

Klasifikasi Kelayakan Penerima Program Indonesia Pintar Siswa MIS NU Dusun III Pinangripan

Juwita Sari¹, Sri Wulandari², Yollanda Putry³, Wiwin Handoko⁴

^{1,2,3,4} Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Indonesia

juwi1909@gmail.com*

* Email Koresponden

DOI : 10.56427/jcbd.v3i1.239

INFO ARTIKEL

Histori Artikel

Diterima : 11 Desember 2023

Ditinjau : 19 Desember 2023

Disetujui : 29 Januari 2024

ABSTRAK

Peran pendidikan sangat penting dalam meningkatkan kecerdasan anak bangsa baik secara emosional, spiritual maupun intelektual. Akan tetapi, faktor ekonomi menjadi penyebab anak putus sekolah. Dan untuk mengatasi hal tersebut, pemerintah telah membuat Program Indoneais Pintar (PIP). Program Indonesia Pintar (PIP) melalui Kartu Indonesia Pintar (KIP) adalah pemberian bantuan tunai pendidikan kepada anak usia sekolah (6-21 tahun). Program ini dimaksud untuk meringankan biaya peserta didik baik langsung maupun tidak langsung. Salah satu sekolah di Asahan yang telah menjalankan program ini adalah MIS NU Dusun III Pinangripan. Adapun kriteria dari untuk pengklasifikasian ini berupa nilai rata-rata siswa, penghasilan orangtua, dan tanggungan orangtua. Namun, pihak sekolah masih memiliki kesalahan dalam pengklasifikasian kelayakan siswa penerima Program Indonesia Pintar (PIP), yang mana hal tersebut dapat merugikan pihak sekolah dan juga siswa yang terlibat. Untuk mengatasi hal tersebut, dibutuhkanlah sebuah sistem Klasifikasi Kelayakan Penerima program Indonesia Pintar (PIP) Siswa MIS NU Dusun III Pinangripan untuk meminimalisir masalah tersebut. Adapun salah satu teknik yang digunakan dalam mengelola kelayakan penerima Program Indonesia Pintar (PIP) adalah *Naïve Bayes* yang merupakan metode dari *Data Mining*. *Naive Bayes* digunakan untuk proses prediksi dengan menggunakan probabilitas dan statistik yang dapat memprediksi peluang dimasa depan berdasarkan pengalaman sebelumnya. Pemilihan Metode ini karena *Naive Bayes* memiliki kelebihan diantaranya dapat memprediksi jumlah siswa yang layak mendapatkan Program Indonesia Pintar (PIP) berdasarkan data-data yang konkrit agar hasilnya dapat dipertanggung jawabkan dan digunakan untuk prediksi selanjutnya. Dataset dalam penelitian ini adalah data Siswa MIS NU Dusun III Pinangripan Tahun 2023/2024 Ganjil sebanyak 130 siswa dengan *data training* 91 siswa dan *data testing* 39 siswa. Dengan menggunakan Algoritma *Naïve Bayes* dan *Model Confusion Matrix*, didapatkan nilai *precision* 92%, nilai *recall* 100%, nilai *f1-score* 96%, dan nilai akurasi sebesar 97% yang masuk kedalam kategori *Good Classification*.

The role of education is very important in increasing the intelligence of the nation's

