Jurnal Publikasi Ilmu Komputer dan Multimedia

E-ISSN: 2808-8999
P-ISSN: 2808-9375

Implementasi Algoritma Naive Bayes untuk Sistem Pendukung Keputusan Penerima Bantuan Pendidikan

Chandra Wisnu Nugroho [1*], Sigrid Tamriesfatno [2], dan Sulkifli [3]

1 Institut Kesehatan dan Teknologi Graha Medika ; email : chandra.nugroho7@gmail.com
2 Institut Kesehatan dan Teknologi Buton Raya ; email : sigridaaris@gmail.com
3 Universitas Megarezky ; email : sulkifli@unimerz.ac.id
***** Penulis Korespondensi

Abstract: Education plays a crucial role in shaping excellent human resources, which are the main foundation of national development. However, economic limitations remain a significant barrier for some students in accessing proper education. To address this issue, the government has implemented various educational assistance programs. SMKN 9 Bulukumba, as one of the vocational high schools, also faces challenges in selecting scholarship recipients efficiently, objectively, and transparently.

This study aims to develop a decision support system based on the Naive Bayes Classifier algorithm to assist the selection process of educational aid recipients. Naive Bayes is a simple yet effective probabilistic classification algorithm that performs well on small to medium-sized datasets. This research employs a data mining approach to build a classification model based on several attributes, such as type of residence, means of transportation, and parents' income.
The results of testing on 10 student data samples showed that 6 students were classified as eligible for assistance, while 4 were not. The dominant factors determining eligibility were the low income of the father and the absence of income from the mother. Interestingly, all students using public transportation were not classified as aid recipients, indicating that transportation is not always a primary indicator of eligibility.
Keywords: Education, Data Mining, Naive Bayes, Classification

Abstrak: Pendidikan memiliki peran krusial dalam membentuk sumber daya manusia unggul yang menjadi fondasi utama pembangunan bangsa. Namun, keterbatasan ekonomi masih menjadi hambatan signifikan bagi sebagian siswa untuk memperoleh akses pendidikan yang layak. Untuk mengatasi permasalahan ini, pemerintah telah mengimplementasikan berbagai program bantuan pendidikan. SMKN 9 Bulukumba, sebagai salah satu sekolah menengah kejuruan, turut menghadapi tantangan dalam menyeleksi calon penerima bantuan secara efisien, objektif, dan transparan.

Penelitian ini bertujuan untuk mengembangkan sistem pendukung keputusan berbasis algoritma Naive Bayes Classifier guna membantu proses seleksi penerima bantuan pendidikan. Naive Bayes merupakan algoritma klasifikasi probabilistik yang sederhana namun efektif, serta mampu bekerja dengan baik pada data berukuran kecil hingga sedang. Penelitian ini menggunakan pendekatan data mining untuk membangun model klasifikasi berdasarkan sejumlah atribut, seperti jenis tempat tinggal, alat transportasi, dan penghasilan orang tua.

Hasil pengujian terhadap 10 data siswa menunjukkan bahwa 6 siswa diklasifikasikan sebagai layak menerima bantuan, sedangkan 4 siswa tidak layak. Faktor dominan yang menentukan kelayakan adalah rendahnya penghasilan ayah dan tidak adanya penghasilan dari ibu. Menariknya, seluruh siswa yang menggunakan angkutan umum tidak diklasifikasikan sebagai penerima bantuan, yang mengindikasikan bahwa variabel transportasi tidak selalu menjadi indikator utama kelayakan

Kata kunci: Pendidikan, Data Mining, Naive Bayes, Klasifikasi

Received: 01 Januari 2025
Revised: 10 Februari 2025
Accepted: 30 Mei 2025
Published: Juni 2025

Hak cipta: © 2025 oleh penulis. Diserahkan untuk kemungkinan publikasi akses terbuka berdasarkan syarat dan ketentuan lisensi Creative Commons Attribution (CC BY SA) (https://creativecommons.org/licenses/by-sa/4.0/)

