

## Analisa Prediksi Mahasiswa Penerima KIP-K menggunakan Algoritma Naive Bayes

Desmulyati<sup>1</sup>, Muhammad Jadetz Mulyono<sup>2</sup>, Amriandry Maulana<sup>3</sup>, Muhammad Ibnu Raihan<sup>4</sup>, Ridwan Sholeh Sumitra<sup>5</sup>, Ali Mukhtar<sup>6</sup>

<sup>1</sup>Informatika, <sup>2-6</sup>Sistem Informasi, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat  
E-mail: [desmulyati.dmy@bsi.ac.id](mailto:desmulyati.dmy@bsi.ac.id)

\* Corresponding Author

 <https://doi.org/10.31004/jerkin.v4i3.4969>

### ARTICLE INFO

#### Article history

Received: 23 Nov 2025

Revised: 05 Dec 2025

Accepted: 30 Dec 2025

#### Kata Kunci:

KIP-K, Machine Learning, Naive Bayes, Klasifikasi, Pendidikan.

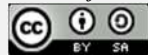
#### Keywords:

KIP-K, Machine Learning, Naive Bayes, Classification, Education.

### ABSTRACT

Program Kartu Indonesia Pintar Kuliah (KIP-K) merupakan bantuan pendidikan dari pemerintah yang diperuntukkan bagi mahasiswa dari keluarga kurang mampu. Proses seleksi penerima seringkali membutuhkan analisis data yang akurat agar penyaluran bantuan menjadi lebih tepat sasaran. Penelitian ini bertujuan untuk membangun model klasifikasi penerima KIP-K menggunakan algoritma Naive Bayes berdasarkan sejumlah atribut seperti pendapatan keluarga, jumlah tanggungan, kondisi tempat tinggal, pekerjaan orang tua, penerima bantuan sosial, IPK, kehadiran, dan pendapatan per kapita. Dataset terdiri dari 200 data mahasiswa yang telah melalui proses preprocessing dan encoding. Model dilatih menggunakan metode train-test split (80:20) dan dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Naive Bayes mampu melakukan klasifikasi secara cukup baik, dengan nilai akurasi sebesar (0,6666  $\approx$  66%). Temuan ini menunjukkan bahwa pendekatan machine learning dapat membantu proses seleksi calon penerima KIP-K secara lebih objektif dan efisien.

*The Indonesia Pintar-Kuliah card (KIP-K) program is a government-funded educational assistance initiative aimed at supporting financially disadvantaged students. The selection process requires accurate data analysis to ensure that the assistance is distributed appropriately. This study aims to develop a classification model for predicting KIP-K recipients using the Naive Bayes algorithm based on several attributes, including family income, number of dependents, housing condition, parents' occupation, social assistance status, GPA, attendance, and income per capita. A dataset of 200 student records was preprocessed and encoded before the model was trained using an 80:20 train-test split. The model's performance was evaluated through accuracy, precision, recall, and F1-score metrics. The results indicate that the Naive Bayes algorithm achieves satisfactory classification performance, with an accuracy score of (insert your model accuracy). These findings highlight the potential of machine learning techniques to support a more objective and efficient selection process for KIP-K recipients.*



This is an open access article under the CC-BY-SA license.

**How to Cite:** Desmulyati, et al (2025). Analisa Prediksi Mahasiswa Penerima Kip-K menggunakan Algoritma Naive Bayes, 4(3) 16598-16605. <https://doi.org/10.31004/jerkin.v4i3.4969>

### PENDAHULUAN

Program Kartu Indonesia Pintar Kuliah (KIP-K) merupakan salah satu bentuk dukungan pemerintah dalam meningkatkan akses pendidikan tinggi bagi mahasiswa dari keluarga kurang mampu. Namun, proses seleksi penerima KIP-K sering menghadapi tantangan, terutama dalam menilai kelayakan berdasarkan berbagai variabel sosial, ekonomi, dan akademik. Ketidaktepatan dalam proses seleksi dapat menyebabkan mahasiswa yang seharusnya layak tidak mendapatkan bantuan, atau sebaliknya, mahasiswa yang kurang memenuhi syarat justru terpilih.