Kata Kunci

Data Mining
Program Indonesia Pintar (PIP)
Naïve Bayes

Keywords

Data Mining
Smart Indonesia Program (PIP)
Naïve Bayes

children, both emotionally, spiritually and intellectually. However, economic factors are the cause of children dropping out of school. And to overcome this, the government has created the Smart Indonesia Program (PIP). The Smart Indonesia Program (PIP) through the Smart Indonesia Card (KIP) is the provision of educational cash assistance to school age children (6-21 years). This program is intended to reduce student costs both directly and indirectly. One of the schools in Asahan that has implemented this program is MIS NU Dusun III Pinangripan. The criteria for this classification are students' average grades, parents' income, and parents' dependents. However, the school still has errors in classifying the eligibility of students who receive the Smart Indonesia Program (PIP), which can be detrimental to the school and the students. which are involved. To overcome this, a Eligibility Classification system for Recipients of the Smart Indonesia Program (PIP) for MIS NU Dusun III Pinangripan Students is needed to minimize this problem. One of the techniques used in managing the eligibility of Smart Indonesia Program (PIP) recipients is Naïve Bayes, which is a Data Mining method. Naive Bayes is used for the prediction process using probability and statistics which can predict future opportunities based on previous experience. This method was chosen because Naive Bayes has advantages, including being able to predict the number of students who are eligible for the Smart Indonesia Program (PIP) based on concrete data so that the results can be accounted for and used for further predictions. The dataset in this research is MIS NU Dusun III Pinangripan Student data for 2023/2024 Odd, totaling 130 students with training data of 91 students and testing data of 39 students. By using the Naïve Bayes Algorithm and Confusion Matrix Model, we obtained a precision value of 92%, a recall value of 100%, an f1-score value of 96%, and an accuracy value of 97% which is included in the Good Classification category.

1. Pendahuluan

Pendidikan merupakan suatu sistem dimana terdiri dari banyaknya komponen yang saling berkaitan erat dan mempunyai tujuan yang sama yaitu menghasilkan sumber daya yang berkualitas. Pendidikan memberikan pengaruh terhadap kualitas sumber daya manusia (SDM) yang sangat menentukan nasib bangsa[1].

MIS NU Dusun III Pinangripan merupakan sekolah yang didirikan oleh Lembaga Ma'arif Nahdlatul Ulama untuk mengembangkan sistem pendidikan yang mandiri dan mewujudkan pengembangan bagi anak-anak bangsa. Namun, banyak anak-anak yang kurang berminat untuk melanjutkan sekolah mereka dikarenakan faktor ekonomi keluarga yang membuat mereka tidak sanggup untuk membayar biaya pendidikan yang setiap tahunnya semakin meningkat. Untuk mengatasi masalah tersebut, pemerintah telah membuat sebuah program untuk memberikan bantuan pendidikan yang dapat memberikan kesempatan belajar juga dapat membantu para siswa yang berasal dari keluarga miskin atau rentan miskin. Program ini disebut Program Indonesia Pintar (PIP)[2].

MIS NU Dusun III Pinangripan telah menjalankan program ini agar dapat membantu para siswa yang kurang mampu untuk tetap melanjutkan sekolah mereka. Namun, pihak sekolah masih memiliki kesalahan dalam pengklasifikasian kelayakan siswa penerima Program Indonesia Pintar(PIP), yang mana hal tersebut dapat merugikan pihak sekolah dan juga siswa yang terlibat. Untuk mengatasi hal tersebut, dibutuhkanlah sebuah sistem Klasifikasi Kelayakan Penerima program Indonesia Pintar (PIP) Siswa MIS NU Dusun III Pinangripan untuk meminimalisir masalah tersebut[3].

Salah satu teknik klasifikasi yang menggunakan pendekatan teorema bayes yaitu metode *Naïve Bayes Classifier*, teorema tersebut merupakan prinsip dari peluang statistika untuk mengkombinasikan pengetahuan sebelumnya dengan pengetahuan yang baru[4]. *Naïve Bayes* membuat dan dapat menilai model dengan sangat cepat dan skala secara linier dalam jumlah prediksi dan baris serta kolom[5]. *Naïve Bayes Classifier* merupakan sebuah metode dari *Data mining*. *Data mining* merupakan suatu proses dimana sejumlah besar data digali dan dianalisis untuk mendapatkan sesuatu yang valid, baru dan bermanfaat, hasilnya bisa menemukan pola atau formula dalam data tersebut[6].

Dengan dilakukannya penelitian ini, dapat membantu pihak sekolah dalam menentukan kelayakan penerima Program Indonesia Pintar(PIP) secara lebih akurat. Hal ini dibuktikan dengan adanya penelitian terdahulu dengan menggunakan algoritma Naïve Bayes dalam melakukan klasifikasi kelayakan penerima Program Indonesia Pintar(PIP).

