International Journal of Basic and Applied Science 10 (3) (2021) 69-78

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Application Of Naive Bayes Algorithm In Classification Of Child Nutrition At The Simalungun Health Office

Susi Septi Hardiani[1], M. Safii[1], Dedi Suhendro[2]

[1]Informatics Engineering Study Program, STIKOM Tunas Bangsa, Pematangsiantar

[2]Computerized Accounting Study Program, Amik Tunas Bangsa, Indonesia

Article Info

***Article history*:**

Received Sep 9, 2021

Revised Sep 20, 2021

Accepted Oct 09, 2021

***Keywords*:**

Data Mining;
Child Nutrition;
Nave Bayes;
Classification.

ABSTRACT

Toddlers are among the most vulnerable groups to nutritional problems when viewed from the point of view of health and nutrition problems, while at this time they are experiencing a cycle of relatively rapid growth and development 7% is quite high where the number of births is relatively large. Researchers try to classify 10 toddlers using WEKA to find out whether they have nutritional disorders or are normal by using 5 attributes as system input and a class namely nutrition which divides this class into 4 namely bad, less, good and more with the amount of training data 219 data then data compared with the actual nutritional conditions and obtained an accuracy of 60% and an error of 40% with these results it can be concluded that the accuracy is not too good. Based on this, it is hoped that the results of this classification can help further research in classifying the nutrition of children under five.

Corresponding Author:

M. Safii

Informatics Engineering Study Program,
STIKOM Tunas Bangsa,
Jl. Jend. Sudirman Blok A No.1-2 Pematangsiantar, North Sumatera-Indonesia
Email: m.safii@amiktunasbangsa.ac.id

1. INTRODUCTION

Children are an investment and hope for the future of the nation as well as the next generation in the future. In the life cycle, childhood is a phase where children experience growth and development that determines their future. It is necessary to optimize the development of children, because besides being crucial at that time children need attention and affection from parents or families so that fundamentally the rights and needs of children can be fulfilled properly[1]. Children must be able to grow and develop into human beings who are physically and mentally healthy, intelligent, happy, of high moral character and commendable, because in the future they are assets that will determine the quality of the nation's civilization[2]. At that age the child is in a period of human growth and development called the golden age. Based on several studies, it is stated that the golden age is present in the period of conception, namely since humans are still in the mother's womb until the first few years of their birth which is termed early age[3]. After the child is 24 months old, there is no more addition of new neuron cells as happened at the previous age, but maturation still continues until the child is four or five years old[4].

In the early stages of life, starting at the age of 3 years, children begin to be able to accept skills as the basis for forming knowledge and thinking processes. The role of parenting includes care and education, providing adequate food for the development of intellectual intelligence, non-material provision for the development of emotional and spiritual intelligence greatly determines the quality of children in the future[5]. According to UNICEF, children's nutritional status is directly affected by food intake and infectious diseases[6]. Malnutrition is the basic cause of children's growth disorders, therefore, it must be prevented so that growth disorders do not occur, although children's physical growth disorders can still be corrected in the future by increasing good nutritional intake, but not so with the development of intelligence[7] .

In computer science there are many branches of science and methods contained therein such as Decision Support Systems, Expert Systems, Data Mining and others. Data mining is a process that employs one or more computer learning techniques (machine learning) to analyze and extract knowledge automatically[8],[9],[10]. There are 5 roles of data mining, namely estimation, prediction, classification, clustering, association[11],[12]. There are many algorithms in classification data mining, one of which is Nave Bayes. Naive Bayes algorithm is one of the algorithms contained in the classification technique[13],[14]. Naive Bayes is a classification with probability and statistical methods proposed by British scientist Thomas Bayes, which predicts future opportunities based on previous experience, so it is known as Bayes' theorem[15],[16]. The theorem is combined with Naive where it is assumed that the conditions between attributes are independent[17]. Naive Bayes classification assumes that the presence or absence of certain characteristics of a class has nothing to do with the characteristics of other classes.

