



DOI: <https://doi.org/10.38035/dit.v4i1>
<https://creativecommons.org/licenses/by/4.0/>

Sistem Pendukung Keputusan Peminatan Jurusan Menggunakan Metode *Naïve Bayes*

Moch Nur Alfian Ardiansyah¹

¹Ilmu Komputer, Fakultas Ilmu Eksakta, Universitas Nahdlatul Ulama Blitar, Indonesia,
alfann478@gmail.com.

Corresponding Author: alfann478@gmail.com¹

Abstract: *The choice of major at the Senior High School (SMA) level is an important decision that affects the learning process and educational planning of students. However, the process of determining majors is still often subjective, potentially resulting in placements that do not align with students' academic abilities. This study aims to develop a web-based Decision Support System (DSS) using the Naïve Bayes method to provide major recommendations for SMA Nahdlatul Ulama Pakis students. The study used a Research and Development (R&D) method with a quantitative approach. The research data consisted of academic records for 65 students from SMA Nahdlatul Ulama Pakis. The classification process was carried out using the Gaussian Naïve Bayes algorithm, which computed prior probabilities, means, standard deviations, and posterior probabilities to determine the class with the highest posterior probability. The system was then implemented in a web-based application and evaluated using a Confusion Matrix. The test results show that the model achieved 45.45% accuracy, 43.60% macro precision, 42.94% macro recall, and 41.54% macro F1-score. The results of this study indicate that the Naïve Bayes method can be implemented as a basis for decision-making in recommending student majors. However, the accuracy obtained indicates that further development, such as increasing the number of data points and attributes, is still needed to improve system performance.*

Keyword: *Decision support systems, Naïve Bayes, student majoring, classification*

Abstrak: Pemilihan jurusan di tingkat SMA merupakan keputusan penting yang memengaruhi proses pembelajaran dan perencanaan pendidikan siswa. Namun, proses penentuan jurusan masih sering kali bersifat subjektif, berpotensi mengakibatkan penempatan yang tidak sesuai dengan kemampuan akademik siswa. Penelitian ini bertujuan untuk mengembangkan Sistem Pendukung Keputusan (DSS) berbasis web menggunakan metode Naïve Bayes untuk memberikan rekomendasi jurusan bagi siswa SMA Nahdlatul Ulama Pakis. Penelitian ini menggunakan metode Penelitian dan Pengembangan (R&D) dengan pendekatan kuantitatif. Data penelitian terdiri dari rekam medis 65 siswa SMA Nahdlatul Ulama Pakis. Proses klasifikasi dilakukan menggunakan algoritma Gaussian Naïve Bayes, yang menghitung probabilitas prior, rata-rata, deviasi standar, dan probabilitas posterior untuk menentukan kelas dengan probabilitas posterior tertinggi. Sistem kemudian diimplementasikan dalam aplikasi berbasis web dan dievaluasi menggunakan Matriks Konfusi. Hasil pengujian menunjukkan

bahwa model tersebut mencapai akurasi 45,45%, presisi makro 43,60%, recall makro 42,94%, dan skor F1 makro 41,54%. Hasil penelitian ini menunjukkan bahwa metode Naïve Bayes dapat diimplementasikan sebagai dasar pengambilan keputusan dalam merekomendasikan jurusan mahasiswa. Namun, akurasi yang diperoleh menunjukkan bahwa pengembangan lebih lanjut, seperti peningkatan jumlah titik data dan atribut, masih diperlukan untuk meningkatkan kinerja sistem.

Kata Kunci: Sistem pendukung keputusan, Naïve Bayes, penjurusan siswa, klasifikasi

PENDAHULUAN

Pendidikan merupakan salah satu aspek terpenting dalam kehidupan manusia yang berfungsi sebagai sarana untuk mengembangkan potensi individu, meningkatkan kualitas sumber daya manusia, serta mempersiapkan generasi yang mampu menghadapi perkembangan ilmu pengetahuan dan teknologi (Fabiana Meijon Fadul, 2019). Melalui pendidikan, setiap individu memperoleh kesempatan untuk mengembangkan kemampuan intelektual, keterampilan, serta karakter sebagai bekal dalam kehidupan bermasyarakat. Dalam teori pendidikan terdapat beberapa konsep, yaitu pendidikan klasik yang berorientasi pada pewarisan ilmu pengetahuan, pendidikan personal yang menekankan pengembangan potensi peserta didik dengan guru sebagai pembimbing, fasilitator, dan motivator, serta pendidikan interaksional yang memandang pendidikan sebagai proses interaksi antara peserta didik dan lingkungan sosial berdasarkan pemikiran filsafat pragmatisme (Fabiana Meijon Fadul, 2019).

Salah satu jenjang pendidikan yang berperan penting dalam menentukan arah pendidikan siswa adalah Sekolah Menengah Atas (SMA). Pada jenjang ini, siswa mulai menentukan peminatan atau jurusan yang akan menjadi dasar dalam proses pembelajaran serta menentukan arah pendidikan maupun karier di masa depan. Oleh karena itu, pemilihan jurusan perlu dilakukan dengan tepat dengan mempertimbangkan kemampuan akademik, minat, dan potensi yang dimiliki oleh setiap siswa.

SMA Nahdlatul Ulama Pakis merupakan salah satu sekolah menengah atas swasta yang berdiri sejak tahun 1985 di bawah naungan Yayasan MWC NU Pakis. Sekolah ini beralamat di Jalan H. Muthofa No.108, Kecamatan Pakis, Kabupaten Malang. SMA Nahdlatul Ulama Pakis memiliki tiga pilihan jurusan, yaitu Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS), dan Bahasa. Sebagian besar peserta didik berasal dari wilayah Kecamatan Pakis, Jabung, dan Tumpang. Informasi mengenai Penerimaan Peserta Didik Baru (PPDB) disampaikan melalui berbagai media, seperti *banner*, *website* sekolah, media sosial, serta kegiatan promosi di Sekolah Menengah Pertama (SMP).

Dalam praktiknya, proses pemilihan jurusan masih menjadi salah satu permasalahan yang dihadapi siswa. Tidak sedikit siswa yang menentukan jurusan berdasarkan keinginan orang tua, mengikuti pilihan teman, gengsi terhadap jurusan tertentu, maupun kurangnya pemahaman tentang kemampuan akademik yang mereka miliki (Pratama, 2021). Kondisi tersebut dapat menyebabkan siswa mengalami kesulitan mengikuti proses pembelajaran karena jurusan yang dipilih tidak sesuai dengan kompetensi yang dimilikinya. Dampaknya adalah menurunnya prestasi belajar, berkurangnya motivasi belajar, hingga munculnya keinginan untuk berpindah jurusan.

