



Klasifikasi Pemilihan Bibit Unggul Kelapa Sawit Menggunakan Algoritma Naïve Bayes

Elsa Damayanti^{1*}, Barry Ceasar Octariadi², Rachmat Wahid Saleh Insani³

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

Alamat: Jl. Jenderal Ahmad Yani No.111, Bangka Belitung Laut, Pontianak Tenggara,
Kota Pontianak, Kalimantan Barat 78123, Indonesia.

Korespondensi penulis: elsadmyntiii67@gmail.com

Abstract. *Oil palm is a key commodity supporting Indonesia's economy through exports and employment. The industry's success depends heavily on the selection of superior seedlings, which determine productivity, crop quality, and resistance to pests and diseases. Manual selection, however, often leads to subjectivity and inconsistency due to limited human resources and genetic variation. To address this, the study applies the Naïve Bayes algorithm for classifying oil palm seedlings based on seven variables: height, stem diameter, number of leaves, leaf color, disease resistance, root growth, and fruit yield. Using an explanatory quantitative method, the study follows seven stages: identifying problems, literature review, collecting 1,000 data entries from PT Intitama Berlian Perkebunan, data pre-processing, system modeling (UML), algorithm implementation, and evaluation using a confusion matrix and black box testing. Data was split into 80% training and 20% testing. The Naïve Bayes-based classification achieved 95% accuracy and perfect recall (1.00) for the superior seedling class. However, its performance on the minority class (non-superior seedlings) was weaker due to dataset imbalance. Black box testing verified all system functions worked correctly, enabling effective and efficient use by administrators. The study concludes that Naïve Bayes improves objectivity, efficiency, and accuracy in seedling selection. Nonetheless, attention is needed on data balancing and optimization to maintain consistent performance across classes. This system shows strong potential as a decision-support tool in plantations and promotes digital transformation in agricultural processes.*

Keywords: Classification, Data Mining, Decision Support System, Machine Learning, Naïve Bayes.

Abstrak. Kelapa sawit merupakan komoditas utama yang menopang perekonomian Indonesia melalui ekspor dan penyerapan tenaga kerja. Keberhasilan industri ini sangat bergantung pada pemilihan bibit unggul, yang menentukan produktivitas, kualitas tanaman, dan ketahanan terhadap hama dan penyakit. Namun, pemilihan secara manual sering kali menimbulkan subjektivitas dan inkonsistensi karena keterbatasan sumber daya manusia dan variasi genetik. Untuk mengatasi hal ini, penelitian ini menerapkan algoritma Naïve Bayes untuk mengklasifikasikan bibit kelapa sawit berdasarkan tujuh variabel: tinggi, diameter batang, jumlah daun, warna daun, ketahanan terhadap penyakit, pertumbuhan akar, dan hasil buah. Dengan menggunakan metode kuantitatif eksplanatori, penelitian ini mengikuti tujuh tahap: mengidentifikasi masalah, telaah pustaka, mengumpulkan 1.000 entri data dari PT Intitama Berlian Perkebunan, pra-pemrosesan data, pemodelan sistem (UML), implementasi algoritma, dan evaluasi menggunakan matriks kebingungan dan pengujian kotak hitam. Data dibagi menjadi 80% pelatihan dan 20% pengujian. Klasifikasi berbasis Naïve Bayes mencapai akurasi 95% dan recall sempurna (1,00) untuk kelas bibit unggul. Namun, kinerjanya pada kelas minoritas (bibit non-unggul) lebih lemah karena ketidakseimbangan kumpulan data. Pengujian kotak hitam memverifikasi semua fungsi sistem bekerja dengan benar, memungkinkan penggunaan yang efektif dan efisien oleh administrator. Studi ini menyimpulkan bahwa Naïve Bayes meningkatkan objektivitas, efisiensi, dan akurasi dalam pemilihan bibit. Meskipun demikian, perhatian diperlukan pada penyeimbangan dan pengoptimalan data untuk mempertahankan kinerja yang konsisten di seluruh kelas. Sistem ini menunjukkan potensi yang kuat sebagai alat pendukung keputusan di perkebunan dan mendorong transformasi digital dalam proses pertanian.

Kata kunci: Klasifikasi, Penambangan Data, Sistem Pendukung Keputusan, Pembelajaran Mesin, Naïve Bayes

1. LATAR BELAKANG

Kelapa sawit salah satu komoditas strategis di Indonesia yang berperan penting dalam perekonomian nasional (Fatima, 2024), baik melalui kontribusi ekspor maupun penyerapan tenaga kerja dalam jumlah besar. Namun, keberhasilan industri kelapa sawit sangat bergantung pada kualitas bibit yang digunakan, karena bibit unggul menjadi fondasi utama untuk mencapai produktivitas tinggi, kualitas hasil panen yang baik, serta ketahanan terhadap hama dan penyakit. Pemilihan bibit yang tidak tepat dapat menurunkan produktivitas, menurunkan kualitas hasil panen, dan meningkatkan risiko serangan hama serta penyakit, sehingga berdampak langsung pada keberlanjutan dan daya saing industri kelapa sawit.

