Jurnal Mantik, 6 (1) (2022) 364-373

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik/index

ANALYSIS OF NAIVE BAYES METHODS ON SCHOLARSHIP ADMISSIONS AT SMKS AL-WASLIYAH 2 MERBAU

Hanggi Kurniawan[1], Ibnu Rasyid Munthe[2], Budianto Bangun[3]

[1,2,3]Faculty of Science and Technology, Labuhanbatu University, Jln. Sisingamangaraja No. 126 A, Kab. Labuhanbatu, North Sumatera, 21418, Indonesia

E-mail: hanggi.kurnia098@gmail.com[1] ibnurasyidmunthe@gmail.com[2] budiantobangun44@gmail.com[3]

ARTICLE INFO

Article history:
Received: April, 02 2022
Revised: May, 10 2022
Accepted:May, 30 2022

Keywords:
Confusion Matrix, Datamining, Naïve Bayes, ROC, Scholarships.

ABSTRACT

Scholarships currently play an important role in the world of education. Scholarships are financial aid given to students to complete their education. Scholarships are awarded by governments, foundations, institutions, or other institutions. Each scholarship has certain requirements. Scholarship programs are highly desired by every student, especially high-achieving but economically constrained students. The form of participation of the community, institutions, and students themselves in participating in building the nation, especially in the field of education, is crucial. By using the nave Bayes method, it is expected that the determination of scholarship recipients will be more targeted and accurate. This method serves to classify scholarship recipients into SMKs Alwasliyah 2 merbau. The Nave Bayes methods have no rules for their use. This method is used using benchmarks from the classification of pre-existing data.

1. Introduction

SMKS Al-Wasliyah 2 In Merbau, one of the favorite schools in Marbau subdistrict, has quite a lot of students, and not all students belong to the group of people who can afford it. In teaching and learning activities, not all students are able to meet their needs. Scholarships can help students meet the needs of learning activities, but because many students often make mistakes in the selection of students who deserve to receive scholarships due to the way of neglect, that is not accurate. So the method of calculation used must be replaced with a more accurate method to minimize scholarship target errors.

Scholarships are financial assistance provided to students that is used for educational purposes only. Scholarships are defined as a form of appreciation given to individuals in order to continue their education to a higher level. The award can be certain access to an institution or in the form of financial assistance. Basically, scholarships are income for those who receive them. In accordance with article 4 paragraph (1) of the PPh Law / 2000. The definition of income is an additional economic ability with the name and in any form received or obtained from Indonesian sources or outside Indonesia that can be used for consumption or increase the wealth of taxpayers. Because scholarships can be interpreted as increasing economic ability for the recipient, it means that scholarships are income"[1]. Scholarships are given to Students with the aim of reducing the economic burden of underprivileged students to finance the education being undertaken. The government through the Directorate General of Higher Education of the Ministry of National Education aims to allocate scholarship assistance funds to outstanding students and underprivileged students.[2]

Data mining is a process to gain knowledge to gain knowledge from a large set of data stored using statistical and mathematical techniques.[3] Data mining is a process that uses statistical, mathematical, artificial intelligence, and machine elerning techniques to extract and identify useful information and related knowledge from various large databases.[4] Data mining algorithms come from a form of algorithm development in various fields of machine learning, statistics, artificial intelligence, and artificial neural networks. Data mining serves to handle data with a large enough size, so one of the directions of research in the field of data mining is to develop the algorithm in order to handle very large data sizes.[5]

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

The Naïve Bayes method is a way of classifying the fastest by calculating the possibilities of existing data by summing the frequencies that arise from the sample data provided. In another sense Naïve Bayes is a classification by probability and statistical methods to determine the odds that will arise using existing possibilities..[6] *Naïve Bayes* is one of those methods that has no rules in its use. This method uses probability theory by looking for the greatest opportunities by using classification methods, namely by looking at the chances of the emergence of each classification in *training data.*.[7]

Confusion matrix is a method that uses matrix tables. If the data set consists of two classes, then one will be considered positive and the other will be considered negative. In the graphic between False Positive to True Positive is called roc curve .[8] The ROC (Receiver Operating Characteristic) curve is a way to visually evaluate accuracy results, to display graphics that determine which one is better for looking for Probability. Roc will display a two-dimensional graph with false positive as horizontal line and true as vertical line.[9]

