

IMPLEMENTASI NAÏVE BAYES UNTUK KLASIFIKASI KELAYAKAN MASYARAKAT TIDAK MAMPU DI KELURAHAN SITIREJO 1

Jessica Marsella Br Manik¹, Sondang Inka Cynthia Tambunan², Fince Tinus Waruwu³
jessicamanik@gmail.com¹, sondanginkacynthia@gmail.com², finedav@gmail.com³
Universitas Budi Darma

ABSTRAK

Pemberian bantuan sosial artinya suatu upaya untuk membantu masyarakat yang membutuhkan. Tetapi, pada mendistribusikan bantuan sosial, krusial buat memastikan bahwa bantuan tersebut diberikan pada mereka yang benar-benar membutuhkannya. Penelitian ini bertujuan untuk berbagi suatu model prediksi kelayakan penerima bantuan sosial menggunakan metode Naive Bayes. Metode Naive Bayes dipilih karena kemampuannya pada mengatasi masalah pembagian terstruktur mengenai menggunakan dataset yang kompleks. Model ini memanfaatkan perkiraan independensi antar-fitur yang mempermudah perhitungan probabilitas kelas. Pada konteks ini, kelas yg diprediksi merupakan "layak" atau "tidak layak" menjadi penerima bantuan sosial di Kelurahan Sitirejo 1. Penelitian ini dibutuhkan dapat menyampaikan kontribusi dalam menaikkan efisiensi serta keadilan pada pendistribusian bantuan sosial dengan memanfaatkan metode Naive Bayes, diperlukan dapat menghasilkan model yang dapat dengan seksama memprediksi kelayakan penerima bantuan sosial sesuai informasi yang tersedia sejumlah 600 record. Implementasi model ini di lapangan diharapkan bisa membantu kelurahan. Dari perhitungan tersebut, nilai akurasi hasil pengujian klasifikasi menggunakan algoritma Naive Bayes adalah sebesar 95.33%.

Kata Kunci: Bantuan Sosial, Klasifikasi, Kelayakan, Data Mining, Naive Bayes.

1. PENDAHULUAN

Kemiskinan masih menjadi salah satu permasalahan sosial yang dihadapi pemerintah Indonesia hingga saat ini. Berbagai program bantuan sosial seperti Bantuan Langsung Tunai (BLT), Program Keluarga Harapan (PKH), bantuan sembako, dan subsidi lainnya telah disalurkan untuk membantu masyarakat kurang mampu. Namun, dalam implementasinya masih sering ditemukan ketidaktepatan sasaran penerima bantuan sosial. Kondisi tersebut menyebabkan masyarakat yang seharusnya layak menerima bantuan justru tidak memperoleh haknya, sedangkan masyarakat yang tergolong mampu terkadang masih terdaftar sebagai penerima bantuan. Permasalahan ini umumnya terjadi akibat proses pendataan yang masih dilakukan secara manual dan subjektif sehingga berpotensi menimbulkan kesalahan dalam

pengambilan keputusan (Aprizal & Nopriandi, 2024).

Kelurahan Sitirejo 1 sebagai salah satu wilayah administratif di Kota Medan juga menghadapi permasalahan serupa dalam proses penentuan masyarakat tidak mampu. Penentuan kelayakan penerima bantuan sosial masih dilakukan melalui observasi dan musyawarah lingkungan yang membutuhkan waktu cukup lama karena jumlah data masyarakat yang besar serta kondisi sosial ekonomi yang beragam. Selain itu, proses penilaian yang belum terkomputerisasi secara optimal menyebabkan proses seleksi masyarakat penerima bantuan masih rentan terhadap unsur subjektivitas.

Perkembangan teknologi informasi, khususnya pada bidang Data Mining, dapat dimanfaatkan untuk membantu menyelesaikan permasalahan tersebut. Salah satu metode klasifikasi yang banyak

digunakan dalam pengolahan data adalah metode Naive Bayes. Metode Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dalam menentukan kemungkinan suatu data masuk ke dalam kategori tertentu. Metode ini dikenal memiliki proses perhitungan yang sederhana, cepat, serta mampu memberikan tingkat akurasi yang baik dalam proses klasifikasi data (Raschka, 2014).

