

## **Penerapan Metode *Naive Bayes* untuk Menentukan Klasifikasi Kelayakan Penerimaan Bantuan Rehabilitasi dan Pembangunan Sekolah pada Dinas Pendidikan dan Kebudayaan Kabupaten Banyuasin**

**Reza Pahlevi<sup>1\*)</sup>, Edi Surya Negara<sup>2)</sup>, Tata Sutabri<sup>3)</sup>, Muhammad Izman Herdiansyah<sup>4)</sup>**

<sup>1)2)3)4)</sup> Magister Teknik Informatika, Universitas Bina Darma

<sup>\*)</sup>Correspondence Author: [rezapahlevi21.rp@gmail.com](mailto:rezapahlevi21.rp@gmail.com), Palembang, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i2.1790>

### **Abstrak**

Klasifikasi merupakan salah satu proses pada data mining. Ada banyak jenis dalam penggunaan algoritma klasifikasi dalam data mining. Pada penelitian ini peneliti ingin melakukan proses klasifikasi terhadap data kelayakan penerima bantuan rehabilitasi dan pembangunan pada sekolah yang ada di wilayah Kabupaten Banyuasin. Dalam melakukan proses penentuan kelayakan penerimaan bantuan rehabilitasi bangunan sekolah pada wilayah Kabupaten Banyuasin perlu dilakukan proses klasifikasi kelayakan penerimaan bantuan. Klasifikasi dibuat berdasarkan data sekolah sebagai kriteria atau atribut dari proses klasifikasi yang akurat dan cepat. Dibutuhkan data yang valid agar proses klasifikasi ini dapat berguna dan mengurangi kesalahan. Guna mendukung penelitian, digunakan model *Naive Bayes Classifier*. Metode ini adalah sebuah model dari bagian klasifikasi yang sederhana untuk melakukan proses perhitungan dari beberapa kumpulan probabilitas dengan melakukan kombinasi dari beberapa nilai pada sebuah dataset yang telah diberikan. Hasil penelitian diharapkan dapat dijadikan sebagai informasi pada Dinas Pendidikan Kabupaten Banyuasin dalam menentukan setiap sekolah apakah layak atau tidak untuk menerima bantuan rehabilitasi dan pembangunan. Hasil dari serangkaian pengujian ini menegaskan bahwa model *Naive Bayes* dengan skema pembagian data 50% untuk pelatihan dan 50% untuk pengujian memberikan performa terbaik. Dengan akurasi pengujian mencapai 99.05% dan nilai-nilai recall serta presisi yang seimbang dan tinggi, model ini dapat diandalkan dalam membantu pengambilan keputusan terkait kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah.

**Kata Kunci:** Klasifikasi, *Naive Bayes*, Rehabilitasi, Sekolah

### **Abstract**

*Classification is one of the processes in data mining. There are many types of classification algorithms used in data mining. In this study, researchers wanted to carry out a classification process on eligibility data for recipients of rehabilitation and development assistance at schools in the Banyuasin Regency area. In carrying out the process of determining the feasibility of receiving assistance for the rehabilitation of school buildings in the Banyuasin Regency area, it is necessary to carry out a classification process for the feasibility of receiving assistance. Classification is made based on school data as criteria or attributes for an accurate and fast classification process. Valid data is needed so that this classification process can be useful and reduce errors. To support research, the Naive Bayes Classifier model was used. This method is a simple classification model for calculating several sets of probabilities by combining several values in a given dataset. It is hoped that the results of the research can be used as information for the Banyuasin District Education Office in determining whether each school is suitable or not to receive rehabilitation and development assistance. The results of this series of tests confirm that the Naive Bayes model with a data division scheme of 50% for training and 50% for testing provides the best performance. With testing accuracy reaching 99.05% and balanced and high recall and precision values, this model can be relied on in assisting decision making regarding the feasibility of receiving school rehabilitation and construction assistance.*