2. Research Methods

Naïve Bayes is a classifier method that has high accuracy and speed on databases with large data to predict the likelihood of membership of a class. [10] This method was first discussed first about the basic concepts and definitions in Bayes' Theorem and subsequently this torema was used to classify and data mining. This method uses the probability classification method of each class in the document. [11] *Naïve Bayes*' algorithm is a way of classifying data using probability and statistical methods by predicting subsequent results based on pre-existing results. [12] Common forms of the *Naïve Bayes* Method are as follows:

$$\mathbf{P(A \mid B)} = \frac{P(B \mid A)\, P(A)}{P(B)} \quad (1)$$

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

X……..……[13]

Confusion Matrix is one of the measurement methods by calculating the true level of classicalization used to perform calculations looking for *accuracy*, *precision*, and *recall* values in *Data Mining*.[14]

TABLE 1
CONFUSION MATRIX TESTING

Clasification	Predicated Class	
	True	False
Actual : True	True Positif (TP)	False Negatif (FN)
Actual : False	False Negatif (FP)	True Negativei (TN)

Where table 1 contains :

1. TP is *True Positive*, which is the correct amount of positive data.
2. TN is *True Negative*, which is the correct amount of negative data.
3. FN is *False Negative*, which is the amount of negative but incorrect data.
4. FP is *False Positive*, which is the amount of positive but incorrect data.

Accuracy value is the value that can be obtained from the comparison of clarified data with the overall data. Accuracy Value Formula:

$$\boldsymbol{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad (2)$$

Precision value can be obtained from the amount of correctly classified posituf category data divided by the total positive classified data, precision can be obtained using the formula:

$$\boldsymbol{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Recalls are used to show the number of persers from the correctly classified positive category data. To find the Recall value using the formula:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \quad (4)[15]$$

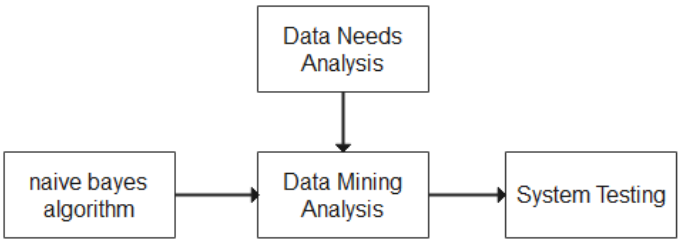

Figure 1. Design of research methods

At Figure 1. Design research methods to perform the stages in accordance with what will be done, among others. Data Needs Analysis is analyzing existing data to find the needs of the data in. Naïve Bayes algorithm is an arithmetic calculation method that will be used in calculating the opportunities that will be obtained from existing training data. Data Mining Analysis is analyzing from the processing of existing training data to find patterns or rules for the naïve bayes system. System Testing is to test the system whether there are still miscalculations and see weaknesses in the system.