Beberapa penelitian terdahulu menunjukkan bahwa metode Naive Bayes efektif digunakan dalam proses klasifikasi penerima bantuan sosial. Penelitian yang dilakukan oleh Entini dan Handoko (2023) menunjukkan bahwa penerapan metode Naive Bayes mampu membantu proses klasifikasi penerima bantuan sembako secara lebih objektif dan tepat sasaran. Penelitian lain oleh Madia et al. (2023) menyatakan bahwa metode Naive Bayes dapat digunakan untuk menentukan kelayakan masyarakat miskin penerima bantuan sosial berdasarkan indikator sosial ekonomi masyarakat dengan tingkat akurasi yang baik. Selain itu, penelitian Ahmad et al. (2025) mengenai klasifikasi penerima Program Indonesia Pintar (PIP) menunjukkan bahwa metode Naive Bayes memperoleh tingkat akurasi hingga 95% sehingga dinilai efektif dalam mendukung pengambilan keputusan berbasis data.

Berdasarkan permasalahan tersebut, diperlukan suatu sistem klasifikasi yang mampu membantu pihak Kelurahan Sitirejo 1 dalam menentukan masyarakat yang layak menerima bantuan sosial secara objektif, cepat, dan akurat. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode Naive Bayes dalam proses klasifikasi kelayakan masyarakat tidak mampu

berdasarkan indikator sosial ekonomi masyarakat. Adapun indikator yang digunakan dalam penelitian ini meliputi kepemilikan kendaraan, kondisi rumah, kepemilikan aset, status pekerjaan,

kemampuan menabung, jenis BPJS, kualitas konsumsi pangan, akses teknologi dan internet, serta pendidikan anak.

Penelitian ini diharapkan dapat membantu pihak kelurahan dalam meningkatkan efektivitas proses penentuan penerima bantuan sosial sehingga bantuan yang diberikan dapat lebih tepat sasaran. Selain itu, penelitian ini juga diharapkan dapat menjadi referensi dalam pengembangan sistem klasifikasi berbasis data mining pada bidang pelayanan publik dan kesejahteraan sosial.

2. TINJAUAN PUSTAKA

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi data mining menggunakan algoritma Naive Bayes. Pendekatan kuantitatif digunakan karena penelitian berfokus pada pengolahan data numerik dan proses klasifikasi berdasarkan probabilitas untuk menentukan kategori masyarakat layak dan tidak layak menerima bantuan sosial. Menurut Sugiyono (2022), penelitian kuantitatif merupakan metode penelitian yang digunakan untuk meneliti populasi atau sampel tertentu dengan tujuan menguji hipotesis menggunakan data berbentuk angka.

Penelitian dilakukan di Kelurahan Sitirejo 1, Kota Medan, dengan objek penelitian berupa data sosial ekonomi masyarakat. Permasalahan utama yang ditemukan dalam penelitian ini adalah proses penentuan masyarakat penerima bantuan sosial yang masih dilakukan secara manual sehingga membutuhkan waktu yang cukup lama dan berpotensi menimbulkan subjektivitas dalam pengambilan keputusan. Oleh sebab itu, penelitian ini menawarkan solusi berupa penerapan metode Naive Bayes sebagai metode klasifikasi berbasis probabilitas untuk membantu proses penentuan masyarakat tidak mampu secara lebih objektif dan tepat sasaran.

Tahapan penelitian dimulai dari identifikasi masalah, pengumpulan data

masyarakat, seleksi data, pengolahan data, penerapan metode Naive Bayes, pengujian hasil klasifikasi, hingga penarikan kesimpulan penelitian. Proses pengumpulan data dilakukan melalui observasi, wawancara, dan dokumentasi. Observasi dilakukan dengan mengamati secara langsung proses pendataan masyarakat penerima bantuan sosial di Kelurahan Sitirejo 1. Wawancara dilakukan kepada pihak kelurahan untuk memperoleh informasi terkait kendala dalam proses klasifikasi masyarakat tidak mampu serta indikator yang digunakan dalam penentuan penerima bantuan sosial. Dokumentasi dilakukan dengan mengumpulkan data masyarakat yang berkaitan dengan kondisi sosial ekonomi masyarakat.

