



Komparasi Algoritma

Naïve Bayes dan K-NN

dalam PEMBAGIAN BANTUAN DESA



Nurhikmah Wulandari
Ibnu Rasyid Munthe, S.T., M.Kom
Angga putra Juledi, S.Kom., M.Kom
Marnis Nasution, S.Kom., M.Kom

Komparasi Algoritma

Naïve Bayes dan K-NN

dalam **PEMBAGIAN
BANTUAN DESA**

Nurhikmah Wulandari
Ibnu Rasyid Munthe, S.T., M.Kom
Angga putra Juledi, S.Kom., M.Kom
Marnis Nasution, S.Kom., M.Kom



**KOMPARASI ALGORITMA NAÏVE BAYES DAN K-NN
DALAM PEMBAGIAN BANTUAN DESA**

Ditulis oleh:

Nurhikmah Wulandari
Ibnu Rasyid Munthe, S.T., M.Kom.
Angga putra Juledi, S.Kom., M.Kom.
Marnis Nasution, S.Kom., M.Kom.

Diterbitkan, dicetak, dan didistribusikan oleh
PT Literasi Nusantara Abadi Grup
Perumahan Puncak Joyo Agung Residence Blok B11 Merjosari
Kecamatan Lowokwaru Kota Malang 65144
Telp : +6285887254603, +6285841411519
Email: literasinusantaraofficial@gmail.com
Web: www.penerbitlitnus.co.id
Anggota IKAPI No. 340/JTI/2022



Hak Cipta dilindungi oleh undang-undang. Dilarang mengutip atau memperbanyak baik sebagian ataupun keseluruhan isi buku dengan cara apa pun tanpa izin tertulis dari penerbit.

Cetakan I, Maret 2025

Perancang sampul: Rosyiful Aqli
Penata letak: Noufal Fahriza

ISBN : 978-634-206-993-6

x + 108 hlm. ; 15,5x23 cm.

©Maret 2025



PRAKATA

Puji dan syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga buku monograf yang berjudul “Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor untuk Klasifikasi Penerima Bantuan Sosial di Desa Marbau Selatan” ini dapat diselesaikan dengan baik. Buku ini disusun sebagai bagian dari upaya akademik dalam mengembangkan metode klasifikasi berbasis machine learning guna meningkatkan efektivitas dan efisiensi dalam penentuan penerima bantuan sosial.

Dalam penelitian ini, kami membandingkan dua algoritma klasifikasi, yakni Naïve Bayes dan K-Nearest Neighbor (KNN), dalam mengolah dan menganalisis data penerima bantuan sosial di Desa Marbau Selatan. Pemilihan algoritma ini didasarkan pada karakteristik masing-masing metode yang memiliki keunggulan dalam menangani data dengan distribusi probabilistik (Naïve Bayes) serta berbasis kedekatan antar data (KNN). Hasil penelitian ini diharapkan dapat memberikan wawasan bagi akademisi, praktisi, serta pemangku kebijakan terkait penerapan teknologi dalam pengelolaan data sosial, khususnya dalam konteks pemberian bantuan sosial yang lebih tepat sasaran.

Kami menyadari bahwa dalam proses penyusunan buku ini masih terdapat berbagai kekurangan. Oleh karena itu, kami sangat mengharapkan kritik dan saran yang membangun dari para pembaca demi penyempurnaan penelitian dan karya ilmiah di masa mendatang. Ucapan terima kasih kami sampaikan kepada rekan-rekan akademisi,

pemerintah desa, serta semua pihak yang telah berkontribusi dalam penelitian ini. Semoga buku ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan aplikasi teknologi di bidang data mining serta sistem pengambilan keputusan berbasis kecerdasan buatan.

Penulis



PENGANTAR

Perkembangan teknologi informasi dan kecerdasan buatan telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam pengelolaan data sosial. Salah satu tantangan utama dalam sistem distribusi bantuan sosial adalah memastikan bahwa bantuan tersebut diterima oleh kelompok yang benar-benar membutuhkan. Proses klasifikasi penerima bantuan yang masih dilakukan secara manual sering kali memiliki kendala, seperti subjektivitas, ketidaktepatan data, serta keterbatasan dalam pengolahan informasi dalam skala besar. Oleh karena itu, penerapan metode berbasis machine learning menjadi solusi yang potensial dalam meningkatkan akurasi dan efisiensi proses ini.

Buku monograf ini berjudul “Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor untuk Klasifikasi Penerima Bantuan Sosial di Desa Marbau Selatan”, yang bertujuan untuk mengeksplorasi penerapan dua algoritma klasifikasi dalam proses seleksi penerima bantuan sosial. Naïve Bayes dikenal sebagai algoritma berbasis probabilitas yang memiliki keunggulan dalam menangani data berskala besar dengan asumsi independensi antar variabel, sedangkan K-Nearest Neighbor (KNN) merupakan algoritma berbasis jarak yang mempertimbangkan kedekatan karakteristik antara satu individu dengan individu lainnya dalam proses klasifikasi. Studi kasus dalam buku ini berfokus pada Desa Marbau Selatan, dengan menggunakan data penerima bantuan sosial sebagai objek penelitian. Perbandingan kedua algoritma ini bertujuan untuk mengidentifikasi

metode yang paling optimal dalam menentukan siapa yang berhak menerima bantuan berdasarkan berbagai faktor yang relevan. Melalui analisis komparatif, buku ini akan memberikan wawasan mengenai kelebihan dan kekurangan masing-masing metode serta rekomendasi mengenai algoritma yang lebih sesuai untuk diterapkan dalam sistem pengelolaan data bantuan sosial di tingkat desa.

