Implementation of The Naïve Bayes Method in the COVID-19 Self-Assessment of Cianjur Regency Government Officials

Muhammad Nasir¹*, Andy Ramadhany Rahayu², I Putu Robin Sunjaya³, Mey Sri Widiastari⁴, Agus Prayitno⁵

¹,³,⁴,⁵ Program Magister Sistem Informasi, STMIK LIKMI Jl. Ir. H. Juanda No.96, Kota Bandung, 40132, Indonesia
² Sekretariat DPRD Kabupaten Cianjur, Jl. KH. Abdullah Bin Nuh, 43215, Indonesia

Email: siap86.nsr@gmail.com¹, andyramadhany28@gmail.com², iputurobins@gmail.com³, meysiwi15@gmail.com⁴, masagusdotid@gmail.com⁵
* corresponding author

Article history:
Received: Feb 24, 2022
Revised: Mar 24, 2022
Accepted: Mar 29, 2022
Available online Mar 30, 2022

Keywords:
Expert System;
Work From Home;
Naïve Bayes;
COVID-19;
Cianjur Regency.

ABSTRACT
The impact of Covid-19 in Indonesia has penetrated into all fields of human activity including in the government sector, efforts to implement work from home for government agencies, especially in the Cianjur district to suppress the positive number of COVID-19 have been carried out. However, in practice the determination of employees to work from home is not appropriate, resulting in a decrease in the performance of government employees in Cianjur Regency, and an increase in positive numbers in the government environment. The method used in this research is an expert system approach with Naïve Bayes which is the fastest and most accurate classification method for determining the problem. Based on the classification of the Naïve Bayes method, samples were taken from Cianjur Regency government employees with symptoms of fever, cough, muscle aches, and loss of sense of smell, they had the highest probability of being classified as unhealthy and eligible for a swab test compared to other classifications, which was 80% percent. An expert system with a naive Bayes approach can be implemented to determine the health status of Cianjur Regency employees related to Covid-19, the suitability of the swab test, and the determination to work from home. For further research, it is suggested that it can be integrated with the existing institution’s attendance system, and if necessary it can be tested with other methods.

© 2022 JTI.C.I.T. All rights reserved.

1. Introduction
Coronavirus 2019 (COVID-19) is an infectious disease caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), which is a novel virus that has never been identified in humans. There are at least two types of coronaviruses known to cause diseases with severe symptoms such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS)[1][2][3][4]. Covid-19 symptoms are mostly non-specific (uncertain) symptoms like fever, cough, myalgia (muscle pain), anosmia, and even relatively without any symptoms[5] [6][7]. The average incubation period is about 5-6 days with the longest incubation period being 14 days [8][9][10]. In severe cases of COVID-19, it can cause pneumonia, acute respiratory syndrome, kidney failure, and even death[11][12][13][14].

As one of the countries affected by COVID-19, Indonesia has put in place a number of measures to prevent the virus from spreading throughout the country. As of July 29, there were 197,126,377 positive confirmed cases of COVID-19 worldwide. At least 0.005543 percent of confirmed cases in Cianjur, West Java [15]. Government policies, such as large-scale social restrictions, lockdown, and most recently, Enforcement of Community Activity Restrictions or known as PPKM, have also been socialized and...
implemented for all aspects of social life, particularly in activities and work processes in offices, agencies, factories, and business entities that have recently implemented work from home (WFH). However, during its implementation, it became a boomerang for the government, particularly in the Cianjur Regency Government, where employees in Cianjur Regency who carried out WFH (Work From Home) did not provide maximum results, but there was an increase in the number of positive confirmations, resulting in a decrease in the performance employees in the Cianjur Regency Environment. Government Cianjur Regency has at least 57 Regional Apparatus Organizations (OPD) with various sectors ranging from Government, Health Office, Industry and Trade, Fisheries, and others[16]. As reported by [17] at the beginning of 2022, there was a spike in the Omicron variant from one of the Covid-19 variants, which resulted in a re-appeal for WFH in all sectors, including the government. Likewise in Cianjur, as reported by [18], ASN officials from the Cianjur district government who were exposed to the Omicron variant made a new rule to carry out WFH again. One of the reasons for the increase in the number of positive COVID-19 patients is the Indonesian people’s lack of awareness of the application of large-scale social restrictions, lockdown, and the lack of awareness of early detection of COVID-19 symptoms, which has led to a wider spread of the virus, particularly in the government sector. In dealing with COVID-19 cases, rapid diagnosis is required to determine the next step. Supporting examinations are one of the tools used to help diagnose COVID-19. PCR swabs, rapid tests, and antigen swabs are examples of laboratory-supported examinations that are currently available[19].