2. RESEARCH METHOD

The research method is a technique or seeking, obtaining, collecting data, both in the form of primary data and secondary data used for the purpose of compiling a scientific work and analyzing factors related to the main issues so that there will be a truth of the data to be obtained. obtained. The settlement method used in this study is the data mining method with the Naive Bayes algorithm[18]. This research was conducted to implement the Naive Bayes algorithm in the classification of children's nutrition at the Simalungun Health Office. From the results of the classification of Child Nutrition data, the Simalungun Health Office can find out information so that it can be handled with Child Nutrition. In conducting research to obtain data, the method used is the Observation method, namely seeing and studying problems that exist in the field related to the object under study and the Literature Review method to find materials that support the definition of problems through books, papers, the internet, which closely related to the object of the problem[19].

The flow diagram of the research work activities carried out in this study is shown in Figure1 below:

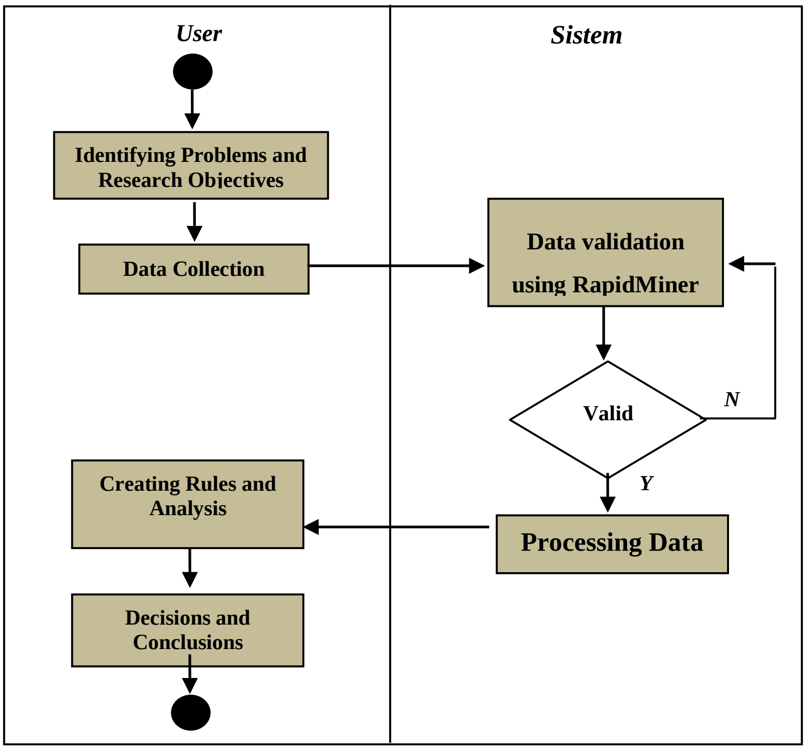

Figure 1. Research Activity Work Diagram

The data to be processed is 50 samples of child data in the Simalungun district health office as training data and testing data taken from the last 15 data from training data.

Table 1. Testing Data

No	Sample	Gender	Age	Weight	Height	Status
1	A36	Boy	Age >12 and <=24	Very thin	Short	Good Nutrition
2	A37	Boy	Age >12 and <=24	Thin	Short	Malnutrition
3	A38	Boy	Age >12 and <=24	Very thin	Very Short	Good Nutrition
4	A39	Boy	Age <=12	Very thin	Very Short	Malnutrition
5	A40	Girl	Age >12 and <=24	Very thin	Very Short	Malnutrition
6	A41	Girl	Age <=12	Very thin	Normal	Good Nutrition
7	A42	Boy	Age <=12	Very thin	Normal	Good Nutrition
8	A43	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
9	A44	Boy	Age <=12	Thin	Short	Malnutrition
10	A45	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
11	A46	Boy	Age >12 and <=24	Thin	Short	Malnutrition
12	A47	Girl	Age >12 and <=24	Thin	Short	Malnutrition
13	A48	Girl	Age >12 and <=24	Normal	Short	Good Nutrition
14	A49	Boy	Age >12 and <=24	Very thin	Very Short	Malnutrition
15	A50	Boy	Age >12 and <=24	Thin	Normal	Good Nutrition

