Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1584

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 31, 2022

Classification of Covid-19 vaccine data screening with Naive Bayes algorithm using Knowledge Discovery in database method

Syariful Alam[1)*], Mochzen Gito Resmi[2)], Nunung Masripah[3)]
[1,2,3)]Sekolah Tinggi Teknologi Wastukancana, Indonesia
[1)]Syarifulalam@wastukancana.ac.id, [2)]Mochzen@wastukancana.ac.id, [3)]nunungmasripah0001@gmail.com

ABSTRACT

Acute Respiratory Syndrome Coronavirus-2 (SARS-Cov-2), known as covid-19, was detected and caused many deaths due to a mysterious respiratory disease. With the death toll continuing to rise, the government was forced to take swift action to break the chain of spread and reduce the number of deaths by taking vaccinations. An adequate vaccine against Covid-19 is expected to vaccinate at least 70% of the population. Therefore, this study was carried out as a step to help break the chain of the spread of the Covid-19 virus by classifying the Covid-19 vaccine screening data. The research method applied in this study is the Knowledge Discovery in Database (KDD) method, in which there are several processes, namely selection, pre-processing, transformation, data mining, and evaluation. The application of the Naive Bayes method is expected to be able to classify Covid-19 vaccine screening data with vaccine class values, no, and delay. The research results on the classification of the Naive Bayes method show that there are 959 data with Vaccine data 695, No 200, and Delay 64. Processed using the Rapidminer application, the accuracy is 96.56%, precision is 92.46%, and recall is 92.13%.

Keywords : Covid-19, Vaccination and Naive Bayes.

INTRODUCTION

Covid-19 is a new series of viruses known as Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-Cov-2), which was detected and caused many deaths due to a mysterious respiratory disease. This virus can be transmitted from human to human and becomes a universal pandemic (NCDC, 2021). The rapid development of vaccines to prevent Covid-19 has become a global imperative. Adequate vaccines against Covid-19 are expected to be produced by 2021 to vaccinate at least 70% of each country's population.

With the development of information technology, the ability to collect and process data is also growing. The utilization of information and knowledge contained in the data at this time is called data mining. Data mining is finding interesting patterns and knowledge from large amounts of data. Data sources can be databases, warehouses, web, repositories and other information, or data that is flowed into the system dynamically (Coding & Untan, 2018). The method used in this research is KDD (Knowledge Discovery in Database). Data mining and knowledge discovery in databases (KDD) are often used interchangeably to describe the process of extracting hidden information in an extensive database. The two terms have different concepts but are related to each other. And one of the stages in the KDD process is data mining (Asroni et al., 2018).

One part of data mining is classification. Classification is a way of grouping objects based on the characteristics of the object of classification. In the process, classification can be done in many ways, either manually or with the help of technology. A classification done manually is a classification that humans carry out without the help of intelligent computer algorithms. Meanwhile, with the help of technology, the classification has several algorithms, including Nave Bayes (Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, 2018). A naive Bayes Classifier is a statistical classification method for predicting the probability (possibility) of members of a class. The accuracy and speed of the Naive Bayes Classifier method are very high when used in a database application that has a large amount of data. The Naive Bayes Classifier algorithm minimizes the error rate compared to other classification algorithms (Titimeidara & Hadikurniawati, 2021).

The advantage of using Naive Bayes is that this method only requires a small amount of training data to determine the estimated parameters needed in the classification process. Naive Bayes often performs much better in most complex real-world situations than expected (Manalu et al., 2017).

In this study, the author will process the COVID-19 vaccine screening dataset at the UPTD of the Bojong Health Center. By referring to the data obtained based on the patient's medical history with the standard provision of vaccines that have been determined, the authors will use it as research material in processing datasets using the

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

Naïve Bayes algorithm for predicting the results of Covid-19 vaccine screening.

LITERATURE REVIEW

Data Mining

Data mining is the process of extracting and identifying information or knowledge from piles of data that are many and scattered in different sources. In the process of data mining using statistical techniques, mathematics, and artificial intelligence. Other terms that refer to data mining are knowledge extraction, pattern analysis, information harvesting, and data archaeology. Data mining results can predict a problem, find new information, find unknown patterns, and assist in decision-making (Firdaus et al., 2021). There is another understanding of Data Mining. Namely, Data mining is a technique used to build machine learning models. Machine learning is a modern artificial intelligence technique that learns to build models using empirical data. Data mining is used to find patterns in large sets of raw data. Data Mining applies Machine Learning techniques to draw knowledge from data (Liliana et al., 2021).