Selain itu, dalam buku ini juga dibahas potensi pengembangan metode hybrid serta kemungkinan penerapan algoritma lain seperti Decision Tree dan Random Forest guna meningkatkan akurasi sistem klasifikasi. Dengan demikian, buku ini tidak hanya memberikan kajian teknis mengenai performa algoritma, tetapi juga membuka ruang bagi penelitian lebih lanjut dalam pengembangan sistem berbasis data mining untuk keperluan sosial dan pemerintahan.

Kami berharap buku ini dapat menjadi referensi yang bermanfaat bagi akademisi, peneliti, serta praktisi yang tertarik dalam bidang machine learning, data mining, dan sistem pengambilan keputusan berbasis kecerdasan buatan. Semoga buku ini dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan dan implementasi teknologi yang lebih inovatif dalam mendukung program sosial yang lebih transparan dan tepat sasaran.

Penulis



DAFTAR ISI

Prakata.....	iii
Pengantar	v
Daftar Isi.....	vii

BAB 1

PENDAHULUAN.....	1
Latar Belakang.....	1
Pentingnya Algoritma Pembelajaran Mesin.....	3
Permasalahan dalam Metode Konvensional	4
Tujuan Penelitian	4
Dampak Potensial terhadap Kebijakan Sosial dan Teknologi	5

BAB 2

KERANGKA TEORETIS.....	7
Definisi dan Konsep Dasar Bantuan Social.....	7
Knowledge Discovery in Database (KDD).....	9
Data Mining.....	11
Konsep Dasar Klasifikasi.....	13
Algoritma Naïve Bayes.....	14
Algoritma K-Nearest Neighbor	17
Microsoft Excel.....	20
Rapid Minner	20

Penelitian Terdahulu Distribusi Bantuan Sosial	21
Keunggulan dan Keterbatasan Kedua Algoritma.....	23

BAB 3

METODOLOGI PENELITIAN27

Deskripsi Lokasi.....	27
Metode Pengumpulan Data	29
Proses Knowledge Discovery in Database (KDD)	31
Pembagian Dataset, Data Training dan Data Testing.....	49
Evaluasi Model	51

BAB 4

ANALISIS ALGORITMA NAÏVE BAYES & K-NEAREST NEIGHBOR 55

Proses Penerapan Algoritma Naïve Bayes	55
Evaluasi Model Algoritma Naïve Bayes	66
Proses Penerapan Algoritma K-Nearest Neighbor	69
Evaluasi Model Klasifikasi dengan KNN.....	78

BAB 5

HASIL IMPLEMENTASI RAPIDMINER DAN PEMBAHASAN..... 83

Interpretasi Hasil Klasifikasi Naïve Bayes.....	83
Analisis Efektivitas Model Naïve Bayes.....	86
Interpretasi Hasil Klasifikasi K-Nearest Neighbors (KNN).....	87
Analisis Efektivitas Model K-Nearest Neighbors (KNN)	89
Perbandingan dari Penggunaan Kedua Algoritma	90
Faktor-faktor yang Memengaruhi Akurasi Algoritma	92

BAB 6

KESIMPULAN DAN REKOMENDASI.....	95
Ringkasan Efektivitas Algoritma.....	95
Rekomendasi Penerapan Berbasis Machine Learning	97
Potensi Penelitian Lanjutan dengan Menerapkan Algoritma Lain.....	99
Daftar Pustaka.....	101
Tentang Penulis.....	105



PENDAHULUAN

Latar Belakang

Pemberian bantuan sosial oleh pemerintah merupakan salah satu instrumen penting dalam upaya untuk meningkatkan kesejahteraan masyarakat, terutama bagi kelompok yang kurang mampu secara ekonomi. Bantuan sosial bertujuan untuk memberikan dukungan kepada mereka yang mengalami kesulitan dalam memenuhi kebutuhan dasar, seperti pangan, kesehatan, dan pendidikan. Meskipun demikian, dalam pelaksanaannya, distribusi bantuan sosial sering kali menghadapi masalah serius terkait ketepatan sasaran. Ketidaktepatan sasaran ini menjadi masalah yang sangat krusial, karena bantuan sosial yang tidak sampai kepada masyarakat yang benar-benar membutuhkan justru memperburuk ketimpangan sosial dan ekonomi yang ada.

Di banyak wilayah, terutama di daerah pedesaan, proses seleksi penerima bantuan sosial masih dilakukan secara tradisional, yang bergantung pada data administrasi yang sering kali tidak akurat atau

tidak lengkap. Hal ini mengakibatkan banyak rumah tangga yang seharusnya menerima bantuan tidak terdaftar sebagai penerima, sementara mereka yang tidak membutuhkan malah mendapatkan bantuan tersebut. Fenomena ini sering terjadi di desa-desa dengan tingkat kemiskinan yang tinggi, di mana akses terhadap informasi dan teknologi sangat terbatas.

