FEASIBILITY TEST OF POOR RICE RECIPIENTS IN BENCOY SUKABUMI VILLAGE USING NAIVE BAYES

Taufik Hidayatulloh [1]; Ardi Winardi [2]; Lestari Yusuf [3*)] ; Satia Suhada[4]; Saeful Bahri[5]

Sistem Informasi Kampus Kota Sukabumi[12]
Universitas Bina Sarana Informatika
www.bsi.ac.id[12]
taufik.tho@bsi.ac.id[1]; ardiwinardiv@gmail.com [2]

Sistem Informasi[3,4,5];
Universitas Nusa Mandiri [3,4,5]
www.nusamandiri.ac.id [3,]
lestari.lyf@nusamandiri.ac.id [3*)]; satia.sqh@nusamandiri.ac.id [4]; saeful.sel@nusamandiri.ac.id [5]
(*) corresponting author

Abstract— A regional head must have a work plan every regional head must have a work plan which is sure to be of benefit to the community. Assisting is a definite work plan in every region. A lot of assistance is usually given from the government to the community and must be managed by the village government so that the aid gets to the right hands. And to improve food security, the people in each region have activities to distribute Poor Rice as a subsidy from the government. In the distribution method, sometimes there are constraints in data collection so that poor rice or what we usually call Raskin is not suitable for distribution. Because of this, a way is needed so that the distribution is appropriate or not in the community in accepting the Raskin so that government assistance can be delivered properly and on target. By using secondary data obtained from Bencoy Village, 205 data were obtained containing the attributes of the eligibility category of Raskin recipients, and 6 categories of attributes were found with the classification method of the Naïve Bayes algorithm. The accuracy value obtained is 96.59%, proving that the prediction using the Naive Bayes algorithm has a good performance. The next results obtained are in the form of AUC value which after being calculated produces a value of 0.999 and this results in an application which is an implementation with a flow that is adjusted to the calculation algorithm in the form of a web-based application.

Keywords: Raskin, Naïve Bayes, Data Mining

Abstrak— *Setiap kepala daerah pasti memiliki rencana kerja yang tetntunya pasti bermanfaat untuk masyarakatnya. Pemberian bantuan merupakan rencana kerja yang pasti ada disetiap daerah. Banyak bantuan yang biasanya diberikan dari pemerintah kepada masyarakat dan harus dimanage oleh pemerintah desa, agar bantuan sampai ke tangan yang tepat. Dan untuk meningkatkan ketahanan pangan masyarakatnya setiap deaerah memiliki kegiatan membagikan Beras Miskin sebagai subsidi dari pemerintah. Dalam cara pembagiannya terkadang terkendala dalam pendataan sehingga beras miskin atau yang biasa kita sebut Raskin ini tidak sesuai tujuan pembagiannya. Karena hal tersebut maka diperlukan cara agar pembagian sesuai dengan layak tidaknya masyarakat dalam menerima raskin tersebut agar bantuan adari pemerintah dapat tersampaikan dengan baik dan tepat sasaran. Dengan menggunakan data sekunder yang diperoleh dari Desa Bencoy, didapatkan 205 data yang berisikan tentang atribut kategori kelayakan penerima raskin dan di temukan 6 kategori atribut dengan metodenya klasifikasi algoritma Naïve Bayes. Nilai accuracy yang didapat sebesar 96.59% membuktikan prediksi menggunkan algoritma naive bayes ini memiliki performa yang baik. Hasil yang di peroleh berikutnya berupa nilai AUC yang setelah dihitung menghasilkan nilai sebesar 0,999 dan ini menghasilkan aplikasi yang merupakan implementasi dengan alur yang disesuaikan dengna algoritma yang sudah di hitung berupa apliaksi berbasis web.*

Kata kunci: *Raskin, Naïve Bayes, Data Mining*

INTRODUCTION

Government assistance is a program that can help improve the welfare of the community, one of which is Poor Rice (Raskin), whose distribution is also a social assistance program(Tone, 2016). Problems that often occur in the distribution of poor rice assistance usually still use estimates only and there is no prediction at the time of distribution of the distribution of poor rice recipients. So that few or many people sometimes protest because the people

who should have received assistance did not receive assistance(Suryeni et al., 2015)

Bencoy Village is one of the Five Villages in the Cireunghas District, and its address is on Jalan Raya Cireunghas KM 15, Sukabumi Regency. Bencoy Village has an area of 831,750 Km^2, in general, the topography of Bencoy Village also has 4 hamlets, 10 Rw and 54 Rt. The total population of Bencoy Village consists of 446 heads of households 1,816 Poor households include: 4,110 people Male + female: 4,029 people = 8,139 people.