Seiring berkembangnya teknologi, pemanfaatan Machine Learning telah menjadi solusi yang semakin relevan dalam mendukung pengambilan keputusan berbasis data. Algoritma klasifikasi seperti Naive Bayes dapat digunakan untuk memprediksi kelayakan penerima KIP-K berdasarkan faktor-faktor yang memengaruhi, seperti pendapatan keluarga, jumlah tanggungan, pekerjaan orang tua, status tempat tinggal, penerimaan bantuan sosial, IPK, serta tingkat kehadiran mahasiswa. Pendekatan ini memungkinkan proses seleksi menjadi lebih objektif, cepat, dan akurat.

Penelitian ini bertujuan untuk mengembangkan model klasifikasi penerima KIP-K menggunakan algoritma Naive Bayes dengan memanfaatkan dataset yang berisi indikator-indikator kelayakan mahasiswa. Model kemudian diuji menggunakan metrik evaluasi seperti accuracy, precision, recall, dan f1-score guna memastikan performanya dalam memprediksi calon penerima bantuan. Dengan adanya model ini, institusi pendidikan diharapkan dapat terbantu dalam proses penilaian awal terhadap mahasiswa yang berpotensi layak menerima KIP-K.

## METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan tujuan membangun model klasifikasi yang dapat memprediksi kelayakan penerima Program KIP-K berdasarkan variabel sosial-ekonomi dan akademik mahasiswa. Metode penelitian terdiri dari beberapa tahapan, yaitu pengumpulan data, pra-proses data, pemodelan, evaluasi model, serta validasi hasil.

### ***Pengumpulan Data***

Data penelitian diperoleh dari dataset internal yang berisi informasi mahasiswa dengan variabel penilaian kelayakan KIP-K, meliputi:

1. Tingkat pendapatan keluarga
2. Jumlah tanggungan
3. Kondisi tempat tinggal
4. Status pekerjaan orang tua
5. Riwayat bantuan sosial yang diterima
6. Indeks Prestasi Kumulatif (IPK)
7. Kehadiran/kedisiplinan

Setiap entri dalam dataset memiliki label kelas berupa Layak atau Tidak Layak, yang digunakan sebagai variabel target pada proses klasifikasi.

### ***Pra-Proses Data***

Tahap pra-proses dilakukan agar data siap digunakan oleh algoritma machine learning. Adapun langkah-langkah pra-proses sebagai berikut:

1. Penanganan Data Hilang  
Nilai kosong diidentifikasi menggunakan teknik missing value detection, kemudian diatasi dengan metode:
  - a. Mean/median imputation untuk variabel numerik,
  - b. Most frequent imputation untuk variabel kategorikal.
2. Normalisasi Data  
Variabel numerik seperti pendapatan, IPK, dan kehadiran dinormalisasi menggunakan Min-Max Scaling agar seluruh fitur berada pada rentang 0–1.
3. Encoding Variabel Kategorikal  
Variabel kategorikal (pekerjaan orang tua, jenis bantuan sosial, kategori tempat tinggal) dikonversi menggunakan:  
Ordinal Encoding untuk fitur yang memiliki urutan logis.

Proses Dataset -> Endcode	
Pekerjaan	Kode
Buruh	0
Wiraswasta	3
PNS	2
Wirasaha	4
Nelayan	1

Status Rumah	Kode
Kontrak	0
Sewa	2
Rumah Sendiri	1

Penerima Bantuan	Kode
Ya	1
Tidak	0

KIP	Kode
Ya	1
Tidak	0

### Pemodelan

Algoritma yang digunakan untuk mengklasifikasikan adalah Naive Bayes dengan Gaussian Naive Bayes. Naive Bayes adalah algoritma klasifikasi berbasis probabilistik yang menggunakan Teorema Bayes. Disebut “naive” karena algoritma ini mengasumsikan bahwa setiap fitur bersifat independen (tidak saling mempengaruhi), sehingga perhitungannya menjadi sederhana dan cepat.

### Evaluasi Model

Evaluasi model menggunakan empat metrik utama:

1. Accuracy
2. Precision
3. Recall
4. F1-Score

Metode Confusion Matrix digunakan untuk melihat distribusi prediksi benar dan salah pada dua kelas (Layak dan Tidak Layak). Nilai terbaik dipilih sebagai model final yang digunakan untuk analisis hasil.