Desa Marbau Selatan, sebagai lokasi penelitian dalam naskah ini, merupakan contoh nyata dari masalah tersebut. Desa ini memiliki karakteristik demografi dengan sebagian besar penduduknya berprofesi sebagai petani yang bergantung pada hasil pertanian sebagai sumber penghidupan. Namun, meskipun mayoritas penduduknya tergolong dalam golongan ekonomi lemah, sistem distribusi bantuan sosial di desa ini masih belum efektif. Hal ini disebabkan oleh kurangnya pemahaman dan kemampuan dalam mengidentifikasi calon penerima bantuan secara tepat. Ketidakakuratan sistem seleksi penerima bantuan menyebabkan banyaknya kelompok masyarakat yang seharusnya mendapat dukungan, namun tidak memperoleh bantuan yang diperlukan.

Untuk mengatasi permasalahan tersebut, algoritma pembelajaran mesin seperti Naïve Bayes dan K-Nearest Neighbor (K-NN) dapat menjadi solusi yang sangat potensial. Kedua algoritma ini menawarkan pendekatan berbasis data yang lebih objektif dan efisien dalam mengklasifikasikan penerima bantuan sosial. Dengan menggunakan data sosial-ekonomi yang lebih rinci, algoritma ini dapat membantu meningkatkan akurasi dalam menentukan siapa yang seharusnya menerima bantuan. Algoritma Naïve Bayes dapat memperkirakan probabilitas kelayakan penerima bantuan berdasarkan variabel-variabel yang ada, sementara K-NN dapat menangani distribusi data yang tidak merata, yang sering menjadi masalah dalam analisis data sosial-ekonomi. Kombinasi kedua algoritma ini diharapkan dapat memberikan solusi yang lebih tepat dalam memitigasi masalah ketepatan sasaran dalam pemberian bantuan sosial.



KERANGKA TEORETIS

Definisi dan Konsep Dasar Bantuan Social

Bantuan sosial adalah bentuk intervensi yang diberikan oleh pemerintah atau lembaga terkait untuk mendukung kesejahteraan masyarakat, khususnya bagi mereka yang tergolong dalam kelompok rentan atau ekonomi lemah. Bantuan sosial bertujuan untuk memastikan akses masyarakat terhadap kebutuhan dasar, seperti pangan, kesehatan, pendidikan, dan tempat tinggal. Bentuk bantuan sosial yang paling umum meliputi bantuan langsung tunai, bantuan pangan, dan subsidi kebutuhan dasar. Pemberian bantuan sosial ini menjadi instrumen yang sangat penting untuk mengurangi tingkat kemiskinan dan meningkatkan kualitas hidup, khususnya di daerah-daerah yang memiliki tingkat ketimpangan ekonomi yang tinggi.

Namun, meskipun tujuan dari bantuan sosial adalah untuk mengurangi ketimpangan sosial dan ekonomi, sistem distribusinya sering kali menemui berbagai tantangan. Salah satu tantangan terbesar adalah ketidaktepatan sasaran. Hal ini disebabkan oleh berbagai

faktor, seperti data yang tidak akurat, kurangnya transparansi, dan proses seleksi penerima yang tidak efisien. Misalnya, di banyak daerah, proses seleksi penerima bantuan masih dilakukan dengan cara yang sangat manual, yang mengandalkan data yang sudah lama atau tidak lengkap. Sistem yang berbasis pada pendataan administratif yang tidak terkelola dengan baik sering kali menghasilkan penerima bantuan yang tidak membutuhkan, sementara mereka yang benar-benar membutuhkan tidak mendapatkan bantuan yang seharusnya mereka terima.

Ketidaktepatan dalam mendistribusikan bantuan ini tidak hanya berisiko menambah kesenjangan sosial, tetapi juga dapat menurunkan kepercayaan masyarakat terhadap pemerintah atau lembaga penyedia bantuan. Jika masyarakat merasa bahwa bantuan tidak sampai ke tangan yang membutuhkan, hal ini dapat mengurangi efektivitas kebijakan bantuan sosial secara keseluruhan. Selain itu, sistem yang tidak efisien juga berdampak pada penggunaan sumber daya yang tidak optimal, baik dari segi waktu, tenaga, maupun dana. Oleh karena itu, penting untuk memiliki sistem yang lebih transparan, efisien, dan berbasis data yang akurat dalam melakukan seleksi penerima bantuan.

Untuk mengatasi tantangan ini, banyak penelitian menunjukkan bahwa penerapan teknologi informasi dan pembelajaran mesin (machine learning) dapat menjadi solusi yang sangat efektif. Dengan menggunakan algoritma seperti Naïve Bayes atau K-Nearest Neighbor (KNN), seleksi penerima bantuan sosial dapat dilakukan dengan lebih objektif, berdasarkan data yang lebih akurat dan terkini. Teknologi ini memungkinkan analisis data sosial-ekonomi yang lebih mendalam dan dapat mengklasifikasikan penerima bantuan dengan lebih tepat, sehingga distribusi bantuan dapat lebih tepat sasaran. Dengan penerapan teknologi ini, diharapkan bantuan sosial dapat diberikan kepada mereka yang benar-benar membutuhkan, mengurangi pemborosan, dan meningkatkan kepercayaan publik terhadap kebijakan sosial yang ada.