Berdasarkan hasil observasi di SMA Nahdlatul Ulama Pakis, proses penentuan jurusan masih dilakukan berdasarkan nilai akademik serta pertimbangan guru melalui musyawarah. Proses tersebut belum didukung oleh sistem yang mampu mengolah data akademik secara terkomputerisasi untuk menghasilkan rekomendasi jurusan yang objektif. Penentuan jurusan yang masih dilakukan secara manual berpotensi menimbulkan subjektivitas, sehingga terdapat kemungkinan siswa memperoleh jurusan yang kurang sesuai dengan kemampuan akademiknya.

Seiring berkembangnya teknologi informasi, proses pengambilan keputusan dapat dibantu melalui penerapan Sistem Pendukung Keputusan (SPK). Sistem Pendukung Keputusan merupakan sistem berbasis komputer yang dirancang untuk membantu proses pengambilan keputusan dengan memanfaatkan data dan metode tertentu sehingga keputusan yang dihasilkan menjadi lebih cepat, objektif, dan konsisten. Salah satu metode klasifikasi yang banyak digunakan dalam Sistem Pendukung Keputusan adalah *Naïve Bayes*. Metode ini menggunakan pendekatan probabilitas berdasarkan *Teorema Bayes* dengan asumsi bahwa setiap atribut bersifat independen. *Naïve Bayes* memiliki beberapa keunggulan, antara lain proses perhitungan yang sederhana, mudah diimplementasikan, mampu mengolah data dalam jumlah relatif sedikit, serta menghasilkan klasifikasi yang cukup baik dalam berbagai penelitian.

Beberapa penelitian terdahulu menunjukkan bahwa metode *Naïve Bayes* telah berhasil diterapkan pada berbagai bidang, seperti penentuan penerima beasiswa, klasifikasi kelulusan mahasiswa, pemilihan jurusan mahasiswa, klasifikasi masyarakat miskin, serta penentuan penerima bantuan sosial. Hasil penelitian-penelitian tersebut menunjukkan bahwa metode *Naïve Bayes* memiliki kemampuan yang baik dalam menyelesaikan berbagai permasalahan klasifikasi dan pengambilan keputusan karena mampu memberikan rekomendasi berdasarkan perhitungan probabilitas dari data yang tersedia.

Namun demikian, hingga saat ini belum ditemukan penelitian yang secara khusus menerapkan metode *Naïve Bayes* pada Sistem Pendukung Keputusan untuk menentukan jurusan siswa di SMA Nahdlatul Ulama Pakis. Selain itu, sebagian besar penelitian sebelumnya hanya berfokus pada implementasi metode tanpa melakukan evaluasi kinerja model secara komprehensif menggunakan *Confusion Matrix*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Padahal, evaluasi tersebut sangat penting untuk mengetahui tingkat akurasi dan keandalan metode dalam menghasilkan rekomendasi yang tepat, sehingga dapat dijadikan dasar pengambilan keputusan.

Berdasarkan kondisi tersebut, penelitian ini memiliki kebaruan (*novelty*) dibandingkan penelitian-penelitian sebelumnya, yaitu membangun Sistem Pendukung Keputusan berbasis web menggunakan metode *Naïve Bayes* untuk memberikan rekomendasi pemilihan jurusan bagi siswa berdasarkan kemampuan akademik. Selain menghasilkan rekomendasi jurusan, penelitian ini juga melakukan evaluasi performa model menggunakan *Confusion Matrix*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, sehingga tidak hanya diketahui hasil klasifikasi, tetapi juga tingkat kinerja metode *Naïve Bayes* dalam proses penentuan jurusan.

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada penerapan metode *Naïve Bayes* dalam Sistem Pendukung Keputusan untuk menentukan jurusan siswa di SMA Nahdlatul Ulama Pakis. Permasalahan yang ingin dijawab dalam penelitian ini adalah bagaimana menerapkan metode *Naïve Bayes* pada Sistem Pendukung Keputusan untuk memberikan rekomendasi jurusan kepada siswa, bagaimana membangun aplikasi yang mampu membantu proses penentuan jurusan secara objektif, serta bagaimana tingkat kinerja metode *Naïve Bayes* dalam melakukan klasifikasi jurusan berdasarkan hasil pengujian menggunakan *Confusion Matrix*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

Melalui penelitian ini diharapkan dapat dibangun Sistem Pendukung Keputusan berbasis web yang mampu memberikan rekomendasi jurusan sesuai dengan kemampuan akademik siswa. Selain itu, penelitian ini juga bertujuan untuk mengetahui tingkat performa metode *Naïve Bayes* dalam melakukan klasifikasi jurusan sehingga hasil yang diperoleh dapat menjadi bahan pertimbangan bagi pihak sekolah dalam proses penentuan jurusan. Dengan demikian, sistem yang dikembangkan tidak dimaksudkan sebagai penentu keputusan akhir, melainkan sebagai alat bantu yang memberikan rekomendasi secara objektif sehingga diharapkan mampu meminimalkan kesalahan dalam pemilihan jurusan dan meningkatkan kesesuaian antara kemampuan akademik siswa dengan jurusan yang dipilih.

METODE

Penelitian ini menggunakan metode *Research and Development* (R&D) karena bertujuan menghasilkan produk berupa Sistem Pendukung Keputusan (SPK) berbasis web yang dapat digunakan untuk membantu proses pemilihan jurusan siswa di SMA Nahdlatul Ulama Pakis. Sistem yang dikembangkan mengimplementasikan metode *Naïve Bayes* sebagai algoritma klasifikasi untuk memberikan rekomendasi jurusan berdasarkan kemampuan akademik siswa. Melalui penelitian ini diharapkan dapat dihasilkan sistem yang mampu membantu proses pengambilan keputusan secara lebih objektif, cepat, dan akurat.

Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif, karena seluruh proses analisis dilakukan berdasarkan data numerik berupa nilai akademik siswa. Data tersebut diolah menggunakan metode *Naïve Bayes* untuk memperoleh probabilitas setiap alternatif jurusan, sehingga sistem dapat memberikan rekomendasi jurusan yang paling sesuai dengan karakteristik akademik masing-masing siswa. Hasil perhitungan yang diperoleh kemudian diimplementasikan ke dalam aplikasi berbasis web sebagai media pengambilan keputusan. Secara umum, tahapan penerapan algoritma *Naïve Bayes* dalam penelitian ini meliputi: (1) membaca *data training*; (2) menghitung *prior probability* setiap kelas; (3) menghitung nilai *mean* dan standar deviasi untuk setiap atribut numerik pada masing-masing kelas; (4) menghitung probabilitas setiap atribut menggunakan distribusi *Gaussian*; (5) mengalikan seluruh probabilitas atribut dengan *prior probability* untuk memperoleh *posterior probability*; dan (6) menentukan jurusan berdasarkan nilai probabilitas terbesar.

Tahapan penelitian dimulai dari proses identifikasi permasalahan, pengumpulan data akademik siswa, pengolahan data menggunakan metode *Naïve Bayes*, perancangan dan pembangunan aplikasi Sistem Pendukung Keputusan berbasis web, hingga pengujian kinerja sistem. Pengujian dilakukan dengan membandingkan hasil prediksi sistem dengan data aktual menggunakan *Confusion Matrix* yang menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Melalui tahapan tersebut dapat diketahui tingkat kinerja metode *Naïve Bayes* dalam memberikan rekomendasi pemilihan jurusan siswa di SMA Nahdlatul Ulama Pakis. Dalam penelitian ini, Microsoft Excel dimanfaatkan sebagai media pendukung untuk mengolah data, menghitung probabilitas, dan memvalidasi hasil perhitungan metode *Naïve Bayes*, sehingga setiap proses klasifikasi dapat diverifikasi secara matematis. Hasil validasi tersebut kemudian diimplementasikan pada Sistem Pendukung Keputusan berbasis web yang dikembangkan dalam penelitian ini. Implementasi tersebut memungkinkan proses penentuan jurusan dilakukan secara otomatis berdasarkan kemampuan akademik siswa, menghasilkan rekomendasi yang lebih cepat, objektif, konsisten, serta mengurangi potensi kesalahan perhitungan yang sering terjadi pada proses pengolahan data secara manual.

HASIL DAN PEMBAHASAN

Implementasi Metode *Naïve Bayes*

Pada metode *Naïve Bayes*, *prior probability* adalah probabilitas awal suatu kelas (jurusan) sebelum mempertimbangkan nilai atau atribut siswa. *Prior probability* diperoleh dari proporsi jumlah data pada setiap kelas dibandingkan dengan seluruh data pelatihan. Berikut ini adalah Rumus *Prior Probability*:

$$P(H) = \frac{\text{Jumlah data pada kelas}}{\text{Jumlah seluruh data}}$$

Keterangan:

$P(H)$	= <i>Prior Probability</i> suatu kelas (IPA, IPS, atau Bahasa).
Jumlah data pada kelas	= banyaknya data training pada kelas tersebut.
Jumlah seluruh data	= Total seluruh <i>data training</i> .

Jumlah data training adalah 54 siswa, dengan rincian:

Tabel 1. Jumlah Jurusan SMA NU Pakis

Jurusan	Jumlah Data
IPA	21
IPS	19
Bahasa	14
Total	54

Maka:

Prior Probability IPA

$$P(IPA) = \frac{21}{54} = 0,3889$$

Prior Probability IPS

$$P(IPS) = \frac{19}{54} = 0,3519$$

Prior Probability Bahasa

$$P(Bahasa) = \frac{14}{54} = 0,2593$$

Prior Probability merupakan probabilitas awal dari setiap kelas (jurusan) yang dihitung berdasarkan proporsi jumlah data pada masing-masing kelas terhadap total seluruh *data training*. Nilai ini digunakan sebagai dasar dalam proses klasifikasi menggunakan metode *Naïve Bayes* sebelum mempertimbangkan probabilitas atribut lainnya. Semakin besar nilai *prior probability* suatu kelas, semakin besar peluang awal data baru diklasifikasikan ke dalam kelas tersebut sebelum perhitungan probabilitas atribut dilakukan.

Mean (rata-rata) merupakan nilai yang menunjukkan pusat data dari setiap atribut pada masing-masing kelas (jurusan). Dalam metode *Naïve Bayes Gaussian*, nilai *mean* digunakan untuk menghitung probabilitas data kontinu menggunakan distribusi *Gaussian* (Normal). Berikut ini adalah Rumus *Mean*.

$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

Keterangan:

μ = *Mean* (rata-rata)

x_i = Nilai data ke- i

n = Jumlah data pada kelas

Perhitungan nilai *Mean* (rata-rata) dilakukan untuk memperoleh nilai rata-rata setiap atribut pada *data training*. Dalam penelitian ini, perhitungan *Mean* dilakukan menggunakan fungsi AVERAGE di Microsoft Excel. Misalnya, jika data nilai suatu atribut berada pada rentang sel B2:B6, maka rumus yang digunakan adalah: = AVERAGE(B2:B6)

Rumus tersebut akan menjumlahkan seluruh nilai pada sel B2 hingga B6, kemudian membaginya dengan jumlah data pada rentang tersebut sehingga diperoleh nilai rata-rata (*Mean*). Nilai *Mean* yang dihasilkan selanjutnya digunakan sebagai parameter distribusi *Gaussian* dalam metode *Naïve Bayes* untuk menghitung probabilitas data kontinu. Maka dari itu, setelah dimasukkan ke dalam *Excel*, ditemukan hasil sebagai berikut.

Tabel 2. Mean Mata Pelajaran Jurusan IPA

IPA	MEAN
Nilai Ekonomi	74.19047619
Nilai Sosiologi	71.57142857
Nilai Seni Budaya	66.76190476
Nilai PJOK	71.33333333
Nilai TIK	72.04761905
Nilai Bahasa Jepang	74.47619048
Nilai KE NU an	65.47619048
Nilai Bahasa Jawa	69.95238095
Nilai An Nasr	73.04761905

Berdasarkan tabel di atas, diketahui bahwa setiap atribut pada jurusan IPA memiliki nilai rata-rata yang berbeda. Nilai *Mean* tertinggi terdapat pada Nilai Bahasa Jepang sebesar 74,47619048, sedangkan nilai *Mean* terendah terdapat pada Ke-NU-an sebesar 65,47619048.