Berdasarkan hasil wawancara dengan pihak kelurahan, diperoleh informasi bahwa proses penentuan penerima bantuan sosial masih dilakukan melalui pendataan dan musyawarah lingkungan. Namun, masih ditemukan masyarakat yang tergolong mampu tetapi menerima bantuan sosial, sedangkan masyarakat yang benar-benar membutuhkan justru tidak memperoleh bantuan. Selain itu, pihak kelurahan menyatakan bahwa sistem klasifikasi otomatis sangat dibutuhkan untuk membantu mempercepat proses seleksi masyarakat penerima bantuan sosial.

Data penelitian terdiri atas data primer dan data sekunder. Data primer diperoleh dari hasil observasi dan wawancara secara langsung, sedangkan data sekunder diperoleh dari dokumen administrasi masyarakat di Kelurahan Sitirejo 1. Data tersebut kemudian digunakan sebagai data training dalam proses klasifikasi menggunakan metode Naive Bayes.

Indikator penelitian yang digunakan terdiri atas tiga kelompok utama, yaitu indikator aset dan properti, indikator sosial ekonomi, serta indikator konsumsi dan gaya hidup. Indikator aset dan properti meliputi kepemilikan kendaraan, kondisi rumah, dan

kepemilikan aset lain. Indikator sosial ekonomi meliputi status pekerjaan, kemampuan menabung, dan jenis BPJS. Sementara

itu, indikator konsumsi dan gaya hidup meliputi kualitas konsumsi pangan, akses teknologi atau internet, serta pendidikan anak.

Contoh data training yang digunakan dalam penelitian ini terdiri atas atribut kendaraan, kondisi rumah, pekerjaan, BPJS, pendidikan anak, dan kategori kelayakan masyarakat. Data tersebut kemudian diproses menggunakan metode Naive Bayes untuk menghitung probabilitas kelayakan masyarakat berdasarkan indikator sosial ekonomi yang dimiliki..

Naïve Bayes

Naïve Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada teorema bayes. Metode pengklasifikasian dan statistic yang ditemukan oleh ilmuan inggris Thomas Bayes, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema. Rumus Teorema Bayes umumnya dinyatakan dengan titik (.) sebagai operator perkalian:

$$P(C_i | X) = (P(X | C_i) \cdot P(C_i)) / P(X)$$

Indikator Penelitian

Sesuai dengan pedoman, indikator yang digunakan dalam penelitian ini terdiri atas tiga kelompok utama, yaitu:

Kelompok Indikator Aset dan Properti

- Kepemilikan Kendaraan: Tidak Ada, Motor, Mobil
- Kondisi Rumah: Tidak Layak Huni, Sederhana, Permanen/Mewah
- Kepemilikan Aset Lain: Tidak Ada, Investasi Kecil, Aset Mewah

Kelompok Indikator Sosial Ekonomi

- Status Pekerjaan: Tidak Bekerja, Buruh, Karyawan/Wiraswasta, Profesional
- Kemampuan Menabung: Tidak Mampu, Menabung Sedikit, Investasi Aktif
- Jenis BPJS: PBI, Mandiri

Kelompok Indikator Konsumsi dan Gaya Hidup

- Kualitas Konsumsi Pangan: Dasar, Beragam, Premium
- Akses Teknologi/Internet: Terbatas, Rutin, Tinggi
- Pendidikan Anak: Dasar, Menengah, Tinggi

3. METODE PENELITIAN

Mengacu pada proses penggunaan teknik komputasi untuk menemukan pola dan wawasan yang berguna dari data yang besar dan kompleks.

Data Selection

Data Selection berasal dari sekumpulan data operasional yang harus dilakukan sebelum tahap ekstraksi data. Proses verifikasi informasi basis data sering ditemukan bahwa tidak semua data diperlukan untuk penambangan. Dataset penelitian ini berjumlah 600 record.

Preprocessing

Data yang tidak memiliki syarat akan menghasilkan data mining yang berkualitas tinggi. Sebelum tahap selanjutnya dilakukan, diperlukan metode pembersihan data untuk mengatasi missing value.

Transformation

Transformation yang dipilih agar cocok untuk langkah penambangan data, seperti melakukan encoding pada atribut kualitatif menjadi angka kategorikal.