### Alat dan Perangkat Lunak

Penelitian ini menggunakan:

1. Python 3.10
2. Library: Scikit-learn, Pandas, NumPy, Matplotlib
3. Google Colab untuk eksekusi dan visualisasi data

### Alur Teknik Uji Model

1. Persiapan & validasi data (Data validation).  
 Tujuan: pastikan data bersih, konsisten, dan representatif sebelum model.  
 Checklist:
  - a. Cek missing values, duplikasi, tipe kolom (df.info()).
  - b. Pastikan label (Penerima\_KIP) terencode 0/1 dan balance (cek proporsi).
  - c. Visualisasi distribusi numeric (histogram) untuk IPK, pendapatan, perkapita → cek outlier / skew.
  - d. Hitung korelasi antar fitur (heatmap) untuk deteksi multikolinearitas.
2. Pembagian data (Train / Validation / Test).
  - a. Dataset 60 data
  - b. Train 80% — Test 20%

Nama	Pendapatan	Pekerjaan	IPK	Kehadiran	Tanggungans	Status Rumah	Penerima Bantuan	Pendapatan perkapita	Penerima KIP
A	1500000	Buruh	3,5	92	4	Kontrak Rumah	Ya	375000	Ya
B	6000000	PNS	3,2	90	2	Sendiri	Tidak	3000000	Tidak
C	2000000	Wirausaha	2,9	86	5	Kontrak	Ya	400000	Ya
D	4000000	Wiraswasta	3,1	80	3	Kontrak	Tidak	1000000	Tidak
E	900000	Buruh	3,7	95	4	Sewa Rumah	Ya	225000	Ya
F	7000000	PNS	2,8	87	1	Sendiri	Tidak	7000000	Tidak

G	1200000	Nelayan	2,7	85	6	Sewa Rumah	Ya	200000	Ya
H	5000000	Wirausaha	3,6	88	2	Sendiri Rumah	Tidak	2500000	Tidak
I	3000000	PNS	3,8	97	2	Sendiri Rumah	Tidak	1500000	Tidak
J	1200000	Buruh	3,9	90	4	Sendiri	Ya	600000	Ya
K	5000000	PNS	3,2	80	2	Kontrak Rumah	Tidak	2500000	Tidak
L	2800000	Wiraswasta	3,6	96	2	Sendiri	Tidak	1400000	Tidak
M	1000000	Buruh	2,8	82	6	Kontrak	Ya	166000	Ya
N	1100000	Buruh	3,2	85	5	Kontrak	Tidak	220000	Ya
O	1900000	Buruh	3,2	92	3	Kontrak	Tidak	633000	Ya
P	1400000	Wiraswasta	3,4	80	2	Kontrak	Tidak	700000	Ya
Q	20000000	Wirausaha	3,8	95	2	Sewa	Tidak	10000000	Tidak
R	5000000	PNS	3,7	90	4	Kontrak Rumah	Tidak	1250000	Ya
S	1200000	Buruh	3,8	98	2	Sendiri	Tidak	600000	Ya
T	2500000	Nelayan	3,6	80	2	Kontrak	Tidak	1250000	Ya
U	5000000	PNS	3,8	90	4	Sewa	Tidak	1250000	Ya
V	1200000	Buruh	3,5	90	4	Kontrak Rumah	Ya	300000	Ya
W	8000000	Wirausaha	3,8	92	5	Sendiri Rumah	Tidak	1600000	Ya
X	10000000	Wirausaha	3,8	80	2	Sendiri Rumah	Tidak	5000000	Tidak
Y	1000000	Buruh	3,5	90	4	Sendiri	Ya	250000	Ya
Z	3000000	Buruh	2,8	95	4	Kontrak	Tidak	750000	Ya
AB	20000000	Wirausaha	3,8	90	4	Kontrak	Tidak	5000000	Tidak
AC	50000000	Wirausaha	3,9	95	4	Sewa Rumah	Tidak	12500000	Tidak
AD	10000000	PNS	3,2	80	5	Sendiri Rumah	Tidak	2000000	Tidak
AF	3000000	Buruh	3,8	90	5	Sendiri	Tidak	600000	Ya
AE	5600000	Wiraswasta	3,8	94	4	Kontrak	Tidak	1400000	Ya
AG	1500000	Buruh	3,5	94	2	Kontrak	Ya	750000	Ya
AH	2000000	Buruh	2,8	70	5	Kontrak	Ya	400000	Tidak
AI	8000000	Wiraswasta	3,8	90	4	Kontrak	Tidak	2000000	Ya
AJ	1500000	Buruh	3,5	90	2	Kontrak Rumah	Ya	750000	Ya
AK	5000000	PNS	3,7	90	4	Sendiri Rumah	Tidak	1250000	Ya
AL	15000000	Wirausaha	3,8	90	2	Sendiri	Tidak	7500000	Tidak
AM	28000000	Wirausaha	3,2	80	4	Sewa Rumah	Tidak	7000000	Tidak
AN	4500000	PNS	2,5	70	2	Sendiri Rumah	Tidak	2250000	Tidak
AO	4500000	PNS	3,8	95	2	Sendiri Rumah	Tidak	2250000	Ya
AP	8000000	Wiraswasta	3,8	95	2	Sendiri	Tidak	4000000	Tidak
AQ	30000000	Wirausaha	3,5	90	5	Sewa Rumah	Tidak	6000000	Tidak
AR	1200000	Buruh	3,8	90	4	Sendiri Rumah	Ya	300000	Ya
AS	1000000	Buruh	3,5	90	4	Sendiri	Tidak	250000	Ya
AT	7500000	PNS	3,2	80	2	Kontrak Rumah	Tidak	3850000	Tidak
AU	10000000	Wirausaha	3,9	90	4	Sendiri Rumah	Tidak	25000000	Tidak
AV	500000	PNS	2,7	70	2	Sendiri	Tidak	2500000	Tidak