3. Result And Discussion

3.1 Analysis Data

No	Name	parents' job	parent's income	total liabilities	electrical power	average value	Status
1	Junaidi	Petani	1.000.000	4	450	70-80	Feasible
2	Indah Permata sari	PNS	4.000.000	4	1200	70-80	Not Feasible
3	Rini Pratiwi	Wirausaha	3.000.000	2	900	70-80	Not Feasible
4	Angga Syahputra	Buruh	1.000.000	3	450	81-90	Feasible
5	Muhammad Ali	Petani	4.000.000	1	1200	70-80	Not Feasible
6	Syahputri	Buruh	1.000.000	4	450	81-90	Feasible
7	Putri Kumalasari	Pegawai Swasta	2.000.000	3	450	70-80	Feasible
8	Dewi Rumini	Wirausaha	3.000.000	4	900	81-90	Feasible
9	Wahyuni	Pegawai Swasta	3.000.000	2	900	65-69	Not Feasible
10	Anggraini Putri	Petani	2.000.000	4	450	81-90	Feasible
11	Indra Pratama	Petani	4.000.000	1	450	81-90	Feasible
12	Dedi iskandar	Buruh	2.000.000	4	450	70-80	Feasible
13	M.Fadli	Wirausaha	2.000.000	3	900	81-90	Feasible
14	Dea fadilah	PNS	3.000.000	1	900	81-90	Not Feasible
15	Ari Ditio	Wirausaha	4.000.000	6	900	81-90	Feasible
16	Muhammad Rio	Pegawai Swasta	3.000.000	4	900	81-90	Feasible
17	Juliana	Pegawai Swasta	4.000.000	3	900	81-90	Feasible
18	Ersi Mutiasari	Buruh	2.000.000	4	450	70-80	Feasible
19	Nur Cahaya	Buruh	3.000.000	3	900	81-90	Feasible
20	Nurummi Hidayah	Buruh	2.000.000	1	450	70-80	Feasible
21	Selamet Wahyono	Petani	3.000.000	4	900	81-90	Feasible
22	Dedek Andayani	Buruh	2.000.000	2	450	81-90	Feasible
23	Alex Candra	Petani	2.000.000	2	900	81-90	Feasible
24	Muhammad Arif	Petani	4.000.000	1	1200	81-90	Not Feasible
25	Khairani	Petani	3.000.000	3	900	81-90	Feasible
26	Kurnia Pratiwi	PNS	3.000.000	4	900	81-90	Not Feasible
27	Fanny Albaniah	Petani	3.000.000	2	450	81-90	Feasible
28	Dony Pradana	Buruh	2.000.000	2	450	81-90	Feasible
29	Muhammad Fazar	PNS	4.000.000	3	900	70-80	Not Feasible
30	Didi Pamungkas	Petani	3.000.000	2	450	81-90	Feasible
31	Sri Rezeki	Petani	2.000.000	4	900	81-90	Feasible
32	Muhammad Al Farizi	PNS	5.000.000	3	1200	81-90	Feasible
33	Titin Afrianti	Buruh	3.000.000	3	450	70-80	Not Feasible
34	Pria Andika	Petani	2.000.000	1	450	70-80	Feasible

Figure 2. Dataset Siswa

In Figure 2.Data ser students contain from the student dataset SMKs Al-Wasliyah 2 Merbau which contains from the dataset of students who successfully obtained as many as 34 students which is data from 1 Class. The figure above consists of Student Name, Parent's Job, Parent Income, Number of Parental Dependents, Electrical Power used by Student Average Score.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

TABLE 2
STUDENT DATA AT TRIBUTES

No	Attribute	Type	Description
1	Name	Text	Student Name
2	Parents Job	Categorical	Parental Work
3	Parents Income	Numeic	Parental Income
4	Total Liabilities	Numeric	Number of dependents in the family
5	Electrical Power	Categorical	Electrical Power used
6	Average Value	Categorical	Student Average Score
7	Status	Categorical	Admission Status

In Table 2. Research Attribute is a collection of data that was successfully obtained from SMKs AL-Wasliyah 2 Merbau. The attribute data that was successfully selected in the table above has been explained each is still the type of existing attributes and has been equipped with a description to make it easier to understand it.

3.2 Data Training

Data Training is sample data that has been obtained from SMKs AL-Wasliyah 2 Merbau in the form of .xlsx files that are used as a measuring pattern to find information from new data that will be searched for the status of the eligibility for scholarship assistance.