Since 2020, application development has been carried out, but the results have not been seen or utilized optimally by users or consumers. When the COVID outbreak struck, many application innovations were made to support consumer needs by presenting various features that took into account security issues, price, capacity, speed, reliability (various OS and devices), updates, and user convenience[20]. Expert system technology based on the Naive Bayes algorithm is one of the appropriate technologies to be developed to handle the COVID-19 diagnosis. Which is a system that uses an expert’s knowledge to solve a problem and produce a result that is fed into a system. The advantages of expert systems are to increase reliability, reduce errors, reduce costs, use a variety of skills, provide intelligence databases, and reduce hazards. The disadvantage of an expert system is that there is no common sense and no change in the environment[21]. Several previous studies have implemented this system, but it only determines the results of a suspected diagnosis or without further action[22][23]. Other research conducted on [24][25] concluded that it only provided information on the early symptoms of the COVID-19 disease, the research conducted by [26] resulted in a high level of confidence regarding the Covid-19 case and suggested expert testing and Bayesian methods, and research conducted at [27] with the results of his research got the level of accuracy of the certainty factor method more accurate than the Naive Bayes method. For that current research will be carried out re-testing the Naive Bayes method in this covid-19 case. Whereas, the application of this research will produce the health status of each employee in Cianjur Regency and determine whether the employee is eligible for a diagnosis via Antigen Swab/PCR, as well as determine the status of work absences (WFH/WFO) which can be done independently (Self-assessment). Furthermore, the expert system in this study is useful for tracing in the Cianjur Regency civil service area. One doctor from the COVID-19 task force is involved in this expert system, which includes a decision-making system based on the Naive Bayes algorithm.

2. Methods

At this stage, there are research methods and the naive Bayes methods.

2.1 Method of Naïve Bayes

The Naïve Bayes algorithm is one of the algorithms in classification technology that is easy to implement and fast in processing speed [28]. The Naïve Bayes algorithm is based on conditional probability. It uses Bayes Theorem, that formula calculates the probability by calculating the frequency of values and combinations of values in the history of the data. Bayes’ theorem finds the probability of an event occurring given the probability of another event that has occurred [29]. The Naïve Bayes method uses a statistical model to perform the data classification process. This method calculates the probability value of the test data based on case data that has occurred[30].

JTI C.I.T, Vol. 14, No. 1 March 2022: 16-26
Hermanto in [31], that Naïve Bayes works very well compared rate than the other classifier models because it has a better accuracy level. This Bayes method uses a classification based on probability and statistics, if there are two separate events, then the Bayes theorem equation in (1) and the Bayes value taken is the highest percentage of all possible events.

\[ P(A|B) = \frac{P(B|A)}{P(B)} P(A) \]  

Equation description:
- **A**: Hypothesis data is a specific class
- **B**: Data with an unknown class
- **P(A|B)**: The probability of hypothesis A based on condition B (posterior probability)
- **P(A)**: The probability of hypothesis A (prior probability)
- **P(B|A)**: Probability of B based on the conditions on the hypothesis A
- **P(B)**: Probability B

To know how the flow of Naïve Bayes, can be seen in Figure 1 below.

![Naïve Bayes Scheme](image)

The explanation for Figure 1 above is as follows:
- a. Read training data
- b. Calculate numbers and probabilities, but if the data are numbers:
  1) Find the mean and standard deviation of each parameter, which is numerical data.
  2) The probability value is found by calculating the number of appropriate data from the same category divided by the number of data in that category.
- c. Get the values in the Mean, Standard Deviation, and Probability tables.
- d. Then make the solution.

2.2 Research Methods

These Research methods used in research studies can be seen in figure 2. Research Stages. As for techniques for data collection, those were:

- a) Interview, is a data collection is done by asking questions answer or dialogue directly with the parties related to the research conducted. In this study, interviews were conducted with
stakeholders which is in every OPD in Cianjur Regency, as well as an expert from medical science (doctor), dr. Irfan Nur Fauzy, M.Kes. as acting head of the health department Cianjur Regency and the Handling Task Force COVID-19 in Cianjur Regency.

b) Observation, is a collection method data by conducting a direct review to the object under study. In this study look at the document Analysis of the COVID-19 situation in Cianjur Regency.

c) Literature Review, to get data theoretically, the authors collect data by reading and studying books, papers or other references related to the problems discussed.

Fig 2. Research Stages

3. Results and Discussion

3.1 Symptoms Determination and Weighting

To identify the extent of infection with the corona virus in patients, the weight of the symptoms that will be used as an indicator is determined beforehand. These symptoms are selected according to the symptoms experienced. Determination of symptoms is shown in Table 1. List of Symptoms.

