

Analysis of the Naïve Bayes Method for Determining Social Assistance Eligibility Public

Adinda Pratiwi Siregar^{1)*}, Deci Irmayani²⁾, Mila Nirmala Sari³⁾

^{1,2,3)}Universitas Labuhanbatu,

¹⁾ adindapратиwi42183@gmail.com, ²⁾ deacyirmayani@gmail.com,

³⁾ milanirmalasari7@gmail.com

Submitted : Mar 15, 2023 | Accepted : Mar 27, 2023 | Published : Apr 1, 2023

Abstract: Economic needs are community needs that are used to meet daily needs. Therefore, economic needs are very important for the life of every society. There is a gap in the economic needs of the community, the government created a social assistance program which is assistance provided to the community in the form of cash or non-cash. The help is made for welfare society from inequality, especially economic inequality. So researchers will carry out a data classification of people who are eligible for social assistance. The classification will be carried out using the Naïve Bayes method. The Naïve Bayes method is a simple classification method for calculating the probability of a combination of certain data. The data to be used by researchers is community data as much as 62 community data. research done by using the Naïve Bayes method aims to classify community data that is feasible to forget social assistance. The first stage of this classification is the process of collecting community data and determining community data that will be used as a filtered sample cleaned, furthermore preprocessing data and then designing the Naïve Bayes Algorithm model. The results of data classification using the Naïve Bayes method show that the number of people who are eligible for social assistance is 14 community data and people who are not eligible for social assistance are 48 community data. These results can be a reference for determining the eligibility of the community to receive social assistance.

Keywords: Confusion Matrix, Data Mining, Naïve Bayes, Orange, Roc Analysis, Social Assistance

INTRODUCTION

Sosopan Hamlet is an area located in Kotapinang District, South Labuhanbatu Regency, North Sumatra Province. Before we go deeper into Sosopan Village, we will first discuss the history of Kotapinang District. Initially Kotapinang District had the name Pinang Awan Sultanate. The Pinang Awan Sultanate was founded by Batara Sinomba (Batara Gurga) Pinayungan Tuanku Raja Nan Sakti, he was the son of Sultan Alamsyah Syaifuddin who came from the Pagaruyung Kingdom. Sosopan Village is one of the villages located in Kotapinang District. Sosopan Village has several hamlets. One of them is Dusun Sosopan. Dusun Sosopan is an area that is visited by many people, both from within the city and outside the city. This happened because the Regent's office was located in Sosopan Hamlet. Until now, more and more developments have taken place in Sosopan Hamlet. But sometimes there are still many people who lack economic needs.

Economic need is a definite need which is one of the factors to keep living life. Many people lack economic needs, hampering their health. Because sometimes, due to a lack of economic needs, people will reduce their portion of food and that's what causes them to get sick. Therefore the government has

*name of corresponding author



created a social assistance program for the less fortunate. This program was created to reduce the level of poverty in Indonesia. This program has been made since 2007 and is well known in the international world with the term Conditional Cash Transfers (CCT). The existence of this social assistance program has proven to be successful and can reduce the level of poverty in Indonesia. For social assistance programs, in fact there are still many that can be used as parameters for people who deserve assistance, namely below average income, age already reaching the elderly, namely ± 70 years, disability, below average IQ, breastfeeding mothers, below average salary, and people who have cards (Noerkaisar, 2021).

But in this study the author will make a study of people who are eligible to get assistance with the parameters being for people who have a Social Protection Card (KPS), Smart Indonesia Card (KIP), Prosperous Family Card (KKS), and finally the Program Hope Family (PKH). This research was made in order to provide assistance to people who deserve help socially. basically the state should be able to prosper its people. Because welfare can ease the burden on society in terms of the economy (Susanto, Puspaningrum, & Neneng, 2021) (Rahmansyah, Qadri, Sakti, & Ikhsan, 2020). Assistance given to the community is an allocation of funds from the government which is given to people who are deficient in an economic sense (Fadhli & Fahimah, 2021). Therefore this research needs to be made in order to provide appropriate data for people who are eligible for social assistance. This research was not done manually, but was done with the help of an orange application, in this application community data will later be entered into the data mining process to classify community data that is feasible and not eligible for social assistance using the naïve Bayes method.