Name	parents' job	parent's income	total liabilities	electrical power	average value	Status
Junaidi	Petani	1.000.000	4	450	70-80	Feasible
Indah Permata sari	PNS	4.000.000	4	1200	70-80	Not Feasible
Rini Pratiwi	Wirausaha	3.000.000	2	900	70-80	Not Feasible
Angga Syahputra	Buruh	1.000.000	3	450	81-90	Feasible
Muhammad Ali	Petani	4.000.000	1	1200	70-80	Not Feasible
Syahputri	Buruh	1.000.000	4	450	81-90	Feasible
Putri Kumalasari	Pegawai Swasta	2.000.000	3	450	70-80	Feasible
Dewi Rumini	Wirausaha	3.000.000	4	900	81-90	Feasible
Wahyuni	Pegawai Swasta	3.000.000	2	900	65-69	Not Feasible
Anggraini Putri	Petani	2.000.000	4	450	81-90	Feasible
Indra Pratama	Petani	4.000.000	1	450	81-90	Feasible
Dedi iskandar	Buruh	2.000.000	4	450	70-80	Feasible
M.Fadli	Wirausaha	2.000.000	3	900	81-90	Feasible
Dea fadilah	PNS	3.000.000	1	900	81-90	Not Feasible
Ari Ditio	Wirausaha	4.000.000	6	900	81-90	Feasible
Muhammad Rio	Pegawai Swasta	3.000.000	4	900	81-90	Feasible
Juliana	Pegawai Swasta	4.000.000	3	900	81-90	Feasible
Ersi Mutiasari	Buruh	2.000.000	4	450	70-80	Feasible
Nur Cahaya	Buruh	3.000.000	3	900	81-90	Feasible
Nurummi Hidayah	Buruh	2.000.000	1	450	70-80	Feasible
Selamet Wahyono	Petani	3.000.000	4	900	81-90	Feasible
Dedek Andayani	Buruh	2.000.000	2	450	81-90	Feasible
Alex Candra	Petani	2.000.000	2	900	81-90	Feasible
Muhammad Arif	Petani	4.000.000	1	1200	81-90	Not Feasible
Khairani	Petani	3.000.000	3	900	81-90	Feasible
Kurnia Pratiwi	PNS	3.000.000	4	900	81-90	Not Feasible
Fanny Albaniah	Petani	3.000.000	2	450	81-90	Feasible

Figure 3. Data Training

Figure 3 contains from the data of prospective scholarship admissions that were successfully obtained from SMKs Al-Wasliyah 2 Merbau which contains all the data needed in selecting scholarship acceptance candidates.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

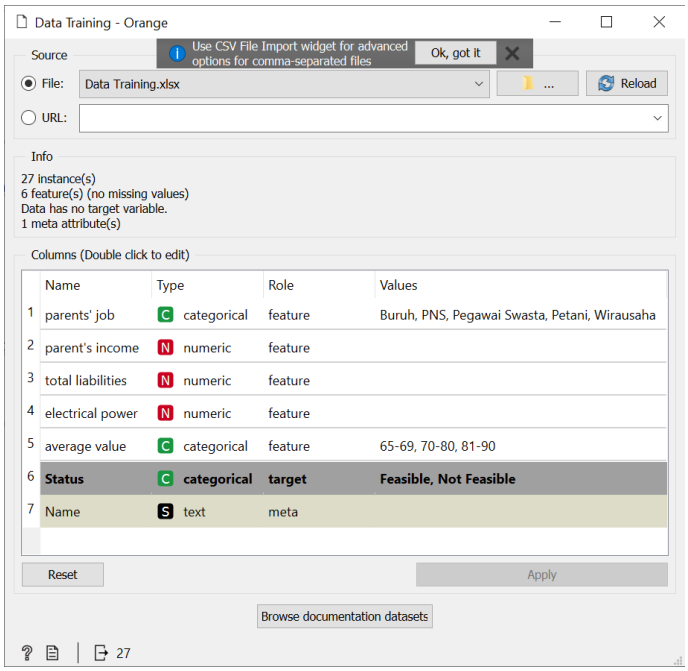

Figure 4. Selection of training data targets

Figure 4. is the process of selecting data that will be searched with the target attribute is Description, the function of the step above aims so that in data testing can be found the attribute we are looking for by using the Naïve bayes method, namely the Description attribute. by changing the *Role* from *future* to *target* so that we can find results in accordance with existing information.

3.3 Data Selection Process/ Prepocessing

The Select Columns Process aims to complete the Sisea SMKs Al-Wasliyah 2 Merbau dataset process with the aim of finding information attribute values that do not yet exist so that the target column is not filled in.