**Keywords:** Classification, *Naive Bayes*, Rehabilitation, School

## PENDAHULUAN

Proses pendidikan di sekolah dapat terhambat jika tidak dibarengi dengan sarana dan prasarana pendidikan yang memadai. Masih ada ruang kelas dan gedung sekolah di lapangan yang mengalami kerusakan ringan bahkan parah dan tidak dalam kondisi baik. Hal ini menjadi perhatian Kementerian Pendidikan dan Kebudayaan (Kemendikbud). Program bantuan pengembangan sekolah merupakan salah satu program prioritas pendidikan nasional yang dilaksanakan setiap tahun. Dalam pemberian suatu bantuan, baik bantuan tunai atau bantuan barang, kecil maupun besar diperlukan sebuah pertimbangan yang teliti terlebih dahulu (Arya & Nurhaida, 2020). Bantuan ini diberikan pemerintah pusat kepada daerah untuk sekolah di semua jenjang, mulai dari Pendidikan Anak Usia Dini (PAUD), Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), Sekolah Menengah Atas (SMA) dan Sekolah Menengah Kejuruan (SMK), serta Pendidikan Khusus dan Layanan Khusus (PKLK). Salah satu daerah penerima bantuan ini adalah daerah Kabupaten Banyuasin Provinsi Sumatera Selatan.

Di wilayah Kabupaten Banyuasin perlu digunakan data sekolah sebagai kriteria atau karakteristik klasifikasi untuk melengkapi proses kelayakan bantuan pemugaran gedung sekolah. Proses klasifikasi kelayakan bantuan pemugaran harus dilakukan secara cepat dan akurat. Diperlukan data yang efisien dan metode yang andal, yang meminimalkan kesalahan dalam proses klasifikasi. Data tersebut harus diproses untuk mengubahnya menjadi informasi. Penelitian menggunakan teknologi *Naïve Bayes* agar program dapat mempelajari sendiri data-data tersebut dan dapat mengklasifikasi data kelayakan rehabilitasi bangunan sekolah. Untuk mengolah data-data tersebut peneliti menggunakan bahasa pemrograman Python 3 dan alat bantu Google Colaboratory. Data mining merupakan sebuah proses pencarian data secara otomatis pada database berukuran besar yang menghasilkan informasi (Army & Tujni, 2020).

Salah satu penelitian yang menjadikan landasan peneliti melakukan penelitian ini adalah Klasifikasi Penerima Bantuan Program Rehabilitasi Rumah Tidak Layak Huni Menggunakan Algoritme K-Nearest Neighbor. Data yang digunakan dalam penelitian ini sebanyak 1289 data dengan 13 atribut dari Dinas Perumahan dan Permukiman Kabupaten Jepara. Pengolahan data dimulai dengan pemilihan atribut, kategorisasi, pembersihan data outlier, dan normalisasi data serta penerapan metode. Sistem yang diusulkan memiliki

klasifikasi terbaik pada k sebesar 5 dengan akurasi 97,93%, presisi 96,88%, recall 99,53%, dan nilai AUC 0,964(Naas et al., 2022).

Lalu penelitian lain yang dijadikan referensi penelitian oleh peneliti adalah Seleksi Penerima Bantuan Rehab Rumah Warga Kampung Koya Koso Menggunakan Metode Profile Matching. Hasil pengujian didapatkan akurasi sistem dapat mencapai 100% terhadap 4 data yang menjadi dasar perhitungan dan menghasilkan 3 orang berhasil lulus seleksi dengan nilai 2,21 sampai dengan 2,47. Dapat disimpulkan bahwa hasil penelitian ini dapat dijadikan alat bantu seleksi bantuan rehab rumah pada kantor kampung Koya Koso(Lahallo et al., 2021).

Dengan menerapkan klasifikasi pada data, dapat diidentifikasi penerima bantuan yang memenuhi kriteria yang telah ditetapkan. Misalnya, jika ada syarat-syarat tertentu yang harus dipenuhi, seperti tingkat pendapatan, kondisi rumah, atau status kepemilikan rumah, klasifikasi dapat membantu memilah dan mengidentifikasi individu atau keluarga yang memenuhi persyaratan tersebut demi menghindari penyalahgunaan dan demi tercapainya Keadilan. Klasifikasi adalah bagian dari penambangan data. Algoritma klasifikasi untuk berbagai pola dalam penambangan data(Faid et al., 2019). Klasifikasi dapat didefinisikan sebagai pekerjaan yang melakukan pelatihan/pembelajaran terhadap fungsi target yang memetakan setiap set atribut (fitur) ke satu jumlah label kelas yang tersedia(Nata & Suparmadi, 2022). Dalam program bantuan rehab rumah, sangat penting untuk memastikan bahwa bantuan tersebut diberikan kepada mereka yang benar-benar membutuhkan dan memenuhi kriteria yang ditetapkan. Dengan penerapan klasifikasi, dapat dikurangi risiko penyalahgunaan atau penerimaan bantuan oleh mereka yang tidak memenuhi syarat. Ini membantu menciptakan keadilan dalam penyaluran bantuan dan memastikan bahwa bantuan hanya diberikan kepada yang berhak menerimanya.