METODOLOGI PENELITIAN

Deskripsi Lokasi

Desa Marbau Selatan terletak di Kecamatan Marbau, Kabupaten Labuhanbatu Utara, Provinsi Sumatera Utara. Secara geografis, desa ini berada pada koordinat 2°50' Lintang Utara dan 99°25' Bujur Timur, dengan ketinggian daratan mulai dari 0 hingga 700 meter di atas permukaan laut. Wilayah ini memiliki iklim tropis dengan dua musim utama: musim hujan dan musim kemarau. Curah hujan rata-rata bulanan di Kabupaten Labuhanbatu Utara adalah 255,42 mm, dengan hari hujan rata-rata sebanyak 12,33 hari per bulan. Musim hujan biasanya terjadi pada bulan Oktober dengan curah hujan tertinggi mencapai 443 mm, sedangkan musim kemarau terjadi pada bulan Maret dengan curah hujan terendah sekitar 82 mm.

Berdasarkan data dari Badan Pusat Statistik Kabupaten Labuhanbatu Utara, jumlah penduduk Kabupaten Labuhanbatu Utara tercatat sebesar 347.456 jiwa dengan laju pertumbuhan penduduk 2,98 persen. Jumlah rumah tangga sebanyak 80.520 dengan rata-rata

jumlah anggota rumah tangga adalah 4 orang. Komposisi penduduk masih didominasi oleh kelompok usia muda, yang menyebabkan rasio ketergantungan (dependency ratio) yang cukup tinggi, yaitu sebesar 61,61 persen. Jumlah penduduk laki-laki di Kabupaten Labuhanbatu masih lebih banyak dibandingkan perempuan, dengan nilai sex ratio sebesar 101, yang berarti untuk setiap 100 orang penduduk perempuan terdapat 101 orang penduduk laki-laki.

Desa Marbau Selatan, sebagai bagian dari Kecamatan Marbau, memiliki karakteristik demografis yang mencerminkan komposisi penduduk Kabupaten Labuhanbatu Utara secara umum. Masyarakat desa ini mayoritas beragama Islam, dengan sejarah keagamaan yang kaya, termasuk adanya masjid tua yang menjadi pusat kegiatan keagamaan dan sosial masyarakat. Pendidikan di desa ini masih dalam tahap pengembangan, dengan upaya peningkatan angka harapan lama sekolah yang tercatat pada tahun 2014 sebesar 11,80 tahun, sedikit meningkat dibandingkan tahun 2013 yang mencapai 11,30 tahun. Namun, angka ini masih di bawah rata-rata provinsi Sumatera Utara yang mencapai 12,61 tahun pada tahun 2014.

Dari segi ekonomi, Desa Marbau Selatan memiliki potensi sumber daya alam yang melimpah, terutama dalam sektor pertanian dan perkebunan. Masyarakat desa banyak yang bekerja sebagai petani dan pekebun, dengan komoditas utama seperti kelapa sawit dan karet. Namun, sektor pertanian di desa ini masih menghadapi tantangan dalam hal akses terhadap teknologi pertanian modern, pemasaran hasil pertanian, dan infrastruktur pendukung lainnya. Upaya pengembangan ekonomi desa melalui Badan Usaha Milik Desa (BUMDes) telah dilakukan untuk memperkuat perekonomian lokal dan meningkatkan kesejahteraan masyarakat.

Secara keseluruhan, Desa Marbau Selatan memiliki karakteristik demografis dan ekonomi yang mencerminkan dinamika masyarakat pedesaan di Kabupaten Labuhanbatu Utara. Meskipun terdapat tantangan dalam bidang pendidikan dan ekonomi, potensi sumber



ANALISIS ALGORITMA NAÏVE BAYES & K-NEAREST NEIGHBOR

Proses Penerapan Algoritma Naïve Bayes

Naïve Bayes adalah algoritma pembelajaran mesin berbasis probabilitas yang digunakan untuk klasifikasi data. Algoritma ini mengasumsikan bahwa setiap fitur dalam data adalah independen, yang memungkinkan perhitungan probabilitas untuk setiap kelas berdasarkan fitur yang ada. Dalam konteks penelitian ini, algoritma Naïve Bayes diterapkan untuk klasifikasi penerima bantuan sosial di Desa Marbau Selatan, berdasarkan data sosial-ekonomi seperti nama, gender kepala keluarga, usia, penghasilan, jumlah tanggungan, dan kepemilikan asset/rumah.

Algoritma Naïve Bayes kemudian menghitung probabilitas untuk setiap kelas dan memilih kelas dengan probabilitas tertinggi sebagai hasil klasifikasi. Misalnya, jika kita memiliki data untuk sebuah

keluarga dengan pendapatan tertentu, jumlah tanggungan, dan status pekerjaan tertentu, model akan menghitung probabilitas kelas “layak” dan “tidak layak” menerima bantuan sosial, dan memilih kelas dengan nilai probabilitas tertinggi.

Penerapan Naïve Bayes pada dataset sosial-ekonomi ini memungkinkan klasifikasi yang cepat dan efisien. Algoritma ini sangat cocok untuk dataset yang memiliki variabel kategori, karena setiap fitur diperlakukan secara independen dalam perhitungannya. Namun, meskipun asumsi independensi ini jarang berlaku secara sempurna dalam dunia nyata, Naïve Bayes tetap memberikan hasil yang baik dalam banyak aplikasi, termasuk dalam klasifikasi penerima bantuan sosial.