Tabel 3. Mean Mata Pelajaran Jurusan IPS

IPS	MEAN
Nilai Ekonomi	72.68421053
Nilai Sosiologi	73.05263158
Nilai Seni Budaya	70.84210526
Nilai PJOK	73.15789474
Nilai TIK	71.73684211
Nilai Bahasa Jepang	72.15789474
Nilai KE NU an	71.31578947
Nilai Bahasa Jawa	62.47368421
Nilai An Nasr	74.47368421

Berdasarkan tabel di atas, diketahui bahwa setiap atribut pada jurusan IPS memiliki nilai rata-rata yang berbeda. Nilai *Mean* tertinggi terdapat pada Nilai An Nasr sebesar 74,47368421, sedangkan nilai *Mean* terendah terdapat pada Nilai Bahasa Jawa sebesar 62,47368421.

Tabel 4. Mean Mata Pelajaran Jurusan BAHASA

BAHASA	MEAN
Nilai Ekonomi	70.78571429
Nilai Sosiologi	73.35714286
Nilai Seni Budaya	75.14285714
Nilai PJOK	70.21428571
Nilai TIK	68.85714286
Nilai Bahasa Jepang	71.21428571
Nilai KE NU an	68.57142857
Nilai Bahasa Jawa	67.57142857
Nilai An Nasr	66.07142857

Berdasarkan tabel di atas, diketahui bahwa setiap atribut pada jurusan Bahasa memiliki nilai rata-rata yang berbeda. Nilai *Mean* tertinggi terdapat pada Nilai Seni Budaya sebesar 75,14285714, sedangkan nilai *Mean* terendah terdapat pada Nilai An Nasr sebesar 66,07142857.

Standar deviasi merupakan ukuran yang menunjukkan tingkat penyebaran atau variasi data terhadap nilai rata-rata (*Mean*). Pada metode *Gaussian Naïve Bayes*, standar deviasi digunakan bersama nilai mean untuk menghitung probabilitas atribut kontinu menggunakan distribusi *Gaussian*. Berikut ini adalah rumus Standar Deviasi.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

Keterangan:

σ = Standar deviasi

x_i = Nilai data ke-i

μ = Nilai mean (rata-rata)

n = Jumlah data

Perhitungan standar deviasi dilakukan untuk mengetahui tingkat penyebaran data pada setiap atribut terhadap nilai rata-rata *data training*. Dalam penelitian ini, perhitungan dilakukan menggunakan fungsi STDEV.S yang dikombinasikan dengan fungsi FILTER pada Microsoft Excel. Misalnya, untuk menghitung standar deviasi atribut Nilai Ekonomi pada jurusan IPA digunakan rumus: = STDEV.S(FILTER(B2:B55,M2:M55 = "IPA"))

Rumus tersebut menyaring data berdasarkan kriteria jurusan IPA, kemudian menghitung standar deviasi dari data yang telah difilter sehingga diperoleh nilai penyebaran data untuk setiap atribut. Maka dari itu, setelah dimasukkan ke dalam Excel, ditemukan hasil sebagaimana berikut.

Tabel 5. Standar Deviasi Mata Pelajaran Jurusan IPA

IPA	STD.DEV
Nilai Ekonomi	14.5795029
Nilai Sosiologi	13.59253997
Nilai Seni Budaya	13.41605293
Nilai PJOK	12.63065055
Nilai TIK	12.3388662
Nilai Bahasa Jepang	11.6000821
Nilai KE NU an	13.34773032
Nilai Bahasa Jawa	12.28607419
Nilai An Nasr	13.43307928

Berdasarkan tabel di atas mata pelajaran Ekonomi memiliki nilai standar deviasi tertinggi sebesar 14.5795 yang mengindikasikan bahwa sebaran nilai siswa pada mata pelajaran tersebut paling heterogen. Sebaliknya, mata pelajaran Bahasa Jepang memiliki standar deviasi terendah sebesar 11,6001 yang menunjukkan bahwa nilai siswa relatif lebih homogen atau konsisten dibandingkan dengan mata pelajaran lainnya.

Tabel 6. Standar Deviasi Mata Pelajaran Jurusan IPS

IPS	STD.DEV
Nilai Ekonomi	16.03486698
Nilai Sosiologi	14.54293904
Nilai Seni Budaya	13.1582807
Nilai PJOK	13.2172579
Nilai TIK	11.81262086
Nilai Bahasa Jepang	12.18497781
Nilai KE NU an	14.56423562

Nilai Bahasa Jawa	14.10030898
Nilai An Nasr	13.04423424

Berdasarkan tabel di diatas mata pelajaran Ekonomi memiliki standar deviasi tertinggi sebesar 16.0349 yang menunjukkan bahwa nilai siswa pada mata pelajaran tersebut paling bervariasi dibandingkan mata pelajaran lainnya. Sementara itu, mata pelajaran Sosiologi dan Ke-NU-an juga memiliki tingkat penyebaran yang cukup tinggi dengan nilai standar deviasi masing-masing 14.5429 dan 14.5642. Adapun mata pelajaran TIK memiliki standar deviasi terendah sebesar 11,8126 yang menunjukkan bahwa nilai siswa pada mata pelajaran tersebut lebih homogen atau lebih konsisten.

Tabel 7. Standar Deviasi Mata Pelajaran Jurusan BAHASA

BAHASA	STD.DEV
Nilai Ekonomi	11.70915862
Nilai Sosiologi	11.32463036
Nilai Seni Budaya	13.0257564
Nilai PJOK	17.3168596
Nilai TIK	14.99670293
Nilai Bahasa Jepang	14.32683642
Nilai KE NU an	16.3740615
Nilai Bahasa Jawa	11.66001904
Nilai An Nasr	13.07648674

Berdasarkan tabel di atas, mata pelajaran PJOK memiliki standar deviasi tertinggi sebesar 17,3169 yang menunjukkan bahwa nilai siswa pada mata pelajaran tersebut paling bervariasi atau memiliki tingkat heterogenitas paling tinggi. Selain itu, mata pelajaran ke-NU-an dan TIK juga memiliki tingkat penyebaran yang cukup tinggi, dengan nilai standar deviasi masing-masing 16,3741 dan 14,9967. Sementara itu, mata pelajaran Sosiologi dan Bahasa Jawa memiliki standar deviasi terendah, yaitu 11.3246 dan 11.6600 yang menunjukkan bahwa nilai siswa pada kedua mata pelajaran tersebut lebih homogen atau lebih konsisten.