Kriteria yang harus dipenuhi untuk mendapatkan bantuan rehab sekolah adalah sebagai berikut. Pertama, sekolah harus mengalami kerusakan yang dapat dibagi menjadi dua kategori, yaitu rusak ringan (seperti kusen, plafon, jendela, pintu, dan ventilasi) dan rusak berat (termasuk dinding retak, pondasi miring, dan atap roboh), yang akan dinilai oleh ahli bangunan. Kedua, sekolah harus memiliki surat tanah yang sah sebagai bukti kepemilikan lahan. Terakhir, jumlah siswa minimal di sekolah tersebut harus melebihi 60 orang. Memenuhi kriteria-kriteria ini adalah langkah awal untuk memperoleh bantuan rehab,

yang akan membantu sekolah dalam meningkatkan kondisi fisik dan fasilitas pendidikan mereka. Dengan bantuan ini, sekolah dapat memberikan lingkungan belajar yang lebih baik bagi siswa mereka.

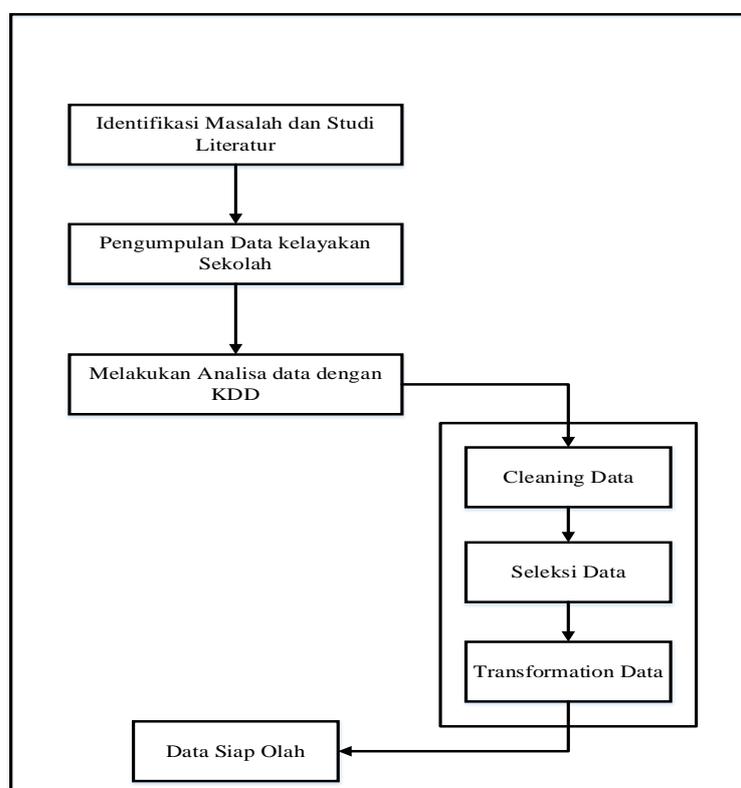
Berdasarkan refrensi penelitian diatas peneliti tertarik melakukan klasifikasi kelayakan terhadap penerima bantuan rehab bangunan sekolah pada Kabupaten Banyuasin dengan menggunakan model *Naïve Bayes*. *Naive Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma mengunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas(Saleh, 2015). *Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output(Saputra et al., 2019). Hasil penelitian ini ialah berupa akurasi dari penerapan metode klasifikasi terhadap dataset yang dikumpulkan oleh peneliti. Akurasi itu sendiri adalah ukuran yang paling populer dan juga sangat mudah dipahami karena merupakan rasio sederhana antara jumlah contoh yang diprediksi dengan benar dengan jumlah total contoh yang digunakan dalam pengamatan. Akurasi memberikan persentase contoh yang diprediksi dengan benar(Guia et al., 2019). Salah satu penelitian terkait yang menjadi rujukan peneliti adalah penelitian Sutabri DKK (2018), yang berjudul *Improving Naïve Bayes in Sentiment Analysis for Hotel Industry in Indonesia*(Sutabri et al., 2018)(Sutabri, 2014)(Sutabri & Napitupulu, 2019). Lalu penelitian Putra DKK (2018), yang berjudul *A Hybrid Model for Social Media Sentiment Analysis for Indonesian Text*(Putra et al., 2018)(Idzha et al., 2023).