The distribution of rice distribution activities for the poor in Bencoy Village is which devoted to poor or insufficient families is certainly very far from what was expected because poor rice which was originally devoted to poor families, now well-off and high-income families also receive. Some poor families do not receive assistance, while families who can afford rice assistance are poor. The major obstacle encountered was the distribution of Raskin due to the mistargeting of Raskin recipients (Nasir, 2019).

One thing that can used in predicting a receipt of assistance is by using data mining. As it is well known that mining data is the decomposition of database findings using mathematical techniques, Intelligent artificial statistics, machine learning to extract information (Kaesman, 2016). To make decisions with a large enough amount of data, data ming is a technique used to get new patterns (Wulandari et al., 2020). While the data mining method used in this study is the Naïve Bayes algorithm. The naïve Bayes algorithm assumes a simple estimated probability classification method for calculating a set of probabilities by summing the frequencies and combinations of data (Hidayat et al., 2017).

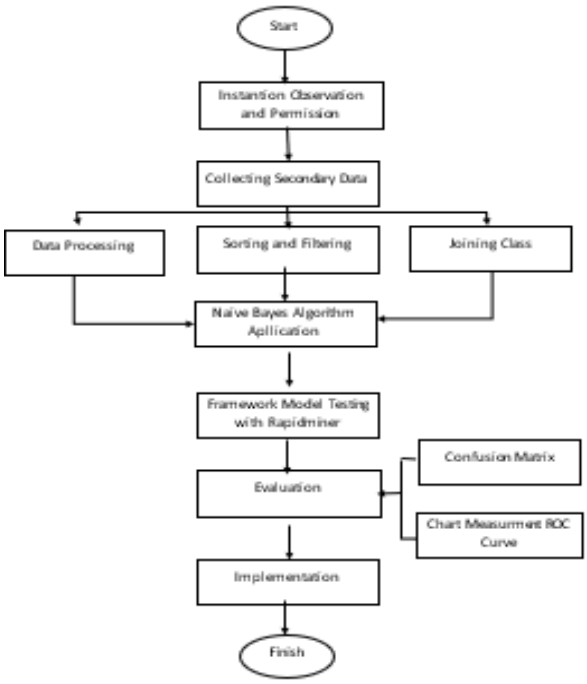
Not only that but the assessment of the distribution of Raskin was also implemented using the top-down approach (Nisak, 2014). Then the previous researchers had also applied the product weighted method to determine recipients of Raskin assistance (Firdyana et al., 2017). Even in previous studies, the calculation of the distribution of Raskin rice was carried out using the transportation method (Simbolon et al., 2014).

The feasibility of Raskin recipients using the Naive Bayes algorithm has previously been used to get an accuracy of 35.6% for a decent class and 36, 32% for an unfeasible class (Fadlan et al., 2018). In previous research, Raskin recipients were determined by a decision-making system using the C.45 algorithm with the results concluding that the largest entropy for Raskin recipients was from the Farmer category (Ermawati & Hidayatulloh, 2016). In the classification of a wood image, 50 data are comparing the two methods, Naive Bayes gets a higher accuracy value of 82.7% compared to K-NN with an accuracy of 75%(Waliyansyah & Fitriyah, 2019). With 1335 data, the accuracy rate of Naive Bayes even outperformed K-NN with an accuracy value of 83.83% while the accuracy of K-NN was 82.34%.(Maricar & Dian Pramana, 2019).

So this study aims to use data mining classification to calculate the accuracy of the Naive Bayes algorithm in determining the eligibility of Raskin recipients and implementing it into a program.