1. Pendahuluan

Pendidikan merupakan fondasi utama dalam pembangunan suatu bangsa. Pendidikan merupakan proses kompleks yang melibatkan pembentukan karakter, pengembangan kemampuan kritis, dan persiapan untuk hidup bermasyarakat, yang mendorong inklusi sosial dan pembangunan berkelanjutan (Sriatun et al., 2024). Melalui pendidikan yang berkualitas, sumber daya manusia yang unggul dapat dilahirkan, yang pada gilirannya akan mendorong kemajuan di berbagai sektor kehidupan.

Pemerintah Indonesia menyadari betul pentingnya pendidikan dan terus berupaya untuk meningkatkan aksesibilitas dan kualitas pendidikan bagi seluruh warga negara. Salah satu wujud nyata dari upaya ini adalah melalui berbagai program bantuan pendidikan yang ditujukan untuk meringankan beban ekonomi keluarga kurang mampu agar anak-anak mereka tetap dapat mengenyam pendidikan yang layak.

SMKN 9 Bulukumba, sebagai salah satu lembaga pendidikan kejuruan di Kabupaten Bulukumba, memiliki peran strategis dalam menghasilkan lulusan yang siap kerja dan kompeten di bidangnya. Namun, seperti halnya sekolah-sekolah lain, SMKN 9 Bulukumba juga menghadapi tantangan dalam mengelola dan mendistribusikan bantuan pendidikan secara efektif dan efisien. Proses seleksi penerima bantuan yang masih manual seringkali membutuhkan waktu yang lama, rentan terhadap subjektivitas, dan berpotensi menimbulkan permasalahan dalam hal transparansi dan akuntabilitas.

Dalam mengatasi permasalahan tersebut, pemanfaatan teknologi informasi menjadi solusi yang menjanjikan. Salah satu cabang ilmu komputer yang dapat diimplementasikan adalah data mining. Penambangan data memiliki peran penting dalam berbagai aktivitas manusia dengan mengekstraksi pola-pola berguna yang sebelumnya tidak diketahui (atau pengetahuan), dan sangat penting dalam berbagai bidang aplikasi seperti perbankan, ritel, medis, asuransi, bioinformatika, dan lainnya (Manoj, et al., 2020). Dalam konteks sistem pendukung keputusan, data mining dapat digunakan untuk membangun model klasifikasi yang mampu memprediksi atau mengklasifikasikan siswa yang paling layak menerima bantuan pendidikan berdasarkan data-data historis penerima bantuan sebelumnya.

Naive Bayes Classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes (Felicia Watratan et al., 2020a). Algoritma Naive Bayes merupakan algoritma klasifikasi probabilistik yang sederhana namun efektif dan seringkali memberikan hasil yang akurat dalam berbagai aplikasi, termasuk dalam bidang pendidikan. Keunggulan utama dari algoritma ini adalah kemampuannya untuk bekerja dengan baik pada dataset berukuran kecil hingga sedang, relatif cepat dalam proses pelatihan dan pengujian, serta mudah diinterpretasikan hasilnya. Dengan memanfaatkan probabilitas bersyarat, algoritma Naive Bayes dapat memprediksi probabilitas seorang siswa termasuk dalam kategori penerima bantuan atau tidak berdasarkan atribut-atribut yang relevan.

Berdasarkan latar belakang permasalahan dan potensi solusi yang ditawarkan oleh algoritma Naive Bayes, penelitian ini bertujuan untuk mengimplementasikan algoritma tersebut dalam membangun sebuah sistem pendukung keputusan untuk membantu pihak SMKN 9 Bulukumba dalam menyeleksi calon penerima bantuan pendidikan secara lebih objektif, efisien, dan transparan. Studi kasus pada SMKN 9 Bulukumba dipilih untuk memberikan konteks implementasi yang spesifik dan relevan dengan tantangan yang dihadapi oleh sekolah tersebut dalam pengelolaan program bantuan pendidikan. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi praktis bagi sekolah dan menjadi referensi bagi penelitian serupa di masa mendatang.