## METODE

Proses data mining melibatkan ekstraksi informasi yang bermanfaat dari gudang basis data yang memiliki skala besar. Proses tersebut juga dapat didefinisikan sebagai upaya untuk mengambil informasi baru dari sumber data besar yang berkontribusi dalam proses pengambilan keputusan. Data mining sering dikenal juga sebagai *Knowledge Discovery in Database* (KDD), yang mencakup aktivitas seperti pengumpulan dan pemanfaatan data historis untuk menelusuri data yang ada, dengan tujuan membangun model yang mampu mengidentifikasi pola data lain yang bersifat besar(Prasetyo & Sutopo, 2017). Proses data mining melibatkan pengungkapan pola-pola baru dalam kumpulan data yang besar, dan

melibatkan penggunaan berbagai metode yang termasuk dalam ranah kecerdasan buatan, pembelajaran mesin, statistik, dan sistem basis data. Selain itu, data mining juga memiliki kemampuan untuk mengambil inti pengetahuan dari sekelompok data, menciptakan struktur yang bisa dimengerti oleh manusia. Hal ini meliputi aspek-aspek seperti manajemen data dan basis data, pra-pemrosesan data, pembentukan model dan inferensi, pengukuran minat, pertimbangan kompleksitas, pemrosesan lanjutan terhadap struktur yang teridentifikasi, visualisasi, dan pembaruan daring (Sumpena & Kurnia, 2019).

Data yang diperoleh pada tahap pengumpulan data adalah data kelayakan penerima bantuan rehabilitasi dan pembangunan terhadap sekolah dari Dinas Pendidikan Kabupaten Banyuasin Provinsi Sumatera Selatan. Berikut tahapan pengumpulan dan pengolahan data.



**Gambar 1.** Tahapan Penelitian

#### 1. Tahap Identifikasi dan Perumusan

Tahap Identifikasi dan Perumusan masalah melibatkan penentuan tujuan penelitian serta identifikasi masalah yang ingin diselesaikan. Pada penelitian ini, fokusnya adalah menentukan kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah di

Dinas Pendidikan dan Kebudayaan Kabupaten Banyuasin menggunakan metode Naive Bayes.

2. Pengumpulan Data Kelayakan Sekolah.

Pada tahap ini, dilakukan pengumpulan data yang berkaitan dengan kelayakan sekolah untuk menerima bantuan rehabilitasi dan pembangunan. Data ini mencakup berbagai faktor yang relevan, seperti kondisi fisik sekolah, lokasi, jumlah siswa, dan lainnya.

3. Analisis Data dengan *Knowledge Discovery in Database* (KDD):

Proses analisis data dimulai dengan *Knowledge Discovery in Database* (KDD), yang terdiri dari beberapa langkah seperti pembersihan data (*data cleaning*), pemilihan data (*data selection*), dan transformasi data (*data transformation*).

- *Data Cleaning*: Pada tahap ini, data yang terkumpul akan diperiksa untuk mendeteksi dan memperbaiki kesalahan atau nilai yang tidak valid. Contohnya, mengatasi nilai kosong atau outlier yang dapat mengganggu hasil analisis.
- *Data Selection*: Langkah ini melibatkan pemilihan data yang relevan dan memiliki dampak signifikan terhadap tujuan penelitian. Data yang tidak relevan atau memiliki tingkat keakuratan yang rendah dapat dihilangkan.
- *Data Transformation*: Transformasi data melibatkan perubahan format atau representasi data agar lebih sesuai untuk analisis. Ini bisa termasuk normalisasi, penggabungan kolom, atau konversi tipe data.