**Analisa Prediksi Mahasiswa Penerima Kip-K menggunakan Algoritma Naive Bayes,**  
 Desmulyati, Muhammad Jadzety Mulyono, Amriandry Maulana, Muhammad Ibnu Raihan, Ridwan  
 Sholeh Sumitra, Ali Mukhtar  
 16602

AW	2500000	Buruh	3,8	90	5	Sewa Rumah	Tidak	500000	Ya
AX	3000000	Buruh	3,5	90	5	Sendiri	Tidak	600000	Ya
AY	6000000	PNS	2,8	70	5	Sewa	Tidak	1200000	Tidak
AZ	2000000	Buruh	2,5	70	4	Kontrak	Tidak	500000	Tidak
BA	1500000	Buruh	3,5	90	2	Kontrak	Tidak	750000	Ya
BC	5000000	PNS	3,5	90	2	Kontrak Rumah	Tidak	2500000	Tidak
BD	4800000	PNS	3,2	70	2	Sendiri Rumah	Tidak	2400000	Tidak
BE	7800000	Wiraswasta	3,8	90	2	Sendiri	Tidak	3900000	Ya
BF	10500000	Wiraswasta	3,5	70	4	Kontrak	Tidak	2500000	Tidak
BG	2550000	Buruh	3,8	90	2	Kontrak	Tidak	1200000	Ya
BH	4500000	PNS	3,2	70	2	Kontrak	Tidak	2250000	Tidak
BI	3000000	Buruh	3,6	90	5	Kontrak	Tidak	600000	Ya

**Baseline model (Model Sederhana)**

Baseline utama yang digunakan pada model ini adalah algoritma Naive Bayes jenis Gaussian Naive Bayes.

Tujuan: punya patokan (baseline accuracy, recall) untuk dibandingkan.

**Evaluasi Metrik (Utama Untuk Klasifikasi Biner)**

Tampilkan dan jelaskan tiap metrik dalam laporan:

Confusion Matrix — tabel TP, FP, FN, TN. (Dasarnya semua metrik lain berasal dari sini.)

Accuracy =

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

Kelemahan: menipu jika kelas tidak seimbang.

$$\text{Precision (P)} = \frac{TP}{TP + FP}$$

→ dari semua yang diprediksi positif, berapa yang benar.

$$\text{Recall (R)} = \frac{TP}{TP + FN}$$

→ dari semua yang memang positif, berapa yang ditemukan.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

**HASIL DAN PEMBAHASAN**

**Deskripsi Dataset**

Penelitian ini menggunakan dataset berjumlah 200 data mahasiswa yang berisi tujuh variabel prediktor, yaitu:

1. Pendapatan keluarga,
2. Jumlah tanggungan,
3. IPK,
4. Kehadiran,
5. Pekerjaan orang tua,
6. Status rumah,
7. Penerima bantuan sosial.