Berdasarkan penelitian sebelumnya dengan judul “Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar”, menunjukkan bahwa model klasifikasi Naïve Bayes memiliki kinerja yang baik dengan akurasi sebesar 93,3%[2]. Berdasarkan penelitian dengan judul “Implementasi Algoritme Naïve Bayes Untuk Menentukan Kelayakan Calon Penerima Beasiswa” , mendapatkan hasil akurasi sebesar 90% dan dapat diaplikasikan ke dalam sistem pendukung keputusan untuk mendapatkan calon penerima beasiswa lebih baik[7]. Berdasarkan penelitian dengan judul “Prediksi Penerimaan Bantuan Pip Pada Smks Al-Furqon Batubara Dengan Metode Naïve Bayes” mendapatkan hasil bahwa Algoritma Naïve Bayes dapat menentukan kelayakan penerimaan Bantuan PIP dengan baik dengan akurasi sebesar 97%[8].

2. Metodologi Penelitian

Data Mining merupakan suatu kegiatan atau proses pengumpulan suatu data yang kemudian diekstrak atau diolah untuk mendapatkan suatu ilmu pengetahuan yang baru. *Data Mining* juga merupakan gabungan dari berbagai bidang ilmu diantaranya, *database system*, *data warehouse*, *statistics*, *machine learning*, *i retrival*, dan komputasi tingkat tinggi yang berjalan pada data yang cenderung terus membesar dan teknik terbaik yang digunakan kemudian berorientasi kepada data berukuran sangat besar untuk mendapatkan kesimpulan dan keputusan yang paling layak[9].

Data mining merupakan teknik menggali informasi berharga, terpendam, tersembunyi didalam data (*database*) dengan jumlah besar dengan begitu ditemukannya pola menarik yang sebelumnya belum diketahui, yang memiliki tujuan menyaring data yang sangat besar dan didapatkan sifat baru, bermanfaat, benar terhadap pemahaman suatu pola atau variasi dari database[10]. Menurut Larose, *Data Mining* memiliki enam fase CRISP-DM (*Cross Industry Standard Process for Data Mining*).



Gambar 1. Tahapan CRISP-DM

Berdasarkan gambar 1 dapat dilihat terdapat 6 fase CRISP-DM yaitu:

Bussines Understanding

Program Indonesia Pintar (PIP) merupakan sebuah program yang diberikan kepada siswa kurang mampu dengan bantuan uang saku untuk memberikan kesempatan bagi para siswa dalam melanjutkan pendidikan mereka. MIS NU Dusun III Pinangripan telah menjalankan program ini untuk membantu para siswa yang ingin bersekolah tetapi memiliki ekonomi yang rendah. Tujuan penelitian ini untuk menganalisis bagaimana pengklasifikasian kelayakan siswa dalam penerimaan Program Indonesia Pintar (PIP) ini secara lebih tepat dan akurat.

Data Understanding

Sumber data yang didapat berasal dari MIS NU Dusun III Pinangripan, Kabupaten Asahan. Data tersebut berisikan atribut antara lain : Nama, Nilai Rata-Rata, Penghasilan Orangtua, Jumlah Tanggungan dan Keterangan.

```
#menampilkan info dataset
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 130 entries, 0 to 129
Data columns (total 6 columns):
#   Column                Non-Null Count  Dtype
---  -
0   NO                     130 non-null    int64
1   NAMA                   130 non-null    object
2   NILAI RATA-RATA        130 non-null    int32
3   PENGHASILAN ORANGTUA   130 non-null    int64
4   JUMLAH TANGGUNGAN      130 non-null    int64
5   KETERANGAN             130 non-null    object
dtypes: int32(1), int64(3), object(2)
memory usage: 5.7+ KB
```

Gambar 2. Deskripsi Data

Data Preparation

Ada beberapa tahapan-tahapan yang dilakukan dalam persiapan data diantaranya yaitu memilih data, membangun data, mengintegrasikan data dan pembersihan data.