Data mining is a technique for extracting large amounts of knowledge using a combination of statistics and mathematics (Uçar & Karahoca, 2021). Data mining has been widely used in various ways, such as risk analysis, interest, consumer satisfaction (Uçar & Karahoca, 2021). Data mining is part of the analysis which is an activity to check such as Classification, cluster and associations (Hassan, El-Mashade, & Aboshosha, 2022). Many text data mining techniques can be used by combining several techniques to obtain logical and valuable results (Yassir et al., 2020). In this mining process, we will classify the data later. Data mining techniques have been widely used in several applications, but this time we are using an orange application to perform data mining (Patil & Tamane, 2018). Data Classification is grouping data by differentiating and dividing the data into different groups according to their respective provisions. In this case to do a data classification, we will use the Naïve Bayes method. This method will classify data according to predetermined categories.

METHOD

The naïve Bayes method is a simple classification method for calculating the probability of a combination of certain data (Murwantara, Yugopuspito, & Hermawan, 2020). This method is structured based on certain accuracy (Santoso et al., 2020). This method is used to classify data with simple probability and is designed to be used with assumptions between variables using statistical and mathematical techniques. Statistical techniques will later be used to calculate certain probabilities according to existing assumptions. This naïve Bayes method will later be applied using the orange application. The orange application will be used to design widgets with the rules of the Naïve Bayes method. This is because each method has its own processing technique. Therefore, to carry out this Classification, we must adapt the method to the contents of the pattern or widget, which we will later apply to the Orange application.

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$

Information:

A : hypothesis of data A (specific class)

B : data with unknown classes

$P(A | B)$: Probability of hypothesis based on condition B

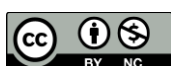
$P(A)$: Probability of hypothesis A

$P(B | A)$: Probability B when condition A

$P(B)$: Probability

(Di & Duan, 2014)

*name of corresponding author



The naïve Bayes method is a probability by calculating other related probabilities. After being applied to the Naive Bayes algorithm, this formula produces a basic assumption. In looking at a feature, this algorithm always assumes that the feature is independent, equal, and has a contribution to the result. The way it works is that we try to find the probability of event A, if event B is true. Event B is also referred to as evidence. For P(A) is the a priori of A (the prior probability, i.e. the probability of the event before the proof is seen) and the proof is the attribute value of the unknown instance (event B). For P(A|B) is the posteriori probability of B, that is, the probability that it will occur after the evidence is shown.

Confusion Matrix

Confusion matrix is an easy and effective tool to use to show the performance of a Classification and very easy to use to determine the results (Yun, 2021). Confusion matrix can be used to evaluate the work results of a model and can be used to determine the results of a data mining. The confusion matrix has several calculations, namely as follows.

Table 1. Confusion Matrix

Confusion Matrix	True Class (Actual)		
		P	N
Predicted class	Y	True Positive (TP)	False Positive (FP)
	N	False Negative (FN)	True Negative (TN)

To determine the calculation of the confusion matrix, we can do it by calculating accuracy, precision and recall.

Accuracy is calculated with the condition that the prediction number (TP + TN) is divided by the number of samples available. To calculate accuracy, the following formula can be seen (Yun, 2021) :

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Precision is used to identify positive cases with a high false positive rate, which can be calculated as follows (Normawati & Prayogi, 2021) :

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%$$

In contrast to precision, recall serves to identify positive cases with high false negative values. Recall can be calculated in this way (Agustina, Adrian, & Hermawati, 2021) :

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

RESULT

Data Analysis

In the picture below is the data from the people of Dusun Sosopan which will be used as a sample to carry out a data mining classification using the naïve Bayes method. In the sample data, the eligibility qualifications for people who deserve social assistance are people who have a Social Protection Card (KPS), Smart Indonesia Card (KIP), Prosperous Family Card (KKS), and the last is the Family Hope Program (PKH).). The card is a card given by the government to help less fortunate people, especially in the community's economy. Therefore, people who have these four cards deserve assistance, be it cash or groceries.