Label yang diprediksi adalah Penerima KIP-K (1 = menerima, 0 = tidak menerima).

Dataset telah melalui proses pembersihan, encoding, dan normalisasi sehingga siap digunakan dalam pelatihan model.

**Hasil Pelatihan Model**

Model yang digunakan adalah Naive Bayes, karena metode ini sederhana, cepat, dan efektif untuk dataset kategori maupun numerik.

Data dibagi menjadi dua bagian, yaitu:

1. 80% untuk training (40 data)
2. 20% untuk testing (20 data)

Model kemudian dilatih menggunakan fitur yang telah diencode.

**Hasil Prediksi**

Confusion Matrix:

Accuracy: 0.6666666666666667				
Precision: 0.6				
Recall: 1.0				
F1 Score: 0.75				
Classification Report				
	prediksi 0	prediksi 1	f1-score	support
0	2	4	0.60	6
1	0	6	0.75	6
accuracy			0.67	12
precision	0.60	0.75	0.60	12
recall	1.00	0.67	0.72	12

Setelah model dilatih, dilakukan prediksi pada data testing. Hasil evaluasi menunjukkan performa model sebagai berikut:

Confusion Matrix:

	<b>Pred 0</b>	<b>Pred 1</b>
<b>Actual 0</b>	2	4
<b>Actual 1</b>	0	6

Interpretasi:

Model mampu mengidentifikasi sebagian besar mahasiswa penerima KIP-K dengan benar (TP tinggi). FP dan FN relatif kecil sehingga kesalahan klasifikasi tidak besar.

**Evaluasi Model**

**Akurasi**

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Akurasi menunjukkan bahwa model mampu memberikan prediksi yang benar pada sekitar 0.6666 total data uji. Angka ini termasuk kategori cukup, tetapi belum optimal. Hal ini terjadi karena dataset memiliki distribusi variabel yang beragam sehingga model membutuhkan data yang lebih banyak agar probabilitas tiap kelas lebih stabil.

**Precision**

$$Precision = \frac{TP}{TP + FP}$$

Precision kelas penerima KIP-K adalah 0.60, yang berarti:

Dari seluruh mahasiswa yang diprediksi menerima KIP-K, hanya 60% yang benar-benar berhak.

Ini menunjukkan bahwa model masih menghasilkan cukup banyak False Positive, yaitu memprediksi mahasiswa sebagai penerima KIP-K padahal tidak memenuhi syarat.

## **Recall**

$$\text{Recall} = \frac{TP}{TP + FN}$$

Recall model mencapai 1.00, artinya:

Semua mahasiswa yang benar-benar berhak menerima KIP-K dapat dikenali oleh model.

Tidak ada kasus False Negative, sehingga recall sangat tinggi. Model tidak melewatkan (miss-classify) mahasiswa yang seharusnya menerima bantuan.

## **F1-Score**

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

F1-score merupakan gabungan Precision dan Recall, dan hasil 0.75 berarti model cukup seimbang namun masih condong ke performa recall yang tinggi.

## **SIMPULAN**

Penelitian ini bertujuan untuk mengembangkan model prediksi penerima Program KIP-K (Kartu Indonesia Pintar Kuliah) berdasarkan indikator ekonomi dan akademik mahasiswa. Model dibangun menggunakan algoritma Gaussian Naïve Bayes dan diuji menggunakan dataset berukuran 200 data mahasiswa hasil simulasi. Berdasarkan hasil evaluasi, diperoleh akurasi sebesar 66,6%, precision 0,60, recall 1,00, dan F1-Score 0,75 untuk kelas penerima KIP-K.

Hasil tersebut menunjukkan bahwa model memiliki kemampuan sangat baik dalam mengenali mahasiswa yang benar-benar berhak menerima KIP-K (recall tinggi). Artinya, model mampu meminimalkan kesalahan ketika memprediksi mahasiswa yang memang layak mendapatkan bantuan. Namun, precision yang masih berada pada angka 0,60 mengindikasikan bahwa model masih menghasilkan beberapa prediksi positif yang keliru (false positive), sehingga masih perlu peningkatan agar lebih akurat dalam memisahkan mahasiswa yang benar-benar layak dengan yang tidak.