3. RESULTS AND DISCUSSIONS

To solve the problem in this research, data mining techniques with the Naïve Bayes algorithm are used as follows[20] :

$$P(H\backslash X) = \frac{P(H\backslash X)*(P(H)}{P(X)} \qquad (1)$$

Where:
X : Data with unknown class
H : Hypothesis data is a specific class
P(H|X) : Probability of hypothesis H based on condition X (posteriori probability)
P(H) : Hypothesis probability H (prior probability)
P(X|H) : Probability of X based on the conditions on the hypothesis H
P(X) : Probability X

3.1. Solution with *Naïve Bayes*

The amount of data used is 50 data, Good Nutrition data is 27 and Malnutrition data is 23 data. In step 1, an indication of Nutritional Status has been obtained which is predicted based on Gender, Age, Body Weight and Body Height, according to the decision table 2.

Tabel 2. Data *Training*

No	Sample	Gender	Age (0 - 24)	Weight	Height	Status
1	A1	Boy	Age <=12	Thin	Normal	Malnutrition
2	A2	Girl	Age >12 and <=24	Fat	Short	Good Nutrition
3	A3	Boy	Age >12 and <=24	Thin	Very Short	Malnutrition
4	A4	Boy	Age >12 and <=24	Very thin	Normal	Good Nutrition
5	A5	Boy	Age >12 and <=24	Thin	Normal	Good Nutrition
6	A6	Girl	Age >12 and <=24	Very thin	Very Short	Malnutrition
7	A7	Girl	Age >12 and <=24	Normal	Short	Malnutrition
8	A8	Girl	Age <=12	Fat	Very Short	Malnutrition
9	A9	Boy	Age <=12	Very thin	Very Short	Malnutrition
10	A10	Boy	Age <=12	Very thin	Normal	Good Nutrition
11	A11	Boy	Age >12 and <=24	Very thin	Very Short	Malnutrition
12	A12	Boy	Age <=12	Normal	Normal	Good Nutrition
13	A13	Girl	Age <=12	Very thin	Normal	Malnutrition
14	A14	Girl	Age >12 and <=24	Normal	Short	Good Nutrition
15	A15	Girl	Age >12 and <=24	Fat	Normal	Good Nutrition
16	A16	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
17	A17	Boy	Age <=12	Normal	Normal	Good Nutrition
18	A18	Boy	Age <=12	Very thin	Short	Malnutrition
19	A19	Boy	Age >12 and <=24	Very thin	Short	Malnutrition
20	A20	Boy	Age >12 and <=24	Thin	Short	Malnutrition
21	A21	Boy	Age <=12	Thin	Short	Malnutrition
22	A22	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
23	A23	Boy	Age <=12	Normal	Normal	Good Nutrition
24	A24	Girl	Age <=12	Thin	Normal	Good Nutrition
25	A25	Girl	Age >12 and <=24	Fat	Normal	Good Nutrition
26	A26	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
27	A27	Boy	Age >12 and <=24	Thin	Very Short	Malnutrition
28	A28	Boy	Age >12 and <=24	Normal	Normal	Good Nutrition
29	A29	Girl	Age >12 and <=24	Fat	Normal	Good Nutrition
30	A30	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
31	A31	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
32	A32	Boy	Age >12 and <=24	Very thin	Very Short	Malnutrition
33	A33	Girl	Age <=12	Very thin	Very Short	Malnutrition
34	A34	Girl	Age <=12	Very thin	Short	Malnutrition