Tabel. Data Latih algoritma Naïve Bayes

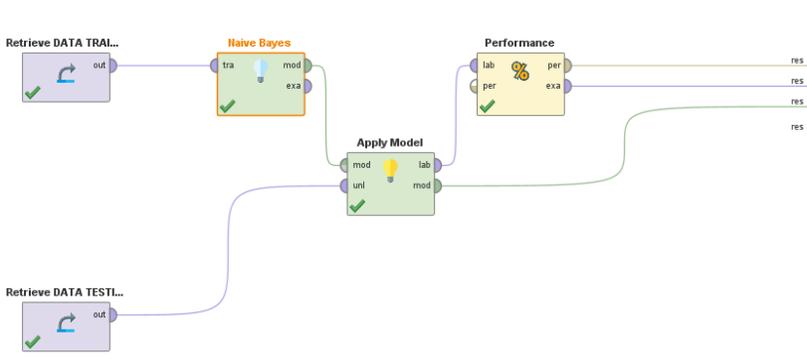
Nama	Gender Kepala Keluarga	Usia	Penghasilan	Tanggungan	Kepemilikan Rumah	Status
KK1	Laki-Laki	Dewasa	Diatas UMK	Kecil	Milik Orang Lain	Tidak Layak
KK2	Laki-Laki	Muda	Diatas UMK	Kecil	Milik Pribadi	Tidak Layak
KK3	Perempuan	Lansia	Dibawah UMK	Kecil	Milik Pribadi	Tidak Layak
KK4	Perempuan	Dewasa	Diatas UMK	Kecil	Milik Pribadi	Tidak Layak
KK5	Perempuan	Dewasa	Diatas UMK	Kecil	Milik Orang Lain	Layak
KK6	Laki-Laki	Dewasa	Diatas UMK	Kecil	Milik Orang Lain	Tidak Layak
KK7	Laki-Laki	Dewasa	Dibawah UMK	Kecil	Milik Orang Lain	Tidak Layak
KK8	Laki-Laki	Dewasa	Diatas UMK	Kecil	Milik Orang Lain	Tidak Layak
KK9	Laki-Laki	Dewasa	Diatas UMK	Sedang	Milik Orang Lain	Layak
KK10	Perempuan	Lansia	Dibawah UMK	Kecil	Milik Pribadi	Layak



HASIL IMPLEMENTASI RAPIDMINER DAN PEMBAHASAN

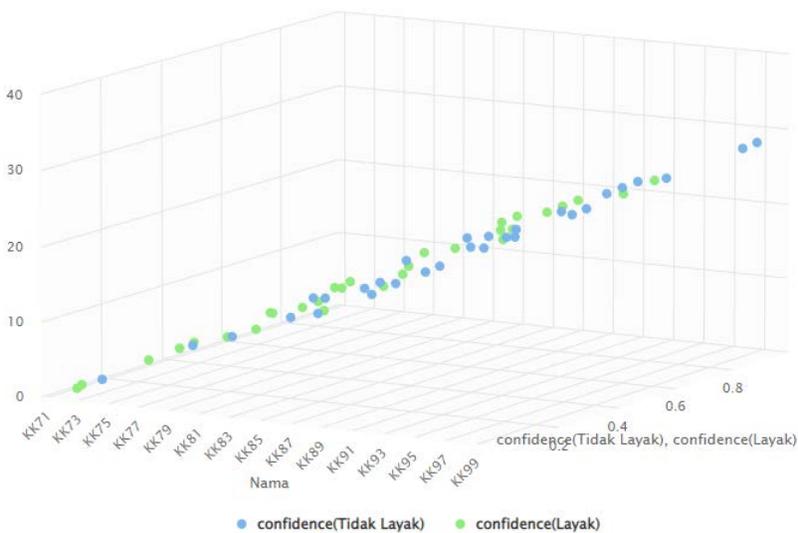
Interpretasi Hasil Klasifikasi Naïve Bayes

Metode Naïve Bayes telah banyak digunakan dalam klasifikasi data karena kemampuannya dalam menangani dataset dengan fitur yang bersifat independen. Pada penelitian ini, algoritma Naïve Bayes diterapkan dalam sistem seleksi penerima bantuan sosial untuk mengelompokkan individu ke dalam kategori “Layak” atau “Tidak Layak” menerima bantuan. Implementasi dilakukan menggunakan RapidMiner, sebuah perangkat lunak berbasis visual yang memungkinkan pengguna untuk membangun, melatih, dan mengevaluasi model machine learning secara efisien.



Gambar 4.1 Rancangan Model Naïve Bayes

Gambar diatas merupakan rancangan model machine learning dengan menggunakan metode *Naïve Bayes*. Implementasi ini dilakukan dengan menggunakan software data mining yaitu RapidMiner Studio.