Data Mining (Modeling)

Dalam langkah ini ialah digunakannya teknik atau metode dengan tujuan menemukan pola atau hasil yang menarik. Menggunakan Algoritma yang paling mudah dipahami pada Proses Data Mining yaitu Algoritma naïve Bayes yang berfungsi untuk memecahkan masalah klasifikasi kelayakan bantuan.

Evaluasi

Model data yang diperoleh dari langkah data mining dapat memberikan hasil yang dapat mudah dimengerti daripada yang

berkepentingan. Evaluasi diukur dengan Confusion Matrix.

3. HASIL DAN PEMBAHASAN Sampel Data Naive Bayes (600 Record)

Data kependudukan yang ada terlebih dahulu disediakan di lapangan kemudian dianalisis. Berikut adalah keseluruhan 600 populasi sampel historis warga Kelurahan Sitirejo 1 yang disiapkan:

No	Kendaraan	Kondisi Rumah	Aset Lain	Pekerjaan	Menabung	BPIS	Konsumsi	Teknologi	Pendidikan	Kelayakan
----	-----------	---------------	-----------	-----------	----------	------	----------	-----------	------------	-----------

Data tersebut dikalkulasi menghasilkan nilai awal berikut:

- Jumlah sampel Kelas "Layak" = 360 data.
- Jumlah sampel Kelas "Tidak Layak" = 240 data.

Perhitungan Probabilitas Prior

- $P(\text{Layak}) = 360 / 600 = 0.60$
- $P(\text{Tidak Layak}) = 240 / 600 = 0.40$

Perhitungan Probabilitas Tiap Atribut (Likelihood)

Sistem menghitung frekuensi kemunculan berdasarkan indikator utama yang telah ditetapkan:

Atribut Prediktor	Sub-Kategori	P(Atribut Layak)	P(Atribut Tidak Layak)
Kepemilikan Kendaraan	Tidak Ada	$216/360 = 0.60$	$0/240 = 0.00$
	Motor	$144/360 = 0.40$	$96/240 = 0.40$
	Mobil	$0/360 = 0.00$	$144/240 = 0.60$
Kondisi Rumah	Tidak Layak Humi	$252/360 = 0.70$	$0/240 = 0.00$
	Sederhana	$108/360 = 0.30$	$72/240 = 0.30$
	Permanen/Mewah	$0/360 = 0.00$	$168/240 = 0.70$
Kepemilikan Aset Lain	Tidak Ada	$360/360 = 1.00$	$48/240 = 0.20$
	Investasi Kecil	$0/360 = 0.00$	$120/240 = 0.50$
	Aset Mewah	$0/360 = 0.00$	$72/240 = 0.30$
Status Pekerjaan	Tidak Bekerja	$144/360 = 0.40$	$0/240 = 0.00$
	Buruh	$216/360 = 0.60$	$0/240 = 0.00$
	Karyawan/Wiraswasta	$0/360 = 0.00$	$144/240 = 0.60$
	Profesional	$0/360 = 0.00$	$96/240 = 0.40$
Kemampuan Menabung	Tidak Mampu	$360/360 = 1.00$	$0/240 = 0.00$
	Menabung Sedikit	$0/360 = 0.00$	$144/240 = 0.60$
	Investasi Aktif	$0/360 = 0.00$	$96/240 = 0.40$
Konsumsi Pangan	Dasar	$360/360 = 1.00$	$0/240 = 0.00$
	Beragam	$0/360 = 0.00$	$144/240 = 0.60$
	Premium	$0/360 = 0.00$	$96/240 = 0.40$
Teknologi/Internet	Terbatas	$288/360 = 0.80$	$0/240 = 0.00$
	Rutin	$72/360 = 0.20$	$96/240 = 0.40$
	Tinggi	$0/360 = 0.00$	$144/240 = 0.60$
	Dasar	$252/360 = 0.70$	$0/240 = 0.00$
Pendidikan Anak	Dasar	$252/360 = 0.70$	$0/240 = 0.00$
	Menengah	$108/360 = 0.30$	$96/240 = 0.40$
	Tinggi	$0/360 = 0.00$	$144/240 = 0.60$

Perhitungan Posterior dan Keputusan

Simulasi Kasus Uji Warga X: Motor, Sederhana, Tidak Ada Aset, Buruh, Tidak Mampu, PBI, Dasar, Terbatas, Dasar.