MATERIALS AND METHODS

To solve the problem in this study, the following methods will used by the author, methods will describe in figure 1:



Source: (Hidayatulloh, 2021)

Figure 1. Research Stages

In this study, the method used is Naïve Bayes to classification the data. In figure 1 starts with observation and permission in this study to get the data. Then after got the secondary data we start for data processing, sorting dan filtering the data, and divide the data into 2 classes to application the naive Bayes algorithm. After that stage, next is testing using the framework model with Rapidminer. Evaluate our calculation and will get the confusion matrix and ROC Curve. And the last stage is implementation according to the algorithm.

Datasets

By looking at previous research, the observation data used the Naïve Bayes algorithm, and then the test data were documented and

Rank 3 Accredited Journal based on Decree No. 85/M/KPT/2020
DOI: 10.33480/pilar.v17i1.2227

categorized (Sugiharti et al., 2017). The steps to obtain the dataset this time are to analyze the poor rice distribution system, then conduct interviews with staff who entrusted with managing the poor rice distribution data. Next, with a population belonging to the Bencoy Village of 205 residents, the sample data dataset obtained as described in the following table 1:

Table 1. Sample of Data

Raskin Reception Year	Recipient Cases		Total Sample
	Feasible	Not feasible	
2020	116	89	205

Source : (Winardi, 2021)

Data analysis was carried out using data on the feasibility of poor rice is which has 2 classes that are feasible and unfit. The data processing uses data mining techniques by applying the naïve Bayes algorithm and tested using the cross-validation method the evaluation of the an algorithm testing can to seen using the Confusion Matrix method to produce accuracy value while the ROC (Receiver Operating Characteristics) curve is used to measure. This research consists of six stages based on the CRISP-DM experimental model, along with the explanation:

Bussines Understanding

Based on the observations made. From this data, there is feasibility data with 6 attributes as predictor attributes and 1 result attribute. This research was conducted to determine the community that can receive assistance for receiving rice for the poor (Raskin). So it is necessary to have a study that applies a method to predict eligibility of acceptance which will make it easier for Bencoy Village officers to provide poor rice assistance.

Data Understanding

The data understanding stage is carried out in determining the purpose of a study, at this stage the process of collecting, analyzing, and evaluating the feasibility data is carried out such as understanding each predictor attribute, these attributes become parameters that can be seen in table 2.

Table 2. Attributes of Poor Rice Eligibility Acceptance Category

Attribute	Category
Type of House Floor	Board
	Bilik Bambu
	Ceramics
	Cement
Type of House Wall	Board
	Bilik Bambu
	Cement
Home Lighting	Personal electricity
	Connect to other
Head of Family's work	Seller
	Labor
	PNS
	Retired
	Driver
	Doctor
	Farmer
	Unemployment
Income	Middle
	Small
	High
Ownership of Asset	Yes
	No

Source : (Winardi, 2021).

Preparation

At the Data Preparation stage, feasibility data processing is carried out, where the data is processed first with the help of sort & filter aims to simplify the process of calculating the probability later.

Modeling

This modeling stage implements the naïve Bayes algorithm, for processing the dataset, then using the 10 fold cross-validation method on the testing data to produce an evaluation model that is measured by the Accuracy and ROC Curve values.

Evaluation

Evaluation and validation use the cross-validation method as a testing method, confusion matrix as an evaluation of this research which will produce Accuracy, where this Accuracy will be a measure of the success of a method because the higher the accuracy value the better the method used. The method used this time is the naïve Bayes method. Meanwhile, for measuring the performance of the Naïve Bayes algorithm, the ROC curve is used.

Deployment

This deployment stage carries out a process of making a report on the results of the research that has been carried out and will implemented on a website.