<table>
<thead>
<tr>
<th>SYMPTOMS ID</th>
<th>NAME OF SYMPTOMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G01</td>
<td>Fever</td>
</tr>
<tr>
<td>G02</td>
<td>Cough / Cold</td>
</tr>
<tr>
<td>G03</td>
<td>Headache</td>
</tr>
<tr>
<td>G04</td>
<td>Muscleache</td>
</tr>
<tr>
<td>G05</td>
<td>Lost Sense of Taste</td>
</tr>
<tr>
<td>G06</td>
<td>Lost Sense of Smell</td>
</tr>
<tr>
<td>G07</td>
<td>Out of Breath</td>
</tr>
<tr>
<td>G08</td>
<td>Sore Throat</td>
</tr>
<tr>
<td>G09</td>
<td>Diarrhea</td>
</tr>
</tbody>
</table>

Table 1. List Of Symptoms
In determining the health of the OPD employees in Cianjur Regency, several classifications are made that will be used as a reference for determining the results of this expert system, which are listed in Table 2. Classification List.

### Table 2
Classification list

<table>
<thead>
<tr>
<th>CLASSIFICATION ID</th>
<th>NAME OF CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>K01</td>
<td>Healthy</td>
</tr>
<tr>
<td>K02</td>
<td>Unhealthy</td>
</tr>
<tr>
<td>K03</td>
<td>Unhealthy, Eligible for Swab</td>
</tr>
</tbody>
</table>

After determining the classification list and the list of symptoms, the next step is to determine the limiting rules for each symptoms that appears in each classification, the following table is shown the list of rule restrictions.

### Table 3
Limits of role

<table>
<thead>
<tr>
<th>CLASSIFICATION ID</th>
<th>SYMPTOMS ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>K01</td>
<td>G09, G10, G11, G12, G13, G14</td>
</tr>
<tr>
<td>K02</td>
<td>G01, G03, G09, G10, G11, G12, G13, G14</td>
</tr>
<tr>
<td>K03</td>
<td>G01, G02, G03, G04, G05, G06, G07, G08, G09, G10, G11, G12, G13, G14</td>
</tr>
</tbody>
</table>

3.2 **Manual Calculation Using Naïve Bayes**

The first step in calculating Bayes in the expert system is to determine the probability value of the classification and its symptoms. The probability value is determined based on the symptoms filled in by the Cianjur district government officials in the system. After that, the Naïve Bayes calculation is carried out.

The following is an example of a Naïve Bayes calculation for the symptoms filled in by the Cianjur district administration officer who experience symptoms of numbers G01. Fever, G02. Cough/Cold, G04. Muscleache, and G06 for Lost Sense of Smell.

Based on the symptoms experienced by the Cianjur district government officials, the symptoms that appear are in K02 for Healthy and K03 for Unhealthy, Eligible for Swab (see Table 3). Furthermore, the probability calculation for each classification is carried out based on the symptoms mentioned above by referring to equation (1).

**Symptoms description:**
- G01. Fever
- G02. Cough / Cold
- G04. Muscleache
- G06. Lost Sense of Smell

**Probability of K02**

\[
\text{Probability of } K02 = \frac{\text{Probability of appearing classification}}{\text{Sum of all classifications}} = \frac{1}{3} = 0.33
\]

Where 1 is the minimum prediction of the classification that appears and 3 is the sum of all classifications.

Then the probability calculation for all symptoms (G) against K01 is carried out with the following conditions:
Number of symptoms G01 that occurs = 1/3 = 0.33

Where 1 is the number of symptoms G01/G02/G04/G06 that appeared on K01. The same applies to other symptoms, as follows:

G02 = 0/3 = 0.00
G04 = 0/3 = 0.00
G06 = 0/3 = 0.00

In the same way:

**Probability of K03**
K03 = 1/3 = 0.33

Probability G against K03
G01 = 1/3 = 0.33
G02 = 1/3 = 0.33
G04 = 1/3 = 0.33
G06 = 1/3 = 0.33

Next, the Naï ve Bayes calculation is carried out for every K.

**For K02**:  
\[
P(K02|G01) = \frac{P(G01|K02) \times P(K02)}{P(G01|K02) \times P(K02) + P(G01|K03) \times P(K03)}
\]
\[
= \frac{(0.33 \times 0.33)}{(0.33 \times 0.33) + (0.33 \times 0.33)} = 0.5
\]

\[
P(K02|G02) = \frac{P(G02|K02) \times P(K02)}{P(G02|K02) \times P(K02) + P(G02|K03) \times P(K03)}
\]
\[
= \frac{(0.00 \times 0.33)}{(0.00 \times 0.33) + (0.00 \times 0.33)} = 0.00
\]

\[
P(K02|G04) = \frac{P(G04|K02) \times P(K02)}{P(G04|K02) \times P(K02) + P(G04|K03) \times P(K03)}
\]
\[
= \frac{(0.00 \times 0.33)}{(0.00 \times 0.33) + (0.00 \times 0.33)} = 0.00
\]

\[
P(K02|G06) = \frac{P(G06|K02) \times P(K02)}{P(G06|K02) \times P(K02) + P(G06|K03) \times P(K03)}
\]
\[
= \frac{(0.00 \times 0.33)}{(0.00 \times 0.33) + (0.00 \times 0.33)} = 0.00
\]