No	Sample	Gender	Age (0 - 24)	Weight	Height	Status
35	A35	Boy	Age <=12	Fat	Normal	Good Nutrition
36	A36	Boy	Age >12 and <=24	Very thin	Short	Good Nutrition
37	A37	Boy	Age >12 and <=24	Thin	Short	Malnutrition
38	A38	Boy	Age >12 and <=24	Very thin	Very Short	Good Nutrition
39	A39	Boy	Age <=12	Very thin	Very Short	Malnutrition
40	A40	Girl	Age >12 and <=24	Very thin	Very Short	Malnutrition
41	A41	Girl	Age <=12	Very thin	Normal	Good Nutrition
42	A42	Boy	Age <=12	Very thin	Normal	Good Nutrition
43	A43	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
44	A44	Boy	Age <=12	Thin	Short	Malnutrition
45	A45	Boy	Age >12 and <=24	Normal	Short	Good Nutrition
46	A46	Boy	Age >12 and <=24	Thin	Short	Malnutrition
47	A47	Girl	Age >12 and <=24	Thin	Short	Malnutrition
48	A48	Girl	Age >12 and <=24	Normal	Short	Good Nutrition
49	A49	Boy	Age >12 and <=24	Very thin	Very Short	Malnutrition
50	A50	Boy	Age >12 and <=24	Thin	Normal	Good Nutrition

a. Counting the Number of Classes/Labels
Based on the training data above, it can be calculated by the following equation formula:

Status Good Nutrition Status = $\frac{Number\ of\ Possible\ Good\ Nutrition}{Number\ of\ Possible\ Nutritional\ Status} = \frac{27}{50} = 0{,}54\ or\ 54\%$

Malnutrition Status = $\frac{Number\ of\ Possible\ Malnutrition}{Number\ of\ Possible\ Nutritional\ Status} = \frac{23}{50} = 0{,}46\ or\ 46\%$

b. Counting the number of cases with the same class
After looking for the probability value of each criterion, obtained from the training data, the probability value for each criterion can be seen in the following table 2 With the following equation formula:

P. Good Nutrition, boys = $\frac{Same\ Number\ of\ Cases}{Total\ Cases\ (class)\ are\ the\ same} = \frac{19}{27} = 0{,}70$

P. Good Nutrition, boys = $\frac{Same\ Number\ of\ Cases}{Total\ Cases\ (class)\ are\ the\ same} = \frac{8}{27} = 0{,}30$

P. Malnutrition, Girls = $\frac{Jumlah\ Kasus\ Yang\ sama}{Total\ Cases\ (class)\ are\ the\ same} = \frac{15}{23} = 0{,}65$

P. Malnutrition, Girls = $\frac{Same\ Number\ of\ Cases}{Total\ Cases\ (class)\ are\ the\ same} = \frac{8}{23} = 0{,}35$

Table 3. Table of Gender Probability Values

P=Gender	Nutrition Status		Probilitas	
	Good Nutrition	Malnutrition	Good Nutrition	Malnutrition
Boy	19	15	0,70	0,65
Girl	8	8	0,30	0,35
Total	27	23	1	1

Table 4. Age Probability Value Table

P=Age (0 - 24)	Nutrition Status		Probilitas	
	Good Nutrition	Malnutrition	Good Nutrition	Malnutrition
Age <=12	8	10	0,30	0,43
Age >12 and <=24	19	13	0,70	0,57

Total	27	23	1	1

Table 5. Table of Probability Values for Body Weight

P=Weight	Nutrition Status		Probility	
	Good Nutrition	Malnutrition	Good Nutrition	Malnutrition
Fat	5	1	0,19	0,04
Normal	13	1	0,48	0,04
Thin	3	9	0,11	0,39
Very thin	6	12	0,22	0,52
Total	27	23	1	1

Tabel 6. Tabel Nilai Probabilitas Tinggi Baand

P=Height	Nutrition Status		Probility	
	Good Nutrition	Malnutrition	Good Nutrition	Malnutrition
Normal	15	2	0,56	0,09
Short	11	10	0,41	0,43
Very Short	1	11	0,04	0,48
Total	27	23	1	1

c. Multiply All Results Variable Good Nutrition and Malnutrition

The next step is to multiply all the results of the Good Nutrition and Malnutrition variables using the test data. The following is the test data used in the following table 7.