RESULT AND DISCUSSION

The results of this study aim to determine the accuracy value of the naïve Bayes algorithm in the eligibility of receiving Raskin by processing sample data. The results of this study are in the form of a

Rank 3 Accredited Journal based on Decree No. 85/M/KPT/2020
DOI: 10.33480/pilar.v17i1.2227

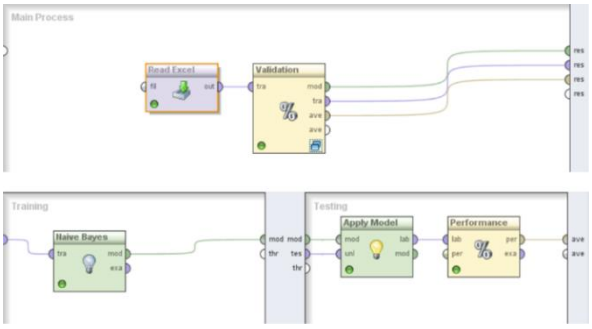
calculation process based on the naïve Bayes model and calculations using cross-validation, then the results of these calculations are implemented on a website.

Dataset

Before entering the Naïve Bayes algorithm process, we must first sort (sort & filter) the data, the amount of data collected is 205 records. From this data, we get 6 criteria, including the type of house walls, the lighting used, the work of the head of the household, the amount of income, and asset ownership.

Naïve Bayes Algorithm

After the data is sorted, the next step is to calculate the available data into Naïve Bayes. The results of the calculation of the probability value above can be made of the Naïve Bayes shown below:



Source: (Winardi, 2021)

Figure 2. Stages of Naïve Bayes Algorithm Model

The results of this figure 2 can be seen that this modeling produces two classes, namely an inadequate class with a probability value of 0.434 or 43.4% and a feasible class with a probability value of 0.566 or 56.6% so that when the two classes are added together will produce a value of 100%. This prior probability value is usually used to determine the class in a new case. Examples of cases can be seen in Table 4.

Table 4. example case

Attributes	Parameter	Feasible	unfeasible
Type of house floor	Board	0,060344828	0,033707865
Type of house wall	Board	0,172413793	0,08988764
Home Lighting	Personal Electricity	0,810344828	0,921348315
Head of Family's work	Seller	0,206896552	0,404494382
Income	Middle	0,155172414	0,516853933
Ownership of Asset	Yes	0,344827586	1

Source: (Winardi, 2021)

After calculating the probability of each class, then we will calculate the total probability based on the above calculations, using the following formula:

$$\boldsymbol{P(X|Ci) : P(X|result = fiesable} \quad \textbf{(1)}$$
$$\boldsymbol{P(X|Ci) : P(X|result = unfiesable} \quad \textbf{(2)}$$

And from the results of the total calculation that has been done, It can be seen that the P-value (X | Eligible) is 5.28148E-05, while the P-value (X | Unfeasible) is 0.00025338. This means that the value of P (X | Eligible) is smaller than the value of P (X | Not Feasible), so it can be concluded that for this case it will be classified as an Inadequate class.

Evaluation Model with Confusion Matrix

The results of the trials carried out are to produce the accuracy value and the AUC (Area Under Curve) value. The confusion matrix model will form a matrix consisting of true positive, true negative, false positive, and false negative. Based on the image above, there are details of the number of True Positives (TP) 83, False Negative (FN) 1, False Positive (FP) 6, and True Negative (TN) 115. These data can be calculated for the value of accuracy, precision, sensitivity, specificity, PPV. And NPV. The test results of the testing data using the Naïve Bayes algorithm produce the AUC value in the ROC graph which can be seen in the image below:

Source: (Winardi, 2021)

Figure 3. The AUC Value of the Naïve Bayes Algorithm in the ROC Graph

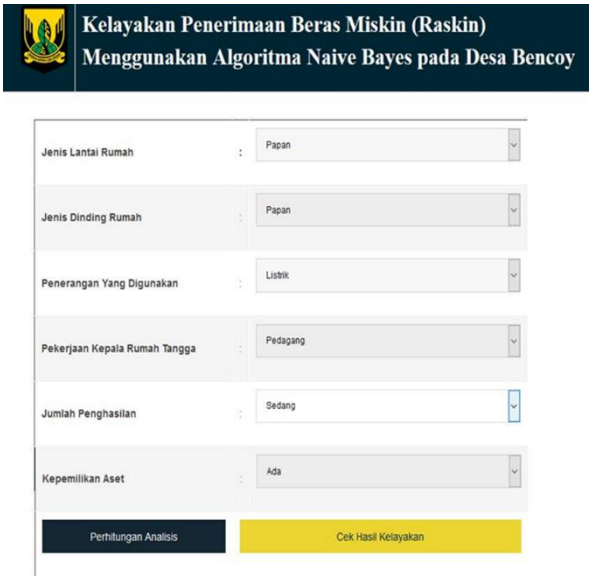
Based on figure 3 the AUC value of 0.999 is shown in the image above and when the curve number is rounded to 100 then the accuracy has this classification level is very good.