*name of corresponding author



Name	Gender	KIP	KPS	KKS	PKH	Address	Category
P1	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P2	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P3	Man	No	No	No	No	Sosopan	Not Feasible
P4	Man	No	No	No	No	Sosopan	Not Feasible
P5	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P6	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P7	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P8	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P9	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P10	Woman	Yes	Yes	Yes	Yes	Desa Sosopan	Feasible
P11	Man	No	No	No	No	Sosopan	Not Feasible
P12	Man	No	No	No	No	Jl. Pam - Kotapinang	Not Feasible
P13	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P14	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P15	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P16	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P17	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible
P18	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P19	Woman	Yes	Yes	Yes	Yes	Gonting Gajah	Feasible
P20	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P21	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P22	Man	No	No	No	No	Sosopan	Not Feasible
P23	Man	No	No	No	No	Sosopan	Not Feasible
P24	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P25	Man	No	No	No	No	Jl. Bukit	Not Feasible
P26	Man	No	No	No	No	Sosopan	Not Feasible
P27	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P28	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P29	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P30	Man	No	No	No	No	Sosopan	Not Feasible
P31	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P32	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P33	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P34	Man	No	No	No	No	Dsn Sosopan	Not Feasible
P35	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P36	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P37	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P38	Woman	No	No	No	No	Perum Bangun Jadi	Not Feasible
P39	Man	No	No	No	No	Sosopan Dusun I	Not Feasible
P40	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P41	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P42	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P43	Man	No	No	No	No	Dusun Bangun Jadi	Not Feasible
P44	Man	No	No	No	No	Dusun Sosopan	Not Feasible

*name of corresponding author



P45	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P46	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P47	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P48	Man	No	No	No	No	Sosopan	Not Feasible
P49	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible
P50	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P51	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P52	Man	No	No	No	No	Sosopan	Not Feasible
P53	Man	No	No	No	No	Sosopan	Not Feasible
P54	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P55	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P56	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P57	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P58	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P59	Man	No	No	No	No	Sosopan	Not Feasible
P60	Man	No	No	No	No	Sosopan	Not Feasible
P61	Man	No	No	No	No	Sosopan	Not Feasible
P62	Man	No	No	No	No	Sosopan	Not Feasible

Figure 1. People Data

In Figure 1, the table data above is data from the Sosopan Hamlet community obtained from the Sosopan Dusun Kadus. The data was taken from one of the hamlets in Sosopan Village, namely data from the Sosopan Hamlet Community. The data obtained as many as 62 community data. The picture above consists of community names, gender, Social Protection Card (KPS) cards, Smart Indonesia Cards (KIP), Prosperous Family Cards (KKS), Family Hope Program (PKH) and the last is the address.

Table 2. Attributes of community data

No	Attribute	Type	Descriptions
1	Name	Text	Community full name
2	Gender	Category	Gender Society
3	KIP	Category	Student identification number
4	KPS	Category	Value of student knowledge
5	KKS	Category	Assess students' skills
6	PKH	Category	Assess students' attitudes
7	Address	Category	Community address

In table 2, the research attributes are data that has been obtained from the Sosopan Dusun Kadus. The attribute data is equipped with the type and description of each attribute.

Training Data

The training data is the data that we will use as our research sample, we have obtained this data from the Sosopan Hamlet Kadus. Initially, this data was only a family card that was compiled and entered into the Microsoft Excel application with the .xlsx file format. the data that we have obtained will be an attribute to determine the eligibility of the community to get social assistance.

Name	Gender	KIP	KPS	KKS	PKH	Address	Category
P1	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P2	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P3	Man	No	No	No	No	Sosopan	Not Feasible

*name of corresponding author



P4	Man	No	No	No	No	Sosopan	Not Feasible
P5	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P6	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P7	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P8	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P9	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P10	Woman	Yes	Yes	Yes	Yes	Desa Sosopan	Feasible
P11	Man	No	No	No	No	Sosopan	Not Feasible
P12	Man	No	No	No	No	Jl. Pam - Kotapinang	Not Feasible
P13	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P14	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P15	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P16	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P17	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible
P18	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P19	Woman	Yes	Yes	Yes	Yes	Gonting Gajah	Feasible
P20	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P21	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P22	Man	No	No	No	No	Sosopan	Not Feasible
P23	Man	No	No	No	No	Sosopan	Not Feasible
P24	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P25	Man	No	No	No	No	Jl. Bukit	Not Feasible
P26	Man	No	No	No	No	Sosopan	Not Feasible
P27	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P28	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P29	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P30	Man	No	No	No	No	Sosopan	Not Feasible
P31	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P32	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P33	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P34	Man	No	No	No	No	Dsn Sosopan	Not Feasible
P35	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P36	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P37	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P38	Woman	No	No	No	No	Perum Bangun Jadi	Not Feasible
P39	Man	No	No	No	No	Sosopan Dusun I	Not Feasible
P40	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P41	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P42	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P43	Man	No	No	No	No	Dusun Bangun Jadi	Not Feasible
P44	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P45	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P46	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P47	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P48	Man	No	No	No	No	Sosopan	Not Feasible
P49	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible

*name of corresponding author



P50	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P51	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P52	Man	No	No	No	No	Sosopan	Not Feasible
P53	Man	No	No	No	No	Sosopan	Not Feasible
P54	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P55	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P56	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P57	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P58	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P59	Man	No	No	No	No	Sosopan	Not Feasible
P60	Man	No	No	No	No	Sosopan	Not Feasible
P61	Man	No	No	No	No	Sosopan	Not Feasible
P62	Man	No	No	No	No	Sosopan	Not Feasible

Figure 2. Data Training

Figure 2 contains the data of the Sosopan Hamlet community that we have obtained from the Sosopan Dusun Kadus and already contains all the data we will need in the process of selecting the eligibility of the community to receive social assistance.

Table 3. People Data Columns

No	Attribute	Type	Role	Values
1	Nama	Text	Meta	
2	Gender	Categorical	Feature	Man, Woman
3	KIP	Categorical	Feature	No, Yes
4	KPS	Categorical	Feature	No, Yes
5	KKS	Categorical	Feature	No, Yes
6	PKH	Categorical	Feature	No, Yes
7	Address	Categorical	Feature	Dusun Sosopan, Perum Pulomas, Sosopan
8	Category	Categorical	Target	Feasible, Not Feasible

In table 3 above are column data from training data that will be used to determine the eligibility classification of the community to get social assistance. By using the Naïve Bayes method, namely by changing the role from feature to target. This is done so that we find appropriate results.

Data Selection Process (Preprocessing)

The data selection process is a process to determine the data needed or to be used to carry out a data classification. The data selection process will select the data that will be needed before entering the data mining process that will be carried out withuse naïve bayes method. So in this preprocessing, the data that will be entered in data mining is name, gender, Social Protection Card (KPS), Smart Indonesia Card (KIP), Prosperous Family Card (KKS), Family Hope Program (PKH) and lastly the address. The Social Protection Card (KPS), Smart Indonesia Card (KIP), Prosperous Family Card (KKS), and Family Hope Program (PKH) are cards that are owned by every community that will become research parameters that will be carried out using the orange application.

Data Mining Process

The data mining process is carried out using a classification model using the orange application using the Naïve Bayes method to determine the eligibility of the community to receive social assistance in Sosopan Hamlet.

*name of corresponding author



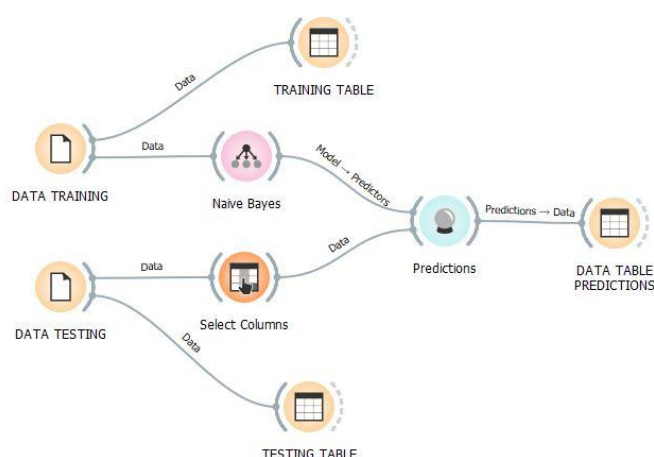


Figure 3. Data Mining Process

In figure 3, the data mining process is designing a classification pattern using the orange application. This is done using the Naïve Bayes method, namely the Classification method which will classify people who are eligible and not eligible for social assistance. In the picture above there is a curved line that connects widget 1 with other widgets. The line is curved because in 1 widget it is connected to 2 widgets, meaning that the line can only be straight when 1 widget is connected to 1 widget as well. So if 1 widget is connected with 2 widgets, the line will be curved.

Model Classification Testing Process

In the process of testing the Naïve Bayes method is an algorithm that is used to carry out a classification that must have training data and test data. This was done to find the results of the data being tested. Like the picture below. Below is a data mining process that was carried out to determine and find the results of the Classification which was carried out using the naïve Bayes method (Watratan, B, Moeis, Informasi, & Makassar, 2020).