Data mining has several stages: data cleansing, data integration, data transformation, data mining techniques, pattern evaluation, and knowledge presentation. (Hayuningtyas, 2019)

There are 6 stages of data mining with the following explanation:

1. Data Cleaning is the process of removing inconsistent data or irrelevant data
2. Data Integration Merging or combining data from several sources.
3. Data Selection of data is from a set of operational data before obtaining knowledge and discovering information in the database.
4. Data Transformation Data is converted or combined into a format suitable for processing in data mining.
5. Data Mining Process The primary process is when methods are applied to find knowledge or information from data.
6. Pattern Evaluation To identify exciting patterns in the knowledge-based found.
7. Knowledge presentation Knowledge of the methods used to acquire user knowledge.

Naive Bayes

Naive Bayes is a simple probabilistic classifier that calculates a set of probabilities by adding up the frequencies and combinations of values from a given dataset. The algorithm uses the Bayes theorem and assumes all attributes are independent or not interdependent, given by the value of the class variable (Manalu et al., 2017).

At the classification stage, the probability value of each existing class label will be calculated for the given input. The class label that has the most significant probability value will be used as the class label for the input data. Naive Bayes is the most straightforward Bayes theorem calculation because it can reduce the computational complexity to a simple multiplication of probabilities. In addition, the Naive Bayes algorithm is also able to handle data sets that have many attributes (Sartika & Indra, 2017)

The advantage of using this method is that it only requires a small amount of training data to determine the parameter estimates needed in the classification process. Because it is assumed as an independent variable, only the variance of a variable in a class is needed to determine the classification, not the entire covariance matrix (Alvina Felicia Watratan et al., 2020).

Naive Bayes formula, according to (Hozairi et al., 2021):

$$\mathrm{P(H|X)} = \frac{P(X|H).P(H)}{P(X)}$$

X = Data with unknown class
H = Hypothesis data is a specific class.
P(H|X) = Probability of hypothesis H based on condition X (*Posterior Probability*)
P(H) = Hypothesis probability H (*Prior Probability*)
P(X|H) = Probability X based on hypothesis H
P(X) = Probability X

* Corresponding author

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

METHOD

KDD Method Research Stages

In this study, the author uses the method or stages of KDD (Knowledge Discovery in Database). In the KDD stage, there are 5 stages of the process: selection, pre-processing, transformation, data mining, and evaluation. The following stages of the KDD process consist of:

Data Selection

Data Selection analyzes relevant data from the database because it is often found that not all data is needed in the data mining process. The data was selected from the dataset to be analyzed; the data used in this study was in the form of COVID-19 vaccine screening data in 2021 at the UPTD of the Bojong Health Center. The data consists of name, NIK, date of birth, age, address, temperature, blood pressure, autoimmune disease, medication, heart, Covid-19 history and Screening result label. With the Data Selection, the processing process will be better by achieving the research objectives with 7 attributes, namely age, blood pressure, autoimmune disease, blood treatment, heart disease, and Covid-19.

Pre-Processing

After the selection stage, the next stage is Pre-processing; before carrying out the data mining process, it is necessary to do data cleaning or cleaning on the selection data. The cleaning process includes, among others, removing duplicate data, checking for inconsistent data, and correcting errors in data, such as typographical or printing errors (Luluk Elvitaria, 2017). The Covid-19 vaccine screening data cleared 1000 data to 959 data. The data to be cleaned is empty data.

Transformation

In this Transformation stage, the data is converted into a form adapted for the data mining process. Data that was previously processed using Microsoft Excel 2019 will be processed and further processed using rapid miner studio 9.8.001 software; the data file format has not changed because Rapidminer tudio 9.8.001 can directly import files with excel.xlsx format. The data transformation occurs by alternating the Age Attribute and the Blood Pressure Attribute. The following is the process of the data transformation:

Table 3. 1 Age Attribute Transformation

Age	Transformation
5-11 Years	Child
12-25 Years	Teenager
26-45 Years	Mature
>46	seniors

