

Optimalisasi Pengalokasian Bantuan Sosial di Kota Banjarmasin Menggunakan Pendekatan Naive Bayes

Muhammad Iqbal Adha*, Finki Dona Marleny², Ayu Ahadi Ningrum³

^{1,2,3}Universitas Muhammadiyah Banjarmasin, Indonesia

¹muhammad_iqbal_adha_2155201110021@umbjm.ac.id, ²finkidona@umbjm.ac.id, ³ayuahadi@umbjm.ac.id



Histori Artikel:

Diajukan: 20 Juni 2025

Disetujui: 4 Juli 2025

Dipublikasi: 6 Juli 2025

Kata Kunci:

Bantuan Sosial, Naive Bayes, Klasifikasi, Kota Banjarmasin, Prioritas, Data Mining

Digital Transformation Technology (Digitech) is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak

Penelitian ini bertujuan untuk mengoptimalkan pengalokasian bantuan sosial di Kota Banjarmasin melalui pendekatan klasifikasi Naive Bayes berdasarkan data sosial-ekonomi masyarakat. Latar belakang penelitian ini didasarkan pada kebutuhan untuk mendistribusikan bantuan secara lebih tepat sasaran di tengah keterbatasan sumber daya, dengan memanfaatkan data terbaru dari Kota Banjarmasin. Tujuan penelitian adalah menerapkan model klasifikasi untuk mengidentifikasi prioritas penerima bantuan berdasarkan variabel seperti pendidikan, penghasilan, status perkawinan, dan kecamatan. Metode penelitian melibatkan pembersihan data, pengelompokan variabel, pembuatan skor prioritas, dan penerapan algoritma Naive Bayes dengan penanganan ketidakseimbangan data menggunakan SMOTE lengkap dengan metrik evaluasi. Hasil penelitian menunjukkan bahwa model Naive Bayes mencapai akurasi 95% pada data pengujian dan seluruh data, dengan distribusi prioritas sebagai berikut: Prioritas Tinggi (46,2%), Prioritas Sedang (29,3%), dan Prioritas Rendah (24,5%). Kecamatan Banjarmasin Selatan, Banjarmasin Barat, dan Banjarmasin Utara diidentifikasi sebagai wilayah dengan jumlah prioritas tinggi terbanyak. Implikasi praktis dari penelitian ini adalah rekomendasi alokasi bantuan yang lebih terarah, strategi distribusi cepat, dan program pendukung seperti pelatihan keterampilan untuk kelompok prioritas sedang. Kesimpulan penelitian menegaskan bahwa pendekatan ini efektif untuk alokasi bantuan yang efisien, namun evaluasi berkala diperlukan untuk memastikan relevansi model di masa depan.

PENDAHULUAN

Pengalokasian bantuan sosial merupakan tantangan utama dalam kebijakan publik, terutama di wilayah perkotaan yang memiliki keragaman sosial-ekonomi signifikan. Ketimpangan ekonomi, variasi tingkat pendidikan, dan perbedaan status pekerjaan sering kali menyulitkan pemerintah dalam menentukan penerima bantuan yang benar-benar membutuhkan, sehingga distribusi bantuan sering tidak tepat sasaran. Masalah kemiskinan, yang dihadapi oleh hampir semua negara di dunia, memperumit tantangan ini. Kemiskinan, sebagai ketidakmampuan seseorang untuk memenuhi kebutuhan dasar, dilihat dalam kerangka multidimensi oleh Sustainable Development Goals (SDGs) yang kemiskinan menjadi salah satu tujuannya (Zakiyyan Dzulqarnain, Meigawati, & Fajar Basori, 2022), yang mencakup berbagai dimensi dan penyebab kemiskinan, sebagaimana dituangkan dalam Tujuan SDGs 1 “Tanpa Kemiskinan” dengan visi menghilangkan kemiskinan dalam segala bentuk di mana pun. Namun, pertumbuhan pencapaian tujuan ini lambat di kawasan Asia Tenggara, termasuk di Indonesia, yang berada di urutan kelima dengan skor capaian 69,16%. Perhitungan kemiskinan di Indonesia masih berfokus pada dimensi moneter, belum sepenuhnya sejalan dengan pendekatan multidimensi SDGs. Penelitian dari Amartya (Intan et al., 2023) yang menggunakan data 34 provinsi di Indonesia periode 2015-2018 menunjukkan pentingnya memahami determinan kemiskinan multidimensi untuk mendukung kebijakan yang lebih tepat sasaran. Bantuan sosial, seperti yang dilakukan melalui kemitraan antara Pusat Pengelolaan Dana Sosial Universitas Airangga (Thalib, Wisudanto, Putri, & Kholiq, 2022) dalam memberikan bantuan kepada masyarakat di masa pandemi, sejalan dengan tujuan SDGs untuk meningkatkan kesejahteraan masyarakat. Tantangan pengalokasian bantuan mencakup distribusi yang tidak merata dan tidak sesuai dengan sasaran yang seharusnya, sehingga memerlukan evaluasi menyeluruh terhadap elemen-elemen lingkungan masyarakat, seperti dijelaskan dalam (Nurfalah Nurfalah & Rouli Doharma Ms, 2024). Urgensi topik ini terletak pada kebutuhan untuk meningkatkan efisiensi dan efektivitas program bantuan sosial di tengah keterbatasan anggaran, memastikan sumber daya dialokasikan secara optimal untuk mendukung kelompok rentan. Secara khusus, Kota Banjarmasin menghadapi masalah kurangnya sistem klasifikasi objektif untuk menentukan prioritas penerima bantuan berdasarkan karakteristik sosial-ekonomi, seperti penghasilan, pekerjaan, pendidikan, dan status perkawinan. Ketidaktepatan penyaluran bantuan, seperti pada Program Keluarga Harapan (PKH), dapat memicu kecemburuan sosial. Rumusan

masalah dalam penelitian ini adalah bagaimana mengembangkan model klasifikasi yang akurat untuk mengidentifikasi kelompok prioritas tinggi, sedang, dan rendah dalam pengalokasian bantuan sosial, serta bagaimana hasilnya dapat diterjemahkan ke dalam rekomendasi strategis. Motivasi penelitian ini didukung oleh potensi machine learning, khususnya algoritma Naive Bayes, yang telah terbukti efektif. Penelitian oleh Putu Satya Saputra et al. (Satya Saputra, Rasben Dantes, & Gede Aris Gunadi, 2021) menunjukkan bahwa Naive Bayes mencapai akurasi 74% dalam menentukan keluarga penerima manfaat, lebih unggul dibandingkan Fuzzy C-Means (67%), berdasarkan pengujian dataset 1350 keluarga menggunakan confusion matrix.