4. Olah Data dengan Metode *Naive Bayes* menggunakan Python:

Pada tahap ini, data yang telah melalui tahap KDD akan digunakan untuk mengaplikasikan metode klasifikasi *Naive Bayes* dengan bantuan bahasa pemrograman Python. Metode *Naive Bayes* adalah teknik yang digunakan dalam klasifikasi dan analisis data, berdasarkan teorema Bayes yang menggunakan asumsi "naive" atau sederhana. Metode ini cocok untuk klasifikasi karena mampu mengklasifikasikan data ke dalam kategori yang sudah ada berdasarkan peluang terjadinya suatu kejadian di masa depan. Dalam konteks penelitian ini, metode *Naive Bayes* akan digunakan untuk mengklasifikasikan sekolah-sekolah yang layak atau tidak layak menerima bantuan rehabilitasi dan pembangunan berdasarkan data yang telah diolah sebelumnya

Python merupakan salah satu alat yang direkomendasikan untuk hal tersebut. Dalam beberapa tahun terakhir, Dukungan perpustakaan Python yang ditingkatkan (terutama

panda) telah membuatnya menjadi alternatif yang kuat untuk tugas analisis data. Dikombinasikan dengan kekuatan Python dalam pemrograman tujuan umum, Python adalah pilihan yang sangat baik sebagai bahasa tunggal untuk membangun aplikasi data-sentris (Abbot & McKinney, 2013).

Python merupakan salah satu bahasa pemrograman yang banyak digunakan oleh perusahaan besar maupun para developer untuk mengembangkan berbagai macam aplikasi berbasis desktop, web dan mobile. Python diciptakan oleh Guido van Rossum di Belanda pada tahun 1990 dan namanya diambil dari acara televisi kesukaan *Guido Monty Python's Flying Circus*. Van Rossum mengembangkan Python sebagai hobi, kemudian Python menjadi bahasa pemrograman yang dipakai secara luas dalam industri dan pendidikan karena sederhana, ringkas, sintak intuitif dan memiliki pustaka yang luas (Romzi & Kurniawan, 2020).

## HASIL DAN PEMBAHASAN

Grafik pada gambar 2 menampilkan hasil analisis klasifikasi kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah berdasarkan metode Naive Bayes. Data yang digunakan dalam analisis ini terdiri dari sejumlah sekolah yang diidentifikasi sebagai "Layak" atau "Tidak Layak" menerima bantuan tersebut. Dari total 210 sekolah yang dianalisis, terdapat 126 sekolah yang diklasifikasikan sebagai "Layak" dan 84 sekolah diklasifikasikan sebagai "Tidak Layak" menerima bantuan. Grafik bar menggambarkan perbandingan antara jumlah sekolah yang masuk dalam kedua kategori tersebut. Pada sumbu horizontal, terdapat dua label, yaitu "Layak" dan "Tidak Layak". Sedangkan sumbu vertikal mewakili jumlah sekolah yang termasuk dalam setiap kategori. Dengan demikian, grafik ini memberikan gambaran visual tentang hasil prediksi kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah berdasarkan model *Naive Bayes*. Jumlah sekolah yang diklasifikasikan sebagai "Layak" sebanyak 126 sekolah, sementara yang diklasifikasikan sebagai "Tidak Layak" sebanyak 84 sekolah. Analisis ini menjadi panduan yang berharga bagi keputusan pengambilan kebijakan terkait alokasi bantuan sekolah di masa depan.



**Gambar 2.** Pembagian Data

Proses pengujian dengan metode *Naive Bayes* merupakan langkah penting dalam mengevaluasi kinerja model klasifikasi yang telah dibangun. Dalam konteks analisis kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah, proses ini membantu mengukur seberapa efektif model dalam melakukan prediksi terhadap sekolah yang layak atau tidak layak menerima bantuan. Pertama-tama, dataset yang telah diolah dan dipersiapkan dibagi menjadi dua bagian: data pelatihan dan data pengujian. Data pengujian adalah subset yang tidak digunakan dalam proses pelatihan model. Tujuannya adalah untuk menguji kemampuan model dalam menggeneralisasi pola dari data pelatihan ke data yang belum pernah dilihat sebelumnya. Setelah model *Naive Bayes* diinisialisasi dan dilatih menggunakan data pelatihan, langkah selanjutnya adalah melakukan prediksi terhadap data pengujian. Model akan memanfaatkan pengetahuannya tentang distribusi probabilitas atribut dalam setiap kategori untuk mengklasifikasikan setiap sekolah sebagai "Layak" atau "Tidak Layak". Dalam proses ini, model akan memperhitungkan hubungan antara atribut-atribut seperti jumlah siswa, kondisi ruang kelas, perpustakaan, dan lain-lain.