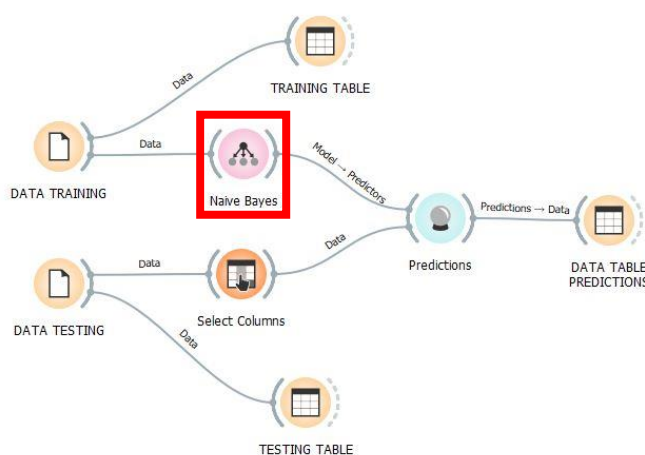


Figure 4. Classification model widget design dataset social assistance receipt status

In Figure 4. The widget classification model in the image above is a widget that will determine the eligibility of the community to get assistance. The widget in the red box is the method used to determine the eligibility of the people of Dusun Sosopan to receive social assistance. The method is Naïve Bayes which is a Classification method.

*name of corresponding author



Classification Model Prediction Process

This process is a process for making a prediction with data classification. Prediction is a function that is performed to determine certain patterns (Damuri, Riyanto, Rusdianto, & Aminudin, 2021). We will use this pattern to carry out a data classification.

Name	Gender	KIP	KPS	KKS	PKH	Address	Category
P1	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P2	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P3	Man	No	No	No	No	Sosopan	Not Feasible
P4	Man	No	No	No	No	Sosopan	Not Feasible
P5	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P6	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P7	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P8	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P9	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P10	Woman	Yes	Yes	Yes	Yes	Desa Sosopan	Feasible
P11	Man	No	No	No	No	Sosopan	Not Feasible
P12	Man	No	No	No	No	Jl. Pam - Kotapinang	Not Feasible
P13	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P14	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P15	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P16	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P17	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible
P18	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P19	Woman	Yes	Yes	Yes	Yes	Gonting Gajah	Feasible
P20	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P21	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P22	Man	No	No	No	No	Sosopan	Not Feasible
P23	Man	No	No	No	No	Sosopan	Not Feasible
P24	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P25	Man	No	No	No	No	Jl. Bukit	Not Feasible
P26	Man	No	No	No	No	Sosopan	Not Feasible
P27	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P28	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P29	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P30	Man	No	No	No	No	Sosopan	Not Feasible
P31	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P32	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P33	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P34	Man	No	No	No	No	Dsn Sosopan	Not Feasible
P35	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P36	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P37	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P38	Woman	No	No	No	No	Perum Bangun Jadi	Not Feasible
P39	Man	No	No	No	No	Sosopan Dusun I	Not Feasible
P40	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible

*name of corresponding author



P41	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P42	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P43	Man	No	No	No	No	Dusun Bangun Jadi	Not Feasible
P44	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P45	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P46	Woman	Yes	Yes	Yes	Yes	Sosopan	Feasible
P47	Man	No	No	No	No	Perumahan Pulo Mas	Not Feasible
P48	Man	No	No	No	No	Sosopan	Not Feasible
P49	Man	No	No	No	No	Perum Mitra B. Jadi	Not Feasible
P50	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P51	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P52	Man	No	No	No	No	Sosopan	Not Feasible
P53	Man	No	No	No	No	Sosopan	Not Feasible
P54	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P55	Woman	Yes	Yes	Yes	Yes	Dsn Sosopan	Feasible
P56	Woman	No	No	No	No	Dusun Sosopan	Not Feasible
P57	Woman	Yes	Yes	Yes	Yes	Dusun Sosopan	Feasible
P58	Man	No	No	No	No	Dusun Sosopan	Not Feasible
P59	Man	No	No	No	No	Sosopan	Not Feasible
P60	Man	No	No	No	No	Sosopan	Not Feasible
P61	Man	No	No	No	No	Sosopan	Not Feasible
P62	Man	No	No	No	No	Sosopan	Not Feasible

Figure 5. Classification Model Prediction Results

In Figure 5, which is the result of Classification using the Naïve Bayes method, it states that the Naïve Bayes method is suitable when used to carry out a Classification. The results obtained for people who are eligible for social assistance are 14 community data (for a representation of 22.58%) and for people who are not eligible for social assistance are 48 community data (for a representation of 77.42%).