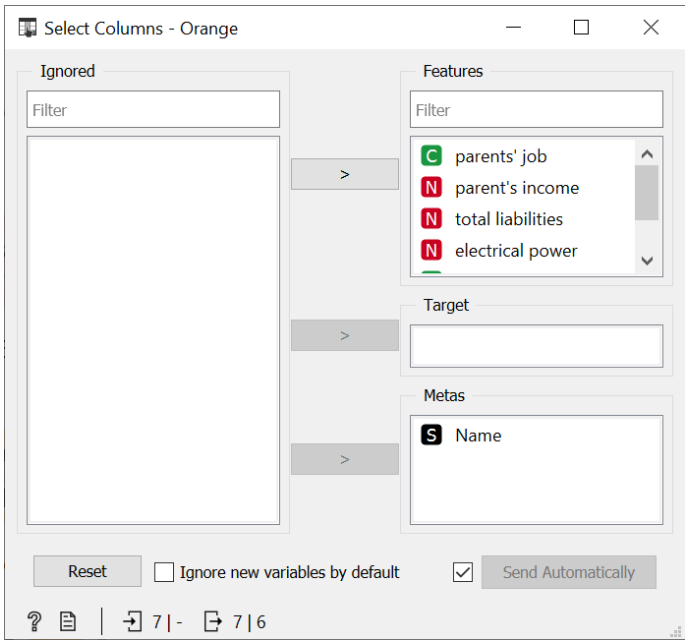

Figure 5. Data Selection Proses/ Preprocessing

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Figure 5 contains the data that we want to find the eligibility status of its acceptance. The data contains from parental income data, Number of Dependents, Electrical Power, Average value With its target attribute Description, in the Target Column is still empty because in the Testing data we do not know the results of the information and the target has been determined in *the Training Data*.

3.4 Data Mining Process

In the Data Mining Process, the Classification model is carried out using the Orange Tool using the Naïve Bayes Method to classify the scholarship eligibility status dataset at SMKs Al-Wasliyah 2 Merbau

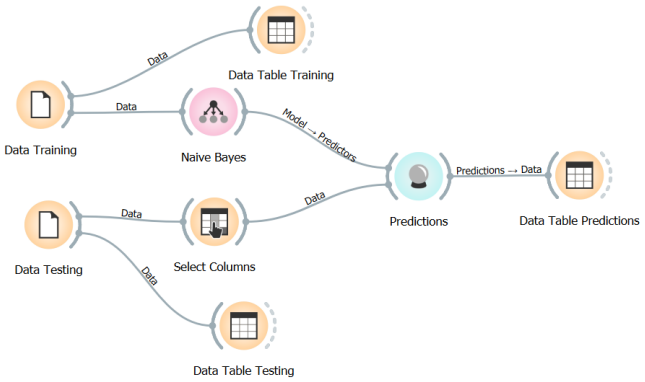

Figure 6. Data Mining Process

At figure 6. Data Mining Process is the design of the Naïve Bayes method wigdet using Orange Software using datasets that have been obtained before it in the form of Data Training and data testing in the process with the Naïve Bayes classification method..

3.5 Classification Model Testing Process

In the Naïve Bayes method testing process that has been made before, training data and data testing are needed to test the results of the classification using the Naïve Bayes Method such as Figure 5.

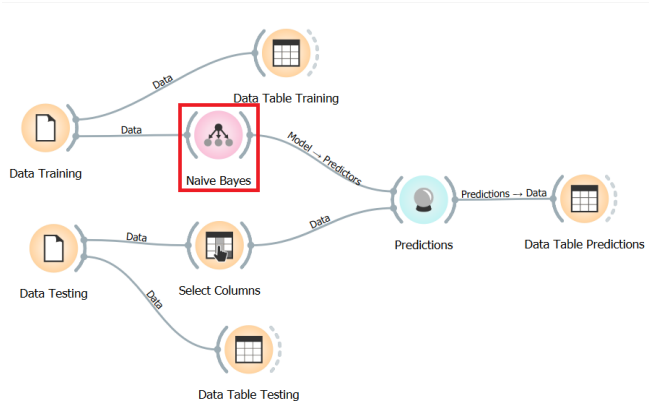

Figure 7. Widget design model classification of scholarship acceptance status dataset

At figure 7. The widget design of the calcification model of the scholarship acceptance status dataset is a widget design that has been added to the classification process. On the red box is the Naïve Bayes method used to test the data that has been obtained to find out the results of eligibility for admission to the SMKs Al-Wasliyah 2 Merbau Scholarship.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

3.6 Classification Model Predictions Process

The next process is to predict the results of the classification using Predictions which are used to determine the eligibility status of scholarship receipts such as Figure 6. Figure 8. The widget design to determine the eligibility status is a widget to determine the eligibility status by using *Preditions* which will then be displayed in Figure 9.