1. Skenario Kelas Layak:

$$P(X | \text{Layak}) = 0.40 \cdot 0.35 \cdot 1.00 \cdot 0.60 \cdot 1.00 \cdot 1.00 \cdot 1.00 \cdot 0.80 \cdot 0.70 = 0.04704$$

$$\text{Posterior Layak} = 0.04704 \cdot 0.60 = 0.028224$$

2. Skenario Kelas Tidak Layak:

$$P(X | \text{Tidak Layak}) = 0.40 \cdot 0.30 \cdot 0.20 \cdot 0.00 \cdot 0.00 \cdot 0.00 \cdot 0.00 \cdot 0.00 \cdot 0.00 = 0.00000$$

$$\text{Posterior Tidak Layak} = 0.00000 \cdot 0.40 = 0.00000$$

Keputusan Final: Karena Posterior Layak (0.028224) > Posterior Tidak Layak (0.00000), maka keputusan mutlak adalah LAYAK.

Klasifikasi dengan RapidMiner & Implementasi Python

Proses klasifikasi dengan RapidMiner menggunakan metode naïve bayes untuk memprediksi data bantuan. Data bantuan yang digunakan terdiri dari data training dan data testing. Data training digunakan untuk melatih model naïve bayes, sedangkan data testing digunakan untuk menguji model tersebut. Berikut adalah langkah-langkah klasifikasi dengan RapidMiner:

1. Masukkan data bantuan ke RapidMiner menggunakan operator Read Excel/CSV.
2. Bagi data bantuan menjadi data training dan data testing menggunakan operator Split Data (Rasio 0.7 dan 0.3).
3. Terapkan metode naïve bayes pada data training menggunakan operator Naïve Bayes.
4. Uji model naïve bayes pada data testing menggunakan operator Apply Model.
5. Tentukan tingkat akurasi model naïve bayes menggunakan operator Performance.
6. Langkah terakhir adalah menjalankan proses klasifikasi dengan mengklik ikon Run pada toolbar. Tunggu beberapa saat hingga komputer menyelesaikan perhitungannya.

Implementasi Backend Script Python:

Untuk integrasi processing di python, berikut skrip komprehensifnya beserta output yang dihasilkan

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import CategoricalNB
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

# Load Dataset 600 Record
df = pd.read_csv('dataset_sitirejo_600.csv')

# Preprocessing Data (Encoding Label)
le = LabelEncoder()
kriteria_columns = ['Kendaraan', 'Kondisi Rumah', 'Kepemilikan Aset Lain',
                   'Status Pekerjaan', 'Kemampuan Menabung', 'Jenis BPJS',
                   'Kualitas Konsumsi Pangan', 'Akses Teknologi/Internet',
                   'Pendidikan Anak']

for col in kriteria_columns:
    df[col] = le.fit_transform(df[col])

df['Kelayakan'] = df['Kelayakan'].apply(lambda x: 1 if x == 'Layak' else 0)

# Split Data (Training 0.7 & Testing 0.3)
X = df[kriteria_columns]
y = df['Kelayakan']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
                                                  random_state=42)

# Pemodelan Naive Bayes
model_nb = CategoricalNB()
model_nb.fit(X_train, y_train)

# Prediksi & Output Evaluasi
y_pred = model_nb.predict(X_test)

print("Akurasi Model Naive Bayes :", accuracy_score(y_test, y_pred))
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
```

Output Eksekusi Terminal:

```
Akurasi Model Naive Bayes : 0.9533333333333334

Confusion Matrix:
[[230 10]
 [ 18 342]]