Application Development

For the validity of the calculation to be applied, the authors made an application which later this application will be developed and applied in the village of Bencoy. The following is an

Rank 3 Accredited Journal based on Decree No. 85/M/KPT/2020
DOI: 10.33480/pilar.v17i1.2227

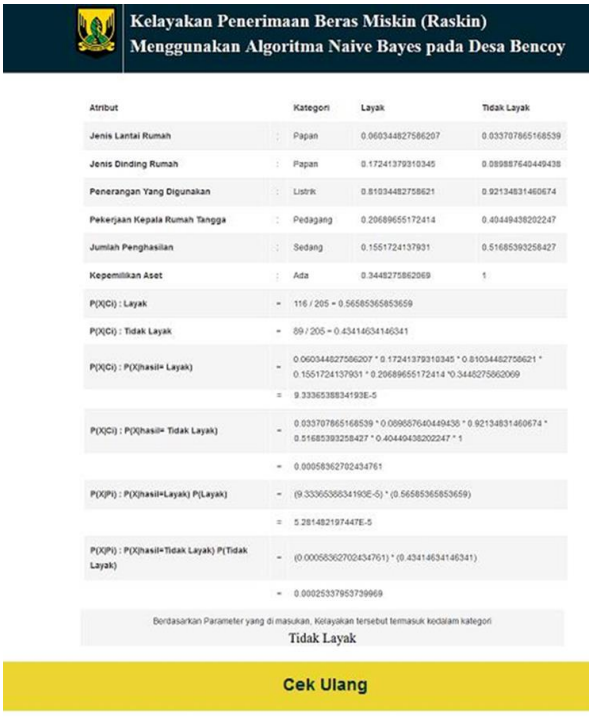
application view of the calculations that have been done.



Source: (Winardi, 2021)

Figure 4. Implementation Model from Naïve Bayes Algorithm

Figure 4 is an implementation of the calculation using the naive Bayes algorithm. This program is made using PHP language. This is the first testing is to input the category by the attribute.



Source: (Winardi, 2021)

Figure 5. Implementation Model from Naïve Bayes Algorithm

Figure 5 shows the results obtained from testing on the application. Based on this test, it can be seen that the predictions generated by the application are the same as the data result's probability.

Source:(Winardi, 2021)

Figure 6. Implementation Model from Algorithm

This application in figure 6, which was made to check the eligibility of receiving poor rice assistance, will be able to help Bencoy Village employees determine the eligibility of receiving poor rice assistance.

CONCLUSION

Based on the results of research using the Naïve Bayes algorithm, in predicting the feasibility of receiving poor rice with 205 datasets of testing data, an accuracy value of 96.59% is obtained, where the accuracy is obtained from the calculation of true positive values of 83 records, true negative of 115 records, false. positive for 6 records, and negative for 1 record. While the resulting AUC value is 0.999 and when the curve number is rounded it will be 100%, then this accuracy has a very good classification level. And with the creation of an application that adapts to calculations that are adjusted to the probability, it is hoped that it will greatly facilitate the Bencoy village officers when they will assist. And last This study uses a dataset of 205 data and 6 predictor attribute variables and 1 result class attribute, it is recommended for further research to add to the predictor attribute variables and the addition of more datasets so that the results seen from the measurement of accuracy and the roc curve can be better. It is also hoped that further research can be developed again by comparing this naïve Bayes method with other data.