2. Tinjauan Literatur

2.1. Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis. Data mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan trend baru yang bermakna dengan menyaring data yang

sangat besar, yang tersimpan dalam penyimpanan, menggunakan Teknik pengenalan pola seperti teknik statistic dan matematika (Aryusya Thamrin & Amanda Putri, 2024).

Data tidak dapat dipisahkan dari kehidupan sehari-hari. Suatu organisasi baik besar maupun kecil dapat dibanjiri dengan berbagai macam data (Yusuf & Rian Pratikto, 2006). Di era digital ini, setiap interaksi, transaksi, dan aktivitas meninggalkan jejak data yang melimpah. Fenomena ini juga berlaku bagi organisasi, baik skala besar maupun kecil, yang sering kali "dibanjiri" oleh berbagai jenis data yang terus bertambah setiap detiknya. Mulai dari data pelanggan, operasional, keuangan, hingga media sosial, semuanya membentuk tumpukan informasi yang sangat besar.

Data Mining merupakan upaya kolaboratif antara manusia dan komputer, yang bertujuan untuk menemukan informasi berharga dan non-sepele dalam sejumlah besar data (S. Gnanapriya, et al., 2010). Tujuannya adalah untuk menemukan informasi berharga dan non-sepele yang tersembunyi dalam sejumlah besar data. Ini bukan sekadar mencari fakta yang sudah jelas, melainkan tentang menggunakan algoritma canggih dan keahlian manusia untuk mengidentifikasi pola, tren, anomali, dan hubungan yang mungkin tidak akan pernah ditemukan secara manual.

Pada proses Knowledge Discovery Database (KDD) di atas terdapat beberapa tahapan yaitu sebagai berikut:

[1] Seleksi Data (Data Selection). Selection (seleksi/pemilihan) data merupakan sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam Knowledge Discovery Database (KDD) dimulai. Data seleksi yang akan digunakan untuk proses Data Mining , disimpan dalam suatu berkas, terpisah dari basis data operasional.

[2] Pemilihan Data (Preprocessing / Cleaning) Proses Preprocessing mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses "memperkaya" data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

[3] Transformasi (Transformation) Pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang valid atau siap untuk dilakukan proses Data Mining.

[4] Data Mining Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

[5] Interpretasi/Evaluasi (Interpratation/Evaluation). Pada fase terakhir ini yang dilakukan adalah proses pembentukan keluaran yang bersumber pada proses Data Mining Pola informasi.

2.2. Klasifikasi

Salah satu teknik data mining yang umum digunakan yaitu teknik klasifikasi. Klasifikasi adalah suatu teknik pembentukan model dari data yang belum terklasifikasi, untuk digunakan mengklasifikasi data baru. Klasifikasi membutuhkan data pelatihan untuk membangun suatu model klasifikasinya (Lestari, 2014)

Karena proses pembentukan modelnya membutuhkan data pelatihan yang sudah berlabel, klasifikasi termasuk dalam kategori supervised learning. Ini berbeda dengan unsupervised learning yang tidak memerlukan data berlabel untuk membangun modelnya.

Ada banyak contoh penerapan klasifikasi dalam kehidupan sehari-hari, seperti:

[1] Penyaringan spam email: Model klasifikasi dilatih dengan email yang sudah ditandai sebagai spam atau bukan spam, lalu digunakan untuk menyaring email masuk yang baru.

[2] Deteksi penipuan transaksi: Bank menggunakan klasifikasi untuk mengidentifikasi transaksi yang mencurigakan berdasarkan pola-pola penipuan yang sudah dikenal.

[3] Diagnosa penyakit: Dalam dunia medis, klasifikasi bisa digunakan untuk memprediksi kemungkinan penyakit berdasarkan gejala pasien dan riwayat medis.