Classification Model Evaluation Results

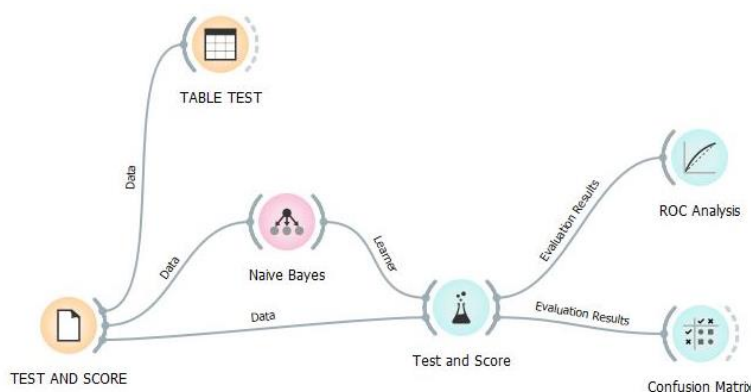


Figure 6. Classification Evaluation Widg

t

Figure 6 contains a classification evaluation consisting of a widget to determine test scores and scores. Confusion Matrix and ROC Analysis will be applied using the Naïve Bayes method. To get the Confusion Matrix and ROC Analysis, namely by using training data and test data that have been combined into test data with 1 attribute as a target. 1 attribute as text, namely the name of the community, 6 attributes as a category, namely gender, address, Social Protection Card (KPS), Smart Indonesia Card

*name of corresponding author



(KIP), Prosperous Family Card (KKS), and the last is the Family Hope Program (PKH). Then the test results and scores will be obtained in table 4.

Tabel 4, Result of Test and Score

Model	AUC	CA	F1	Precision	Recall
Naïve Bayes	1.000	1.000	1.000	1.000	1.000

In table 4 above are the test results and scores from the results of calculations from the 62 community data that have been tested, so we will get results from an AUC of 1,000, a CA of 1,000, F1 of 1,000, Precision of 1,000, Recall of 1,000. These results were obtained using the Naïve Bayes method which was operated on the orange application.

Evaluation Results with Confusion Matrix

Confusion Matrix is a measuring tool for prediction methods by calculating the correctness of data that has been classified using the Naïve Bayes method.

Tabel 5, Result of Confusion Matrix

		Predicted		Σ
		Feasible	Not Feasible	
Actual	Feasible	14	0	14
	Not Feasible	0	48	48
Σ		14	48	62

In table 5 above the results of the True Positive (TP) are 14. True Negative (TN) is 48, False Positive (FP) is 0 and False Negative (FN) is 0. Then the values for accuracy, precision and recall are as follows:

Accuracy = $\frac{14+48}{14+48+0+0} \times 100\%$ Then the accuracy value = 100%

Precision = $\frac{14}{14+0} \times 100\%$ Then the precision value = 100%

Recall = $\frac{14}{14+0} \times 100\%$ Then the recall value = 100%

Evaluation Results with ROC Curve

The Roc Curve is obtained from the true signal (sensitivity) and (1 specificity) over the entire cut off point range to obtain the ROC curve visualized from the Confusion Matrix.