	Name	Naive Bayes	aive Bayes (Feasibl	ve Bayes (Not Feasi	parents' job	parent's income	total liabilities	electrical power	average value
1	Dony Pradana	Feasible	0.989801	0.0101989	Buruh	2000000	2	450	81-90
2	Muhammad Fa...	Not Feasible	0.0995625	0.900438	PNS	4000000	3	900	70-80
3	Didi Pamungkas	Feasible	0.834224	0.165776	Petani	3000000	2	450	81-90
4	Sri Rezeki	Feasible	0.930935	0.0690648	Petani	2000000	4	900	81-90
5	Muhammad Al ...	Not Feasible	0.0355469	0.964453	PNS	5000000	3	1200	81-90
6	Titin Afrianti	Feasible	0.96745	0.0325504	Buruh	3000000	3	450	70-80
7	Pria Andika	Feasible	0.872185	0.127815	Petani	2000000	1	450	70-80

Figure 8. Results from predictions using the Naïve Bayes method

Figure 9 contains the results of predictions that have been obtained by using metoed naïve bayes and displayed using table predictions data.

3.7 Classification Model Evaluation Results

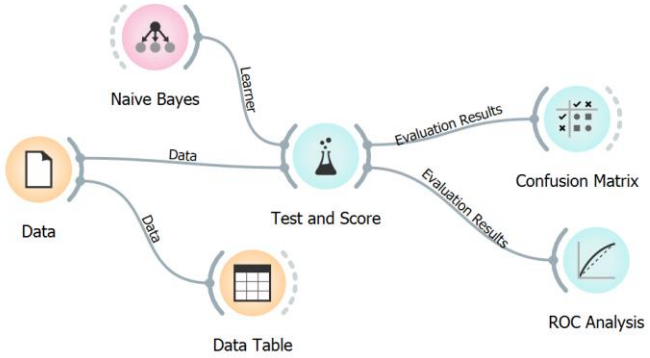

Figure 9. Classification Evaluation Widget

At Figure 10. Classification Evaluation Widget consists of widgets to get test and score, confusion matrix and ROC Analysis results from student data files that have been obtained using the Naïve Bayes method. The results of the classification model simulation are carried out using Training and Data Testing which has been made one into Test Data with 1 attribute as a target, 1 text, namely meta, 2 categorical, namely parental work and average value, 3 numeric namely Income, number of dependents and electrical power. So that the *test score* results as in Figure 11.

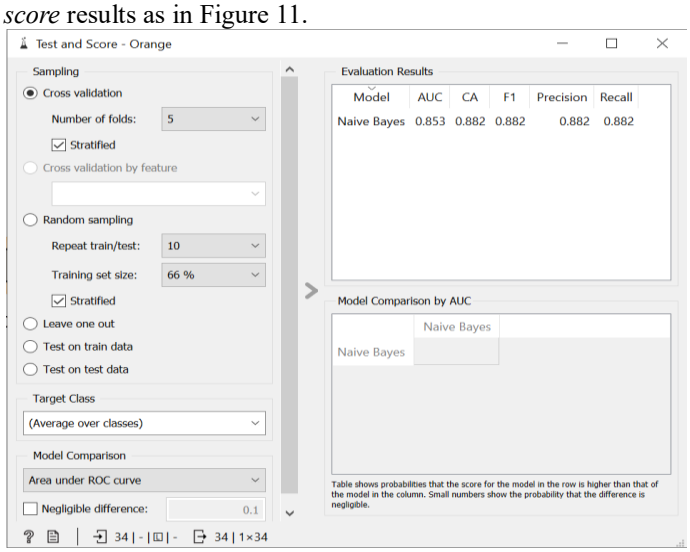

Figure 10. Results of Test and Score

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

At Figure 11. The results of the test and score were obtained from the results of calculations from 34 data that had been tested, obtained the results of the AUC calculation of 0.853 AUC used to measure discriminatory work by estimating the probability obtained randomly from positive and negative populations, the large acclamb of AUC then the good classification results used.