Table 3. 2 Transformation of Blood Pressure Attributes

Blood pressure	Transformation
<140/80	Normal
>=140/80	Hypertension

RESULT

Data Mining Process

At the data mining stage, the Covid-19 vaccine screening data analysis was carried out, the method used was the classification method using the Naive Bayes algorithm, and the analysis was carried out using the Rapidminer 9.8.001 application. This stage will determine the results of implementing the nave Bayes algorithm in classifying. For reference in determining whether the algorithm is correct, manual calculations are also carried out to classify the results using Rapidminer 9.8.001. The following is a calculation using the Naive Bayes and Rapidminer algorithms:

Calculating Probability Prior

Experiments carried out in this study are to calculate the Prior Probability and Posterior Probability. Prior is the distribution of the parameters. Determining priors is done by the availability of previous research information (Dhany, H. W., & Izhari, 2019). The calculation of the prior probability uses a total of 959 data records, of which 3

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

classes are formed, namely Vaccines, No and Delays. The following is a table for calculating prior probabilities:

Table 4. 1 Priority Probability Calculations

Attributes/Variables		Amount of data	Vaccine	Not	Cancel	P(X)Ci Vaccine	P(X)Ci Not	P(X)Ci Cancel
Total		959	695	200	64	0.725	0.209	0.067
Age	Child	118	109	1	8	0.924	0.008	0.068
	Teenager	162	152	5	5	0.938	0.031	0.031
	Mature	264	203	18	43	0.769	0.068	0.163
	seniors	415	231	176	8	0.557	0.424	0.019
Blood pressure	Normal	710	683	26	1	0.962	0.037	0.001
	Hypertension	249	12	174	63	0.048	0.699	0.253
Autoimmune Disease	Yes	26	0	26	0	0.000	1,000	0.000
	Not	933	695	174	64	0.745	0.186	0.069
Blood Treatment	Yes	4	1	3	0	0.250	0.750	0.000
	Not	955	694	197	64	0.727	0.206	0.067
Heart	Yes	10	0	10	0	0.000	1,000	0.000
	Not	949	695	190	64	0.732	0.200	0.067
Covid-19	Yes	7	4	1	2	0.571	0.143	0.286
	Not	952	691	199	62	0.726	0.209	0.065

Calculating Posterior Probability

After calculating the Prior Probability, the next step is calculating the Posterior Probability. The posterior probability calculation is used to determine the class of the new data to be classified (Hayuningtyas, 2019).
Calculating the Vaccine, No and Delay class value is taken from Prior Probability, with Naive Bayes with 959 data records. Calculation of vaccine selection, No and Delay. The data shows 695 Vaccines, 200 No, and 64 Delays using the formula below.

P(H│X)= (P(X|H)P(H))/(P(X))
P(H|X) = P(X|H) x P(H)
P(VACCINE) = 695/959 = 0.725
P(NO) = 200/959 = 0.209
P(DELAY) = 64/959 = 0.067

Calculating the probability value can be calculated by normalizing the class, and by calculating this probability value, you can find out the final result of the calculation using the Naive Bayes Vaccine, No or Delay method:

Table 4.2 Posterior Probability Cases

Age	Pressure Blood	Disease Autoimmune	Treatment Blood	Heart	Covid-19	Screening Results
ELDERLY	NORMAL	Not	Not	Not	Yes	???

It is known that the result class has 3 classifications, namely:
C1 = Screening Result = Vaccine
C2 = Screening Result = No
C3 = Screening Result = Delay

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

In table 4.2 There is a case whose class is not yet known.
The sum of each result class is divided by the total data.
P (Screening Results="VACCINE") = 695/959 = 0.725
P (Screening Result="NO") = 200/959 = 0.209
P(Screening Result="DELAY") = 64/959 = 0.067

Calculate P(X|Ci), which is the probability of each attribute in X data, then divided by the number of Vaccine, No and Delay classes:

- P(Age = "Elderly" | Screening Result = "Vaccine") = 231/695 = 0.332
- P(Age = "Elderly" | Screening Result = "No") = 176/200 = 0.880
- P(Age = "Elderly" | Screening Result = "Delay") = 8/64 = 0.125

- P(Blood Pressure = "Normal" | Screening Result = "Vaccine") = 683/695 = 0.983
- P(Blood Pressure = "Normal" | Screening Result = "No") = 26/200 = 0.13
- P(Blood Pressure = "Normal" | Screening Result = "Delay") = 1/64 = 0.016