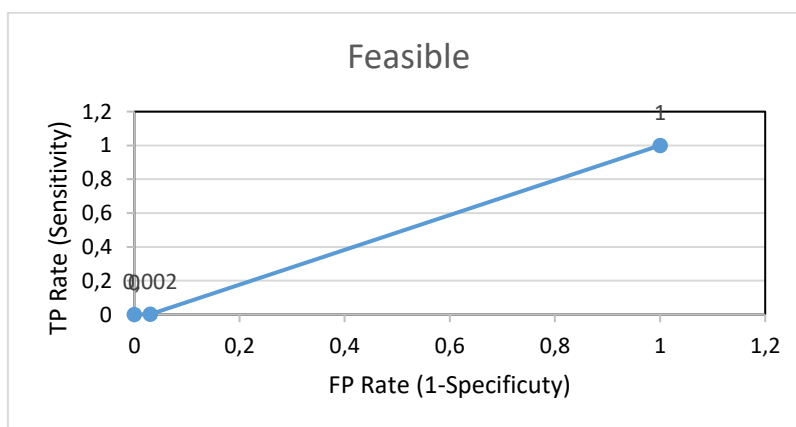


Figure 7. ROC Analysis Targeting Social Assistance Eligible People

*name of corresponding author



Figure 7 states that the results of the ROC Analysis of the eligibility of the people of Dusun Sosopan to qualify for social assistance, were carried out using the Naïve Bayes method with a result of 0.002.

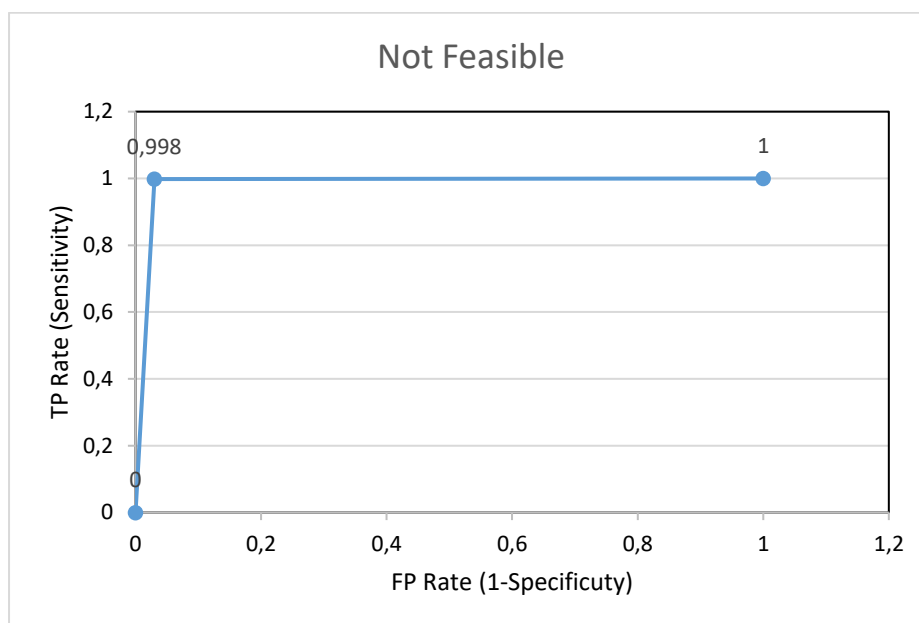


Figure 8. ROC analysis targeting people who are not eligible for social assistance

Figure 8 states that the results of the ROC Analysis of the eligibility of the people of Dusun Sosopan not to be eligible for social assistance, were carried out using the Naïve Bayes method with a result of 0.998.

DISCUSSIONS

This research was conducted using the Naïve Bayes method to classify data on people who are eligible for social assistance. In accordance with government programs that aid made In order to reduce the level of poverty, a data classification of people who are eligible for assistance is therefore made in order to avoid mistakes in selecting people who are eligible for social assistance. The first process that must be carried out is to know the flow of the research to be carried out, starting from data collection to get Classification results, which can be in the form of values or data visualization. The data used as research samples were 62 community data. The data used to carry out a Classification is divided into 2 data, namely training data and testing data. After classifying withuse the naïve Bayes method, the results of the accuracy of the test and score are 100% and the results of the accuracy of the confusion matrix are 100%. The comparison between the results of the test and score with the confusion matrix is 1: 1 meaning that both get perfect results. The conclusion is that the naïve Bayes method has succeeded in being the best method for conducting due diligence by classifying training data and testing data. After obtaining the results of the Classification of eligibility data for people who are eligible to get social assistance, the data can be used as a reference for people who are eligible for social assistance.

CONCLUSION

This research was conducted in order to determine or classify community data that is eligible for social assistance. Lots of people who should get social assistance, but in reality they do not get it. On the other hand, there are also many who shouldn't be able to, but they do get social assistance. Therefore we carry out a Classification to determine people who are eligible for social assistance. This will minimize the occurrence of errors in determining the eligibility of the community to receive social assistance. The Naïve Bayes algorithm is a suitable method for carrying out this classification. This method can classify and group data according to its category.