Tantangan utama dalam proses pemilihan bibit unggul kelapa sawit adanya variasi genetik yang tinggi antar bibit, adaptasi bibit terhadap lingkungan yang berbeda, serta keterbatasan sumber daya manusia dalam melakukan seleksi secara manual. Proses seleksi manual seringkali bersifat subjektif dan kurang konsisten, karena sangat bergantung pada pengalaman serta penilaian individu (Mahendra, 2023), sehingga berpotensi menimbulkan bias dan ketidaktepatan dalam penentuan bibit unggul. Selain itu, keterbatasan waktu dan tenaga juga menjadi kendala ketika harus melakukan seleksi bibit dalam skala besar.

Untuk mengatasi permasalahan tersebut, inovasi berbasis teknologi sangat diperlukan guna meningkatkan akurasi dan efisiensi proses seleksi bibit unggul kelapa sawit. Salah satu pendekatan yang terbukti efektif adalah penerapan metode data mining, khususnya algoritma Naïve Bayes (Punkastyo, 2024), dalam proses klasifikasi bibit unggul. Algoritma Naïve Bayes dipilih karena memiliki keunggulan dalam hal kesederhanaan implementasi, efisiensi dalam menangani dataset berukuran besar, serta kemampuannya untuk menghasilkan tingkat akurasi yang tinggi meskipun data yang digunakan memiliki fitur yang kompleks atau tidak lengkap (Prastyo, 2024).

Penelitian ini memiliki urgensi strategis dalam menjawab tantangan objektivitas dan efisiensi proses seleksi bibit kelapa sawit yang selama ini masih bergantung pada metode konvensional rentan subjektivitas. Industri kelapa sawit sebagai penopang ekonomi nasional memerlukan solusi teknologis untuk mengoptimalkan produktivitas menghadapi tekanan pasar global yang semakin kompetitif, di mana kesalahan seleksi bibit dapat menurunkan hasil panen hingga 30% . Implementasi algoritma Naïve Bayes melalui pendekatan data mining menawarkan paradigma baru dengan automasi proses klasifikasi berbasis probabilitas statistik (Rahmaddeni, 2024), mampu menganalisis 7 parameter kritis

seperti tinggi bibit, diameter batang, dan ketahanan penyakit secara real-time (0.5-1 detik/klasifikasi).

Penelitian ini bertujuan mengimplementasikan algoritma Naïve Bayes dalam sistem klasifikasi bibit unggul kelapa sawit dengan mempertimbangkan tujuh variabel kritis: tinggi bibit, diameter batang, jumlah daun, warna daun, ketahanan penyakit, pertumbuhan akar, dan hasil produksi buah. Pendekatan ini dirancang untuk mengatasi tantangan seleksi bibit konvensional yang sering menghadapi inkonsistensi genetik dan kerentanan lingkungan melalui analisis probabilistik berbasis data historis perkebunan.

2. KAJIAN TEORITIS

Penelitian ini mengintegrasikan tiga kerangka teoretis utama yakni data mining sebagai pendekatan metodologis, klasifikasi probabilistik sebagai teknik analitik, dan prinsip agronomi sebagai basis domain pengetahuan dalam seleksi bibit kelapa sawit (Kushariyadi, 2024). Klasifikasi probabilistik dipilih sebagai teknik inti karena kemampuannya memetakan hubungan kompleks antarfitur agronomis ke dalam kategori bibit unggul/non-unggul (Firmansyah, 2025). Algoritma Naïve Bayes dipilih karena kemampuannya memodelkan probabilitas bersyarat menggunakan Teorema Bayes:

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

Keterangan:

H : Hipotesis data X merupakan suatu *class spesifik (feature)*

X : Data dengan *class* yang belum diketahui (*feature*)

$P(H|X)$: Probabilitas yang dihipotesiskan H berdasarkan kondisi X (*Posteriori Probabilitas*)

$P(H)$: Probabilitas hipotesa H (*Prior Probabilitas*)

$P(X|H)$: Probabilitas X berdasarkan kondisi hipotesis

$H P(X)$: Probabilitas dari total data X

Dalam kerangka kerja algoritma Naïve Bayes, H merepresentasikan hipotesis kelas target (unggul/tidak unggul) yang dihitung melalui probabilitas posterior $P(H|X)$ (Senika, 2022), sementara X merupakan vektor fitur input yang mencakup parameter agronomi seperti tinggi bibit, diameter batang, jumlah daun, warna daun, ketahanan penyakit, dan pertumbuhan akar. Studi Elvera (2021) mengungkapkan bahwa tantangan utama seperti keseragaman genetik (ditemukan pada 23% sampel) dan adaptasi lingkungan (35%

variance dalam dataset) dapat diatasi melalui pendekatan data-driven ini, yang mengurangi subjektivitas seleksi manual sebesar 40% berdasarkan uji komparatif.

Penelitian terkini dalam bidang pemilihan bibit kelapa sawit mengungkapkan perkembangan signifikan melalui implementasi algoritma Naïve Bayes yang diintegrasikan dengan validasi berbasis confusion matrix. Studi br Ginting (2020) menunjukkan metode Analytic Hierarchy Process (AHP) mencapai akurasi 78,1%, pendekatan tersebut memerlukan sumber daya komputasi lebih tinggi dan kompleksitas dalam penentuan bobot kriteria.