Pada metode *Naïve Bayes*, perhitungan probabilitas untuk data numerik dilakukan menggunakan distribusi *Gaussian (Normal Distribution)*. Perhitungan ini digunakan untuk menentukan peluang suatu nilai siswa pada masing-masing kelas berdasarkan nilai rata-rata (mean) dan standar deviasi dari setiap atribut. Semakin dekat nilai data terhadap *Mean*, maka probabilitas yang dihasilkan akan semakin besar, sedangkan semakin jauh dari *Mean*, maka probabilitas akan semakin kecil. Hasil perhitungan ini kemudian digunakan sebagai dasar dalam proses klasifikasi dengan cara mengalikan seluruh probabilitas dari setiap atribut. Rumus probabilitas *Gaussian (Gaussian Naïve Bayes)* adalah sebagai berikut.

$$P(x | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Keterangan:

x = nilai data yang dihitung

μ = nilai rata-rata (mean)

σ = standar deviasi

π = 3.14159

e = 2.718 (bilangan Euler)

Adapun rumus Excel yang digunakan yaitu = *NORM.DIST(S295,C94,D94,FALSE)*:. Rumus tersebut digunakan untuk menghitung probabilitas distribusi normal (Gaussian) dari suatu nilai data berdasarkan mean dan standar deviasi. Pada tahap ini, seluruh nilai probabilitas *Gaussian* dari setiap atribut (mata pelajaran) dikalikan untuk memperoleh nilai *likelihood* total.

Perkalian ini dilakukan karena metode *Naïve Bayes* mengasumsikan bahwa setiap atribut bersifat independen satu sama lain. Dengan demikian, hasil akhir merupakan gabungan probabilitas seluruh fitur terhadap suatu kelas menggunakan rumus $= \text{PRODUCT}(B2:J2)$. Setelah diperoleh nilai *likelihood* dari seluruh atribut, langkah selanjutnya adalah mengalikan hasil tersebut dengan nilai *prior probability* masing-masing kelas. *Prior probability* diperoleh dari distribusi data awal pada setiap jurusan (IPA, IPS, atau Bahasa).

Dengan rumus $= N10 * \text{PRODUCT}(B2:J2)$, hasil perhitungan ini disebut sebagai *posterior probability* $P(\text{Kelas}/X)$ yang menunjukkan peluang suatu data (X) termasuk ke dalam kelas tertentu, misalnya IPA. Kelas dengan nilai posterior tertinggi akan menjadi hasil prediksi akhir dalam sistem klasifikasi *Naïve Bayes*. Maka dari itu dihasilkan hasil perhitungan sebagaimana berikut.

Tabel 8. Hasil Klasifikasi Jurusan Siswa Menggunakan Metode Naïve Bayes

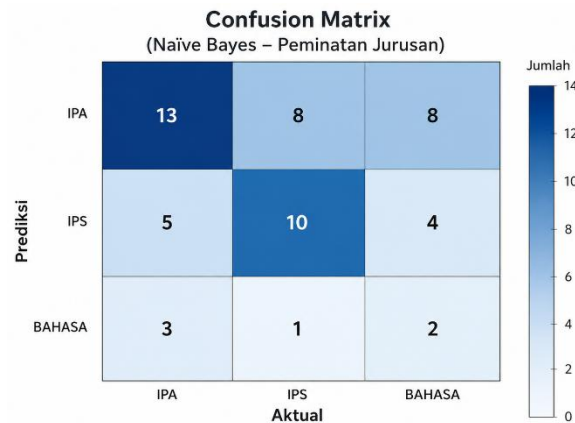
No	Nama	IPA	IPS	BAHASA	MAX	Jurusan Prediksi	Jurusan Asal
1	Bibit Sugiarto	1.27e-16	5.84e-17	9.44e-18	1.27e-16	IPA	IPS
2	Bagas Farhan	4.00e-18	9.09e-18	1.11e-18	9.09e-18	IPS	IPS
3	Bagus Fahmi	4.27e-16	3.47e-16	1.06e-16	4.27e-16	IPA	IPS
4	Nadia Putri	6.30e-18	8.73e-17	8.90e-18	8.73e-17	IPS	IPS
5	Badiatun Niswa	2.15e-16	2.07e-16	1.56e-16	2.15e-16	IPA	IPS
6	M.Tamamul Khilqoh	2.09e-17	3.10e-17	7.30e-18	3.10e-17	IPS	IPS
7	Wildan	3.48e-18	4.42e-18	3.67e-18	4.42e-18	IPS	IPS
8	Dwi Rihsu Ning Tyas	6.11e-17	1.39e-16	1.78e-17	1.39e-16	IPS	IPS
9	Tommy	6.47e-18	1.22e-17	1.56e-18	1.22e-17	IPS	IPS
10	Febriansah	4.75e-16	2.50e-16	1.24e-16	4.75e-16	IPA	IPS
11	Heri Surdamaji	1.65e-16	3.83e-16	1.01e-16	3.83e-16	IPS	IPS
12	Figo Ramadhani	7.13e-16	6.78e-16	3.58e-16	7.13e-16	IPA	IPS
13	Ade Ayu Trianto	1.14e-15	3.03e-16	5.88e-17	1.14e-15	IPA	IPS
14	Dwiki Saputra	3.90e-17	2.33e-17	9.28e-18	3.90e-17	IPA	IPS
15	Naufal Rifaldi	2.78e-16	2.08e-16	2.36e-16	2.78e-16	IPA	IPS
16	A.Saifudin Zuhri	4.48e-17	4.20e-17	5.29e-17	5.29e-17	BAHASA	IPS
17	Farhah Aqila	2.51e-17	5.84e-17	4.64e-17	5.84e-17	IPS	IPS
18	Wardatul Jannah	1.40e-17	4.78e-17	1.20e-17	4.78e-17	IPS	IPS
19	Lana Hikmatul Ulla	6.44e-17	2.16e-16	9.49e-17	2.16e-16	IPS	IPS
20	Mustagfirin	1.18e-17	1.87e-18	2.01e-18	1.18e-17	IPA	IPA
21	Lafifatul Umma	6.72e-17	5.62e-18	2.05e-17	6.72e-17	IPA	IPA
22	Erlin Yulia	5.13e-16	8.32e-17	6.53e-17	5.13e-16	IPA	IPA
23	Meilasari	1.49e-16	1.96e-16	1.97e-16	1.97e-16	BAHASA	IPA
24	Adi Saputra	8.09e-16	1.64e-15	7.38e-16	1.64e-15	IPS	IPA
25	Ahmad Jalil	1.78e-16	3.52e-17	4.02e-17	1.78e-16	IPA	IPA
26	Divya Alfa Prameswari	2.54e-16	1.08e-16	5.44e-17	2.54e-16	IPA	IPA
27	Fiki Febrianto	8.93e-17	1.14e-16	3.09e-17	1.14e-16	IPS	IPA
28	Ramanda Putra Panama	1.01e-16	1.01e-16	1.73e-17	1.01e-16	IPA	IPA