2.3. Naive Bayes

Salah satu teknik data mining yang umum digunakan yaitu teknik klasifikasi. Klasifikasi adalah suatu teknik pembentukan model dari data yang belum terklasifikasi, untuk digunakan

mengklasifikasi data baru. Klasifikasi membutuhkan data pelatihan untuk membangun suatu model klasifikasinya (Lestari, 2014)

Asumsi "naive" (naif) dalam algoritma ini adalah bahwa semua fitur (variabel independen) saling independen secara bersyarat diberikan kelas (target). Meskipun asumsi ini sering tidak sepenuhnya benar dalam data dunia nyata, Naive Bayes seringkali menunjukkan kinerja yang baik, terutama pada dataset dengan dimensi tinggi.

Teorema Bayes dinyatakan sebagai berikut:

$$P(H|X) = 1 + \frac{P(X|H)P(H)}{P(X)}$$

Keterangan:

X	:	Data dengan class yang belum diketahui
H	:	Hipotesis data X merupakan suatu class spesifik
P(H|X)	:	Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
P(H)	:	Probabilitas hipotesis H (prior prob.)
P(X|H)	:	Probabilitas X berdasarkan kondisi tersebut
P(X)	:	Probabilitas dari X

3. Metode

3.1. Pengumpulan Data

Berikut merupakan data calon penerima bantuan pendidikan berserta fitur fitur yang mencakup Jenis Tinggal, Alat Transportasi, Penghasilan Ayah, Penghasilan Ibu dan label yaitu sebagai penerima bantuan atau tidak. Prosedur sistematik yang digunakan untuk mengumpulkan data yaitu dengan metode penelitian kuantitatif yang berfokus pada penggunaan angka, tabel, grafik dan diagram untuk menampilkan hasil data yang diperoleh dan akan digunakan untuk bahan penganalisisan data terhadap metode Naive Bayes (Felicia Watratan et al., 2020).

Tabel 1. Data Training Siswa SMK Negeri 9 Bulukumba

Nama	Jenis Tinggal	Alat Transportasi	Penghasilan Ayah	Penghasilan Ibu	LABEL
Siswa001	Rumah Orangtua	Sepeda motor	Besar	Besar	Tidak
Siswa002	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa003	Wali	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa004	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa005	Wali	Jalan kaki	Sedang	Kecil	Tidak
Siswa006	Rumah Orangtua	Sepeda motor	Kecil	Sangat Kecil	Tidak
Siswa007	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa008	Wali	Jalan kaki	Kecil	Sangat Kecil	Ya
Siswa009	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa010	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa011	Rumah Orangtua	Sepeda motor	Kecil	Kecil	Tidak
Siswa012	Rumah Orangtua	Sepeda motor	Kecil	Kecil	Tidak
Siswa013	Rumah Orangtua	Jalan kaki	Besar	Sangat Kecil	Tidak
Siswa014	Rumah Orangtua	Sepeda motor	Kecil	Tidak Ada	Tidak
Siswa015	Rumah Orangtua	Angkutan umum	Sedang	Tidak Ada	Tidak
Siswa016	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak

Siswa017	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa018	Rumah Orangtua	Jalan kaki	Sedang	Tidak Ada	Tidak
Siswa019	Rumah Orangtua	Jalan kaki	Tidak Ada	Tidak Ada	Ya
Siswa020	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa021	Rumah Orangtua	Jalan kaki	Kecil	Kecil	Tidak
Siswa022	Wali	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa023	Wali	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa024	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Ya
Siswa025	Rumah Orangtua	Sepeda motor	Kecil	Kecil	Tidak
Siswa026	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa027	Rumah Orangtua	Jalan kaki	Sangat Kecil	Tidak Ada	Ya
Siswa028	Rumah Orangtua	Jalan kaki	Tidak Ada	Tidak Ada	Ya
Siswa029	Rumah Orangtua	Jalan kaki	Kecil	Kecil	Tidak
Siswa030	Rumah Orangtua	Jalan kaki	Besar	Tidak Ada	Tidak
Siswa031	Rumah Orangtua	Mobil pribadi	Sedang	Tidak Ada	Tidak
Siswa032	Rumah Orangtua	Jalan kaki	Sangat Kecil	Tidak Ada	Ya
Siswa033	Rumah Orangtua	Jalan kaki	Besar	Tidak Ada	Tidak
Siswa034	Rumah Orangtua	Jalan kaki	Kecil	Kecil	Ya
Siswa035	Rumah Orangtua	Jalan kaki	Sangat Kecil	Tidak Ada	Ya
Siswa036	Wali	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa037	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa038	Rumah Orangtua	Jalan kaki	Sangat Kecil	Tidak Ada	Ya
Siswa039	Rumah Orangtua	Jalan kaki	Sangat Kecil	Sangat Kecil	Ya
Siswa040	Rumah Orangtua	Jalan kaki	Besar	Tidak Ada	Tidak
Siswa041	Rumah Orangtua	Mobil pribadi	Kecil	Tidak Ada	Tidak
Siswa042	Rumah Orangtua	Jalan kaki	Sangat Kecil	Tidak Ada	Ya
Siswa043	Rumah Orangtua	Jalan kaki	Kecil	Kecil	Tidak
Siswa044	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa045	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa046	Rumah Orangtua	Mobil pribadi	Besar	Sedang	Tidak
Siswa047	Rumah Orangtua	Sepeda motor	Kecil	Besar	Tidak
Siswa048	Rumah Orangtua	Sepeda motor	Kecil	Tidak Ada	Tidak
Siswa049	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa050	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa051	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa052	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa053	Rumah Orangtua	Angkutan umum	Sedang	Tidak Ada	Tidak
Siswa054	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa055	Wali	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa056	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya
Siswa057	Wali	Jalan kaki	Sangat Kecil	Tidak Ada	Ya

Siswa058	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	Tidak
Siswa059	Wali	Jalan kaki	Kecil	Tidak Ada	Tidak
Siswa060	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	Ya

Daftar penghasilan orangtua ditulis Tidak ada, sangat kecil, kecil, sedang dan besar dengan keterangan sebagai berikut:
[1] Tidak Ada
[2] Sangat Kecil (Kurang dari IDR 500.000)
[3] Kecil (IDR 500.000 – IDR 999.999)
[4] Sedang (IDR 1.000.000 – IDR 1.999.999)
[5] Besar (IDR 2.000.000 – IDR 4.999.999)

3.2. Menghitung Probabilitas Label

$$P(Ya) = \frac{Jumlah\ Label\ Ya}{Total\ Jumlah\ Data} = \frac{23}{60} = 0{,}3833$$

$$P(Tidak) = \frac{Jumlah\ Label\ Tidak}{Total\ Jumlah\ Data} = \frac{37}{60} = 0{,}6167$$

3.3. Menghitung Probabilitas Atribut

Tabel 2. Probabilitas Masing-Masing Atribut

Aribut		P (Ya)	Nilai	P (Tidak)	Nilai
Jenis Tinggal	Bersama orangtua	18	0.7826	33	0.8919
	Wali	5	0.2174	4	0.1081
Alat Transportasi	Angkutan Umum	0	LC	2	0.0541
	Jalan Kaki	23	1.0000	24	0.6486
	Sepeda Motor	0	LC	8	0.2162
	Mobil Pribadi	0	LC	3	0.0811
Penghasilan Ayah	Tidak ada	2	0.0870	0	LC
	Sangat Kecil	0	LC	0	LC
	Kecil	14	0.6087	26	0.7027
	Sedang	0	LC	5	0.1351
	Besar	0	LC	6	0.1622
Penghasilan Ibu	Tidak ada	19	0.8261	16	0.4324
	Sangat Kecil	0	LC	0	LC
	Kecil	1	0.0435	7	0.1892
	Sedang	0	LC	1	0.0270
	Besar	0	LC	2	0.0541

LC atau sering juga disebut Laplacian Correction adalah sebuah teknik yang digunakan dalam algoritma klasifikasi Naive Bayes untuk mengatasi masalah probabilitas nol (zero probability).