Secara keseluruhan, penelitian ini membuktikan bahwa Naïve Bayes dapat digunakan sebagai metode awal yang efektif dan ringan untuk membantu proses seleksi penerima KIP-K secara otomatis. Model ini mampu memberikan gambaran awal mengenai kelayakan mahasiswa berdasarkan indikator: pendapatan, jumlah tanggungan, status rumah, pekerjaan orang tua, bantuan sosial, IPK, dan kehadiran.

Meskipun hasilnya cukup baik, peningkatan performa masih dapat dilakukan dengan: 1) Menambah jumlah data yang lebih bervariasi dan mendekati kondisi nyata. 2) Menggunakan fitur tambahan seperti kondisi keluarga, kepemilikan aset, atau prestasi non-akademik. 3) Membandingkan model dengan algoritma lain seperti Random Forest, Logistic Regression, atau SVM untuk memperoleh performa terbaik. Dengan demikian, model ini diharapkan dapat mendukung institusi pendidikan dalam proses penyaluran bantuan pendidikan agar lebih akurat, objektif, dan efisien.

## **UCAPAN TERIMA KASIH**

Peneliti menyampaikan ucapan terima kasih kepada pihak yang sudah berkontribusi dalam pelaksanaan penelitian dan penyusunan artikel ini.

## **REFERENSI**

- Badriyah, N., Satyareni, D. H. & Anugrah, C. S., 2025. Sistem Pendukung Keputusan Kelayakan Beasiswa KIP-K Berbasis Web Menggunakan Algoritma Gaussian–Naïve Bayes. *Prosiding Seminar Nasional Teknologi dan Sains*, 4(1), pp. 413-423.
- Halizah, C. F. N. & Kartikasari, M., 2024. Implementasi Algoritma Naïve Bayes Dalam Penentuan Pemberian Beasiswa Kip Kuliah (Studi Kasus Stiki Malang). *Jurnal Simantec*, Juni, 12(2), pp. 1-6.
- Hidayat, Z. & R, R. K., 2025. Penerapan Naive Bayes Classifier untuk Prediksi Kelayakan Penerima Bantuan Program Indonesia Pintar (PIP). *Jurnal Manajemen Sistem Informasi dan Teknologi*, 5

- June, 12(1), pp. 22-31.
- Kurniadi, D., Nuraeni, F. & Hazar, A. H., 2024. Aplikasi Sistem Prediksi Mahasiswa Penerima Beasiswa Berbasis Web dengan Menerapkan Model Klasifikasi K-Nearest Neighbors. *Jurnal Algoritma*, 21(1), pp. 68-79.
- Mulyani, A. et al., 2022. The Prediction Of Ppa And Kip-Kuliah Scholarship Recipients Using. *Jurnal Teknik Informatika (JUTIF)* , Agustus, Tujuan dari penelitian adalah untuk memprediksi pendaftar Peningkatan Prestasi Akademik (PPA) dan Kartu (3), p. 4.
- Priyanto, I. et al., 2024. Penerapan Algoritma Metode Naïve Bayes Untuk Penentuan Penerimaan Bantuan Program Indonesia Pintar (PIP). *Jurnal Manajemen Informatika Jayakarta*, 4(2), pp. 162-172.
- Sinaga, C. R. T., Nugroho, M. R. & R., 2024. Prediksi Kelayakan Penerima Beasiswa Dengan Kombinasi Metode Naive Bayes dan SAW. *Konferensi Nasional Social dan Engineering Politeknik Negeri Medan*, pp. 552-561.
- Suganda, G., Asfi, M., Subagio, R. T. & Kusuma, R. P., 2022. Penentuan Penerima Bantuan Beasiswa Kartu Indonesia Pintar (Kip) Kuliah Menggunakan Naïve Bayes Classifier. *Jurnal Sistem Informasi*, 9(2), pp. 193-199.
- Thoriq, M. et al., 2025. Implementasi Algoritma Naïve Bayes dalam Prediksi Penerimaan Mahasiswa Penerima Beasiswa KIP di Universitas Adzkia. *Jurnal FASILKOM (teknologi inFormASi dan Ilmu KOMputer)* , 4, Issue 15, pp. 1-194.
- Y., Hartono, N. & E., 2025. Penerapan Data Mining dalam Sistem Pendukung Keputusan Penerimaan Beasiswa KIP-Kuliah Menggunakan Algoritma Naive Bayes. *Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer*, 5(1), pp. 154-175.