Gambar 4.2 Visualisasi Klasifikasi Naïve Bayes

Gambar yang ditampilkan merupakan visualisasi hasil klasifikasi dengan metode Naïve Bayes (NB) dalam menentukan status penerima bantuan sosial, yaitu “Layak” dan “Tidak Layak”. Visualisasi ini berbentuk scatter plot 3D, yang menunjukkan bagaimana model



KESIMPULAN DAN REKOMENDASI

Ringkasan Efektivitas Algoritma

Dalam penelitian ini, telah dilakukan analisis terhadap efektivitas berbagai algoritma dalam klasifikasi penerima bantuan sosial, dengan fokus pada dua metode utama, yaitu Naïve Bayes dan K-Nearest Neighbors (KNN). Studi ini menunjukkan bahwa masing-masing algoritma memiliki keunggulan dan kelemahan yang berbeda dalam menangani data klasifikasi, terutama terkait dengan akurasi, efisiensi komputasi, serta kemampuan menangani pola hubungan antar fitur dalam dataset. Oleh karena itu, pemilihan algoritma yang tepat sangat bergantung pada karakteristik data serta tujuan utama dari sistem klasifikasi yang digunakan dalam distribusi bantuan sosial.

Dari hasil analisis yang dilakukan, Naïve Bayes terbukti lebih efisien secara komputasi, dengan waktu pemrosesan yang lebih cepat dibandingkan KNN. Hal ini disebabkan oleh pendekatan probabilistiknya yang memungkinkan pemrosesan data dalam jumlah

besar tanpa memerlukan perhitungan jarak antar sampel. Namun, kekurangan utama dari Naïve Bayes adalah asumsi independensi antar fitur, yang sering kali tidak sesuai dengan kondisi data nyata. Misalnya, dalam sistem distribusi bantuan sosial, faktor seperti penghasilan rumah tangga dan jumlah tanggungan tidak dapat dianggap sepenuhnya independen, sehingga dapat menyebabkan kesalahan dalam klasifikasi.

Di sisi lain, KNN menawarkan fleksibilitas yang lebih besar dalam menangani data dengan pola yang lebih kompleks, karena pendekatan berbasis kedekatan yang digunakan memungkinkan algoritma ini menangkap hubungan non-linear antar variabel. Meskipun demikian, KNN memiliki kelemahan utama dalam hal efisiensi komputasi, terutama ketika jumlah data yang digunakan semakin besar. Selain itu, pemilihan nilai K yang optimal menjadi faktor krusial yang dapat memengaruhi kinerja algoritma ini. Nilai K yang terlalu kecil dapat menyebabkan model terlalu peka terhadap noise dalam data, sedangkan nilai K yang terlalu besar dapat mengurangi sensitivitas model terhadap pola-pola minoritas yang penting.

Berdasarkan hasil perbandingan antara kedua algoritma, pendekatan hybrid antara Naïve Bayes dan KNN telah dieksplorasi sebagai solusi potensial untuk mengatasi keterbatasan masing-masing metode. Pendekatan ini bertujuan untuk menggabungkan efisiensi Naïve Bayes dengan fleksibilitas KNN, sehingga menghasilkan sistem klasifikasi yang lebih akurat dan efisien. Hasil eksperimen menunjukkan bahwa metode hybrid mampu meningkatkan akurasi dibandingkan metode tunggal, khususnya dalam menangani data yang memiliki hubungan kompleks antar fitur serta ketidakseimbangan kelas.



DAFTAR PUSTAKA

1. Hidayat, T., Munthe, I. R., & Juledi, A. P., “Analisis Data Penjualan Menggunakan Algoritma Apriori pada Analisis Kopi,” *INFORMATIKA*, vol. 12, no. 3, p. 56, 2024, <https://doi.org/10.36987/informatika.v12i3.6064>.
2. D. S. O. Panggabean, E. Buulolo, and N. Silalahi, “Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda,” *Jurikom (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, 2020, doi: 10.30865/jurikom.v7i1.1947.
3. I. A. N. Afifah, “Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means,” *Jipi (Jurnal Ilm. Penelit. Dan Pembelajaran Inform.)*, vol. 8, no. 3, pp. 802–814, 2023, doi: 10.29100/jipi.v8i3.3891.
4. D. A. Fakhri, S. Defit, and S. Sumijan, “Optimalisasi Pelayanan Perpustakaan Terhadap Minat Baca Menggunakan Metode K-Means Clustering,” *J. Inf. Dan Teknol.*, pp. 160–166, 2021, doi: 10.37034/jjdt.v3i3.137.
5. N. Asiah, “Evaluasi Penerapan Sistem Manajemen Keselamatan dan Kesehatan Kerja (SMK3) di Rumah Sakit Umum Daerah dr. Zainoel Abidin Banda Aceh,” Universitas Islam Negeri Ar-Raniry, 2020.
6. Amansyah, R., Masrizal, M., & Munthe, I. R., “Pengimplementasian Tingkat Ketepatan Waktu Kelulusan Siswa (Studi Kasus Di MTS Nur Ibarhimy) Menggunakan Algoritma C4.5,” *INFORMATIKA*,