- P(Autoimmune Disease = "No" | Screening Result = "Vaccine") = 695/695 = 1
- P(Autoimmune Disease = "No" | Screening Result = "No") = 174/200 = 0.87
- P(Autoimmune Disease = "No" | Screening Result = "Delay") = 64/64 = 1

- P(Blood Treatment = "No" | Screening Result = "Vaccine") = 694/695 = 0.999
- P(Blood Treatment = "No" | Screening Result = "No") = 197/200 = 0.985
- P(Blood Treatment = "No" | Screening Result = "Delay") 64/64 = 1

- P(Heart = "No" | Screening Result = "Vaccine") = 695/695 = 1
- P(Heart = "No" | Screening Result = "No") = 190/200 = 0.950
- P(Heart = "No" | Screening Result = "Delay") = 64/64 = 1

- P(Covid = "Yes" | Screening Result = "Vaccine") = 4/695 = 0.006
- P(Covid = "Yes" | Screening Result = "No") = 1/200 = 0.01
- P(Covid = "Yes" | Screening Result = "Delay") = 2/64 = 0.31

Calculate each value for each attribute in X data:
P(X|Screening Results="VACCINE")
= 0.332 x 0.983 x 1 x 0.999 x 1 x 0.006
= 0.00188

P(X|Screen Result = "NO")
= 0.880 x 0.13 x 0.87 x 0.985 x 0.950 x 0.01
= 0.000466

P(X|Screen Result="DELAY")
= 0.125 x 0.016 x 1 x 1 x 1 x 0.031
=0.000061

After obtaining the total for each class in X data, then multiply by the class probability in all data:
P(X|Screening Results="VACCINE") P(Screening Results="VACCINE")
= 0.00188 /0.725 = 0.00136
P(X|Screening Result = "No") P(Screening Result = "No")
= 0.000466 /0.209 = 0.000097
P(X|Screening Result="Delay") P(Screening Result="No")
= 0.000061 / 0.067 = 0.000004

With the final value of the vaccine class 0.00136, No 0.000097 and Delay 0.000004, the largest class value is the vaccine class value, then the class contained in the above case is the vaccine.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

Rapidminer

At this stage, data processing is carried out using the Rapidminer application with the Naïve Bayes model. The main purpose of this study was to determine the accuracy of the Naive Bayes method in the classification of COVID-19 vaccine screening data. The following is the data processing process in the Rapidminer application.

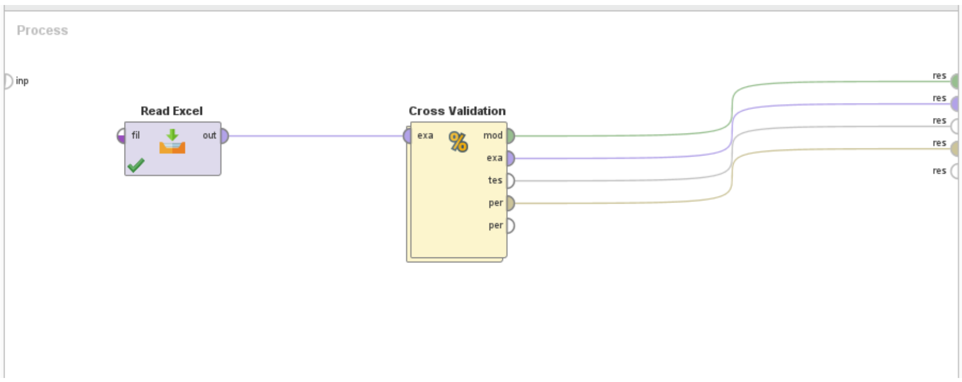

Figure 4. 1 Process validation

Figure 4.1 is a validation process to check your model performance on one dataset you use for training and testing. There are two operators, Read excel and Cross-Validation, where the read excel operator functions to retrieve the dataset in the document. In contrast, the cross-validation operator functions to validate data with an algorithmic model and to maximize the data processing accuracy value.