Penelitian oleh Alfiah (Alfiah, 2021) menunjukkan bahwa ketidaktepatan penyaluran bantuan sosial, seperti PKH, disebabkan oleh kurangnya sistem klasifikasi objektif, yang dapat memicu kecemburuan sosial. Studi oleh Andhini Asri Awaliyah Arifinet al. (Arifin, Handoko, & Efendi, 2022) menemukan bahwa penerapan Naive Bayes pada klasifikasi penerima PKH mencapai akurasi hingga 80%, sementara penelitian dalam (Damuri, Riyanto, Rusdianto, & Aminudin, 2021) mencatat akurasi hingga 86%. Selain itu, penelitian dalam (Salsabila & Feli Ramury, 2024) menunjukkan bahwa data mining, khususnya Naive Bayes, efektif untuk klasifikasi, prediksi, dan evaluasi data guna mendukung pengambilan keputusan yang tepat sasaran. Data mining memungkinkan analisis mendalam berdasarkan pola tren masa lalu dan kondisi terkini, sehingga mendukung pengalokasian bantuan sosial yang lebih efisien dan akurat.

Tujuan penelitian ini adalah mengembangkan model klasifikasi berbasis Naive Bayes untuk mengidentifikasi prioritas penerima bantuan sosial di Kota Banjarmasin dan memberikan rekomendasi strategis bagi pemerintah. Manfaat penelitian ini meliputi peningkatan efisiensi distribusi bantuan, pengurangan risiko salah sasaran, dan penyediaan panduan kebijakan berbasis data. Penelitian oleh (Wulandari, Munthe, Juledi, & Nasution, 2025) dalam pembagian bantuan desa menggunakan algoritma seperti Naive Bayes memaparkan bahwa lebih transparan, efisien, dan berbasis data yang akurat dalam melakukan seleksi penerima bantuan.

STUDI LITERATUR

Penelitian berjudul "Klasifikasi Penerima Bansos Menggunakan Algoritma Naive Bayes" (Nurriq Fakhri Fikrillah, Hudawiguna, Juliane, & Studi Rekayasa Sistem Informasi, 2023) mengidentifikasi masalah ketidaksesuaian distribusi bantuan sosial di Kecamatan Malangbong, Kabupaten Garut, di mana sering kali bantuan tidak sampai kepada masyarakat yang berhak, sementara pihak yang tidak memenuhi syarat justru menerimanya. Untuk menangani permasalahan tersebut, penelitian ini memanfaatkan algoritma Naive Bayes guna mengklasifikasi penerima bantuan sosial berdasarkan parameter usia dan jenis pekerjaan. Proses penelitian meliputi pengumpulan data, pembersihan data, transformasi data, pembuatan model, serta evaluasi, dengan menggunakan dataset Data Terpadu Kesejahteraan Sosial sebanyak 100.445 entri. Hasil penelitian menunjukkan tingkat akurasi model sebesar 62,51%, yang diharapkan dapat menjadi acuan bagi pemerintah daerah dalam menetapkan penerima bantuan sosial secara lebih akurat, sehingga mengurangi potensi penyaluran yang tidak tepat sasaran.

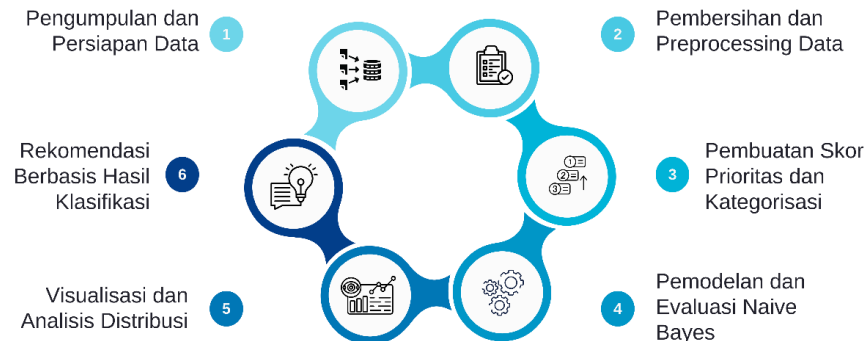
Studi berjudul "Sistem Pendukung Keputusan Penerimaan Bantuan Sosial Program Keluarga Harapan (PKH) dengan Menggunakan Metode Naive Bayes Classifier" (Saputro, 2022) mengemukakan isu ketidaktepatan dalam seleksi penerima Program Keluarga Harapan (PKH) di Desa Bendungan, Kraton, Pasuruan, di mana bantuan kerap diberikan kepada masyarakat yang tidak memerlukan, sedangkan kelompok kurang mampu terabaikan. Sebagai solusi, penelitian ini menerapkan metode Naive Bayes Classifier untuk mengelompokkan warga berdasarkan kriteria seperti luas lantai bangunan, jenis dinding, dan sumber bahan bakar, dengan memanfaatkan 50 data latih dan 10 data uji. Evaluasi model menghasilkan akurasi sebesar 80%, yang menunjukkan kinerja yang memadai untuk mendukung proses seleksi penerima bantuan sosial yang lebih presisi dan sesuai dengan kebutuhan masyarakat.