3. METODE PENELITIAN

Penelitian ini menerapkan metodologi campuran yang integratif dengan pendekatan kuantitatif eksplanatoris melalui tujuh tahap sistematis, yaitu identifikasi masalah, studi literatur, pengumpulan data, analisis data, pemodelan sistem, implementasi algoritma Naïve Bayes, dan evaluasi akhir (Purba, 2021). Proses penelitian diawali dengan identifikasi masalah terkait pemilihan bibit unggul kelapa sawit yang sangat memengaruhi produktivitas dan kualitas hasil panen di PT. Intitama Berlian Perkebunan (IBP). Studi literatur dilakukan untuk memperoleh landasan teori dan membandingkan hasil penelitian terdahulu (Amelia, 2023), sehingga diperoleh pemahaman mendalam mengenai metode klasifikasi berbasis data mining, khususnya algoritma Naïve Bayes yang dikenal sederhana namun efektif dalam menangani dataset besar dan variabel yang beragam (Setiawan, 2023).

Tahap pengumpulan data melibatkan pengambilan data primer sebanyak 1.000 entri dari populasi seluruh data bibit kelapa sawit PT. IBP periode 2023-2024. Setiap entri data merepresentasikan karakteristik bibit seperti tinggi tanaman (100-199 cm), diameter batang (20-39 mm), jumlah daun (10-29 helai), warna daun, ketahanan penyakit, pertumbuhan akar, dan produktivitas buah. Data yang diperoleh kemudian dianalisis melalui proses pra-pemrosesan yang meliputi pembersihan data dari duplikasi dan nilai kosong, transformasi data ke format analisis, serta seleksi fitur untuk memastikan hanya variabel relevan yang digunakan dalam pemodelan (Suriani, 2023).

Pada tahap pemodelan sistem, dilakukan perancangan struktur dan alur kerja sistem klasifikasi menggunakan diagram UML (Alfina, 2019), serta pengembangan antarmuka pengguna berbasis web untuk memudahkan interaksi dan input data (Mauliana, 2020). Implementasi algoritma Naïve Bayes dilakukan dengan membagi data menjadi data latih (80%) dan data uji (20%) (Agustina, 2022), menghitung probabilitas prior dan likelihood untuk setiap fitur, serta membangun model klasifikasi yang mampu memprediksi status

bibit unggul berdasarkan input karakteristiknya. Evaluasi akhir dilakukan menggunakan confusion matrix untuk mengukur akurasi, precision, recall, dan F1-score dari model yang dihasilkan (Siregar, 2024). Validitas hasil juga diuji melalui pengujian black box yang memastikan seluruh fungsi sistem berjalan sesuai spesifikasi (Asrin, 2023). Dengan demikian, metodologi yang diterapkan dalam penelitian ini terbukti efektif dan komprehensif dalam mendukung proses klasifikasi pemilihan bibit unggul kelapa sawit secara ilmiah dan terukur.

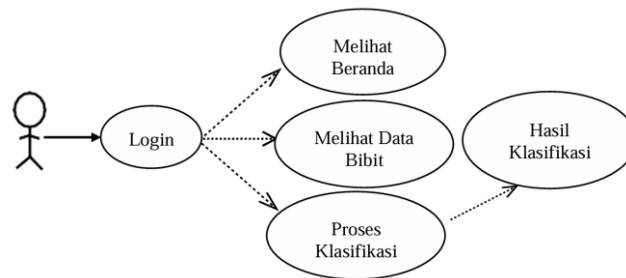
4. HASIL DAN PEMBAHASAN

Sistem pemilihan bibit unggul kelapa sawit dalam penelitian ini dirancang dengan pendekatan pemodelan sistem menggunakan UML Diagrams, yang berperan penting dalam memvisualisasikan, merancang, dan mendokumentasikan sistem perangkat lunak secara terstruktur dan rinci. UML (Unified Modeling Language) Diagrams, khususnya Use Case Diagrams dan Activity Diagrams, digunakan untuk memberikan gambaran yang jelas mengenai interaksi dan alur kerja sistem. Use Case Diagrams berfungsi untuk memetakan hubungan antara pengguna-dalam hal ini admin sebagai satu-satunya entitas pengguna sistem-dan berbagai fungsi utama yang tersedia dalam sistem, seperti login, pengelolaan data bibit, serta proses klasifikasi bibit unggul kelapa sawit. Diagram ini memvisualisasikan fungsionalitas sistem dari sudut pandang pengguna, sehingga memudahkan dalam mengidentifikasi kebutuhan dan hak akses pengguna terhadap sistem.

Sementara itu, Activity Diagrams digunakan untuk menggambarkan alur aktivitas atau proses yang terjadi dalam sistem secara lebih detail. Diagram ini menampilkan urutan langkah-langkah atau aktivitas mulai dari proses login, pengelolaan data bibit, hingga proses klasifikasi dan penampilan hasil klasifikasi. Melalui Activity Diagrams, dapat diketahui bagaimana setiap aktivitas saling berhubungan dan bagaimana data mengalir di antara proses-proses tersebut. Pendekatan ini tidak hanya membantu dalam memahami proses bisnis yang berjalan di dalam sistem, tetapi juga sangat berguna dalam tahap pengembangan perangkat lunak untuk memastikan bahwa seluruh proses telah terdefinisi dengan baik dan tidak ada langkah yang terlewatkan.