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 130 entries, 0 to 129
Data columns (total 4 columns):
#   Column                Non-Null Count  Dtype
---  -
0   NILAI RATA-RATA        130 non-null    int32
1   PENGHASILAN ORANGTUA   130 non-null    int64
2   JUMLAH TANGGUNGAN      130 non-null    int64
3   KETERANGAN             130 non-null    object
dtypes: int32(1), int64(2), object(1)
memory usage: 3.7+ KB
```

Gambar 3. Deskripsi Data yang Digunakan

Modelling

Naïve bayes Classifier merupakan salah satu algoritma pemecahan yang termasuk kedalam metode klasifikasi pada Data Mining. *Naïve bayes Classifier* mengadopsi ilmu statistika yaitu dengan menggunakan teori kemungkinan (Probabilitas) untuk menyelesaikan sebuah kasus *Supervised Learning*, artinya dalam himpunan data terdapat label, Class atau Target sebagai acuan atau gurunya[11]. *Naive Bayes* juga merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan *Teorema Bayes* dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas[12].

Naive Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada *Teorema Bayes* (aturan-aturan *Bayes*) yang mengasumsikan independensi (ketidaktergantungan). Model yang digunakan dalam *Naive Bayes* adalah “model fitur independen”[13].

Bentuk Umum dari *teorema bayes* :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan :

- X : Data dengan kelas yang belum diketahui
H : Hipotesis data x merupakan suatu kelas spesifik
P(H|X) : Probabilitas hipotesis H berdasarkan kondisi X (posterior probability)
P(H) : Probabilitas H (prior probability)
P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
P(X) : Probabilitas X

Cara Kerja *Naïve Bayes* sebagai berikut :

1. Siapkan dataset.
2. Hitung jumlah kelas pada data training.
3. Hitung jumlah kasus yang sama dengan kelas yang sama.
4. Kalikan semua hasil sesuai dengan data testing yang akan dicari kelasnya.
5. Bandingkan hasil perkelas, nilai tertinggi ditetapkan sebagai kelas baru.

Untuk memperoleh dataset dengan jumlah atribut yang lebih sedikit tetapi bersifat *informative*[14]. Maka dari 130 data dibagi menjadi 2 bagian yaitu data *training* dan data *testing* dengan porsi 70:30.

```
x_train.shape
(91, 3)