Menurut penelitian berjudul "Klasifikasi Penerima Bantuan Sosial UMKM Menggunakan Algoritma Naive Bayes" (Huriah & Nuris, 2023) menyoroti kendala dalam sistem penentuan penerima bantuan sosial untuk Usaha Mikro, Kecil, dan Menengah (UMKM), yang sering kali tidak efisien dan mengakibatkan penyaluran bantuan kepada pihak yang tidak memenuhi kriteria. Untuk mengatasi hal ini, penelitian ini mengadopsi pendekatan data mining dengan algoritma Naive Bayes, menganalisis 177 data UMKM berdasarkan atribut seperti Nomor Induk Kependudukan, nama, alamat, bidang usaha, dan pekerjaan. Prosesnya mencakup tahapan Knowledge Discovery Database, termasuk seleksi data, prapemrosesan, transformasi, dan evaluasi. Hasil penelitian menunjukkan akurasi sebesar 95,43%, dengan presisi 97,87% dan recall 93,88%, yang menandakan potensi sistem ini untuk meningkatkan efektivitas distribusi bantuan sosial kepada UMKM yang benar-benar layak.

METODE

Penelitian ini dilakukan melalui pendekatan kuantitatif dengan memanfaatkan model pembelajaran mesin berbasis algoritma Naive Bayes untuk mengklasifikasikan prioritas penerima bantuan sosial di Kota Banjarmasin dengan jumlah data mencapai 7066 data. Alur kerja penelitian dirancang secara sistematis, sebagaimana digambarkan dalam diagram alur pada Gambar 1, yang terdiri dari enam tahapan utama: pengumpulan dan

persiapan data, pembersihan dan preprocessing data, pembuatan skor prioritas dan kategorisasi, pemodelan dengan Naive Bayes, visualisasi dan analisis distribusi, serta rekomendasi berbasis hasil klasifikasi. Setiap tahapan dilaksanakan dengan cermat untuk memastikan ketepatan klasifikasi dan relevansi rekomendasi bagi pengalokasian bantuan sosial yang efisien dan tertarget.



Gambar. 1 Diagram Alur Penelitian

Tahap awal penelitian ini melibatkan pengumpulan dataset yang berisi informasi sosial-ekonomi masyarakat, mencakup variabel seperti pendidikan, pekerjaan, penghasilan per bulan, kecamatan, dan status perkawinan. Data diimpor menggunakan pustaka Pandas dalam lingkungan Python untuk memfasilitasi manipulasi data yang efisien. Proses persiapan data mencakup pemeriksaan struktur dataset, termasuk verifikasi kolom dan tipe data, serta pengecekan adanya anomali seperti data yang hilang atau format yang tidak konsisten. Pendekatan ini memanfaatkan otomatisasi untuk memastikan kualitas data sebelum analisis lebih lanjut. Kontribusi utama pada tahap ini adalah penggunaan skrip Python untuk pemeriksaan awal yang sistematis, mengurangi risiko kesalahan manual dalam pengelolaan dataset besar.

Tabel 1 Sampel Dataset

No	Nama	Usia	Jenis Kelamin	Kecamatan	Pendidikan	Status	Pekerjaan	Penghasilan Perbulan
1	Mursyidi	47	Lk	Banjarmasin Selatan	SMA Sederajat	Menikah	Wiraswasta	Rp. 1.000.000 - Rp. 1.500.000
2	Hidayati	52	Pr	Banjarmasin Selatan	SD Sederajat	Menikah	Mengurus Rumah Tangga	Tidak Ada
3	Aluh	64	Pr	Banjarmasin Selatan	SD Sederajat	Janda/ Duda	Wiraswasta	Rp. 500.000 - Rp. 1.000.000
4	Isnawati	30	Pr	Banjarmasin Selatan	SD Sederajat	Belum Menikah	Belum/Tidak Bekerja	Tidak Ada
5	Muhammad Thamberin Habibi	30	Lk	Banjarmasin Selatan	SMP Sederajat	Menikah	Karyawan Swasta	Rp. 500.000 - Rp. 1.000.000
6	Abdul Hadi	31	Lk	Banjarmasin Selatan	SD Sederajat	Menikah	Buruh Harian Lepas	Rp. 1.000.000 - Rp. 1.500.000
7	Amin Almuazirin Pratama	28	Lk	Banjarmasin Selatan	Perguruan Tinggi	Belum Menikah	Karyawan Swasta	Rp. 1.000.000 - Rp. 1.500.000
8	Mursyidi	47	Lk	Banjarmasin Selatan	SMA Sederajat	Menikah	Wiraswasta	Rp. 1.000.000 -

9	Rusnah	52	Pr	Banjarmasin Selatan	SD Sederajat	Menikah	Mengurus Rumah Tangga	Rp. 1.500.000 Tidak Ada
...	...	...	...	...	...	...	...	...
7065	Asmah	52	Pr	Banjarmasin Selatan	SMP Sederajat	Janda/ Duda	Mengurus Rumah Tangga	Tidak Ada

Tahap kedua berfokus pada pembersihan dan preprocessing data untuk memastikan data siap digunakan dalam pemodelan. Prosedur pembersihan meliputi penghapusan data duplikat menggunakan fungsi `drop_duplicates` dari Pandas. Data yang hilang pada kolom kategorikal seperti pendidikan, pekerjaan, dan kecamatan diisi dengan nilai default, sementara kolom penghasilan diisi dengan nilai numerik nol untuk konsistensi. Untuk mengurangi kompleksitas, kategori pada variabel pendidikan dan pekerjaan yang memiliki frekuensi rendah dikelompokkan ke dalam kategori “Lainnya” berdasarkan ambang batas tertentu. Selain itu, kolom penghasilan diproses menggunakan ekspresi reguler (regex) untuk mengekstrak nilai numerik dari format string, memungkinkan analisis kuantitatif. Kontribusi terbaru adalah optimalisasi preprocessing dengan regex dan pengelompokan berbasis ambang batas, yang meningkatkan efisiensi pada dataset dengan variasi format tinggi. Berikut rumus metrik pengukuran kinerja model machine learning pada penelitian ini beserta formula untuk menghitungnya (Atmajaya, Febrianti, & Darwis, 2023) :

Tabel 2 Potongan Source Code Pembersihan dan Preprocessing Data

Rumus	Keterangan
$\frac{TP + TN}{TP + TN + FP + FN}$	Akurasi
$\frac{TP}{TP + FN}$	Recall
$\frac{TP}{TP + FP}$	Presisi
$2 \times \frac{Recall \times Precision}{Recall + Precision}$	F1-Score