Table 7. Test Data

No	Sample	Gender	Age (0 - 24)	Weight	Height	Status
1	A36	Boy	Age >12 and <=24	Very thin	Short	?
2	A37	Boy	Age >12 and <=24	Thin	Short	?
3	A38	Boy	Age >12 and <=24	Very thin	Very Short	?
4	A39	Boy	Age <=12	Very thin	Very Short	?
5	A40	Girl	Age >12 and <=24	Very thin	Very Short	?
6	A41	Girl	Age <=12	Very thin	Normal	?
7	A42	Boy	Age <=12	Very thin	Normal	?
8	A43	Boy	Age >12 and <=24	Normal	Short	?
9	A44	Boy	Age <=12	Thin	Short	?
10	A45	Boy	Age >12 and <=24	Normal	Short	?
11	A46	Boy	Age >12 and <=24	Thin	Short	?
12	A47	Girl	Age >12 and <=24	Thin	Short	?
13	A48	Girl	Age >12 and <=24	Normal	Short	?
14	A49	Boy	Age >12 and <=24	Very thin	Very Short	?
15	A50	Boy	Age >12 and <=24	Thin	Normal	?

The classification results are not yet known because these variables are predicted results from data calculated from each variable. Here is the calculation process:

Sample (Status) = Gender * Age (0-24) * Weight *Hight

A36 (Good Nutrition) = Boy * Age >12 and <=24 * Very thin * Short
= 0,70 * 0,70 * 0,22 * 0,41
= 0,024

A36 (Malnutrition) = Boy * Age >12 and <=24 * Very thin * Short
= 0,65 * 0,57 * 0,52 * 0,43
= 0,038

A37 (Good Nutrition) = Boy * Age >12 and <=24 * Thin * Short
= 0,70 * 0,70 * 0,11 * 0,41
= 0,012
A37 (Malnutrition) = Boy * Age >12 and <=24 * Thin * Short
= 0,65 * 0,57 * 0,39 * 0,43
= 0,029

A38 (Good Nutrition) = 0,002
A38 (Malnutrition) = 0,042
A39 (Good Nutrition) = 0,001
A39 (Malnutrition) = 0,033
A40 (Good Nutrition) = 0,001
A40 (Malnutrition) = 0,023
.
.
.
A50 (Good Nutrition) = 0,017
A50 (Malnutrition) = 0,006

Table 8. Calculation Results of Each Variable

No	Sample	Gender	Age (0 - 24)	Height	Hight	Good Nutrition	Malnutrition
1	A36	Boy	Age >12 and <=24	Very thin	Short	0,024	0,038
2	A37	Boy	Age >12 and <=24	Thin	Short	0,012	0,029
3	A38	Boy	Age >12 and <=24	Very thin	Very Short	0,002	0,042
4	A39	Boy	Age <=12	Very thin	Very Short	0,001	0,033
5	A40	Girl	Age >12 and <=24	Very thin	Very Short	0,001	0,023
6	A41	Girl	Age <=12	Very thin	Normal	0,006	0,003
7	A42	Boy	Age <=12	Very thin	Normal	0,014	0,006
8	A43	Boy	Age >12 and <=24	Normal	Short	0,052	0,003
9	A44	Boy	Age <=12	Thin	Short	0,005	0,022
10	A45	Boy	Age >12 and <=24	Normal	Short	0,052	0,003
11	A46	Boy	Age >12 and <=24	Thin	Short	0,012	0,029
12	A47	Girl	Age >12 and <=24	Thin	Short	0,005	0,015
13	A48	Girl	Age >12 and <=24	Normal	Short	0,022	0,002
14	A49	Boy	Age >12 and <=24	Very thin	Very Short	0,002	0,042
15	A50	Boy	Age >12 and <=24	Thin	Normal	0,017	0,006

d. Compare Good Nutrition and Malnutrition Class Results
From the results of these calculations, compare the probability values between class (P|Good Nutrition) and class (P|Malnutrition). The highest probability value can be seen in the following table 9:

Table 9. Comparison Results

Sample	Gender	Age (0 - 24)	Weight	Height	Nutrition Status	Prediction	Class Prediction	Good Nutrition	Malnutrition
A36	Boy	Age >12 and <=24	Very thin	Short	Good Nutrition		Malnutrition	0,024	0,038
A37	Boy	Age >12 and <=24	Thin	Short	Malnutrition		Malnutrition	0,012	0,029
A38	Boy	Age >12 and <=24	Very thin	Very Short	Good Nutrition		Malnutrition	0,002	0,042
A39	Boy	Age <=12	Very thin	Very Short	Malnutrition		Malnutrition	0,001	0,033
A40	Girl	Age >12 and <=24	Very thin	Very Short	Malnutrition		Malnutrition	0,001	0,023
A41	Girl	Age <=12	Very thin	Normal	Good Nutrition		Good Nutrition	0,006	0,003
A42	Boy	Age <=12	Very thin	Normal	Good Nutrition		Good Nutrition	0,014	0,006
A43	Boy	Age >12 and <=24	Normal	Short	Good Nutrition		Good Nutrition	0,052	0,003
A44	Boy	Age <=12	Thin	Short	Malnutrition		Malnutrition	0,005	0,022
A45	Boy	Age >12 and <=24	Normal	Short	Good Nutrition		Good Nutrition	0,052	0,003
A46	Boy	Age >12 and <=24	Thin	Short	Malnutrition		Malnutrition	0,012	0,029
A47	Girl	Age >12 and <=24	Thin	Short	Malnutrition		Malnutrition	0,005	0,015
A48	Girl	Age >12 and <=24	Normal	Short	Good Nutrition		Good Nutrition	0,022	0,002
A49	Boy	Age >12 and <=24	Very thin	Very Short	Malnutrition		Malnutrition	0,002	0,042
A50	Boy	Age >12 and <=24	Thin	Normal	Good Nutrition		Good Nutrition	0,017	0,006

Table 10. Confusion Table

	Class	
Prediction	Good Nutrition	Malnutrition
Good Nutrition	6	0
Malnutrition	2	7

Accuracy Count = Prediction (Good Nutrition, Good Class) + Prediction (Malnutrition, Poor Class)/ Total Number of Predictions (Good Nutrition, Good Class) and Prediction (Malnutrition, Poor Class)

$= (6 + 7) / 15$

$= 0{,}87$ atau 87%

From the results of these calculations, comparing the probability value between the class (P|Good Nutrition) and the class (P|Malnutrition) the highest probability value can be concluded that the Nutritional Status has an accuracy of 87%.

4. CONCLUSION

Based on the results of the research conducted, it can be concluded that applying Data Mining using the Naïve Bayes algorithm on the classification of children's nutrition at the Simalungun Health Office can be applied properly. The source of data processing used in this study is data obtained directly at the Simalungun Health Office. The amount of data in this test is 15 samples of data tested from the comparison of Good Nutrition and Malnutrition. From the calculation results of the Naive Bayes Algorithm, the classification with Good Nutrition and Malnutrition is 6 Good Nutrition and 7 Malnutrition. Nutrition as many as 7 records. It can be concluded that the Simalungun Health Service provides good service to the community.

ACKNOWLEDGEMENTS

Thank you to the supervisors as lecturers at AMIK and STIKOM Tunas Bangsa Pematangsiantar who have assisted in the preparation of this research so that it has reached the publication stage. Thanks to all parties, especially Simalungun Health Office and it is hoped that this research can be used as a source of data that can assist in decision making. For the development and progress of this research title, we expect constructive input and criticism so that this research can be continued.

REFERENCES

[1] World Health Organization, "Early Childhood: Development and Disability," *PsycCRITIQUES*, vol. 33, no. 8, p. 269, 2012, [Online]. Available: http://content.apa.org/reviews/025945.

[2] A. L. Costa, "Teaching and Assessing Habits of Mind," vol. 96741, no. 808, 1999.

[3] R. H. Lestari, M. Mudhawaroh, and M. Ratnawati, "Intelligence Optimization in the Golden Age by Stimulating the Right-Brain in Mojokrapak Village, Tembelang District, Jombang Regency," *Nucleus*, vol. 1, no. 2, pp. 58–61, 2020, doi: 10.37010/nuc.v1i2.166.

[4] A. Chapnick, "The golden age," *Int. J.*, vol. 64, no. 1, pp. 205–221, 2008, doi: 10.1177/002070200906400118.