4. Hasil dan Pembahasan

Tabel 3. Data Uji Siswa SMK Negeri 9 Bulukumba

Nama	Jenis Tinggal	Alat Transportasi	Penghasilan Ayah	Penghasilan Ibu	LABEL
Siswa_uji001	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	?
Siswa_uji002	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	?
Siswa_uji003	Rumah Orangtua	Angkutan umum	Sedang	Tidak Ada	?
Siswa_uji004	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	?
Siswa_uji005	Wali	Jalan kaki	Kecil	Tidak Ada	?
Siswa_uji006	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	?
Siswa_uji007	Wali	Jalan kaki	Sangat Kecil	Tidak Ada	?
Siswa_uji008	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil	?
Siswa_uji009	Wali	Jalan kaki	Kecil	Tidak Ada	?
Siswa_uji010	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada	?

4.1 . Menghitung Probabilitas Secara Manual

Data yang akan digunakan adalah data pertama yaitu ID: Siswa_uji0001, sebagai berikut:
Probabilitas Label:
P (Ya) = 0,3833
P (Tidak) = 0,6167

Tabel 4. Probabilitas Data Uji ID: Siswa_uji0001

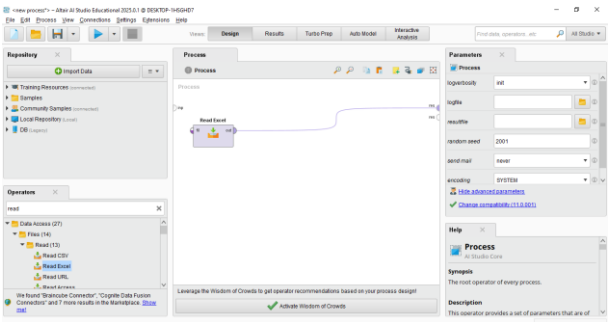
Aribut		P (Ya)	Nilai	P (Tidak)	Nilai
Jenis Tinggal	Bersama orangtua	18	0.7826	33	0.8919
Alat Transportasi	Jalan Kaki	23	1.0000	24	0.6486
Penghasilan Ayah	Kecil	14	0.6087	26	0.7027
Penghasilan Ibu	Tidak ada	19	0.8261	16	0.4324

Nilai P (Ya) = P (Total|Ya) * P (Bersama Orangtua|Ya) * P (Jalan Kaki|Ya) * P (Kecil|Ya) * P (Tidak ada|Ya)
= 0,3833 * 0.7826 * 1.0000 * 0.6087 * 0.8261
= 0.1509

Nilai P (Tidak) = P (Total|Tidak) * P (Bersama Orangtua|Tidak) * P (Jalan Kaki|Tidak) * P (Kecil|Tidak) * P (Tidak ada|Tidak)
= 0,6167 * 0.8919 * 0.6486 * 0.7027 * 0.4324
= 0.1084

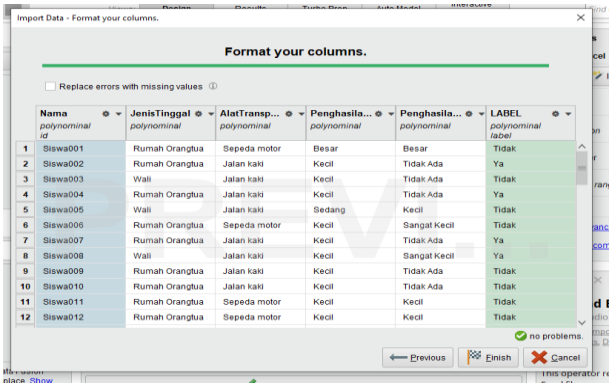
Sehingga karena nilai P (Ya) lebih besar dibanding nilai Nilai P (Tidak), maka ID: Siswa_uji0001 diprediksi sebagai Ya atau sebagai penerima bantuan pendidikan.