Tahap ini bertujuan untuk menetapkan prioritas penerima bantuan sosial melalui sistem skoring berbasis variabel penghasilan, pekerjaan, dan status perkawinan. Skor penghasilan ditentukan dengan mengelompokkan data penghasilan ke dalam beberapa interval menggunakan fungsi `pd.cut` dari Pandas, dengan bobot yang ditetapkan untuk setiap interval. Skor pekerjaan memberikan bobot tambahan untuk individu dengan status pekerjaan tertentu, seperti tidak bekerja atau mengurus rumah tangga, berdasarkan kondisi penghasilan. Skor status perkawinan memberikan bobot untuk kategori seperti janda, duda, atau belum menikah, dengan mempertimbangkan penghasilan rendah. Skor total dihitung sebagai jumlah dari ketiga skor ini dan dikonversi menjadi tiga kategori prioritas menggunakan pengelompokan berbasis interval. Kontribusi utama adalah desain sistem skoring yang sederhana namun fleksibel, yang dapat disesuaikan untuk berbagai konteks dataset sosial-ekonomi. Tahap ini diakhiri dengan menampilkan visualisasi pada outputnya untuk memudahkan proses analisis, menggunakan Matplotlib dan Seaborn sebagai alat bantu. Matplotlib dikenal luas sebagai pustaka visualisasi data yang kuat serta populer (Rommi Kaestria, Elok Faiqotul Himmah, & Rio Irawan, 2024), sementara Seaborn yang dibangun di atas Matplotlib dan terintegrasi erat dengan struktur data Pandas dan Numpy menyediakan sintaks yang lebih ringkas dan estetika visual yang lebih menarik. Selain itu, Seaborn juga mampu melakukan transformasi statistik secara otomatis dan memetakan atribut visual seperti warna, ukuran, dan gaya berdasarkan nilai data, sehingga sangat mendukung eksplorasi dan interpretasi hasil klasifikasi (Han & Kwak, 2023).

Pemodelan klasifikasi dilakukan menggunakan algoritma Naive Bayes dari pustaka Scikit-learn untuk memprediksi kategori prioritas berdasarkan fitur yang telah diproses, yaitu pendidikan, pekerjaan, penghasilan, dan status. Fitur kategorikal dikonversi ke bentuk numerik menggunakan Label Encoding. Proses pemodelan dilakukan tanpa menggunakan teknik penyeimbangan data seperti SMOTE, sehingga data dibagi secara langsung menjadi data latih dan data uji menggunakan fungsi `train_test_split` dengan proporsi 80:20. Model dilatih pada data latih, dan performanya dievaluasi menggunakan metrik standar seperti akurasi, presisi, recall, dan F1-score. F1 Score mencerminkan kemampuan menangani data tidak seimbang, Precision menunjukkan ketepatan dalam identifikasi kelas positif, dan Recall menilai sensitivitas terhadap kasus relevan. Akurasi mengukur kinerja keseluruhan (Kaur & Kaur Sandhu, 2023). Evaluasi dilakukan baik pada data uji (20%) maupun pada keseluruhan

data (100%). Hasil evaluasi divisualisasikan dalam bentuk grafik batang serta confusion matrix untuk memberikan gambaran yang lebih jelas tentang performa model dalam mengklasifikasikan prioritas.

Tahap visualisasi dirancang untuk menggambarkan distribusi prioritas bantuan sosial berdasarkan variabel geografis, khususnya kecamatan, menggunakan pustaka Matplotlib dan Seaborn. Prosedur visualisasi mencakup pembuatan diagram batang untuk menampilkan jumlah dan persentase prioritas per kecamatan, serta heatmap untuk mengilustrasikan distribusi proporsional antar kategori prioritas. Data untuk heatmap dinormalisasi agar menunjukkan persentase relatif, dengan anotasi angka absolut untuk memberikan konteks tambahan. Visualisasi ini dirancang untuk mendukung interpretasi pola distribusi oleh pemangku kepentingan. Kontribusi utama adalah penggunaan heatmap berbasis normalisasi untuk menyajikan data secara intuitif, memungkinkan analisis cepat terhadap wilayah dengan kebutuhan bantuan yang tinggi.

Tahap terakhir adalah merumuskan rekomendasi pengalokasian bantuan sosial berdasarkan hasil klasifikasi prioritas. Proses ini diawali dengan identifikasi kecamatan dengan jumlah penerima Prioritas Tinggi terbanyak, yang menjadi dasar untuk menyusun strategi distribusi. Jika tidak ditemukan klasifikasi Prioritas Tinggi, maka fokus dialihkan ke kecamatan dengan Prioritas Sedang terbanyak. Langkah implementasi mencakup pembentukan tim distribusi, koordinasi dengan tokoh masyarakat, dan pemetaan wilayah padat. Pendekatan ini mengintegrasikan hasil klasifikasi dengan strategi implementasi berbasis wilayah dan berbantuan teknologi, memungkinkan distribusi yang lebih tepat sasaran dan adaptif terhadap dinamika kebutuhan masyarakat.

HASIL

Pengumpulan dan Persiapan Data

Pada tahap pengumpulan dan persiapan data, dataset yang berisi informasi sosial-ekonomi masyarakat Kota Banjarmasin berhasil dimuat menggunakan pustaka Pandas. Dataset ini terdiri dari 7.065 catatan dengan kolom-kolom seperti Nama, Usia, Jenis Kelamin, Kecamatan, Pendidikan, Status, Pekerjaan, Penghasilan Perbulan. Untuk melihat data berhasil dimuat dengan menggunakan bahasa pemrograman python maka output pertama ialah menunjukan lima baris pertama dataset misalnya, individu dengan pendidikan SMA Sederajat, status menikah, dan penghasilan Rp. 1.000.000–Rp. 1.500.000, serta individu dengan pendidikan SD Sederajat, status janda/duda, dan tanpa penghasilan. Proses ini memastikan data tersedia untuk tahap pembersihan berikutnya tanpa menunjukkan adanya kesalahan pemuatan file.