Then:

**Total of K02**  
\[
P(K02|G01) + P(K02|G02) + P(K02|G04) + P(K02|G06)
\]
\[
= 0.5 + 0.00 + 0.00 + 0.00 = 0.5
\]

**For K03**:  
\[
P(K03|G01) = \frac{P(G01|K03) \times P(K03)}{P(G01|K02) \times P(K02) + P(G01|K03) \times P(K03)}
\]
\[
= \frac{(0.33 \times 0.33)}{(0.33 \times 0.33) + (0.33 \times 0.33)} = 0.5
\]

\[
P(K03|G02) = \frac{P(G02|K03) \times P(K03)}{P(G02|K02) \times P(K02) + P(G02|K03) \times P(K03)}
\]
\[
= \frac{(0.33 \times 0.33)}{(0.33 \times 0.33) + (0.33 \times 0.33)} = 0.5
\]

\[
P(K_03|G_04) = \frac{P(G_04|K_03) \times P(K_03)}{P(G_04|K_02) \times P(K_02) + P(G_04|K_03) \times P(K_03)}
\]

\[
= \frac{(0,33 \times 0.33) + (0,33 \times 0.33)}{(0,33 \times 0.33) + (0,33 \times 0.33)}
\]

\[
= 0.11 / 0.22 = 0.5
\]

And Then:

Total of K_03

\[
P(K_03|G_01) + P(K_03|G_02) + P(K_03|G_04) + P(K_03|G_06)
\]

\[
= 0.5 + 0.5 + 0.5 + 0.5
\]

\[
= 2
\]

Then the total of the two P is added up, then Total Bayes

\[
P = Total \ K_02 + Total \ K_03 = 0.5 + 2 = 2.5
\]

The next step is to calculate the percentage of the predicted value for each classification (P), namely:

\[
P_2 = \frac{(0.5 / 2.5) \times 100}{100} = 20%
\]

\[
P_3 = \frac{(2 / 2.5) \times 100}{100} = 80%
\]

If it is seen from the calculation of the percentage of prediction value, the highest predictive value is P3 which is reached for 80%. This means that the classification of symptoms experienced by the Cianjur district government officials (G01. Fever, G02. Cough/Cold, G04. Muscleache, and G06 for Lost Sense of Smell) is predicted to be Unhealthy, Eligible for Swab.

3.3 Analysis of the Built System

At this stage, the author analyzes users who are divided into 3 types of users of this application, including a) Admin, b) Monitoring and c) Employee. Then, do a design of the requirements that will be made, such as making activity diagrams for employees that can describe activities on the system in general, which are shown in the image below:
3.4 User Interface
The expert system in this research is an expert system based on android mobile. To be able to carry out the diagnostic process, the user, in this case the OPD employee of Cianjur Regency, must first log in, if the employee account has not been obtained, then the Personnel Monitoring section in each OPD will add it first. The following are some interfaces on the Cianjur Regency OPD health check application:
Fig 4. Login Interface and Dashboard of Monitoring User

After the user, in this case the Employee, opens the application, a Login Form will be displayed to enter the application, after successfully logging in, the Dashboard is displayed according to the access rights (Admin, Monitoring, and Employee).

Fig 5. Profile and Approval Form

When an employee accesses the Profile menu, the profile page will be displayed, on this page the employee can edit their own profile, change passwords, and change profile photos. When the Health Check menu is accessed, a screening approval form will be displayed as shown in Figure 4. Profile and Approval Form.
Fig 6. Filling in and Diagnostic Results

After that, the results and conclusions of the diagnoses will be displayed, as shown in Figure 5 above.

4. Conclusion

Based on the results of discussions, calculations and comparisons of expert systems for the diagnosis of COVID-19 and for determining the feasibility of swabs and work from home using a naive Bayesian approach, the following conclusions can be drawn: (a). This application can make it easier for Cianjur Regency Government employees to diagnose COVID-19 early based on the symptoms experienced and determine the suitability of swabs and working from home. (b). This time, the system using the Naive Bayes method has an accuracy of 80% for the G01 symptom. Fever, G02. Cough/cold, G04 for muscle pain and G06 for anosmia with unhealthy results, eligible for swab, compared with other classifications. (c). For further research, it can be integrated with the existing government agency attendance system or retested with other methods.

References

2051, 2021.