4.1 . Pengujian Menggunakan Altair AI Studio



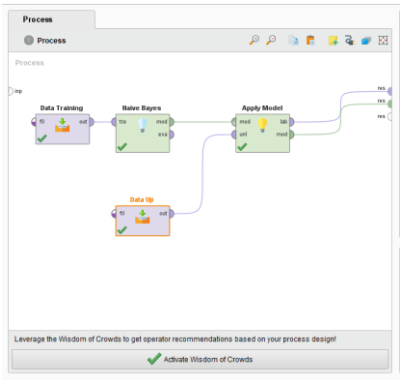
Gambar 1. Import Data Training

Di bagian operator, seret dan letakkan operator "Read Excel" ke area kerja (kanvas), kemudian unggah atau masukkan data pelatihan yang sudah disiapkan



Gambar 2. Pemberian ID dan Label

Tetapkan atribut "Nama siswa" sebagai ID dan tentukan atribut label yang sesuai melalui menu label.



Gambar 3. Proses Klasifikasi Naïve Bayes

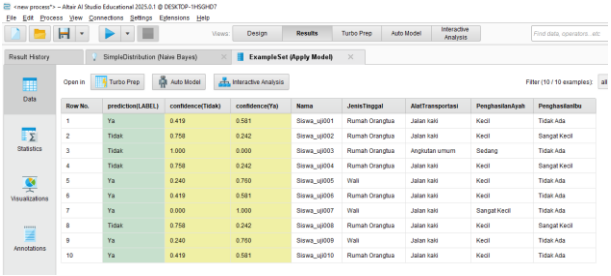
Berikut merpakan alur proses klasifikasi menggunakan algoritma Naive Bayes di dalam platform Altair AI Studio. Proses dimulai dengan memuat data pelatihan, yang kemudian diteruskan ke operator Naive Bayes untuk membangun model klasifikasi. Model yang dihasilkan selanjutnya digunakan oleh operator Apply Model untuk memprediksi label pada data uji. Operator Data Uji bertugas memuat data pengujian yang belum memiliki label. Alur ini menunjukkan tahap pelatihan model dan penerapan model ke data baru.

Setelah model diterapkan pada data uji, hasil prediksi akan ditampilkan sebagai output. Tiga jalur output dari operator Apply Model menunjukkan bahwa pengguna bisa melihat model, hasil label prediksi, serta data terkait lainnya. Proses ini merupakan contoh dasar dalam

penerapan machine learning, di mana data dilatih untuk membentuk model, kemudian model tersebut digunakan untuk mengklasifikasikan data baru yang belum diketahui hasilnya.

4.2 . Hasil Pengujian

Berikut merupakan hasil pengujian menggunakan tools Altair AI Studio.



Row No.	prediction(LABEL)	confidence(Tidak)	confidence(Ya)	Nama	JenisTinggal	AlatTransportasi	PenghasilanAyah	PenghasilanIbu
1	Ya	0.419	0.581	Siswa_uji001	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada
2	Tidak	0.758	0.242	Siswa_uji002	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil
3	Tidak	1.000	0.000	Siswa_uji003	Rumah Orangtua	Angkutan umum	Sedang	Tidak Ada
4	Tidak	0.758	0.242	Siswa_uji004	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil
5	Ya	0.240	0.760	Siswa_uji005	Wali	Jalan kaki	Kecil	Tidak Ada
6	Ya	0.419	0.581	Siswa_uji006	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada
7	Ya	0.000	1.000	Siswa_uji007	Wali	Jalan kaki	Sangat Kecil	Tidak Ada
8	Tidak	0.758	0.242	Siswa_uji008	Rumah Orangtua	Jalan kaki	Kecil	Sangat Kecil
9	Ya	0.240	0.760	Siswa_uji009	Wali	Jalan kaki	Kecil	Tidak Ada
10	Ya	0.419	0.581	Siswa_uji010	Rumah Orangtua	Jalan kaki	Kecil	Tidak Ada

Gambar 4. Hasil Pengujian Data Uji

Berdasakan hasil pengujian menggunakan tools Altair AI Studio, hasilnya menunjukan jawaban yang sama untuk data uji pada ID: Siswa_uji001 yaitu diprediksi sebagai (Ya) atau menerima bantuan Pendidikan.