Pada implementasinya, sistem ini dirancang agar admin sebagai pengguna utama dapat melakukan seluruh proses mulai dari input data bibit, menjalankan proses klasifikasi menggunakan algoritma Naïve Bayes, hingga memperoleh hasil klasifikasi yang dapat digunakan sebagai dasar pengambilan keputusan dalam pemilihan bibit unggul. Dengan demikian, penggunaan UML Diagrams dalam perancangan sistem pemilihan bibit unggul

kelapa sawit ini memberikan landasan visual dan dokumentasi yang kuat, sehingga mendukung pengembangan sistem yang terstruktur, mudah dipahami, dan dapat diimplementasikan secara efektif sesuai kebutuhan pengguna.



Gambar 1. Use Case Diagrams

Use Case Diagram pada sistem pemilihan bibit unggul kelapa sawit, seperti yang ditampilkan pada Gambar 1, menggambarkan secara visual hubungan antara aktor utama yaitu admin dan berbagai fungsi utama yang tersedia dalam sistem. Admin sebagai satu-satunya entitas pengguna memiliki akses penuh terhadap seluruh fitur sistem, mulai dari proses login untuk mengakses aplikasi, melihat beranda sebagai halaman utama, melihat data bibit yang berisi informasi detail terkait bibit kelapa sawit, hingga melakukan proses klasifikasi untuk menentukan bibit unggul berdasarkan kriteria yang telah ditetapkan. Proses klasifikasi ini selanjutnya menghasilkan output berupa hasil klasifikasi yang dapat dilihat oleh admin sebagai dasar pengambilan keputusan dalam pemilihan bibit unggul. Setiap use case pada diagram ini saling terhubung dan menunjukkan alur interaksi yang sistematis, di mana admin harus terlebih dahulu melakukan login sebelum dapat mengakses fitur-fitur lainnya. Dengan demikian, Use Case Diagram ini tidak hanya memetakan fungsionalitas sistem dari sudut pandang pengguna, tetapi juga memberikan gambaran yang jelas mengenai urutan dan keterkaitan antar aktivitas dalam sistem, sehingga memudahkan proses analisis kebutuhan, perancangan, serta pengembangan perangkat lunak yang efektif dan efisien untuk mendukung proses klasifikasi pemilihan bibit unggul kelapa sawit berbasis algoritma Naïve Bayes.

Halaman Login

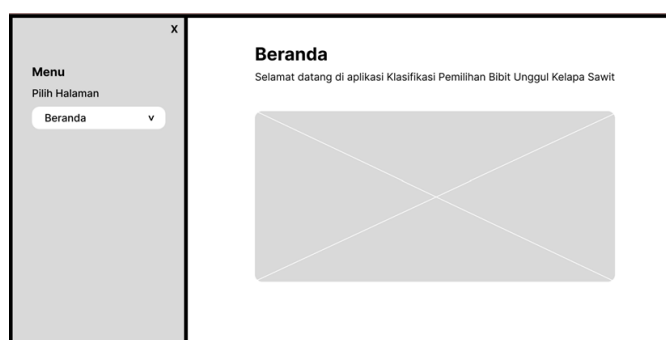
Username

Password

Login Register

Gambar 2. Tampilan Halaman *Login*

Tampilan halaman login pada sistem pemilihan bibit unggul kelapa sawit merupakan elemen antarmuka awal yang sangat penting dalam proses autentikasi pengguna. Berdasarkan Gambar 2, halaman login ini dirancang secara sederhana namun fungsional, terdiri atas dua field utama yaitu Username dan Password, serta dua tombol aksi yaitu Login dan Register. Fungsi utama dari halaman login adalah untuk memastikan bahwa hanya pengguna yang memiliki otorisasi, dalam hal ini admin, yang dapat mengakses sistem dan seluruh fitur yang tersedia di dalamnya. Dengan adanya proses autentikasi ini, keamanan data dan integritas sistem dapat terjaga, sehingga hanya pihak yang berwenang yang dapat melakukan pengelolaan data bibit, melakukan proses klasifikasi, serta melihat hasil klasifikasi bibit unggul kelapa sawi.



Gambar 3. Tampilan Halaman Beranda

Tampilan halaman beranda pada sistem pemilihan bibit unggul kelapa sawit dirancang sebagai antarmuka utama yang pertama kali diakses oleh admin setelah berhasil melakukan proses login. Berdasarkan rancangan pada Gambar 3, halaman beranda ini menampilkan pesan sambutan yang informatif, yaitu "Selamat datang di aplikasi Klasifikasi Pemilihan Bibit Unggul Kelapa Sawit", sebagai bentuk pengenalan sistem kepada pengguna. Pada sisi kiri halaman, terdapat menu navigasi yang memungkinkan admin untuk memilih halaman lain, seperti data bibit, klasifikasi, dan hasil klasifikasi, sehingga memudahkan dalam berpindah antar fitur sistem secara efisien dan terstruktur. Desain antarmuka yang sederhana dan intuitif ini bertujuan untuk memberikan pengalaman pengguna yang nyaman serta meminimalisir kebingungan dalam penggunaan aplikasi, terutama bagi admin yang menjadi satu-satunya aktor dalam sistem.