REFERENCE

Ermawati, E., & Hidayatulloh, T. (2016). Penerapan Algoritma C4 . 5 Pada Sistem Penunjang

Keputusan Penentuan Penerima Raskin (Beras Masyarakat Miskin). *Seminar Nasional Ilmu Pengetahuan Dan Teknologi Komputer Nusa Mandiri*, 123–134.

Fadlan, C., Ningsih, S., & Windarto, A. P. (2018). Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra. *Jurnal Teknik Informatika Musirawas (JUTIM)*, *3*(1), 1. https://doi.org/10.32767/jutim.v3i1.286

Firdyana, S., Cahyadi, D., & Astuti, I. F. (2017). Penerapan Metode Weighted Product Untuk Menentukan Penerima Bantuan Beras Masyarakat Miskin (Raskin). *Prosiding SAKTI (Seminar Ilmu Komputer Dan Teknologi Informasi)*, *2*(1), 336–342. http://e-journals.unmul.ac.id/index.php/SAKTI/article/view/282

Hidayat, R., Marlina, S., & Utami, L. D. (2017). Perancangan Sistem Informasi Penjualan Barang Handmade Berbasis Website Dengan Metode Waterfall. *Simnasiptek 2017*, A-178.

Hidayatulloh, T. et al. (2021). *Feasibility Test Of Poor Rice Recipients In Bencoy Sukabumi Village Using Naive Bayes*.

Kaesman, Y. R. (2016). Penentuan Penerima Beras Raskin di Kelurahan Oesapa Barat Menggunakan Metode K-Nearest Neighbor. *Teknologi Terpadu*, *2*(2).

Maricar, M. A., & Dian Pramana. (2019). Perbandingan Akurasi Naïve Bayes dan K-Nearest Neighbor pada Klasifikasi untuk Meramalkan Status Pekerjaan Alumni ITB STIKOM Bali. *Jurnal Sistem Dan Informatika (JSI)*, *14*(1), 16–22. https://doi.org/10.30864/jsi.v14i1.233

Nasir, jamal A. (2019). Sistem Pendukung Keputusan PEmberian BEras Untuk KEluarga Miskin Dengan MEtode Simple Additive Weigthing. *Jurnal Riset Informatika*, *1*(3), 134–138.

Nisak, A. F. (2014). Implementasi Kebijakan Beras Miskin (Raskin) di Kecamatan Kenjeran Kota Surabaya : Studi Deskriptif pada Kelurahan Tanah Kalikedinding. *Jurnal Politik Muda*, *3*(2), 17–25.

Simbolon, L. D., Situmorang, M., & Napitupulu, N. (2014). Aplikasi Metode Transportasi dalam Optimasi Biaya Distribusi Beras Miskin (Raskin) pada Perum Bulog Sub Divre Medan. *Saintia Matematika*, *2*(3), 299-311.

Sugiharti, E., Firmansyah, S., & Devi, F. R. (2017). Predictive evaluation of performance of computer science students of unnes using data mining based on naÏve bayes classifier (NBC) algorithm. *Journal of Theoretical and Applied Information Technology*, *95*(4), 902–911.

Suryeni, E., Dan, Y. H. A., & Nurfitria, Y. (2015). Sistem Pendukung Keputusan Kelayakan Penerimaan Bantuan Beras Miskin Dengan Metode Weighted Product Di Kelurahan Karikil Kecamatan Mangkubumi Kota Tasikmalaya. *Konferensi Nasional Sistem & Informatika 2015*, 345–350.

Tone, K. (2016). Untuk perancangan proses digambarkan menggunakan DFD (. *Jurnal Instek*, *1*(1), 50–60.

Waliyansyah, R. R., & Fitriyah, C. (2019). Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *5*(2), 157. https://doi.org/10.26418/jp.v5i2.32473

Winardi, A. dkk. (2021). *The Feasibility Test For Beras Miskin In The Village Of Bencoy Sukabumi Using Naïve Bayes*.

Wulandari, D. A. N., Annisa, R., & Yusuf, L. (2020). an Educational Data Mining for Student Academic Prediction Using K-Means Clustering and Naïve Bayes Classifier. *Jurnal Pilar Nusa Mandiri*, 155–160.