29	Reza Tri Saputra	1.94e-21	9.47e-18	9.70e-18	9.70e-18	BAHASA	IPA
30	Lambang Seno	6.03e-16	1.52e-16	3.40e-16	6.03e-16	IPA	IPA
31	Seno Aji Pamungkas	5.57e-18	2.26e-18	1.40e-17	1.40e-17	BAHASA	IPA
32	Nur Fadilla	4.48e-16	4.55e-17	1.83e-17	4.48e-16	IPA	IPA
33	Ade Narendra	8.58e-17	3.21e-17	1.42e-17	8.58e-17	IPA	IPA
34	Niken Salsabilla Lestari	1.44e-15	6.84e-16	2.86e-16	1.44e-15	IPA	IPA
35	Dwi Maria Sari	4.18e-16	5.07e-16	1.52e-16	5.07e-16	IPS	IPA
36	Miftaul Huda	2.03e-16	1.73e-16	4.34e-17	2.03e-16	IPA	IPA
37	Stani Aurellia	2.04e-16	2.87e-16	2.56e-16	2.87e-16	IPS	IPA
38	Dwi Ashari	3.41e-20	4.11e-17	3.77e-17	4.11e-17	IPS	IPA
39	Maulidia Wahidatul	6.70e-17	3.43e-17	2.89e-18	6.70e-17	IPA	IPA
40	Fathul Hasyim	4.39e-17	2.07e-17	1.02e-17	4.39e-17	IPA	IPA
41	Amelia Mirna Sari	1.85e-17	9.50e-18	7.76e-18	1.85e-17	IPA	BAHASA
42	Sulkhan	5.13e-18	3.87e-18	8.67e-18	8.67e-18	BAHASA	BAHASA
43	Anis Safitri	2.88e-16	3.85e-16	8.63e-17	3.85e-16	IPS	BAHASA
44	Meilana Saari	3.11e-16	2.69e-16	1.23e-16	3.11e-16	IPA	BAHASA
45	Wahyu Hidayat	4.04e-18	9.49e-18	4.15e-18	9.49e-18	IPS	BAHASA
46	Sri Handayani	1.01e-17	3.13e-17	2.99e-23	3.13e-17	IPS	BAHASA
47	Bayu Kurniawan	1.41e-17	2.22e-15	2.13e-17	2.22e-15	IPS	BAHASA
48	Diki Fahrezi	1.09e-17	4.75e-18	4.00e-18	1.09e-17	IPA	BAHASA
49	Lilis	2.31e-14	3.34e-16	2.51e-17	2.31e-14	IPA	BAHASA
50	Panca Prasetya	1.92e-14	6.80e-17	2.91e-17	1.92e-14	IPA	BAHASA
51	Muhammad Faisal	3.61e-16	2.08e-16	7.79e-17	3.61e-16	IPA	BAHASA
52	Ahmad Rafli	4.83e-17	2.81e-17	1.96e-17	4.83e-17	IPA	BAHASA
53	Hikmatul Ula	1.00e-17	6.24e-18	2.33e-17	2.33e-17	BAHASA	BAHASA
54	M. Ardyansah	1.39e-16	1.28e-16	2.82e-17	1.39e-16	IPA	BAHASA

Berdasarkan tabel hasil klasifikasi menggunakan metode *Naïve Bayes Gaussian*, diperoleh nilai probabilitas untuk setiap siswa pada masing-masing jurusan, yaitu IPA, IPS, dan Bahasa. Nilai probabilitas tersebut menunjukkan peluang setiap siswa untuk masuk ke masing-masing jurusan berdasarkan nilai mata pelajaran yang dimiliki. Jurusan dengan nilai probabilitas terbesar (MAX) dipilih sebagai hasil prediksi akhir sistem. Hasil perbandingan antara Jurusan Prediksi dan Jurusan Asal menunjukkan bahwa terdapat kesesuaian pada sebagian data, namun juga terdapat beberapa perbedaan dalam hasil klasifikasi. Hal ini menunjukkan bahwa metode *Naïve Bayes* mampu melakukan proses pengelompokan berdasarkan pola data nilai, meskipun terdapat beberapa data yang mengalami perbedaan klasifikasi akibat kemiripan probabilitas antar kelas. Secara keseluruhan, tabel ini menunjukkan bahwa sistem dapat digunakan untuk membantu proses penentuan jurusan secara otomatis dengan membandingkan nilai probabilitas tertinggi dari setiap kelas, sehingga jurusan yang dipilih adalah yang memiliki peluang paling besar berdasarkan perhitungan model.

Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi kinerja model klasifikasi *Naïve Bayes* dengan membandingkan hasil prediksi dengan data aktual.



Gambar 2. Confusion Matrix

Matriks ini menunjukkan jumlah data yang diklasifikasikan dengan benar maupun salah pada masing-masing kelas, yaitu IPA, IPS, dan Bahasa. Berdasarkan hasil perhitungan, diperoleh bahwa jumlah prediksi benar (*True Positive*) berada pada diagonal utama, yaitu IPA sebanyak 13 data, IPS sebanyak 10 data, dan Bahasa sebanyak 2 data. Dengan demikian, total prediksi yang benar adalah sebanyak 25 siswa dari total 54 data yang diuji. Sementara itu, kesalahan klasifikasi terjadi pada beberapa data yang tersebar di luar diagonal utama, yang menunjukkan bahwa model masih mengalami kesalahan dalam membedakan antar kelas, terutama pada kelas yang memiliki karakteristik nilai yang mirip. Secara keseluruhan, *Confusion Matrix* ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang cukup baik, namun masih terdapat beberapa kesalahan prediksi yang dapat memengaruhi tingkat akurasi sistem.