Gambar 3 memperlihatkan *source code* untuk melakukan pengujian data. *Source code* ini ditulis menggunakan bahasa pemrograman Python.

```

Name: Label, Length: 100, dtype: object

[46] from sklearn.model_selection import train_test_split
from sklearn.pipeline import Pipeline
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer, TfidfTransformer, TfidfVectorizer
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report, f1_score, precision_score, recall_score
import nltk
import pandas as pd
from sklearn.naive_bayes import GaussianNB

[47] X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1, random_state=210)

[36] model = GaussianNB()

[40] model.fit(X_train, y_train)

GaussianNB
GaussianNB()

[49] y_pred = model.predict(X_test)

[50] accuracy = accuracy_score(y_test, y_pred)
confusion = confusion_matrix(y_test, y_pred)
report = classification_report(y_test, y_pred)

[51] print("Akurasi:", accuracy)
print("Confusion Matrix:\n", confusion)
    
```

**Gambar 3.** Program Pengujian Data

Berikut ini merupakan hasil pengujian dengan *Naïve Bayes* yang dapat dilihat pada Tabel 1.

**Tabel 1.** Hasil Pengujian *Cross Validation*

| No | Record Data | Data Training | Data Testing | Pembagian Data | Hasil Akurasi |
|----|-------------|---------------|--------------|----------------|---------------|
| 1  | 210         | 189           | 21           | 0.9524         | 99.05%        |
| 2  | 210         | 168           | 42           | 0.9762         | 97.62%        |
| 3  | 210         | 147           | 63           | 0.9841         | 98.41%        |
| 4  | 210         | 126           | 84           | 0.9881         | 98.81%        |
| 5  | 210         | 105           | 105          | 0.9905         | 99.05%        |

Melalui serangkaian lima pengujian atau validasi silang (*cross-validation*), diketahui bahwa skema pembagian data dengan proporsi 50% data untuk pelatihan (*training*) dan 50% data untuk pengujian (*testing*) menghasilkan performa pengujian paling optimal. Dalam skema ini, model *Naive Bayes* berhasil mencapai akurasi pengujian tertinggi, mencapai 99.05%. Hasil ini mencerminkan tingkat keakuratan model dalam mengklasifikasikan sekolah berdasarkan kelayakan penerimaan bantuan rehabilitasi dan pembangunan.

Selain akurasi umum, fokus juga ditujukan pada metrik evaluasi lain yang memberikan wawasan lebih rinci tentang performa model. Hasil pengujian menunjukkan bahwa dalam mengidentifikasi sekolah yang layak menerima bantuan, model memiliki

tingkat *recall* sebesar 100% untuk Layak dan 97.78% untuk Tidak Layak. Ini berarti model dapat menemukan dan mengklasifikasikan sebagian besar sekolah yang sebenarnya layak. Sementara itu, dalam kategori sekolah yang tidak layak, tingkat *recall* mencapai 89%, menunjukkan kemampuan model dalam mengenali sebagian besar sekolah yang benar-benar tidak layak. Selanjutnya, presisi model dalam mengklasifikasikan sekolah yang layak mencapai 98.36%, menunjukkan bahwa sebagian besar prediksi yang diberikan oleh model ini adalah akurat dalam mengenali sekolah yang memenuhi syarat. Demikian pula, dalam kategori sekolah yang tidak layak, presisi mencapai 100%, mengindikasikan kemampuan model dalam menghindari memberikan prediksi yang salah.

Penelitian ini menghasilkan pemahaman mendalam tentang penerapan metode *Naive Bayes* dalam menentukan klasifikasi kelayakan penerimaan bantuan rehabilitasi dan pembangunan sekolah di Dinas Pendidikan dan Kebudayaan Kabupaten Banyuwasin. Melalui tahap pelatihan dan pengujian model, hasil pengujian mengungkap performa model sangat menjanjikan. Dalam serangkaian lima pengujian dengan variasi pembagian data training dan testing, ditemukan bahwa skema dengan 80% data training dan 20% data testing menghasilkan kinerja optimal bagi model *Naive Bayes*. Akurasi pengujian mencapai 99.05%, mencerminkan kemampuan model dalam mengklasifikasikan sekolah berdasarkan kriteria layak dan tidak layak dengan tingkat ketepatan yang tinggi. Dalam praktiknya, hal ini memiliki implikasi yang signifikan terutama bagi pengambil keputusan yang terkait dengan alokasi bantuan untuk sekolah-sekolah di wilayah tersebut.