- 12(2), 280-291,2024, <https://doi.org/10.36987/informatika.v12i2.5767>
7. A. Maulana, A. Nugroho, and I. Romli, "Optimalisasi Support Vector Machine Menggunakan Particle Swarm Optimization Untuk Mendiagnosa Penyakit Kanker Payudara," *J. Pract. Comput. Sci.*, vol. 1, no. 2, pp. 1–11, 2022, doi: 10.37366/jpcs.v1i2.940.
 8. R. Fitra, "Penerapan Metode Algoritma *K-Nearest Neighbor* Menggunakan Rapidminer Studio Pada Klasifikasi Status Sosial Ekonomi Studi Kasus : Kelurahan Kapuk Muara Rt 010 Rw 04," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 11, no. 4, 2022, doi: 10.30591/smartcomp.v11i4.4250.
 9. A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
 10. D. Syaputri, P. H. Noprita, and S. Romelah, "Implementasi Algoritma *K-Means* Untuk Pengelompokan Distribusi Sosial Ekonomi Masyarakat Berdasarkan Demografi Kependudukan," *Malcom Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 1–6, 2021, doi: 10.57152/malcom.v1i1.5.
 11. P. Simanjuntak, C. E. Suharyanto, S. Sitohang, and K. Handoko, "Data Mining Untuk Klasifikasi Status Pandemi Covid 19," *J. Tek. Inf. dan Komput.*, vol. 5, no. 2, p. 327, 2022, doi: 10.37600/tekinkom.v5i2.620.
 12. N. Apriliani, "Analisis Sentimen Review Penggunaan Tiktok Melalui Pendekatan Algoritma Naïve Bayes," *Jati (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3725–3731, 2024, doi: 10.36040/jati.v7i6.8299.
 13. H. Sastypratiwi, Y. Yulianti, and H. Muhardi, "Uji Komparasi Algoritma *Naïve Bayes* Dan Decision Tree Classification

- Menggunakan Covid-19 Dataset,” *J. Edukasi Dan Penelit. Inform.*, vol. 8, no. 1, p. 1, 2022, doi: 10.26418/jp.v8i1.49841.
14. R. A. Rizal, N. O. Purba, L. A. Siregar, K. P. Sinaga, and N. Azizah, “Analysis of Tuberculosis (TB) on X-Ray Image Using SURF Feature Extraction and the *K-Nearest Neighbor* (KNN) Classification Method,” *Jaict*, vol. 5, no. 2, p. 9, 2020, doi: 10.32497/jaict.v5i2.1979.
 15. M. Pagan, “Investigating the Impact of Data Scaling on the K-Nearest Neighbor Algorithm,” *Comput. Sci. Inf. Technol.*, vol. 4, no. 2, pp. 135–142, 2023, doi: 10.11591/csit.v4i2.pp135-142.
 16. G. T. Hariyadi, D. Aqmala, A. B. Setiawan, and I. Farida, “Pelatihan Query Excel Untuk Pengelolaan Data Administrasi Kependidikan Pada TK. Isriati Baiturahman 1 Pandanaran Semarang,” *Abdimasku J. Pengabd. Masy.*, vol. 5, no. 1, p. 20, 2022, doi: 10.33633/ja.v5i1.286.
 17. A. A. A. Ushud, “Pelatihan Microsoft Excel Tingkat Lanjut Karyawan Pt. Nutrisi Juara Asia,” *Kresna J. Ris. Dan Pengabd. Masy.*, vol. 3, no. 1, pp. 86–94, 2023, doi: 10.36080/kresna.v3i1.54.
 18. M. Nasution, A. A. Ritonga, and A. P. Juledi, “Implementasi Rapidminer Dalam Mengklasifikasikan Indeks Demokrasi,” *J. Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 99–106, 2022.
 19. S. A. Nazli, “Coronary Risk Factor Profiles According to Different Age Categories in Premature Coronary Artery Disease Patients Who Have Undergone Percutaneous Coronary Intervention,” *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-53539-6.
 20. A. S. Elbhrawy, “AIRA-ML: Auto Insurance Risk Assessment-Machine Learning Model Using Resampling Methods,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 9, 2023, doi: 10.14569/ijacsa.2023.0140966.



TENTANG PENULIS



Nurhikmah Wulandari lahir di Rantau Prapat pada 22 November 2002. Sejak kecil menempuh pendidikan dasar hingga menengah di Marbau, Labuhanbatu Utara, sebelum akhirnya melanjutkan studi di jenjang perguruan tinggi. Pernah menjadi bagian dari mahasiswa S1 Sistem Informasi di Universitas Labuhanbatu sejak 2021-2025 dengan menemukan minat bakatnya dalam bidang teknologi, komputer dan analisis data.

Ketertarikannya terhadap dunia teknologi, khususnya dalam kecerdasan buatan, data mining, dan sistem pengambilan keputusan berbasis machine learning, mendorongnya untuk terus mengembangkan wawasan dan berkontribusi dalam dunia akademik. Hal ini dibuktikan dengan karya-karya yang telah dihasilkan, salah satunya adalah buku pertamanya yang berjudul “Pengantar Probabilitas”, yang tidak hanya menjadi referensi akademik, tetapi juga mengantarkan penulis meraih penghargaan dalam International Student Competition (ISC) 2023 atas kontribusinya dalam bidang artikel ilmiah.

Sebagai seorang yang selalu ingin memperdalam pemahaman dalam analisis data dan machine learning, penulis kembali menghadirkan buku keduanya yang berjudul “Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor untuk Klasifikasi Penerima Bantuan Sosial di Desa Marbau Selatan”. Buku ini lahir dari minat dan

perhatiannya terhadap pemanfaatan teknologi dalam sektor sosial, khususnya dalam pengelolaan dan klasifikasi data penerima bantuan sosial yang lebih efektif dan akurat.