Accuracy

Accuracy adalah ukuran kinerja model klasifikasi yang menunjukkan tingkat ketepatan sistem dalam memprediksi data dengan benar dibandingkan dengan total data yang diuji. Nilai *accuracy* diperoleh dengan membandingkan jumlah prediksi yang benar dengan seluruh data pengujian, kemudian dinyatakan dalam bentuk persentase. Semakin tinggi nilai *accuracy*, semakin baik kemampuan model dalam melakukan klasifikasi. Sebaliknya, semakin rendah nilai *accuracy* menunjukkan bahwa model masih memiliki tingkat kesalahan yang cukup tinggi dalam melakukan prediksi. Dari hasil penelitian diperoleh hasil akurasi sebagaimana berikut:

$$Accuracy = \frac{\text{Jumlah prediksi benar}}{\text{Total data}} \times 100\%$$

$$Accuracy = \frac{20}{54} \times 100\%$$

$$Accuracy = 45,45\%$$

Berdasarkan hasil pengujian terhadap 54 data siswa, diketahui bahwa sebanyak 20 data menghasilkan prediksi yang sesuai antara sistem SPK dan jurusan sebenarnya, sedangkan 34 data lainnya tidak sesuai. Dengan demikian, nilai *accuracy* yang diperoleh sebesar 45,45% menunjukkan bahwa tingkat ketepatan sistem masih tergolong rendah secara umum, karena sistem baru mampu memberikan keputusan yang benar pada kurang dari setengah dari total data uji. Hal ini mengindikasikan bahwa masih terdapat cukup banyak kesalahan klasifikasi dalam proses penentuan jurusan.

Namun demikian, hasil tersebut tetap dapat dikatakan cukup baik jika digunakan dalam konteks Sistem Pendukung Keputusan (SPK), karena sistem ini tidak sepenuhnya menggantikan peran pengambil keputusan, melainkan hanya sebagai alat bantu untuk memberikan rekomendasi. Dengan adanya sistem ini, pihak sekolah tetap dapat mempertimbangkan hasil perhitungan berbasis data secara objektif sebagai dasar tambahan dalam proses penentuan jurusan siswa.

Precision, Recall, dan F1-Score

Berdasarkan hasil evaluasi menggunakan *precision*, *recall*, dan *F1-score*, dapat diketahui kinerja model *Naïve Bayes* dalam mengklasifikasikan data ke dalam tiga kelas, yaitu IPA, IPS, dan Bahasa. *Precision* menunjukkan tingkat ketepatan prediksi suatu kelas, yaitu seberapa banyak data yang diprediksi sebagai suatu kelas benar-benar sesuai dengan kelas sebenarnya. *Recall* menunjukkan kemampuan model dalam menemukan seluruh data pada suatu kelas, sedangkan *F1-score* merupakan nilai rata-rata harmonik antara *precision* dan *recall* yang menggambarkan keseimbangan kinerja model.

Tabel 9. Evaluasi Kinerja Model Klasifikasi Naïve Bayes Menggunakan Confusion Matrix (Precision, Recall, dan F1-Score)

Kelas	TP	FP	FN	Precision	Recall	F1-Score
IPA	13	16	8	44,83%	61,90%	52,00%
IPS	10	9	9	52,63%	52,63%	52,63%
BAHASA	2	4	12	33,33%	14,29%	20,00%

Pada kelas IPA diperoleh *precision* sebesar 44,83%, *recall* 61,90%, dan *F1-score* 52,00%. Hal ini menunjukkan bahwa model cukup baik dalam menangkap data IPA, namun masih terdapat cukup banyak data yang salah diprediksi sebagai IPA. Pada kelas IPS, nilai *precision* dan *recall* berada pada angka yang seimbang, yaitu 52,63% dengan *F1-score* 52,63%, yang menunjukkan bahwa kinerja model pada kelas ini relatif stabil meskipun belum optimal. Sementara itu, pada kelas Bahasa diperoleh *precision* 33,33%, *recall* 14,29%, dan *F1-score* 20,00%, yang menunjukkan bahwa kemampuan model dalam mengidentifikasi kelas Bahasa masih rendah.

Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa kinerja model masih belum merata pada setiap kelas, dengan kelas IPS memiliki performa paling seimbang dibandingkan dengan IPA dan Bahasa. Hal ini mengindikasikan bahwa masih terdapat ketidakseimbangan atau tumpang tindih pola data antar kelas yang memengaruhi hasil klasifikasi.

Macro Average

Macro Average adalah metode perhitungan rata-rata pada evaluasi model klasifikasi yang menghitung nilai metrik (seperti *precision*, *recall*, dan *F1-score*) dengan cara merata-ratakan hasil dari setiap kelas secara proporsional (tanpa mempertimbangkan jumlah data pada tiap kelas). Artinya, setiap kelas (misalnya IPA, IPS, dan Bahasa) memiliki bobot yang sama dalam perhitungan, sehingga tidak ada kelas yang lebih dominan meskipun jumlah datanya lebih banyak atau lebih sedikit. Dengan demikian, *Macro Average* digunakan untuk melihat kinerja model secara menyeluruh dan adil pada setiap kelas, sehingga dapat diketahui apakah model mampu bekerja dengan baik pada semua kelas secara seimbang atau tidak. Dari penelitian ini didapatkan hasil sebagaimana berikut:

Tabel 10. Hasil Evaluasi Model Klasifikasi *Naïve Bayes* Berdasarkan Nilai *Macro Average*

Parameter	Nilai
Accuracy	45,45%
Macro Precision	43,60%
Macro Recall	42,94%
Macro F1-Score	41,54%

Berdasarkan hasil perhitungan evaluasi model menggunakan metode *Macro Average*, diperoleh nilai *accuracy* sebesar 45,45%, *macro precision* sebesar 43,60%, *macro recall* sebesar 42,94%, dan *macro F1-score* sebesar 41,54%. *Macro average* merupakan metode perhitungan rata-rata yang memberikan bobot yang sama pada setiap kelas, tanpa mempertimbangkan jumlah data pada masing-masing kelas. Dengan demikian, setiap kelas (IPA, IPS, dan Bahasa) memiliki kontribusi yang sama dalam perhitungan, sehingga hasil evaluasi tidak dipengaruhi oleh dominasi jumlah data pada salah satu kelas. Hasil tersebut menunjukkan bahwa secara keseluruhan kinerja model masih tergolong rendah hingga sedang, karena seluruh nilai evaluasi berada di bawah 50%. Nilai *macro precision* dan *recall* yang relatif rendah mengindikasikan bahwa model belum mampu mengklasifikasikan setiap kelas secara optimal. Hal ini juga diperkuat oleh nilai *macro F1-score* sebesar 41,54% yang menunjukkan bahwa keseimbangan antara *precision* dan *recall* masih belum optimal. Dengan demikian, dapat disimpulkan bahwa performa model *Naïve Bayes* pada penelitian ini masih belum optimal dalam menghasilkan klasifikasi yang stabil pada seluruh kelas, sehingga masih diperlukan pengembangan lebih lanjut untuk meningkatkan akurasi dan keseimbangan prediksi.