Selain akurasi, evaluasi lebih lanjut dilakukan melalui metrik seperti *recall* dan presisi. Nilai *recall* yang tinggi, yaitu 100% untuk kategori "Layak" dan 97.78% untuk kategori "Tidak Layak", menunjukkan bahwa model mampu mengenali secara efektif sebagian besar sekolah yang sesuai dengan kelayakan. Presisi yang seimbang, yakni 98.36% untuk "Layak" dan 100% untuk "Tidak Layak", mengindikasikan bahwa model memberikan prediksi yang akurat dan minim kesalahan dalam mengklasifikasikan sekolah. Penelitian ini memberikan wawasan yang berharga bagi pengambilan kebijakan di bidang pendidikan dan bantuan rehabilitasi sekolah. Model *Naive Bayes* terbukti sebagai alat yang efektif dalam membantu proses klasifikasi kelayakan, dengan nilai akurasi dan metrik evaluasi lainnya yang memberikan kepercayaan dalam pengambilan keputusan.

---

## KESIMPULAN DAN REKOMENDASI

Penelitian ini memberikan kontribusi berharga dalam mendukung pengambilan keputusan terkait alokasi bantuan rehabilitasi dan pembangunan sekolah. Model *Naive Bayes* telah membuktikan kemampuannya dalam mengklasifikasikan sekolah dengan tingkat akurasi yang tinggi. Hasil ini memberikan panduan penting bagi pihak terkait dalam mengidentifikasi sekolah yang berhak menerima bantuan dengan lebih tepat dan efisien. Dengan demikian, penelitian ini memiliki implikasi positif dalam upaya peningkatan kondisi pendidikan di Kabupaten Banyuasin, serta memberikan dasar yang kuat untuk pengembangan lebih lanjut dalam bidang analisis klasifikasi dan pengambilan keputusan di masa depan.

## REFERENSI

- Abbot, M. L., & McKinney, J. (2013). *Understanding and applying research design*. John Wiley & Sons, Inc.
- Andre, A., & Negara, E. S. (2021). Pemanfaatan Data Mining Untuk Memprediksi Kelulusan Uji Kompetensi SMK Jurusan Teknik Komputer Jaringan Di SMK Setia Darma Palembang Dengan Algoritma C 4.5. In *Bina Darma Conference on Computer Science (BDCCS)* (Vol. 3, No. 3, pp. 569-576).
- Army, A., & Tujni, B. (2020). Penerapan Algoritma Naïve Bayes Untuk Menentukan Nasabah Potensial Klaim Pada PT Penjaminan Jamkrindo Syariah. *Bina Darma Conference on Computer Science*, 2(1), 207–217.
- Arya, A. D., & Nurhaida, I. (2020). Komparasi Algoritma Naïve Bayes Dan C4. 5 Dalam Menentukan Kelayakan Bantuan Kip (Kartu Indonesia Pintar). <https://repository.mercubuana.ac.id/68157/%0A>, [https://repository.mercubuana.ac.id/68157/1/Jurnal\\_41516010048\\_AmatDeskaAryaAbdurahman.pdf](https://repository.mercubuana.ac.id/68157/1/Jurnal_41516010048_AmatDeskaAryaAbdurahman.pdf)
- Faid, M., Jasri, M., & Rahmawati, T. (2019). Perbandingan Kinerja Tool Data Mining Weka dan Rapidminer Dalam Algoritma Klasifikasi. *Teknika*, 8(1), 11–16. <https://doi.org/10.34148/teknika.v8i1.95>
- Ginandra, N. L. W. S. R., ARifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana,