Dengan semangat untuk terus belajar dan berbagi ilmu, penulis berharap buku ini dapat menjadi referensi yang bermanfaat bagi akademisi, mahasiswa, serta praktisi di bidang teknologi informasi dan analisis data, serta membuka peluang untuk penelitian lebih lanjut dalam pengembangan sistem berbasis kecerdasan buatan di berbagai sektor.



Ibnu Rasyid Munthe, S.T., M.Kom lahir di Kota Rantau Prapat, Kabupaten Labuhanbatu, pada tahun 1987. Beliau memperoleh gelar Sarjana Teknik (S.T.) dari Universitas Nurtanio Bandung dan melanjutkan pendidikan Magister di Universitas Putra Indonesia (YPTK) Padang dengan gelar Magister Komputer (M.Kom.). Beliau aktif dalam menulis artikel ilmiah serta

buku akademik. Saat ini, beliau berkiprah sebagai dosen di Universitas Labuhanbatu. Selain itu, beliau juga mengelola kanal YouTube @manjaddawajada2022 sebagai media pembelajaran dan sarana berbagi ilmu pengetahuan.



Angga Putra Juledi, S.Kom., M.Kom Lahir di Kota Padang pada tanggal 19 Juli 1994. Dalam menempuh Pendidikan dimulai dari Sekolah Dasar SDN 19 Padang tamat tahun 2006, SMPN 3 Padang tamat pada tahun 2009, dan di SMA Pertiwi 2 Padang tamat pada tahun 2012. Lalu melanjutkan ke pendidikan perguruan tinggi swasta yaitu S1 (Sarjana) Universitas Putra Indonesia “YPTK” Padang lulus pada tahun

2018 dengan jurusan Sistem Informasi, Dan melanjutkan Program

Pascasarjana (S2) di Universitas Putra Indonesia “YPTK” Padang pada tahun 2019 Program Studi Teknik Informatika. Konsentrasi Sistem Informasi. Saya mengabdikan diri sebagai salah satu Dosen di bidang Ilmu Komputer pada Fakultas Sains Dan Teknologi dengan Program Studi Sistem Informasi di Universitas LabuhanBatu dan menjadi dosen tetap pada tahun 2020 pada kampus tersebut. Saat ini menjadi bagian Struktural di Universitas Labuhanbatu sebagai Kepala Bagian Sumber Daya Manusia periode 2023 s/d 2027. Buku pertama terbit pada 31 Desember 2021 dengan judul Internetworking Dan TCP/IP. Buku kedua terbit pada tanggal 17 Oktober 2023 dengan judul Panduan Belajar HTML, CSS, dan JavaScript untuk Pemula. Hingga saat ini masih menulis buku setiap tahunnya.



Marnis Nasution, S.Kom., M.Kom Lahir di Bengkulu 30 maret 1990. Selama sekolah dasar sampai menengah ditempuh di kota Bengkulu. Melanjutkan Pendidikan tinggi strata-1 dan strata-2 di Universitas Putra Indonesia “YPTK” Padang dari tahun 2008 sampai 2024 dengan jurusan Sistem Informasi. Saat ini aktif menjadi Dosen Yayasan di Universitas Labuhanbatu, Sumatera Utara dan menulis beberapa karya Ilmiah dan buku.

Komparasi Algoritma

Naïve Bayes dan K-NN

dalam **PEMBAGIAN BANTUAN DESA**

Buku ini membahas penerapan dan perbandingan dua algoritma populer dalam dunia data mining dan kecerdasan buatan, yaitu Naïve Bayes dan *K-Nearest Neighbors* (K-NN), dalam konteks pembagian bantuan di desa.

Dengan meningkatnya kebutuhan akan distribusi bantuan yang tepat sasaran, teknologi dan analisis data menjadi solusi yang efektif untuk memastikan bantuan diberikan kepada pihak yang benar-benar membutuhkan. Dalam buku ini, pembaca akan diperkenalkan pada dasar teori kedua algoritma, kelebihan dan kekurangannya, serta bagaimana mereka diterapkan dalam studi kasus pembagian bantuan sosial di desa.

Melalui analisis perbandingan, buku ini mengulas akurasi, efisiensi, serta keandalan dari Naïve Bayes dan K-NN dalam mengklasifikasikan penerima bantuan berdasarkan data sosial ekonomi. Pembaca juga akan diajak untuk memahami cara mengimplementasikan algoritma tersebut menggunakan bahasa pemrograman populer seperti Python, serta mengevaluasi performanya dengan berbagai metrik evaluasi.

Buku ini ditujukan bagi mahasiswa, peneliti, akademisi, serta praktisi di bidang data science dan pemerintahan yang ingin memahami bagaimana teknologi dapat dimanfaatkan untuk meningkatkan efektivitas program sosial berbasis data.



✉ literasinusantaraofficial@gmail.com
🌐 www.penerbitlitnus.co.id
📘 Literasi Nusantara
📱 literasinusantara_

☎ 085755971589

Sosial

+17

ISBN 978-634-206-993-6



9 786342 069936