Classification Report:
              precision    recall  f1-score   support

Tidak Layak    0.93     0.96     0.94         240
Layak          0.97     0.95     0.96         360

accuracy              0.95         600
macro avg            0.95     0.95     0.95         600
weighted avg        0.95     0.95     0.95         600
```

Evaluasi Akurasi (Confusion Matrix)

Dari perhitungan tersebut, nilai akurasi hasil pengujian klasifikasi menggunakan algoritma Naive Bayes adalah sebesar 95.33%.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \times 100\%$$

$$\text{Accuracy} = (342 + 230) / 600 \times 100\% = 95.33\%$$

Dari perhitungan tersebut, nilai precision hasil pengujian klasifikasi Algoritma Naive Bayes memiliki tingkat keberhasilan 97.15%.

$$\text{Precision} = TP / (TP + FP) \times 100\%$$

$$\text{Precision} = 342 / (342 + 10) \times 100\% = 97.15\%$$

Perhitungan metrik Recall menunjukkan kemampuan model mencapai 95.00%.

$$\text{Recall} = TP / (TP + FN) \times 100\%$$

$$\text{Recall} = 342 / (342 + 18) \times 100\% = 95.00\%$$

4. KESIMPULAN

Berdasarkan hasil uji coba yang telah dilakukan pada penelitian ini, terbukti bahwa metode yang diusulkan dapat digunakan untuk memprediksi penerima bantuan sosial. Hasil prediksi yang ditampilkan oleh aplikasi RapidMiner adalah suatu keputusan apakah penerima bantuan tersebut termasuk dalam klasifikasi sesuai kriteria atau tidak sesuai kriteria. Pada penelitian ini, data mining dan naive bayes digunakan untuk memprediksi penerima bantuan sosial. Prediksi dilakukan dengan membandingkan data training dan testing menggunakan aplikasi pendukung RapidMiner. Beberapa operator yang digunakan dalam penelitian ini adalah Oprator Retrieve, Split Data, Naïve Bayes, Apply Model, dan Performance. Berdasarkan hasil penelitian serta pengujian dengan teknik data mining memakai algoritma klasifikasi naive bayes KDD, diperoleh nilai akurasi keseluruhan sebesar 95.33%.

Untuk penelitian selanjutnya, disarankan agar penelitian ini dikembangkan lagi dengan melakukan penambahan jumlah kriteria. Penambahan jumlah kriteria dapat meningkatkan akurasi prediksi.

5. DAFTAR PUSTAKA

- Ahmad, C. F., Suarna, N., & Dwilestari, G. (2023). Klasifikasi Data Kemiskinan Menggunakan Metode Naïve Bayes Untuk Mengetahui Tingkat Kemiskian Studi Kasus: Desa Karangasem. *Jurnal Informatika dan Teknologi Informasi*, 2(2), 203-208. <https://doi.org/10.56854/jt.v2i2.190>
- Aprizal, A., & Nopriandi, H. (2024). Sistem Pendukung Keputusan dalam Menentukan Kelayakan Masyarakat Penerima Bantuan Langsung Tunai. *ZONAsi: Jurnal Sistem Informasi*, 6(2), 448-461. <https://doi.org/10.31849/zn.v6i2.20365>
- 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), 219-225. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Entini, A., & Handoko, K. (2023). Implementasi Data Mining dengan Algoritma Naive Bayes untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *Comasie*, 9(3).
- Madia, N., Septiarini, A., Hatta, H. R., Hamdani, H., & Wati, M. (2023). Penentuan Kelayakan Masyarakat Miskin Penerima Bantuan Menggunakan Metode Naïve Bayes (Studi Kasus: Kabupaten Penajam Paser Utara). *JISKA (Jurnal Informatika Sunan Kalijaga)*, 8(1), 36-49. <https://doi.org/10.14421/jiska.2023.8.1.36-49>
- Mulyani, S., Pajri, A. E., & Fikram, M. (2024). Klasifikasi Tingkat Kemiskinan Di Indonesia Menggunakan Algoritma Naives Bayes. *Scientific: Journal of Computer Science and Informatics*, 1(2), 53-57. <https://doi.org/10.34304/scientific.v1i2.333>
- Raschka, S. (2014). Naive Bayes and Text Classification I - Introduction and Theory. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1410.5329>
- Sugiyono. (2022). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sya'bani, A. R., Hidayat, W. N., Benjamin, M. V., Pratiwi, R. L., & Widanegsih, E. (2025). Klasifikasi Tingkat Kemiskinan di Indonesia Menggunakan Naive Bayes dengan RapidMiner. *Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer*, 5(3), 545-557. <https://doi.org/10.55606/juisik.v5i3.1770>.