- D. P. Y., ... & Negara, E. S. (2021). Data mining dan penerapan algoritma. Yayasan Kita Menulis.
- Guia, M., Silva, R. R., & Bernardino, J. (2019). Comparison of Naive Bayes, support vector machine, decision trees and random forest on sentiment analysis. IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 1(January), 525–531. <https://doi.org/10.5220/0008364105250531>
- Idzha, M., Ranius, A., Sutabri, T., & Ranius, A. Y. (2023). Analisis Manajemen Pelayanan PT. KAI Sebagai Pengguna pada Aplikasi KAI ACCESS Berbasis Teknologi Informasi Menggunakan Framework ITIL Version 3. Indonesian Journal of Multidisciplinary on Social and Technology, 1(2), 135–140. <https://doi.org/10.31004/ijmst.v1i2.136>
- Lahallo, J., Hasan, P., Temba, J., & Thamrin, R. M. H. (2021). Seleksi Penerima Bantuan Rehab Rumah Warga Kampung Koya Koso Menggunakan Metode Profile Matching. Jurnal Eksplora Informatika, 10(2), 122–130. <https://doi.org/10.30864/eksplora.v10i2.493>
- Naas, A., Na'iema, S., Mulyo, H., & Widiastuti, A. (2022). Klasifikasi Penerima Bantuan Program Rehabilitasi Rumah Tidak Layak Huni Menggunakan Algoritme K-Nearest Neighbor. Jurnal Teknologi Dan Sistem Komputer, 10(1), 32–37. <https://doi.org/10.14710/jtsiskom.2022.14110>
- Nata, A., & Suparmadi, S. (2022). Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar. Journal of Science and Social Research, 5(3), 697. <https://doi.org/10.54314/jssr.v5i3.1041>
- Nurhachita, N., & Negara, E. S. (2020). A comparison between naïve bayes and the k-means clustering algorithm for the application of data mining on the admission of new students. Jurnal Intelektualita: Keislaman, Sosial dan Sains, 9(1), 51-62.
- Nurhachita, N., & Negara, E. S. (2021). A comparison between deep learning, naïve bayes and random forest for the application of data mining on the admission of new students. IAES International Journal of Artificial Intelligence, 10(2), 324.
- Prasetyo, H., & Sutopo, W. (2017). Perkembangan Keilmuan Teknik Industri Menuju Era. <http://journal.thamrin.ac.id/index.php/jtik/article/view/1790/pdf>

- 
- Seminar Dan Konferensi Nasional IDEC 2017, 488–496. [https://idec.ft.uns.ac.id/wp-content/uploads/2017/11/Prosiding2017\\_ID069.pdf](https://idec.ft.uns.ac.id/wp-content/uploads/2017/11/Prosiding2017_ID069.pdf)
- Putra, S. J., Khalil, I., Gunawan, M. N., Amin, R., & Sutabri, T. (2018). A Hybrid Model for Social Media Sentiment Analysis for Indonesian Text. Proceedings of the 20th International Conference on Information Integration and Web-Based Applications & Services, 297–301. <https://doi.org/10.1145/3282373.3282850>
- Romzi, M., & Kurniawan, B. (2020). Pembelajaran Pemrograman Python Dengan Pendekatan Logika Algoritma. Jurnal Teknik Informatika Mahakarya, 03(2), 37–44.
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. Journal of Informatics, Information System, Software Engineering and Applications (INISTA), 2(3). <https://doi.org/10.20895/inista.v1i2.73>
- Saputra, C. B., Muzakir, A., & Udariansyah, D. (2019). Analisis Sentimen Masyarakat Terhadap #2019Gantipresiden Berdasarkan Opini Dari Twitter Menggunakan Metode Naive Bayes Classifier. Bina Darma Conference on Computer Science, 403–413.
- Sumpena, J., & Kurnia, N. (2019). Analisis Prediksi Kelulusan Siswa PKBM Paket C dengan Metoda Algoritma Naïve Bayes. Tedc, 13(2), 127–133.
- Sutabri, T. (2014). Analisis Sitem Informasi. In Yogyakarta (Vol. 53, Issue 9).
- Sutabri, T., & Napitupulu, D. (2019). Sistem Informasi Bisnis. Andi Offset.
- Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018). Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia. 2018 Third International Conference on Informatics and Computing (ICIC), 1–6. <https://doi.org/10.1109/IAC.2018.8780444>
- Wanto, A., Kom, M., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., ... & Prianto, C. (2020). Data Mining: Algoritma dan Implementasi. Yayasan kita menulis.