The read Excel operator and the cross-validation operator are connected by drawing a line from reading excel to cross-validation and dragging the cross-validation operator line to the result. The cross-validation operator has two sub-processes: the training sub-process and the testing sub-process. The following are the cross-validation sub-processes:

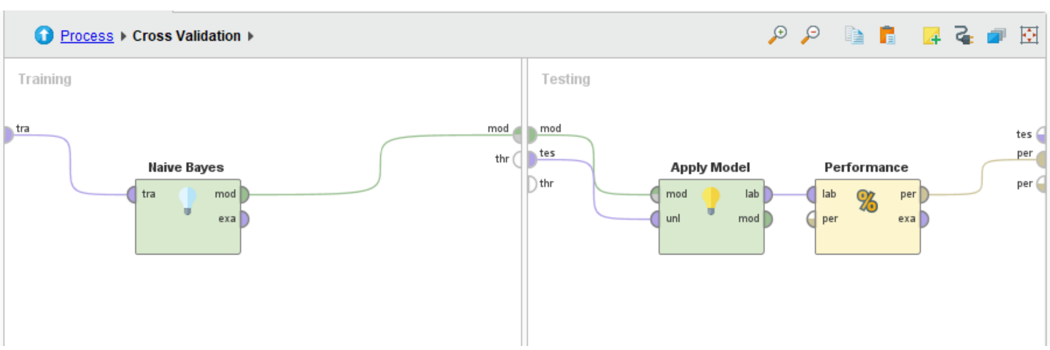

Figure 4. 2 Naive Bayes Model Process with Cross Validation

In Figure 4.2, there are training and testing subprocesses; the training sub-process serves to carry out the model training process by entering the nave Bayes operator, while the testing sub-process is to test the model, which will produce a graph or pattern by entering the Apply Model operator and the Performance operator which will produce Simple Distribution, Accuracy, Precision, and Recall.

Based on the results of the calculation with the Rapidminer with the nave Bayes model, the value of the simple distribution obtained is as shown in the image below; the result is that the Vaccine class value is 0.725, the No class value is 0.209, and the Delay class value is 0.067 where manual calculations with calculations on Rapidminer the results are the same.

Based on the results of the test using the nave Bayes model with the rapid miner application, it can be seen in the following figure 4.4:

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

accuracy: 96.56% +/- 2.46% (micro average: 96.56%)

	true VAKSIN	true TIDAK	true TUNDA	class precision
pred. VAKSIN	682	7	1	98.84%
pred. TIDAK	7	191	10	91.83%
pred. TUNDA	6	2	53	86.89%
class recall	98.13%	95.50%	82.81%	

Figure 4. 4 Accuracy

In Figure 4.4 above is the test result of the nave Bayes method in the form of calculations based on data that has been processed in the Rapidminer application. Of the total 959 data, 682 were classified as vaccines according to predictions, 7 data predicted vaccines but turned out to be No, 1 data predicted vaccines. Still, they turned out to be delayed, 191 data were not classified accordingly, while 7 data predicted No but turned out to be vaccines, 10 data predicted No but turned out to be Cancel. And 53 data of delay were classified accordingly, while 6 data were predicted to be Delayed but turned out to be vaccines, 2 data were predicted to be Delayed but were not.

DISCUSSIONS

Evaluation Process

At this stage, testing and evaluation are carried out using a confusion matrix to determine accuracy, recall, and precision values. Calculations at the evaluation stage are as follows:

Table 5.1 Testing the Confusion Matrix

Table 5.1 Testing the Confusion Matrix

Classification	**Vaccine**	**Not**	**Cancel**
Vaccine	682	7	1
Not	7	191	10
Cancel	6	2	53

By using a 3x3 multiclass confusion matrix table, to calculate accuracy, precision, and recall are as follows:
Accuracy compares all case outcomes with the correct identification value (Subagyo et al., n.d.). The following calculation to find Accuracy:

Accuracy= (682+191+53)/(682+7+1+7+191+10+6+2+53)
=692/959=0.9655 x 100
= 96.55%

The recall is a positive case that is identified as true by the appropriate proportion

Yes (Subagyo et al., n.d.)
682/(682+7+6)=0.981
191/(7+191+2)=0.955
53/(1+10+53)=0.828
Recall= (0.981+0.955+0.828)/3
=0.921 x 100
=92.13%

Precision is the proportion of positive results identified as true (Subagyo et al., n.d.)