3.8 Evaluation Results with Confusion Matrix

Confusion Matrix is a measurement tool for prediction methods by calculating the correctness of data that has been classified using the naïve bayes method. Confusion matrix consists of tables used to invalidate the performance of naïve bayes classifications from existing datasets.

		Predicted		
		Feasible	**Not Feasible**	**Σ**
Actual	**Feasible**	23	2	**25**
	Not Feasible	2	7	**9**
	Σ	**25**	**9**	**34**

Figure 11. Confusion Matrix value of Naïve Bayes method

Figure 12. The result of true positive (TP) is 23. True Negative (TN) is 7, False Positive (FP) is 2 and False Negative (FN) is 2. Then the Accuracy, Precision and Recall Values are as follows:

$$\text{Accuracy} = \frac{23+7}{23+7+2+2} + 100\% \qquad \text{Then the Accuracy value = 88 \%}$$

$$\text{Presisi} = \frac{23}{23+7} + 100\% \qquad \text{Then precision value = 76 \%}$$

$$\text{Recall} = \frac{23}{23+2} + 100\% \qquad \text{Then the Accuracy value = 92 \%}$$

3.9 Evaluation Results with ROC Curve

Roc Curve is obtained from true signal (sensitivity) and (1-specificity) on the entire range of cut-off points to get the ROC curve visualized from the Confusion Matrix. Roc graphic results can be seen in figures 13 and 14.

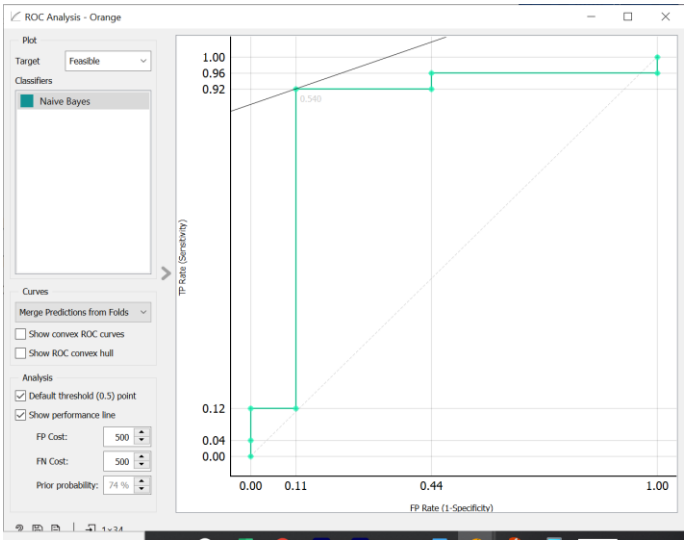

Figure 12. ROC analysis with targeted students who deserve scholarships

Figure 13 states that the results of the ROC analysis of scholarship eligibility in SMKs Al-Wasliyah 2 Merbau are worthy by naïve bayes method as much as 0.540.

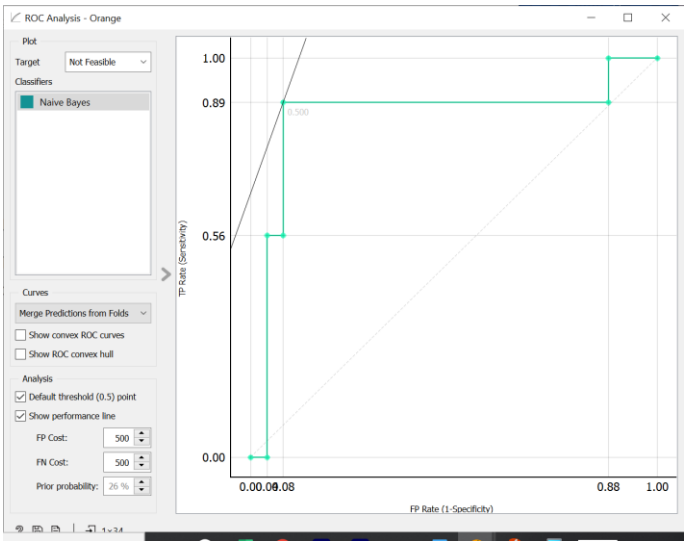

Figure 13. ROC analysis with target students not eligible for scholarships

Figure 14 shows that the results of the ROC analysis of scholarship eligibility at SMKs Al-Wasliyah 2 Merbau are not feasible with the Naïve Bayes method with a result of 0.500.