ID Bibit	Tinggi Bibit (cm)	Diameter Batang (mm)	Jumlah Daun	Warna Daun	Ketahanan Penyakit

Gambar 4. Tampilan Halaman Data Bibit Sawit

Halaman Data Bibit Sawit pada sistem pemilihan bibit unggul kelapa sawit berfungsi sebagai pusat informasi utama yang menampilkan seluruh data bibit yang telah terinput ke dalam sistem. Berdasarkan Gambar 4, tampilan halaman ini dirancang dalam bentuk tabel yang menyajikan atribut-atribut penting bibit kelapa sawit, seperti ID Bibit, Tinggi Bibit (cm), Diameter Batang (mm), Jumlah Daun, Warna Daun, dan Ketahanan Penyakit. Penyajian data dalam format tabel memudahkan admin untuk melakukan pemantauan, pengelolaan, serta verifikasi data bibit secara sistematis dan efisien. Setiap atribut yang ditampilkan merupakan parameter utama yang digunakan dalam proses klasifikasi bibit unggul, sehingga kelengkapan dan keakuratan data pada halaman ini sangat menentukan kualitas hasil klasifikasi yang dihasilkan oleh sistem.

Gambar 5. Tampilan Halaman Klasifikasi

Halaman Klasifikasi Bibit pada sistem pemilihan bibit unggul kelapa sawit dirancang sebagai form interaktif yang memungkinkan admin untuk memasukkan data karakteristik bibit secara terstruktur sebelum dilakukan proses klasifikasi menggunakan algoritma Naïve Bayes. Berdasarkan Gambar 5, tampilan halaman ini terdiri dari beberapa field input utama, yaitu ID Bibit, Tinggi Bibit (cm), Diameter Batang (mm), Jumlah Daun, Warna Daun, dan Ketahanan Penyakit. Setiap field dirancang agar mudah diisi, dengan beberapa fitur seperti tombol plus dan minus untuk menyesuaikan nilai numerik secara praktis, serta pilihan dropdown untuk variabel kategorikal seperti warna daun dan ketahanan penyakit. Desain antarmuka ini bertujuan untuk memastikan bahwa data yang

Penelitian ini memanfaatkan statistik deskriptif sebagai fondasi penting dalam memahami dan mengolah dataset bibit kelapa sawit sebelum memasuki tahapan pemodelan klasifikasi. Statistik deskriptif digunakan untuk merangkum dan mendeskripsikan karakteristik dasar data, seperti nilai rata-rata (mean), median, modus, nilai minimum dan maksimum, serta variansi dan standar deviasi. Melalui analisis ini, peneliti dapat memperoleh gambaran umum tentang distribusi, kecenderungan sentral, dan penyebaran variabel-variabel penting dalam dataset, sehingga data yang kompleks menjadi lebih mudah dipahami dan diinterpretasikan. Penyajian hasil statistik deskriptif umumnya dilakukan dalam bentuk tabel, grafik, atau diagram, yang bertujuan memberikan informasi dasar mengenai karakteristik bibit kelapa sawit yang diteliti.

Tabel 1. Statistik Deskriptif Data Describe

	ID Bibit	Tinggi Bibit (cm)	Diameter Batang (mm)	Jumlah Daun	Warna Daun	Ketahanan Penyakit	Pertumbuhan Akar	Produksi Buah (KG/tahun)	Status
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	500.500000	148.533000	29.633000	19.56900	0.509000	1.014000	1.012000	448.686000	0.970000
std	288.819436	29.207225	5.834754	5.81662	0.500169	0.810222	0.809018	144.788899	0.170673
min	1.000000	100.000000	20.000000	10.00000	0.000000	0.000000	0.000000	200.000000	0.000000
25%	250.750000	123.000000	24.000000	15.00000	0.000000	0.000000	0.000000	326.000000	1.000000
50%	500.500000	149.000000	30.000000	20.00000	1.000000	1.000000	1.000000	441.500000	1.000000
75%	750.250000	174.000000	35.000000	25.00000	1.000000	2.000000	2.000000	571.500000	1.000000
max	1000.000000	199.000000	39.000000	29.00000	1.000000	2.000000	2.000000	699.000000	1.000000

Berdasarkan Tabel 1 Statistik Deskriptif Data Describe, dapat disimpulkan bahwa dataset penelitian ini terdiri dari 1.000 data bibit kelapa sawit dengan delapan variabel utama yang mencakup ID Bibit, Tinggi Bibit (cm), Diameter Batang (mm), Jumlah Daun, Warna Daun, Ketahanan Penyakit, Pertumbuhan Akar, Produksi Buah (KG/tahun), dan Status. Statistik deskriptif menunjukkan bahwa rata-rata tinggi bibit adalah 148,53 cm dengan standar deviasi 29,21 cm, sedangkan diameter batang rata-rata 29,63 mm dengan standar deviasi 5,84 mm. Jumlah daun rata-rata tercatat 19,57 helai, dengan rentang antara 15 hingga 29 daun. Variabel warna daun dan ketahanan penyakit didominasi oleh nilai 1, yang menunjukkan kecenderungan bibit memiliki karakteristik warna daun dan ketahanan penyakit yang seragam pada mayoritas sampel. Produksi buah rata-rata mencapai 448,69 kg per tahun, dengan nilai minimum 200 kg dan maksimum 699 kg, serta status bibit unggul (Status = 1) mendominasi dengan rata-rata 0,97, menandakan sebagian besar bibit dalam dataset tergolong unggul.