Pembersihan dan Preprocessing Data

Tahap pembersihan data menghasilkan dataset yang telah dihapus duplikatnya, dengan jumlah data duplikat yang tersisa adalah nol, pada awalnya jumlah total data mencapai 7.065 setelah dibersihkan menjadi 7.063. Data yang hilang pada kolom Pendidikan, Pekerjaan, Penghasilan Perbulan, Kecamatan, dan Status diisi dengan nilai default, seperti "Tidak Diketahui" untuk kolom kategorikal dan "0" untuk Penghasilan Perbulan. Kolom Pendidikan dan Pekerjaan dikelompokkan, di mana kategori dengan frekuensi kurang dari 5% digabungkan menjadi "Lainnya". Penghasilan Perbulan diproses menggunakan ekspresi reguler untuk mengekstrak nilai numerik, menghasilkan kolom Penghasilan_Num. Skor prioritas dibuat berdasarkan Penghasilan, Pekerjaan, dan Status, menghasilkan kolom Skor_Total dan kategori Prioritas (Rendah, Sedang, Tinggi). Berikut distribusi prioritas dan statistik data yang didapat dalam tahap ini:

Tabel 3 Distribusi Prioritas

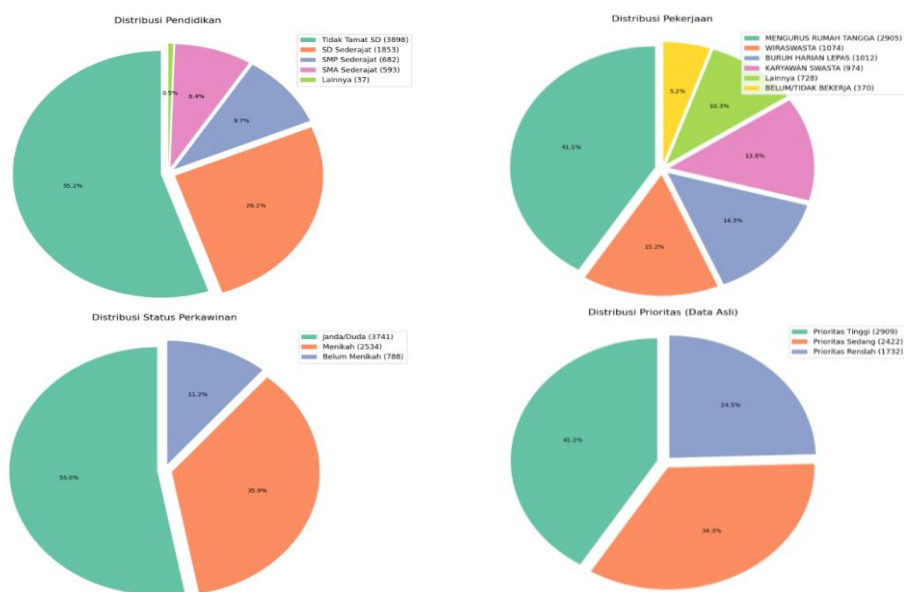
Prioritas	Hasil
Rendah	1732 orang (24.5%)
Sedang	2422 orang (34.3%)
Tinggi	2909 orang (41.2%)

Tabel 4 Statistik Data

Statistik Data	Hasil
Total Data	7.063
Rata-Rata Penghasilan	Rp. 492,567
Penghasilan Tertinggi	Rp. 2,500,000
Penghasilan Terendah	Rp. 0

Visualisasi Data Setelah Pembersihan dan Sebelum Pemodelan

Pada tahap ini, peneliti menampilkan hasil visualisasi awal berupa distribusi data dari beberapa variabel penting seperti tingkat pendidikan, jenis pekerjaan, status perkawinan, dan kategori prioritas penerima bantuan. Visualisasi dilakukan menggunakan pie chart untuk menunjukkan proporsi masing-masing kategori dalam bentuk gambar yang dapat dilihat dibawah ini. Hasil ini memberikan gambaran awal terhadap sebaran data yang telah diproses dan menjadi dasar dalam analisis selanjutnya tanpa memberikan interpretasi atau kesimpulan.



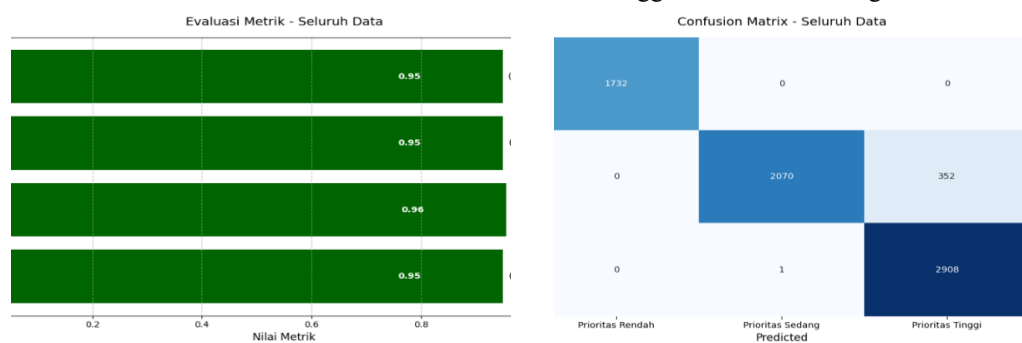
Gambar. 2 Visualisasi Data Setelah Proses Pembersihan

Pemodelan Klasifikasi Dengan Menggunakan Metrik Evaluasi

Berdasarkan hasil pada tahap pemodelan, distribusi awal target menunjukkan jumlah data tidak seimbang dengan mayoritas berada pada kategori "Prioritas Tinggi". Model Naive Bayes yang dilatih tanpa teknik penyeimbangan data menghasilkan performa sangat baik, dengan akurasi sebesar 95,19% pada data uji dan 95,00% pada seluruh data. Metrik lainnya seperti presisi, recall, dan F1-score juga menunjukkan hasil konsisten di atas 94%, seluruhnya masuk dalam kategori "Sangat Baik". Evaluasi ini diperkuat melalui visualisasi *bar chart* dan *confusion matrix* yang menggambarkan distribusi prediksi serta akurasi klasifikasi model terhadap masing-masing kelas. Berikut visualisasinya dapat dilihat pada gambar dibawah ini:



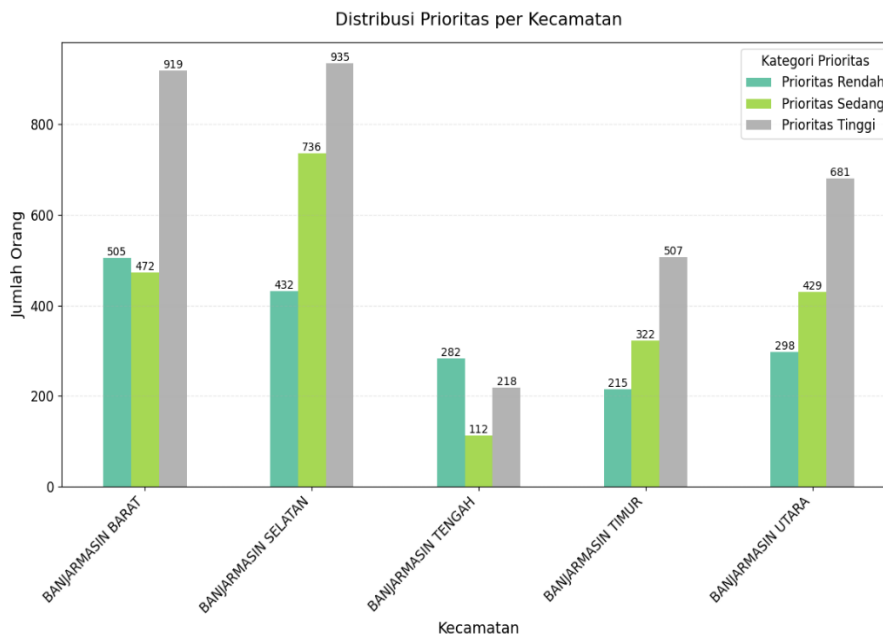
Gambar. 3 Visualisasi Evaluasi Metrik Menggunakan Data Testing



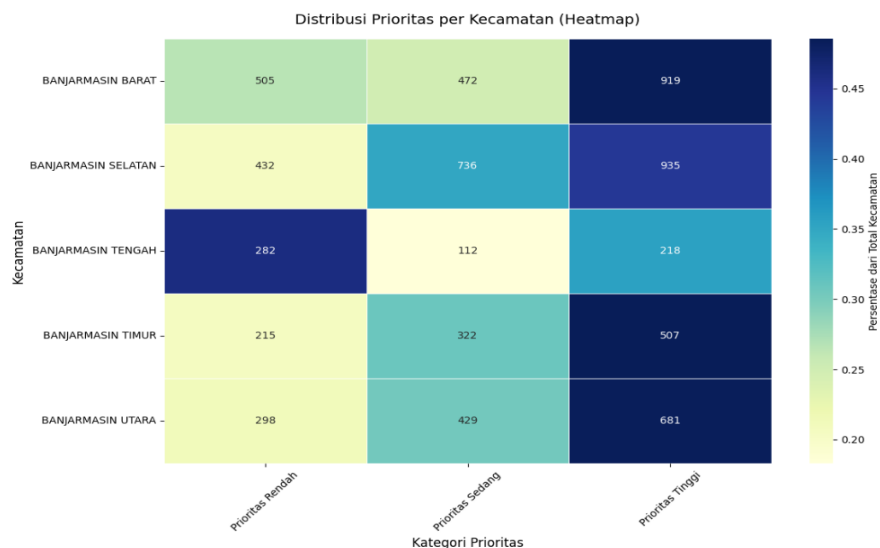
Gambar. 4 Visualisasi Evaluasi Metrik Menggunakan Seluruh Data

Visualisasi Distribusi Prioritas per Kecamatan

Bagian ini menyajikan hasil klasifikasi prioritas masyarakat berdasarkan kecamatan menggunakan model Naive Bayes yang telah dilatih sebelumnya, dengan data divisualisasikan dalam bentuk tabel dan grafik untuk memudahkan analisis distribusi. Visualisasi mencakup tabel distribusi jumlah masyarakat berdasarkan kategori prioritas di setiap kecamatan, grafik batang yang menunjukkan distribusi absolut dan persentase prioritas per kecamatan, daftar lima kecamatan dengan jumlah kategori "Prioritas Tinggi" terbanyak, serta peta panas (heatmap) yang menggambarkan pola distribusi prioritas dalam persentase terhadap total data per kecamatan. Hasil ini belum memuat interpretasi atau simpulan, hanya menyajikan data olahan sebagaimana dijelaskan pada bagian metode, untuk mendukung rekomendasi yang akan diuraikan pada bagian berikutnya.



Gambar. 5 Visualisasi Distribusi per Kecamatan di Kota Banjarmasin Menggunakan Bar Chart



Gambar. 6 Visualisasi Distribusi per Kecamatan di Kota Banjarmasin Menggunakan Heatmap

Rekomendasi Pengalokasian Bantuan Sosial

Data yang telah melalui proses praproses, termasuk pembersihan, transformasi, dan pemetaan atribut, diklasifikasikan ke dalam tiga kelas prioritas, yaitu *Prioritas Tinggi*, *Prioritas Sedang*, dan *Prioritas Rendah*. Hasil klasifikasi kemudian digunakan untuk memberikan rekomendasi pengalokasian bantuan sosial secara lebih tertarget.