4. Conclusion

Scholarships are currently an important role in the world of education. Scholarships are financial aid given to students to complete their education. Scholarships are awarded by governments, foundations, institutions or other institutions. Each scholarship has certain requirements. Scholarship programs are highly desired by every student, especially high-achieving but economically constrained students. The form of participation of the community, institutions and students themselves in participating in building the nation, especially in the field of education. By using the naïve bayes method, it is expected that the determination of scholarship recipients is more targeted and accurate. This method serves to classify scholarship recipients in SMKs Alwasliyah 2 merbau. Naïve bayes methods that have no rules in their use, this method is used using benchmarks from the classification of pre-existing data.

References

[1] Ilham, I. Gede Suwijana, and Nurdin, "Sistem Pendukung Keputusan Penerimaan Beasiswa Pada Smk 2 Sojol Menggunakan Metode Ahp," *J. Elektron. Sistim Inf. Dan Komput.*, vol. 4, no. 2, pp. 48–58, 2018.

[2] V. Tasril, "Decision Supporting System of the Receiption of Acceptance," *J. Inf. Technol. Comput. Sci.*, vol. 1, no. 1, pp. 100–109, 2018.

[3] A. Darmawan, N. Kustian, and W. Rahayu, "Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuya," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 2, no. 3, p. 299, 2018, doi: 10.30998/string.v2i3.2439.

[4] A. Sumiah and N. Mirantika, "Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan", [Online]. Available: https://journal.uniku.ac.id/index.php/buffer

[5] A. Ichlas, I. R. Munthe, and M. N. S. Hasibuan, "Aturan Asosiasi untuk Analisis Data Penjualan Produk Menggunakan Algoritma Apriori," *J. Tek. Inform. UNIKA St. Thomas*, vol. 06, pp. 53–60, 2021, doi: 10.54367/jtiust.v6i1.1280.

[6] M. Ridho Handoko and Neneng, "Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 50–58, 2021, [Online]. Available: http://jim.teknokrat.ac.id/index.php/JTSI

[7] V. Marudut *et al.*, "Menurut Turangan et . al (2017) insentif merupakan salah satu jenis penghargaan yang dikaitkan dengan prestasi kerja . Semakin tinggi prestasi kerja semakin besar pula insentif yang diterima . Sudah menjadi kebiasaan bahwa setiap perusahaan harus meneta," vol. 7, pp. 87–94, 2018.
[8] A. Yani, "Analisa Kelayakan Kredit Menggunakan Artifcial Neural Network dan Backpropogation (Studi Kasus German Credit Data)," *J. Ilm. Komputasi*, vol. 18, no. 4, pp. 385–390, 2019, doi: 10.32409/jikstik.18.4.2672.
[9] S. Widodo, H. Brawijaya, and S. Samudi, "Clustering Kanker Serviks Berdasarkan Perbandingan Euclidean dan Manhattan Menggunakan Metode K-Means," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 687, 2021, doi: 10.30865/mib.v5i2.2947.
[10] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
[11] A. D. Wibisono, S. Dadi Rizkiono, and A. Wantoro, "Filtering Spam Email Menggunakan Metode Naive Bayes," *TELEFORTECH J. Telemat. Inf. Technol.*, vol. 1, no. 1, pp. 9–17, 2020, doi: 10.33365/tft.v1i1.685.
[12] M. Syukri Mustafa, M. Rizky Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Citec J.*, vol. 4, no. 2, pp. 151–162, 2017.
[13] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
[14] E. W. H. Candana, I. G. A. Gunadi, and D. G. H. Divayana, "Perbandingan Fuzzy Tsukamoto, Mamdani Dan Sugeno Dalam Penentuan Hari Baik Pernikahan Berdasarkan Wariga Menggunakan Confusion Matrix," vol. 6, no. November, pp. 14–22, 2021.
[15] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, "Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix," *J. Inform. Upgris*, vol. 6, no. 2, pp. 66–75, 2021, doi: 10.26877/jiu.v6i2.6552.