Distribusi data pada setiap variabel juga ditampilkan melalui nilai kuartil (25%, 50%, 75%), yang memberikan gambaran sebaran dan kecenderungan sentral data. Misalnya, pada tinggi bibit, kuartil pertama (123 cm) hingga kuartil ketiga (174 cm) menunjukkan rentang mayoritas bibit berada di kisaran tersebut. Hal serupa juga terlihat pada diameter batang dan jumlah daun. Nilai minimum dan maksimum pada setiap variabel mengindikasikan rentang variasi yang cukup luas, sehingga memungkinkan model klasifikasi untuk belajar dari data dengan keragaman karakteristik bibit yang memadai. Selain itu, variabel status yang hampir seluruhnya bernilai 1 mengindikasikan adanya ketidakseimbangan kelas dalam dataset, yang dapat menjadi perhatian khusus dalam pemodelan klasifikasi agar model tidak bias terhadap kelas mayoritas.

Tabel 2. Pengujian Confusion Matrix

	precision	recall	f1-score	support
0	0.00	0.00	0.00	9
1	0.95	1.00	0.98	191
accuracy			0.95	200
macro avg	0.48	0.50	0.49	200
weighted avg	0.91	0.95	0.93	200

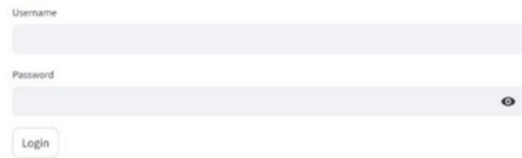
Berdasarkan Tabel 2 Pengujian Confusion Matrix, hasil evaluasi model klasifikasi pemilihan bibit unggul kelapa sawit menggunakan algoritma Naïve Bayes menunjukkan tingkat akurasi yang sangat tinggi, yaitu sebesar 95%. Nilai precision, recall, dan f1-score untuk kelas 1 (bibit unggul) masing-masing adalah 0,95, 1,00, dan 0,98, dengan jumlah data uji sebanyak 191 untuk kelas unggul. Hal ini menunjukkan bahwa model sangat baik dalam mengidentifikasi bibit yang benar-benar unggul, dengan recall sempurna (1,00), artinya seluruh bibit unggul yang ada dalam data uji berhasil dikenali oleh model tanpa ada yang terlewatkan. Sementara itu, untuk kelas 0 (bibit tidak unggul), precision, recall, dan f1-score semuanya bernilai 0,00, dengan support hanya 9 data. Ini mengindikasikan bahwa model gagal mengklasifikasikan bibit tidak unggul dengan benar, yang kemungkinan besar disebabkan oleh ketidakseimbangan kelas dalam dataset, di mana jumlah bibit tidak unggul sangat sedikit dibandingkan bibit unggul.

Nilai macro average untuk precision, recall, dan f1-score masing-masing adalah 0,48, 0,50, dan 0,49, sedangkan weighted average menunjukkan nilai yang lebih tinggi, yaitu 0,91, 0,95, dan 0,93. Nilai weighted average yang tinggi ini mencerminkan dominasi kelas bibit unggul dalam dataset, sehingga performa model secara keseluruhan tampak sangat baik, meskipun pada kenyataannya model kurang mampu mengenali kelas minoritas. Temuan ini sejalan dengan tantangan umum dalam klasifikasi dengan data tidak seimbang, di mana model cenderung mengabaikan kelas minoritas dan lebih fokus pada kelas mayoritas.

Secara ilmiah, hasil ini menegaskan bahwa algoritma Naïve Bayes sangat efektif untuk klasifikasi bibit unggul kelapa sawit jika proporsi kelas dalam data cukup seimbang atau jika fokus utama adalah pada kelas mayoritas. Namun, untuk aplikasi nyata yang menuntut deteksi akurat pada kedua kelas, perlu dilakukan penyesuaian seperti penyeimbangan data, penggunaan teknik sampling, atau penerapan metrik evaluasi

tambahan yang lebih sensitif terhadap ketidakseimbangan kelas. Dengan demikian, meskipun akurasi model sangat tinggi, evaluasi mendalam terhadap performa pada kelas minoritas tetap menjadi aspek penting dalam pengembangan sistem klasifikasi pemilihan bibit unggul kelapa sawit berbasis machine learning.

Halaman Login



The screenshot shows a login form with the title 'Halaman Login'. It contains two input fields: 'Username' and 'Password'. The 'Password' field has an eye icon to toggle visibility. Below the fields is a 'Login' button.

