects of Covid-19 Pandemic on Society and Its Reflections on Education," *Turkish Studies*, vol. 15, no. 4, pp. 1–13, 2020.

[5] Nurdin, U. M. P. Nasution, H. A.-k. Aidilof, and Bustami, "Implementasi Fuzzy C-Means untuk Menentukan Tingkat Kepuasan Mahasiswa dalam Pembelajaran Online Implementation of Fuzzy C-Means to Determine Student Satisfaction Levels," *Jurnal Sistemasi*, vol. 11, no. 1, pp. 121–136, 2022.

[6] A. Susilo, C. M. Rumende, C. W. Pitoyo, W. D. Santoso, M. Yulianti, Herikurniawan, R. Sinto, G. Singh, L. Nainggolan, E. J. Nelwan, L. K. Chen, A. Widhani, E. Wijaya, Wicaksana, Bramantya, M. Maksum, F. Annisa, C. O. Jasirwan, and Y. Evy, "Coronavirus Disease 2019 : Tinjauan Literatur Terkini Coronavirus Disease 2019 : Review of Current Literatures," *Jurnal Penyakit Dalam Indonesia*, vol. 7, no. 1, pp. 45–67, 2020.

[7] S. R. Dubey, "Corona Virus," *International Journal of Nursing Education and Research*, vol. 8, no. 3, pp. 411–414, 2020.

[8] N. I. Fitriani, "Tinjauan Pustaka Covid-19; Virologi, Patogenesis, Dan Manivestasi Klinis," *Journal Medika Malahayati*, vol. 21, no. 1, pp. 1–9, 2020.

[9] J. Irawan, *Buku Pegangan Kuliah Sistem Pakar*. Surabaya: Sekolah Tinggi Manajemen Informatika & Teknik Komputer Surabaya (STIKOM), 2007.

[10] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creative Information Technology Journal*, vol. 2, no. 3, pp. 207–217, 2015.

[11] Zulvandri, "Sistem Pakar Untuk Diagnosa Penyakit Mata Menggunakan Pendekatan Metode Naive Bayes Berbasis Web," Ph.D. dissertation, 2013.

[12] M. T. Kalimatullah, S. Adi Wibowo, and K. Auliasari, "Analisa Perbandingan Metode Certainty Factor Dengan Dempster Shafer Pada Sistem Pakar Untuk Mendiagnosis Penyakit Balita Usia Kurang Dari 5 Tahun," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 4, no. 1, pp. 118–128, 2020.

[13] L. Septiana, "Metode Dempster-Shafer Untuk Sistem Pakar Deteksi Kerusakan Mesin Cuci Berbasis Web," *Jurnal Techno Nusa Mandiri*, vol. 12, no. 2, pp. 137–146, 2015.

[14] R. Ristandi, N. Hidayat, and M. A. Fauzi, “Pemodelan Sistem Pakar Diagnosa Hama-Penyakit pada Tanaman Bawang Merah Menggunakan Metode Naïve Bayes,” Ph.D. dissertation, 2016.

[15] Budianto, I. Fitri, and Winarsih, “Expert System for Early Detection of Disease in Corn Plant Using Naive Bayes Method,” *Jurnal Mantik*, vol. 3, no. 4, pp. 308–317, 2020.

[16] R. S. Wahyuningtyas, Tursina, and H. S. Pratiwi, “Sistem Pakar Penentuan Jenis Kulit Wajah Wanita Menggunakan Metode Naïve Bayes,” *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 4, no. 1, pp. 27–32, 2015.

[17] Chairun Nas, “Sistem Pakar Diagnosa Penyakt Tiroid Menggunakan Metode Dempster Shafer,” *Jurnal Teknologi Dan Open Source*, vol. 2, no. 1, pp. 1–14, 2019.

[18] E. K. Panggabean, “Comparative Analysis Of Dempster Shafer Method With Certainty Factor Method For Diagnose Stroke Diseases,” *International Journal of Artificial Intelligence Research*, vol. 2, no. 1, pp. 37 – 41, mar 2018.

[19] M. Ihsan, F. Agus, and D. M. Khairina, “Penerapan Metode Dempster Shafer Untuk Sistem Deteksi Penyakit Tanaman Padi,” in *Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi (SAKTI)*, vol. 2, no. 1, 2017, pp. 128–135.

[20] M. L. C. Buono, N. Pandiangan, and S. H. D. Loppies, “The Implementation Of An Expert System In Diagnosing Skin Diseases Using The Dempster-Shafer Method,” *Journal of Physics: Conference Series*, vol. 1569, no. 2, pp. 1–11, jul 2020.

[21] R. Kurniawan, M. Z. A. Nazri, S. N. H. S. Abdullah, Z. A. Othman, and S. Abdullah, “Bayesian network and dempster-shafer theory for early diagnosis of eye diseases,” *Compusoft An International Journal of Advanced computer teknologi*, vol. 9, no. 4, pp. 3642–3651, 2020.

[22] Aristoteles, K. Adhianto, R. Andrian, and Y. N. Sari, “Comparative Analysis of Cow Disease Diagnosis Expert System Using Bayesian Network and Dempster-Shafer Method,” *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 4, pp. 227–235, 2019.

[23] M. F. Suryana, F. Fauziah, and R. T. K. Sari, “Implementasi Sistem Pakar Menggunakan Metode Certainty Factor Untuk Mendiagnosa Dini Corona Virus Desease (COVID-19),” *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, pp. 559–566, jul 2020.

[24] H. Nahumury, A. Mulyani, and H. Nurdin, “Sistem Pendukung Keputusan Mendiagnosa Penyakit Virus Corona (Covid-19) Menggunakan Metode Dempster-Shafer JISAMAR (Journal of Information System , Applied , Management , Accounting and Researh) p-ISSN : 2598-8700 (Printed) JISAMAR (Journal of Info,” *JISAMAR, Jurnal Of Information System, Applied, Management, Accounting and Research*, vol. 4, no. 4, pp. 207–214, 2020.

[25] R. A. Dzahabi Yunas, A. Triayudi, and I. D. Sholihati, “Implementasi Sistem Pakar untuk Mendeteksi Virus Covid-19 dengan Perbandingan Metode Naïve Bayes dan Certainty Factor,” *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 5, no. 3, pp. 338–345, jul 2021.

[26] D. Silahudin, Henderi, and A. Holidin, “Model Expert System for Diagnosis of Covid-19 Using Naïve Bayes Classifier,” *IOP Conference Series: Materials Science and Engineering*, vol. 1007, no. 1, pp. 1–7, dec 2020.

[27] S. D. S. Dass, F. Meskaran, and M. Saeedi, “Expert System for Early Diagnosis of Covid-19,” *International Journal of Current Research and Review*, vol. 12, no. 22, pp. 162–165, 2020.

[28] N. Nurdin, M. Suhendri, Y. Afrilia, and R. Rizal, “Klasifikasi Karya Ilmiah (Tugas Akhir) Mahasiswa Menggunakan Metode Naive Bayes Classifier (NBC),” *Jurnal SISTEMASI*, vol. 10, no. 2, pp. 268–279, may 2021.