[5] M. Muslim, H. Ahmad, and S. Rahim, "The effect of emotional, spiritual and intellectual intelligence on auditor professionalism at the inspectorate of South Sulawesi Province," *Indones. Account. Rev.*, vol. 9, no. 1, p. 73, 2019, doi: 10.14414/tiar.v9i1.1416.

[6] UNICEF., *Children, food and nutrition : growing well in a changing world.* 2019.

[7] I. G. Pratiwi and D. A. Restanty, "Pengaruh Metode Think, Pair and Share Terhadap Keterampilan Kader Dalam Pengisian Kms Balita," *JPP (Jurnal Kesehat. Poltekkes Palembang)*, vol. 13, no. 2, pp. 128–135, 2019, doi: 10.36086/jpp.v13i2.237.

[8] W. Yulianti and Salmidi, "Metode Rough Set untuk Menganalisa Problematika Guru Dalam Menggunakan Media Pembelajaran Berbasis Komputer," *J. Teknol. dan Sist. Inf. Univrab*, vol. 1, no. 1, pp. 19–25, 2016.

[9] Y. Cheng, K. Chen, H. Sun, Y. Zhang, and F. Tao, "Data and Knowledge Mining with Big Data towards Smart Production," *J. Ind. Inf. Integr.*, vol. 9, Sep. 2017, doi: 10.1016/j.jii.2017.08.001.

[10] H. Tran, *Survey of Machine Learning and Data Mining Techniques used in Multimedia System*. 2019.

[11] R. Amanda and E. S. Negara, "Analysis and Implementation Machine Learning for YouTube Data Classification by Comparing the Performance of Classification Algorithms," *J. Online Inform.*, vol. 5, no. 1, pp. 61–72, 2020, doi: 10.15575/join.v5i1.505.

[12] S. Suhada and E. Setiawan, "Classification Needs Teachers," no. Icisbc, pp. 408–414, 2013.

[13] A. W. Syaputri, E. Irwandi, and M. Mustakim, "Naïve Bayes Algorithm for Classification of Student Major's Specialization," *J. Intell. Comput. Heal. Informatics*, vol. 1, no. 1, p. 17, 2020, doi: 10.26714/jichi.v1i1.5570.

[14] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 53, 2017, doi: 10.24843/ijeet.2017.v02.i01.p11.

[15] M. J. Sodiq and E. I. Sela, "Perbandingan Metode Naive Bayes Dan K-Nearest Neighbor Pada Klasifikasi Kualitas Udara Di Dki Jakarta," 2019.

[16] R. Rino, "The Comparison of Data Mining Methods Using C4.5 Algorithm and Naive Bayes in Predicting Heart Disease," *Tech-E*, vol. 4, no. 2, p. 44, 2021, doi: 10.31253/te.v4i2.543.

[17] Mustakim *et al.*, "Data Sharing Technique Modeling for Naive Bayes Classifier for Eligibility Classification of Recipient Students in the Smart Indonesia Program," *J. Phys. Conf. Ser.*, vol. 1424, no. 1, 2019, doi: 10.1088/1742-6596/1424/1/012009.

[18] C. Fadlan, S. Ningsih, and A. Windarto, "PENERAPAN METODE NAÏVE BAYES DALAM KLASIFIKASI KELAYAKAN KELUARGA PENERIMA BERAS RASTRA," *J. Tek. Inform. Musirawas*, vol. 3, p. 1, Jun. 2018, doi: 10.32767/jutim.v3i1.286.

[19] M. Collins, "Research methodology," *J. Occup. Heal. Saf. - Aust. New Zeal.*, vol. 6, no. 5, p. 352, 1990.

[20] S. Jyothi and P. Bhargavi, "Applying Naive Bayes Data Mining Technique for Classification of

Agricultural Land Soils," *57. P.Bhargavi , S. Jyothi*, vol. 9, Aug. 2009.

[21] M. Tigor, "Bagaimana Menyatakan Permasalahan Riset (Research Problem/Problem Statement)?," *Rudang Mayang Publ.*, vol. 1, no. 1, pp. 1–2, 2020