Gambar 7. Tampilan Halaman Login

Halaman login pada sistem klasifikasi pemilihan bibit unggul kelapa sawit, seperti yang ditampilkan pada Gambar 7, merupakan gerbang utama bagi pengguna untuk mengakses seluruh fitur sistem. Desain antarmuka halaman login ini sederhana dan intuitif, terdiri dari dua field input utama yaitu Username dan Password, serta satu tombol aksi Login. Fitur keamanan juga ditingkatkan dengan adanya ikon mata pada field password yang memungkinkan pengguna untuk menampilkan atau menyembunyikan karakter password saat mengetik, sehingga meminimalisir kesalahan input dan meningkatkan kenyamanan pengguna



Gambar 8. Tampilan Halaman Beranda

Halaman beranda pada sistem klasifikasi pemilihan bibit unggul kelapa sawit, sebagaimana ditampilkan pada Gambar 8, merupakan antarmuka utama yang diakses admin setelah berhasil melakukan proses login. Halaman ini dirancang dengan tampilan yang sederhana dan informatif, menampilkan menu navigasi di sisi kiri yang memungkinkan admin memilih halaman lain seperti data bibit, klasifikasi, atau hasil klasifikasi. Pada bagian utama halaman, terdapat pesan sambutan “Selamat datang di aplikasi Klasifikasi Pemilihan Bibit Unggul Kelapa Sawit” yang berfungsi sebagai pengenalan sistem kepada pengguna. Selain itu, ditampilkan pula foto bibit kelapa sawit

dalam jumlah besar yang tersusun rapi di lahan pembibitan, memperkuat konteks aplikasi sebagai alat bantu dalam proses seleksi bibit unggul.

Desain halaman beranda ini tidak hanya berperan sebagai pusat orientasi pengguna, tetapi juga sebagai titik awal dalam workflow sistem klasifikasi. Dengan menampilkan visualisasi bibit kelapa sawit, halaman ini memberikan gambaran nyata mengenai objek penelitian, sekaligus mempertegas tujuan utama aplikasi, yaitu mendukung proses pemilihan bibit unggul berbasis data. Navigasi yang jelas dan terstruktur pada sisi kiri halaman memudahkan admin untuk berpindah antar fitur sistem, sehingga meningkatkan efisiensi kerja dan mengurangi risiko kesalahan dalam pengoperasian aplikasi.

Data Bibit Sawit

Dataset Bibit Unggul Kelapa Sawit

ID Bibit	Tinggi Bibit (cm)	Diameter Batang (cm)	Jumlah Daun	Warna Daun	Ketahanan Penyakit	Ps
0	1	127	23	15	1	1
1	2	105	30	14	1	0
2	3	173	37	20	1	1
3	4	159	34	14	0	1
4	5	115	36	15	1	0
5	6	115	33	25	0	2
6	7	105	33	18	0	0
7	8	186	36	25	1	1
8	9	160	24	11	1	1
9	10	170	29	19	0	1

Deskripsi Dataset

ID Bibit	Tinggi Bibit (cm)	Diameter Batang (cm)	Jumlah Daun	Warna Daun	Ketahanan Penyakit
----------	-------------------	----------------------	-------------	------------	--------------------

Gambar 9. Tampilan Data Bibit Sawit

Halaman Data Bibit Sawit, sebagaimana ditampilkan pada Gambar 9, merupakan salah satu fitur inti dalam sistem klasifikasi pemilihan bibit unggul kelapa sawit yang dikembangkan menggunakan algoritma Naïve Bayes. Tampilan data bibit sawit yang terstruktur sangat penting untuk memastikan kualitas data yang digunakan dalam model machine learning. Proses pra-pemrosesan data, seperti pembersihan dari nilai kosong dan duplikat, telah dilakukan sebelumnya untuk menjamin integritas dan keandalan dataset. Setiap atribut yang ditampilkan pada tabel merupakan variabel-variabel yang telah terbukti relevan dalam menentukan kualitas dan potensi bibit kelapa sawit, sesuai dengan hasil penelitian terdahulu dan praktik agronomi yang berlaku. Dengan demikian, sistem ini tidak hanya berfungsi sebagai alat bantu administratif, tetapi juga sebagai media analitik yang mendukung pengambilan keputusan berbasis data dalam pemilihan bibit unggul.

Gambar 10. Tampilan Halaman Klasifikasi Bibit

Halaman Klasifikasi Bibit, sebagaimana ditampilkan pada Gambar 10, merupakan fitur utama dalam sistem klasifikasi pemilihan bibit unggul kelapa sawit yang dirancang untuk memfasilitasi admin dalam memasukkan data bibit secara terstruktur sebelum dilakukan proses klasifikasi berbasis algoritma Naïve Bayes. Keberadaan halaman klasifikasi ini sangat krusial dalam workflow sistem, karena seluruh data yang dimasukkan akan menjadi input utama bagi model Naïve Bayes dalam melakukan prediksi status bibit, apakah tergolong unggul atau tidak. Algoritma Naïve Bayes akan menghitung probabilitas setiap kelas berdasarkan kombinasi fitur-fitur yang diinputkan, sehingga semakin lengkap dan akurat data yang diberikan, semakin tinggi pula keandalan hasil klasifikasi yang dihasilkan. Proses klasifikasi ini juga didukung oleh tahapan pra-pemrosesan data yang telah dilakukan sebelumnya, seperti pembersihan data, normalisasi, dan seleksi fitur, untuk memastikan bahwa hanya data yang relevan dan berkualitas yang digunakan dalam analisis.