x_test.shape
(39, 3)
```

Gambar 4. *Split Data*

Evaluation

Evaluasi bertujuan untuk menguji dan memastikan data telah berfungsi sebagaimana mestinya. Dalam penelitian ini menggunakan *Confusion Matrix* yang biasanya menampilkan akurasi, *precision*, dan *recall*. Dimana *Output*-nya adalah dua kelas yaitu kelas negatif dan kelas positif[15].

Deployment

Tampilan hasil *Data Mining* divisualisasikan dalam bahasa pemrograman agar pihak sekolah dapat memahami dan membacanya[16].

3. Hasil dan Pembahasan

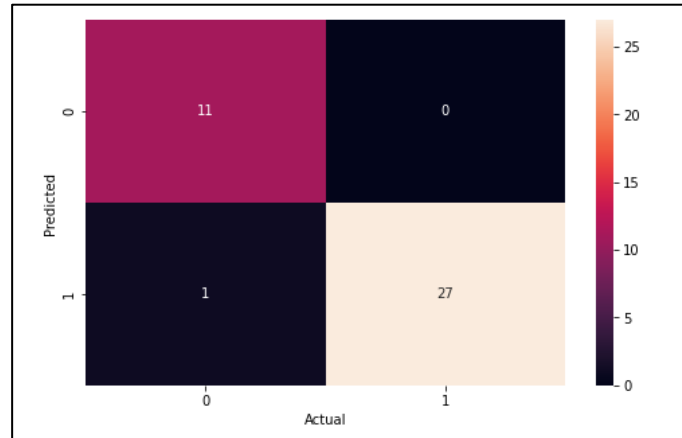
Tahap awal untuk analisis yaitu meng-*import* modul yang diperlukan serta dataset siswa dalam klasifikasi Program Indonesia Pintar (PIP).

Tabel 1. Data siswa dan Kelayakan Penerima PIP

Nilai Rata-Rata	Penghasilan Orangtua	Jumlah Tanggungan	Keterangan
89	1000000	4	Layak
69	2100000	2	Tidak Layak

66	1900000	3	Tidak Layak
94	1500000	2	Layak
85	1000000	1	Layak

Setelah membangun model dengan *data training*, gunakan *data testing* untuk melakukan pengujian. Perolehan *confussion matrix* untuk *metode Naïve Bayes* adalah:



Gambar 5. Visualisasi Confusion Matrix

Pada gambar 5 merupakan hasil perhitungan *visualisasi confusion matrix* menggunakan software *Python 3* dengan bantuan *tools Jupiter Notebook*.

	precision	recall	f1-score	support
Layak	0.92	1.00	0.96	11
Tidak Layak	1.00	0.96	0.98	28
accuracy			0.97	39
macro avg	0.96	0.98	0.97	39
weighted avg	0.98	0.97	0.97	39

Gambar 6. Tingkat Akurasi

Pada gambar 6 menunjukkan hasil perhitungan *visualisasi confusion matrix* menggunakan software *Python 3* dengan bantuan *tools Jupiter Notebook* diperoleh akurasi sebesar 97%, dengan hasil tersebut bahwasanya model yang digunakan dapat dilakukan untuk klasifikasi guna menentukan kelayakan calon penerima PIP[17]. Dan adapun hasil Algoritma menghasilkan nilai *precision* untuk kelas positif 92%, untuk kelas negatif 100%, nilai *recall* untuk kelas positif 100%, untuk kelas negatif 96%, nilai *f1-score* untuk kelas positif 96% dan untuk kelas negatif sebesar 98%.

4. Kesimpulan

Berdasarkan pembahasan menggunakan Algoritma *Naïve Bayes* dan *Model Confusion Matrix*, didapatkan nilai *precision* 92%, nilai *recall* 100%, nilai *f1-score* 96%, dan nilai akurasi sebesar 97% yang masuk dalam kategori *Good Classification*. Oleh karena itu, metode *Naïve Bayes* adalah model yang dikategorikan baik dan diimplementasikan memprediksi peluang berdasarkan pengalaman sebelumnya untuk mempermudah dalam proses klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) siswa MIS NU Dusun III Pinangripan.

Ucapan Terima Kasih

Penulis menyampaikan ucapan terima kasih yang tulus kepada MIS NU Dusun III Pinangripan atas dukungan dan kerjasama yang luar biasa selama penelitian ini. Tanpa bantuan dan partisipasi MIS NU Dusun III Pinangripan, penelitian ini tidak akan terlaksana dengan baik. Penulis juga ingin menyampaikan

rasa terima kasih kepada semua pihak yang terlibat dalam penelitian sebelumnya yang telah memberikan landasan dan pemahaman yang mendalam terkait data penerima program indonesia pintar siswa MIS NU Dusun III Pinangripan. Penelitian sebelumnya telah memberikan wawasan yang berharga dan menjadi pijakan bagi kami dalam merancang solusi yang lebih baik, Kami berharap kerjasama ini dapat berlanjut di masa depan untuk pengembangan solusi-solusi yang lebih baik dan inovatif dalam penerima program indonesia pintar. Terima kasih atas kerjasama, dukungan, dan kesempatan ini.

Referensi

- [1] N. Situmorang and G. Sirait, "Jurnal Comasie," *Comasie*, vol. 6, no. 2, pp. 107–118, 2020.