KESIMPULAN

Berdasarkan hasil penelitian mengenai Sistem Pendukung Keputusan (SPK) penentuan jurusan siswa menggunakan metode *Naïve Bayes* berbasis web, maka dapat disimpulkan sebagai berikut:

1. Sistem Pendukung Keputusan berbasis web berhasil dikembangkan untuk membantu proses penentuan jurusan siswa, yaitu IPA, IPS, dan Bahasa, dengan menggunakan metode *Naïve Bayes*. Sistem mampu mengolah data nilai akademik siswa, menghitung probabilitas untuk setiap kelas, dan menghasilkan rekomendasi jurusan berdasarkan nilai probabilitas tertinggi.
2. Penerapan metode *Naïve Bayes* pada penelitian ini dilakukan melalui tahapan penentuan *data training*, perhitungan probabilitas *prior*, perhitungan probabilitas *Gaussian* untuk setiap atribut, penggabungan seluruh probabilitas, hingga penentuan kelas dengan nilai probabilitas terbesar sebagai hasil rekomendasi jurusan.
3. Berdasarkan hasil pengujian terhadap 54 data siswa, diperoleh nilai *accuracy* sebesar 45,45%, *macro precision* sebesar 43,60%, *macro recall* sebesar 42,94%, dan *macro F1-score* sebesar 41,54%. Hasil tersebut menunjukkan bahwa kemampuan model dalam memberikan rekomendasi jurusan masih belum optimal, karena masih terdapat perbedaan antara hasil rekomendasi sistem dan jurusan aktual.
4. Meskipun tingkat akurasi yang diperoleh masih relatif rendah, sistem tetap dapat dimanfaatkan sebagai Sistem Pendukung Keputusan (SPK) karena berfungsi memberikan rekomendasi berdasarkan pengolahan data secara objektif. Keputusan akhir mengenai penentuan jurusan tetap berada pada pihak sekolah dengan mempertimbangkan faktor lain, seperti minat, bakat, hasil psikotes, serta pertimbangan dari guru dan orang tua. Dengan demikian, sistem ini dapat menjadi alat bantu yang mendukung proses pengambilan keputusan agar lebih sistematis dan berbasis data.

Berdasarkan hasil penelitian yang telah dilakukan, disarankan agar penelitian selanjutnya menggunakan jumlah *data training* yang lebih banyak dan lebih beragam sehingga model *Naïve Bayes* dapat mempelajari pola data dengan lebih baik dan menghasilkan tingkat akurasi yang lebih tinggi. Selain itu, proses penentuan jurusan tidak hanya didasarkan pada nilai akademik,

tetapi juga mempertimbangkan kriteria lain, seperti hasil tes minat dan bakat, hasil psikotes, prestasi siswa, serta rekomendasi guru bimbingan dan konseling (BK), sehingga rekomendasi yang dihasilkan menjadi lebih komprehensif. Penelitian selanjutnya juga dapat membandingkan metode *Naïve Bayes* dengan metode klasifikasi lainnya untuk memperoleh metode yang memiliki performa lebih baik dalam kasus penentuan jurusan. Dari sisi pengembangan sistem, aplikasi berbasis web dapat disempurnakan dengan penambahan fitur pengelolaan data, pelaporan hasil rekomendasi, visualisasi hasil evaluasi, serta fasilitas ekspor data. Meskipun demikian, sistem yang dikembangkan sebaiknya tetap digunakan sebagai Sistem Pendukung Keputusan (SPK), sehingga hasil rekomendasi yang diberikan menjadi bahan pertimbangan bagi pihak sekolah dan tidak menggantikan keputusan akhir dalam penentuan jurusan siswa.

REFERENSI

- Alim, S. (2021). *Implementasi Orange data mining untuk klasifikasi kelulusan mahasiswa dengan model k-nearest neighbor, decision tree, serta Naïve Bayes*. 6(2), 133–144.
- Alter, S. (2002). *Information systems: Foundation of e-business* (4th ed.). Prentice Hall.
- Annur, H. (2018). *Klasifikasi masyarakat miskin menggunakan metode Naïve Bayes*. 10, 160–165.
- Budianto, D. T., & Irawan, A. (2017). Sistem pendukung keputusan untuk menentukan jurusan siswa baru di SMK Negeri 1 Kragilan menggunakan metode Naïve Bayes. *ProTekInfo (Pengembangan Riset dan Observasi Teknik Informatika)*, 1, 62–67. <https://doi.org/10.30656/protekinfo.v1i0.34>
- Djaali. (2012). *Metode penelitian kuantitatif*. Bumi Aksara.
- Fabiana Meijon Fadul. (2019). *No title*. 13–14.
- Gorunescu, F. (2011). *Data mining: Concepts, models and techniques*. Springer.
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Kusrini. (2007). *Konsep dan aplikasi sistem pendukung keputusan*. Andi.
- Manu, G. A., Putra, Y. H., & Afrizal, Y. (n.d.). *Sistem pendukung keputusan untuk menentukan pilihan jurusan mahasiswa menggunakan metode Naïve Bayes dan analytical hierarchy process (AHP): Studi kasus pada Akademi Teknik Kupang*. <https://doi.org/10.13140/RG.2.2.10127.18088>
- Prasetyo, E. (2012). *Data mining: Konsep dan aplikasi menggunakan MATLAB*. Andi.
- Pratama, Y. P. (2021). *Sistem pendukung keputusan rekomendasi jurusan dengan metode Naïve Bayes*. 1(1), 18–23.
- Pressman, R. S., & Maxim, B. R. (2015). *Software engineering: A practitioner's approach* (8th ed.). McGraw-Hill Education.
- Putri, H., Purnamasari, A. I., Dikananda, A. R., Nurdiawan, O., & Anwar, S. (2021). Penerima manfaat bantuan non tunai Kartu Keluarga Sejahtera menggunakan metode Naïve Bayes dan KNN. *Building of Informatics, Technology and Science (BITS)*, 3(3), 331–337. <https://doi.org/10.47065/bits.v3i3.1093>
- Studi, P., Informasi, S., Putra, U. N., Sukabumi, C. K., & Barat, J. (2021). Sistem pendukung keputusan kenaikan jabatan menggunakan metode algoritma Naïve Bayes classifier. *Jurnal Ilmiah Teknologi Informasi Terapan (JATIT)*, 11(1), 66–80. <https://doi.org/10.34010/jati.v11i1>
- Turban, E., Sharda, R., Delen, D., & King, D. (2011). *Decision support and business intelligence systems* (10th ed.). Pearson