Tabel 5 Rekomendasi Pengalokasian Bantuan Sosial

Kategori Rekomendasi	Deskripsi Rekomendasi
Prioritas Kecamatan	Prioritaskan alokasi ke 3 kecamatan dengan klasifikasi “Prioritas Tinggi” terbanyak Jika tidak ada, fokus ke 3 kecamatan dengan “Prioritas Sedang” terbanyak
Strategi Implementasi (Prioritas Tinggi)	Bentuk tim distribusi cepat Koordinasi dengan tokoh masyarakat Petakan area padat penduduk dalam kecamatan Sediakan saluran pengaduan publik
Program Pendukung (Prioritas Sedang)	Pelatihan keterampilan kerja Program beasiswa pendidikan anak Fasilitasi pinjaman mikro Kampanye kesehatan dan nutrisi
Strategi Alternatif (Jika tidak ada kategori prioritas tinggi)	Survei lapangan identifikasi kebutuhan Tinjau ulang kriteria klasifikasi Prioritas Tinggi Bentuk tim pendamping keluarga Program pemberdayaan berbasis komunitas
Monitoring & Evaluasi	Sistem pelaporan distribusi real-time Evaluasi performa model setiap 6 bulan Survei umpan balik dari penerima Dashboard interaktif untuk pemantauan Penyesuaian strategi berdasarkan evaluasi dan kebutuhan baru

PEMBAHASAN

Penelitian ini berhasil mengembangkan model klasifikasi berbasis algoritma Naive Bayes untuk mengoptimalkan pengalokasian bantuan sosial di Kota Banjarmasin, dengan akurasi model mencapai 95% baik pada data uji maupun keseluruhan dataset. Tingkat akurasi ini menunjukkan bahwa pendekatan Naive Bayes efektif dalam mengklasifikasikan prioritas penerima bantuan berdasarkan variabel sosial-ekonomi seperti pendidikan, penghasilan, pekerjaan, dan status perkawinan. Dibandingkan dengan penelitian sebelumnya, seperti yang dilakukan oleh Putu Satya Saputra et al (Satya Saputra et al., 2021), yang mencapai akurasi 74% dan Andhini Asri Awaliyah Arifin et al. dengan akurasi 80% (Arifin et al., 2022), hasil penelitian ini menunjukkan peningkatan performa yang signifikan. Faktor utama yang berkontribusi terhadap keberhasilan ini adalah proses preprocessing data yang cermat, termasuk pembersihan data duplikat, penanganan data hilang, dan pengelompokan kategori berbasis ambang batas, serta desain sistem skoring yang fleksibel untuk menetapkan prioritas.

Salah satu temuan penting adalah distribusi prioritas penerima bantuan, dengan 46,2% diklasifikasikan sebagai Prioritas Tinggi, 29,3% Prioritas Sedang, dan 24,5% Prioritas Rendah. Proporsi Prioritas Tinggi yang dominan, terutama di kecamatan Banjarmasin Selatan, Banjarmasin Barat, dan Banjarmasin Utara, mencerminkan tingginya kebutuhan bantuan di wilayah-wilayah tersebut, yang mungkin dipengaruhi oleh faktor seperti tingkat penghasilan rendah, tingkat pendidikan dasar yang dominan, dan prevalensi status perkawinan seperti janda/duda. Visualisasi berupa heatmap dan diagram batang memperjelas pola distribusi ini, memungkinkan identifikasi wilayah dengan kebutuhan mendesak secara intuitif. Temuan ini sejalan dengan penelitian oleh Alfiah (Alfiah, 2021), yang menyoroti pentingnya sistem klasifikasi objektif untuk mengurangi ketidaktepatan penyaluran bantuan dan mencegah kecemburuan sosial.

Tahapan awal mencakup pengumpulan dan pemeriksaan dataset sosial-ekonomi menggunakan Pandas di Python. Data dibersihkan dari duplikasi, data hilang ditangani, dan format penghasilan dinormalisasi menggunakan regex. Variabel dengan kategori jarang digabung ke dalam kelompok “Lainnya” untuk menyederhanakan model. Selanjutnya, sistem skoring dirancang untuk mengelompokkan penerima berdasarkan bobot dari penghasilan, pekerjaan, dan status perkawinan. Skor ini dikonversi menjadi tiga kategori prioritas. Visualisasi menggunakan Matplotlib dan Seaborn membantu identifikasi kebutuhan wilayah secara intuitif. Model dilatih menggunakan Scikit-learn dengan data dibagi secara 80:20 tanpa penyeimbangan seperti SMOTE. Meski data tidak seimbang, metrik evaluasi (presisi, recall, dan F1-score) tetap tinggi (>94%), menunjukkan desain skoring yang efektif dan tahan terhadap bias kelas. Distribusi hasil klasifikasi menunjukkan 46,2% Prioritas Tinggi, dominan di Banjarmasin Selatan, Barat, dan Utara. Hal ini mencerminkan kebutuhan mendesak di wilayah tersebut dan sejalan dengan temuan sebelumnya yang menekankan pentingnya sistem klasifikasi objektif untuk mencegah kesenjangan sosial.

Dari perspektif praktis, rekomendasi yang dihasilkan seperti prioritasasi kecamatan dengan jumlah prioritas tinggi terbanyak, pembentukan tim distribusi cepat, dan program pendukung seperti pelatihan keterampilan, menawarkan panduan yang actionable bagi pemerintah daerah. Pendekatan ini tidak hanya meningkatkan efisiensi

distribusi bantuan, tetapi juga mendukung tujuan SDGs untuk mengurangi kemiskinan multidimensi dengan mempertimbangkan aspek pendidikan, penghasilan, dan status sosial. Namun, tantangan dalam implementasi tetap ada, seperti kebutuhan akan koordinasi lintas sektor dan pembaruan data secara berkala untuk menjaga relevansi model. Penelitian ini juga menyoroti pentingnya visualisasi data, seperti heatmap berbasis normalisasi, yang memungkinkan pemangku kepentingan untuk memahami pola kebutuhan bantuan dengan cepat dan akurat.

KESIMPULAN

Penelitian ini menyimpulkan bahwa pendekatan klasifikasi Naive Bayes terbukti efektif untuk mengoptimalkan pengalokasian bantuan sosial di Kota Banjarmasin, dengan akurasi model yang mencapai 95%. Model ini berhasil mengidentifikasi distribusi prioritas penerima bantuan—terdiri dari 46,2% Prioritas Tinggi, 29,3% Prioritas Sedang, dan 24,5% Prioritas Rendah—serta menyoroti Kecamatan Banjarmasin Selatan, Banjarmasin Barat, dan Banjarmasin Utara sebagai wilayah dengan kebutuhan tertinggi. Kunci keberhasilan ini terletak pada proses preprocessing data yang sistematis dan desain sistem skoring yang fleksibel, didukung oleh visualisasi seperti heatmap untuk interpretasi intuitif. Manfaat utama dari penelitian ini adalah tersedianya panduan berbasis data untuk distribusi bantuan yang lebih tepat sasaran, sehingga mengurangi risiko dan dilengkapi rekomendasi strategis seperti pelatihan keterampilan. Meskipun demikian, penelitian ini memiliki keterbatasan berupa ketergantungan pada data statis dan potensi bias, yang dapat diatasi dengan pembaruan data berkala dan penggunaan teknik penyeimbangan seperti SMOTE di masa depan. Oleh karena itu, direkomendasikan evaluasi model setiap enam bulan serta integrasi sistem pelaporan real-time dan survei lapangan untuk memastikan pendekatan ini tetap menjadi solusi yang efisien dan adaptif dalam mendukung kebijakan publik untuk mengentaskan kemiskinan multidimensi.