- [2] A. Nata and S. Suparmadi, "Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar," *J. Sci. Soc. Res.*, vol. 5, no. 3, p. 697, 2022, doi: 10.54314/jssr.v5i3.1041.
- [3] D. B. Siswanto and D. Normawati, "Sistem Klasifikasi Monitoring dan Evaluasi Kelayakan Penerima Beasiswa UAD Menggunakan Algoritma Naïve Bayes," *J. SAINTEKOM*, vol. 13, no. 2, pp. 161–172, 2023, doi: 10.33020/saintekom.v13i2.428.
- [4] Gagan Suganda, Marsani Asfi, Ridho Taufiq Subagio, and Ricky Perdana Kusuma, "Penentuan Penerima Bantuan Beasiswa Kartu Indonesia Pintar (Kip) Kuliah Menggunakan Naïve Bayes Classifier," *JSII (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 193–199, 2022, doi: 10.30656/jsii.v9i2.4376.
- [5] F. D. Pratama, I. Zufria, and T. Triase, "Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 1, pp. 77–84, 2022, doi: 10.36341/rabit.v7i1.2217.
- [6] A. Pebdika, R. Herdiana, and D. Solihudin, "Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 452–458, 2023, doi: 10.36040/jati.v7i1.6303.
- [7] H. Jurnal, H. Noviyanto, and B. Mukti, "Jurnal Informatika Dan Teknologi Komputer Implementasi Algoritme Naïve Bayes Untuk Menentukan Kelayakan Calon Penerima Beasiswa," *Juli*, vol. 1, no. 2, pp. 7–12, 2021.
- [8] N. Aini, H. Wiwin, and R. Nurhaliza, "PREDIKSI PENERIMAAN BANTUAN PIP PADA SMKS AL-FURQON BATUBARA DENGAN METODE NAÏVE BAYES," *JUTSI J. Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 11–20, 2024, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI%0APENERAPAN>
- [9] R. A. Iswanto, J. Sahertian, and M. A. D. Widyadara, "Pengembangan Sistem Prediksi Kelulusan Mahasiswa Berdasarkan Nilai Disiplin, Hasil Belajar, Aktivitas Sosial Ekonomi, dan Aktivitas Organisasi Menggunakan Algoritma Naïve Bayes," *Pros. Semin. Nas. Teknol. Dan Sains*, vol. 1, pp. 349–358, 2022.
- [10] O. Rini and S. O. Kunang, "Implementasi Data Mining Menggunakan Metode Naive Bayes Untuk Penentuan Penerima Bantuan Program Indonesia Pintar (Pip) (Studi Kasus : Sd Negeri 9 Air Kumbang)," *Bina Darma Conf. ...*, pp. 714–722, 2021.
- [11] M. M. Effendi and A. Setiawan, "Menentukan Prediksi Kelulusan Siswa Dengan Membandingkan Algoritma C4.5 Dan Naïve Bayes Studi Kasus Smkn. 1 Cikarang Selatan," *SIGMA - J. Teknol. Pelita Bangsa*, vol. 10, no. September, pp. 167–172, 2020.
- [12] R. Marbun, "Implementasi Data Mining Untuk Memprediksi Kelulusan Mahasiswa Menggunakan Algoritma Naive Bayes Classifier Studi Kasus: Poltekkes Kemenkes RI Medan," *JURIKOM (Jurnal Ris. Komputer)*, vol. 6, no. 6, pp. 627–633, 2020, [Online]. Available: <https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom/article/view/1887>
- [13] A. Suwarno *et al.*, "Jurnal Teknologi Pelita Bangsa," *J. Teknol. Pelita Bangsa*, vol. 12, no. 4, pp. 33–40, 2021.
- [14] J. Sumpena and N. Kurnia H., "Analisis Prediksi Kelulusan Siswa PKBM Paket C Dengan Metoda Algoritma Naive Bayes," *Tedc*, vol. 13, no. 2, pp. 127–133, 2019, [Online]. Available: <http://ejournal.poltektedc.ac.id/index.php/tedc/article/view/13>
- [15] Y. Septiani, "Penerapan Algoritma Naive Bayes Menentukan Klasifikasi Tingkat Kelulusan Siswa SMK Media Informatika Jakarta," *Semin. Nas. Mhs. Fak. Teknol. Inf.*, no. September, pp. 607–613, 2022, [Online]. Available: <http://senafiti.budiluhur.ac.id/index.php/senafiti/article/view/361/71>
- [16] A. A. A. Arifin, W. Handoko, and Z. Efendi, "Implementasi Metode Naive Bayes Untuk Klasifikasi Penerima Program Keluarga Harapan," *J-Com (Journal Comput.)*, vol. 2, no. 1, pp. 21–26, 2022, doi: 10.33330/j-com.v2i1.1577.

- [17] A. Hadi, I. Ali, N. Bayes, and R. Miner, "DENGAN METODE ALGORITMA NAIVE BAYES," vol. 7, no. 1, pp. 366–372, 2023.