*name of corresponding author



REFERENCES

- Agustina, N., Adrian, A., & Hermawati, M. (2021). Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media. *Faktor Exacta*, 14(4), 1979–276. <https://doi.org/10.30998/faktorexacta.v14i4.11259>
- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 219. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Di, P., & Duan, L. (2014). New naive Bayes text classification algorithm. *Shuju Caiji Yu Chuli/Journal of Data Acquisition and Processing*, 29(1), 71–75. <https://doi.org/10.11591/telkomnika.v12i2.4180>
- Fadhli, K., & Fahimah, D. A. N. (2021). Pengaruh Pendapatan, Pendidikan, Dan Gaya Hidup Terhadap Kesejahteraan Keluarga Penerima Manfaat (Kpm) Bantuan Sosial Covid-19. *Jurnal Education and Development*, 9(3), 118–124.
- Hassan, A. M., El-Mashade, M. B., & Aboshosha, A. (2022). Deep learning for cancer tumor classification using transfer learning and feature concatenation. *International Journal of Electrical and Computer Engineering*, 12(6), 6736–6743. <https://doi.org/10.11591/ijece.v12i6.pp6736-6743>
- Murwantara, I. M., Yugopuspito, P., & Hermawan, R. (2020). Comparison of machine learning performance for earthquake prediction in Indonesia using 30 years historical data. *Telkomnika (Telecommunication Computing Electronics and Control)*, 18(3), 1331–1342. <https://doi.org/10.12928/TELKOMNIKA.v18i3.14756>
- Noerkaisar, N. (2021). Efektivitas Penyaluran Bantuan Sosial Pemerintah untuk Mengatasi Dampak Covid-19 di Indonesia. *Jurnal Manajemen Perbendaharaan*, 2(1), 83–104. <https://doi.org/10.33105/jmp.v2i1.363>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711. Retrieved from <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- Patil, R., & Tamane, S. (2018). A comparative analysis on the evaluation of classification algorithms in the prediction of diabetes. *International Journal of Electrical and Computer Engineering*, 8(5), 3966–3975. <https://doi.org/10.11591/ijece.v8i5.pp3966-3975>
- Rahmansyah, W., Qadri, R. A., Sakti, R. R. A., & Ikhsan, S. (2020). Pemetaan Permasalahan Penyaluran Bantuan Sosial Untuk Penanganan Covid-19 Di Indonesia. *Jurnal Pajak Dan Keuangan Negara (PKN)*, 2(1), 90–102. <https://doi.org/10.31092/jpkn.v2i1.995>
- Santoso, H. A., Rachmawanto, E. H., Nugraha, A., Nugroho, A. A., Setiadi, D. R. I. M., & Basuki, R. S. (2020). Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization. *Telkomnika (Telecommunication Computing Electronics and Control)*, 18(2), 799–806. <https://doi.org/10.12928/TELKOMNIKA.V18I2.14744>
- Susanto, E. R., Puspaningrum, A. S., & Neneng, N. (2021). Model Rekomendasi Penerima Bantuan Sosial Berdasarkan Data Kesejahteraan Rakyat. *Jurnal Tekno Kompak*, 15(1), 1. <https://doi.org/10.33365/jtk.v15i1.915>
- Uçar, T., & Karahoca, A. (2021). Benchmarking data mining approaches for traveler segmentation. *International Journal of Electrical and Computer Engineering*, 11(1), 409–415. <https://doi.org/10.11591/ijece.v11i1.pp409-415>
- Watraton, A. F., B, A. P., Moeis, D., Informasi, S., & Makassar, S. P. (2020). Implementation of the Naive Bayes Algorithm to Predict the Spread of Covid-19 in Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7–14.
- Yassir, A. H., Mohammed, A. A., Alkhazraji, A. A. J., Hameed, M. E., Talib, M. S., & Ali, M. F. (2020). Sentimental classification analysis of polarity multi-view textual data using data mining techniques. *International Journal of Electrical and Computer Engineering*, 10(5), 5526–5534. <https://doi.org/10.11591/IJECE.V10I5.PP5526-5534>
- Yun, H. (2021). Prediction model of algal blooms using logistic regression and confusion matrix. *International Journal of Electrical and Computer Engineering*, 11(3), 2407–2413. <https://doi.org/10.11591/ijece.v11i3.pp2407-2413>

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.