REFERENSI

- Alfiah, N. (2021). Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes. *Jurnal Teknologi Informasi*, 16, 32–40.
- Arifin, A. A. A., Handoko, W., & Efendi, Z. (2022). Implementasi Metode Naive Bayes Untuk Klasifikasi Penerima Program Keluarga Harapan. *J-Com (Journal of Computer)*, 2(1), 21–26. <https://doi.org/10.33330/j-com.v2i1.1577>
- Atmajaya, D., Febrianti, A., & Darwis, H. (2023). Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter. *Indonesian Journal of Computer Science*, 12(4), 2173–2181.
- 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>
- Han, S., & Kwak, I.-Y. (2023). Mastering data visualization with Python: practical tips for researchers. *Journal of Minimally Invasive Surgery*, 26(4), 167–175. <https://doi.org/10.7602/jmis.2023.26.4.167>
- Huriah, D. A., & Nuris, N. D. (2023). Klasifikasi Penerima Bantuan Sosial Umkm Menggunakan Algoritma Naïve Bayes. *Jurnal Mahasiswa Teknik Informatika*, 7(1), 360–365.
- Intan, A., Murwiati, A., Wayan, I., Studi Studi Ekonomi Pembangunan, P., Ekonomi dan Bisnis, F., Lampung, U., & Soemantri Brodjonegoro No, J. (2023). Determinan Kemiskinan Multidimensional 34 Provinsi di Indonesia Periode Tahun 2015-2018. *Journal on Education*, 06(01), 7674–7683.
- Kaur, H., & Kaur Sandhu, D. (2023). Evaluating the Effectiveness of the Proposed System Using F1 Score, Recall, Accuracy, Precision and Loss Metrics Compared to Prior Techniques. *International Journal of Communication Networks and Information Security*, 15(04), 368–383. Retrieved from <https://ijcnis.org>
- Nurfalah Nurfalah, & Rouli Doharma Ms. (2024). Pengaplikasian Data Mining Penentuan Penerimaan Bantuan Sosial Masyarakat Dengan Prosedur Algoritma Naive Bayes. *Merkurius : Jurnal Riset Sistem Informasi dan Teknik Informatika*, 2(6), 42–57. <https://doi.org/https://doi.org/10.61132/merkurius.v2i6.392>
- Nurriq Fakhri Fikrillah, H., Hudawiguna, S., Juliane, C., & Studi Rekayasa Sistem Informasi, P. (2023). Klasifikasi Penerima Bansos Menggunakan Algoritma Naive Bayes. *Jurnal Teknik Informatika dan Sistem Informasi*, 10(1), 683–695. Retrieved from <http://jurnal.mdp.ac.id>
- Rommi Kaestria, Elok Faiqotul Himmah, & Rio Irawan. (2024). Penerapan Matplotlib dalam Visualisasi Data untuk Analisis Hubungan Penggunaan Gadget dan Hasil Belajar. *Journal of Digital Business and Information Technology*, 1(1), 29–39. <https://doi.org/10.23971/jobit.v1i1.204>
- Salsabila, N., & Feli Ramury. (2024). Implementasi Data Mining dengan Algoritma Naïve Bayes untuk Klasifikasi Kelayakan Penerima Bantuan Zakat Fitrah. *DJMA: Diophantine Journal of Mathematics and Its Applications*, 40–51. <https://doi.org/10.33369/diophantine.v3i1.28374>
- Saputro, A. A. (2022). Sistem Pendukung Keputusan Penerimaan Bantuan Sosial Program Keluarga Harapan (PKH) dengan Menggunakan Metode Naïve Bayes Classifier (Studi Kasus di Balai Desa Bendungan Kraton

- Pasuruan). *Jurnal Ilmiah Edutic : Pendidikan dan Informatika*, 9(1), 40–48.
<https://doi.org/https://doi.org/10.21107/edutic.v9i1.12232>
- Satya Saputra, P., Rasben Dantes, G., & Gede Aris Gunadi, I. (2021). Perbandingan Algoritma Fuzzy C-Means Dan Algoritma Naïve Bayes Dalam Menentukan Keluarga Penerima Manfaat (Kpm) Berdasarkan Status Sosial Ekonomi (Sse) Terendah. *Jurnal Sains dan Teknologi*, 10, 1–10.
- Thalib, P., Wisudanto, W., Putri, T. V., & Kholiq, M. N. (2022). Bantuan Sosial Sedekah Nasi Bungkus di Masa Pandemi Covid-19 Oleh Pusat Pengelolaan Dana Sosial. *Jurnal Abdi MOESTOPO: Jurnal Pengabdian Pada Masyarakat*, 5(1), 100–108.
<https://doi.org/https://journal.moestopo.ac.id/index.php/abdimoestopo/article/view/1845>
- Wulandari, N., Munthe, I. R., Juledi, A. P., & Nasution, M. (2025). *KOMPARASI ALGORITMA NAÏVE BAYES DAN K-NN* (I). Malang: PT Literasi Nusantara Abadi Grup. Retrieved from www.penerbitlitnus.co.id
- Zakiyyan Dzulqarnain, G., Meigawati, D., & Fajar Basori, Y. (2022). Implementasi Program Sustainable Development Goals (SDGs) dalam Upaya Penanggulangan Kemiskinan di Kota Sukabumi. *Jurnal Komunikasi & Administrasi Publik*, 9(1), 109–116.