5. Kesimpulan

Dari data 10 siswa uji, 6 siswa diklasifikasikan sebagai "Ya" (layak menerima bantuan) dan 4 siswa diklasifikasikan sebagai "Tidak" (tidak layak menerima bantuan).

Semua siswa menggunakan yang “jalan kaki” sebagai alat transportasi. Hanya satu siswa yang menggunakan Angkutan umum dan diklasifikasikan "Tidak". Ini mengindikasikan bahwa penggunaan angkutan umum dalam konteks ini tidak menjadi faktor pendukung untuk bantuan.

Siswa yang diklasifikasikan "Ya" memiliki penghasilan ayah "Kecil" dan "Sangat Kecil". Sehingga hal ini menunjukkan bahwa penghasilan ayah yang rendah merupakan faktor penentu utama.sedangkan penghasilan ibu yang diklasifikasikan "Ya" mayoritas memiliki penghasilan ibu "tidak ada". Hal ini sangat menonjol sebagai indikator kelayakan.

Penelitian ini berhasil mengimplementasikan algoritma Naive Bayes sebagai sistem pendukung keputusan untuk menentukan kelayakan penerima bantuan pendidikan. Algoritma Naive Bayes mampu mengklasifikasikan siswa berdasarkan kriteria. Algorima ini dapat membantu pihak terkait dalam membuat keputusan yang lebih objektif dan terukur.

Referensi

[1] Aryusya Thamrin, A., & Amanda Putri, R. (2024). Klasifikasi Tingkat Kecanduan Gadget Pada Balita Menggunakan Algoritma Naïve Bayes Classification of Gadget Addiction Levels in Toddlers Using the Naive bayes Algorithm. In *Kp. Tengah, Kec. Pancur Batu* (Vol. 9, Issue 2). www.jurnal.unimed.ac.id

[2] Felicia Watratan, A., Puspita, A. B., Moeis, D., Informasi, S., & Profesional Makassar, S. (2020a). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. In *JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST)* (Vol. 1, Issue 1). http://journal.isas.or.id/index.php/JACOST

[3] Lestari, M. (2014). PENERAPAN ALGORITMA KLASIFIKASI NEAREST NEIGHBOR (K-NN) UNTUK MENDETEKSI PENYAKIT JANTUNG. In *Faktor Exacta* (Vol. 7, Issue 4).

[4] Sriatun, S., Sugiono, S., Kurniasih, N. B., & Hendrizal, H. (2024). The Essence of Education. *Education Achievement: Journal of Science and Research*, 440–445. https://doi.org/10.51178/jsr.v5i2.1910

[5] Yusuf, Y. W., & Rian Pratikto, F. (2006). PENERAPAN DATA MINING DALAM PENENTUAN ATURAN ASOSIASI ANTAR JENIS ITEM. *Seminar Nasional Aplikasi Teknologi Informasi*.

[6] Aryusya Thamrin, A., & Amanda Putri, R. (2024). Klasifikasi Tingkat Kecanduan Gadget Pada Balita Menggunakan Algoritma Naïve Bayes Classification of Gadget Addiction Levels in Toddlers Using the Naive bayes Algorithm. In Kp. Tengah, Kec. Pancur Batu (Vol. 9, Issue 2). www.jurnal.unimed.ac.id

[7] Felicia Watratan, A., Puspita, A. B., Moeis, D., Informasi, S., & Profesional Makassar, S. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. In JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST) (Vol. 1, Issue 1). http://journal.isas.or.id/index.php/JACOST