682/(682+7+1)=0.988
191/(7+191+10)=0.918
53/(6+2+53)=0.868
Precision= (0.988+0.918+0.868)/3

* Corresponding author

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

=0.9246 X 100
=92.46%

The evaluation results using the confusion matrix shows an accuracy value of 96.55%, recall is 92.13%, and precision is 92.46%.

CONCLUSION

From the results of research on the application of data mining with the Naïve Bayes Classifier method in classifying Covid-19 vaccine screening data at the Bojong Health Center UPTD using a data set that has gone through the early stages of Knowledge Discovery in the database (KDD) as much as 959 data with Vaccines 695, No 200, and Delay 64 using the Rapidminer application for data processing produces an accuracy value of 96.56%, Precision 92.46%, and Recall 92.13%. So the writer can conclude that the Naive Bayes algorithm performs well in classifying COVID-19 vaccine screening data.

REFERENCES

Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokkan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika*, *21*(1), 60–64. https://doi.org/10.18196/st.211211

Coding, J., & Untan, S. K. (2018). *Kata Kunci: Kebakaran Hutan, Data Mining, K-Nearest Neighbor (KNN), Fire Weather Index(FWI)*. *1*. *06*(2).

Dhany, H. W., & Izhari, F. (2019). ANALISIS ALGORITHMS SUPPORT VECTOR MACHINE DENGAN NAIVE BAYES KERNEL PADA KLASIFIKASI DATA. Jurnal Teknik Dan Informatika, 6(2), 30–35. *ANALISIS ALGORITHMS SUPPORT VECTOR MACHINE DENGAN NAIVE BAYES KERNEL PADA KLASIFIKASI DATA. Jurnal Teknik Dan Informatika, 6(2), 30–35.*, *6*.

Elvitaria, L. (2017). MEMPREDIKSI TINGKAT PEMINAT EKSTRAKURIKULER PADA SISWA SMK ANALISIS KESEHATAN ABDURRAB MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS: SMK ANALIS KESEHATAN ABDURRAB). *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, *2*(2), 110-124. https://doi.org/10.36341/rabit.v2i2.212

Firdaus, A., Firdaus, W. I., Studi, P., Informatika, T., Digital, M., & Sriwijaya, P. N. (2021). *Text Mining*. *13*(1), 66–78.

Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita. *Jurnal Informatika*, *6*(1), 18–22. https://doi.org/10.31311/ji.v6i1.4685

Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, *6*(2), 133. https://doi.org/10.21107/nero.v6i2.237

Liliana, D. Y., Maulana, H., & Setiawan, A. (2021). Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes. *Multinetics*, *7*(1), 48–53. https://doi.org/10.32722/multinetics.v7i1.3786

Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan Pada CV. Papadan Mama Pastries. *Jurnal Mantik Penusa*, *1*(2), 16–21. https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ffh&AN=2008-10-Aa4022&site=eds-live&scope=site

NCDC. (2021). COVID-19 Situation Report: Situation Report 95. In *Nigeria Centre for Disease Control* (Vol. 95, Issue June). https://ncdc.gov.ng/themes/common/files/sitreps/fe61a5d7bba46e835a7f6d09bd5343c8.pdf

Sartika, D., & Indra, D. (2017). Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. *Jurnal Teknik Informatika Dan Sistem Informasi*, *1*(2), 151–161.

Subagyo, I., Yulianto, L. D., Permadi, W., & Dewantara, A. W. (n.d.). *Sentiment Analisis Review Film Di IMDB Menggunakan Algoritma SVM Sentiment Analysis of Film Review at IMDB using SVM algorithm Abstrak Pendahuluan Metode Penelitian*. *8*(1), 47–56.

Titimeidara, M. Y., & Hadikurniawati, W. (2021). Implementasi Metode Naïve Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita. *Jurnal Ilmiah Informatika*, *9*(01), 54–59.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 4, Number 2, July 2022
https://doi.org/10.47709/cnahpc.v4i2.1580

Submitted : June 27, 2022
Accepted : July 13, 2022
Published : July 27, 2022

https://doi.org/10.33884/jif.v9i01.3741

Watratan, A. F., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, *1*(1), 7-14.

Wibawa, A. P., Purnama, M. G. A., Akbar, M. F., & Dwiyanto, F. A. (2018). Metode-metode Klasifikasi. In *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi Vol* (Vol. 3, No. 1)

* Corresponding author