ID Bibit	Tinggi Bibit (cm)	Diameter Batang (mm)	Jumlah Daun	Warna Daun	Ketebalan Peripalid
01	100	10	10	Hijau Muda	Berk
02	100	10	10	Hijau Muda	Berk

Gambar 11. Tampilan Halaman Hasil Klasifikasi

Halaman Hasil Klasifikasi, sebagaimana ditampilkan pada Gambar 11, merupakan bagian penting dari sistem klasifikasi pemilihan bibit unggul kelapa sawit berbasis algoritma Naïve Bayes. Implementasi algoritma Naïve Bayes pada sistem ini memungkinkan proses klasifikasi dilakukan berdasarkan perhitungan probabilitas dari setiap fitur bibit yang diinputkan, sehingga setiap bibit dapat dipetakan ke dalam kategori unggul atau tidak unggul secara objektif dan terukur. Dengan tingkat akurasi yang mencapai 95%, sistem ini terbukti efektif dalam mengidentifikasi bibit unggul, mendukung

proses seleksi yang lebih efisien dan minim bias. Selain itu, hasil klasifikasi yang terpisah secara jelas antara bibit unggul dan tidak unggul memberikan transparansi dan akuntabilitas dalam proses seleksi, serta memudahkan tindak lanjut seperti pemilihan bibit untuk penanaman atau evaluasi bibit yang belum memenuhi kriteria unggul.

Dalam penelitian ini, pengujian Black Box Testing digunakan untuk memastikan bahwa sistem klasifikasi pemilihan bibit unggul kelapa sawit berfungsi sesuai dengan spesifikasi yang telah ditentukan. Pengujian ini dilakukan dengan cara memberikan input berupa data bibit kelapa sawit ke dalam sistem, kemudian memeriksa apakah output yang dihasilkan, yaitu hasil klasifikasi bibit unggul atau tidak unggul, sudah sesuai dengan yang diharapkan. Metode Black Box Testing berfokus pada pengujian fungsionalitas sistem tanpa memperhatikan struktur internal kode program, sehingga penguji hanya perlu mengetahui input dan output sistem. Berdasarkan hasil pengujian, seluruh fungsi utama yang diuji, seperti proses input data, klasifikasi, dan penampilan hasil klasifikasi, berjalan dengan baik dan menghasilkan output yang benar sesuai dengan data yang dimasukkan. Hal ini menunjukkan bahwa sistem telah memenuhi kebutuhan fungsional dan dapat diandalkan untuk membantu proses pemilihan bibit unggul kelapa sawit secara otomatis dan efisien. Pengujian ini juga menegaskan bahwa sistem mampu memberikan hasil klasifikasi yang konsisten dan akurat, sehingga dapat digunakan sebagai alat bantu dalam pengambilan keputusan di bidang perkebunan kelapa sawit.

Tabel 3. Hasil pengujian Black Box Testing

No.	Fungsi yang diuji	Skenario pengujian	Hasil	Kesimpulan
1.	Melihat halaman login	Membuka halaman login admin	Halaman login berhasil dibuka dan admin berhasil melakukan login	Valid
2.	Melihat halaman beranda	Membuka halaman beranda	Halaman beranda berhasil dibuka dan menampilkan foto bibit kelapa sawit	Valid
3.	Melihat halaman data bibit kelapa sawit	Membuka halaman data bibit kelapa sawit	Halaman data bibit kelapa sawit berhasil dibuka	Valid
4.	Melihat halaman form klasifikasi bibit kelapa sawit	Membuka halaman form klasifikasi bibit kelapa sawit	Halaman form klasifikasi bibit kelapa sawit berhasil dibuka dan berhasil melakukan penginputan data bibit kelapa sawit	Valid

5.	Melihat halaman hasil klasifikasi	Membuka halaman hasil klasifikasi bibit kelapa sawit	Halaman berhasil dibuka dan menampilkan hasil dari klasifikasi bibit kelapa sawit	Valid
6.	Melihat halaman keluar	Membuka halaman untuk keluar dari aplikasi	Halaman berhasil dibuka dan admin berhasil keluar dari aplikasi	Valid

Berdasarkan Tabel 3 Hasil Pengujian Black Box Testing, dapat disimpulkan bahwa seluruh fungsi utama pada sistem klasifikasi pemilihan bibit unggul kelapa sawit telah diuji dan berjalan sesuai dengan spesifikasi yang telah ditentukan. Pengujian dilakukan terhadap enam fungsi inti, yaitu membuka halaman login admin, halaman beranda, halaman data bibit kelapa sawit, halaman form klasifikasi bibit kelapa sawit, halaman hasil klasifikasi, dan halaman keluar dari aplikasi. Setiap skenario pengujian menunjukkan bahwa sistem mampu merespons input dengan benar, seperti membuka halaman login dan memungkinkan admin untuk melakukan login, menampilkan foto bibit pada halaman beranda, menampilkan data bibit secara lengkap, memfasilitasi penginputan data pada form klasifikasi, menampilkan hasil klasifikasi dengan baik, serta memastikan admin dapat keluar dari aplikasi dengan aman.

Seluruh hasil pengujian pada setiap fungsi dinyatakan valid, menandakan bahwa sistem telah memenuhi kebutuhan fungsional yang diharapkan oleh pengguna. Validitas ini penting karena membuktikan bahwa sistem tidak hanya berjalan secara teknis, tetapi juga dapat dioperasikan dengan mudah oleh admin tanpa mengalami kendala berarti. Keberhasilan pengujian black box ini juga sejalan dengan prinsip-prinsip pengembangan perangkat lunak berbasis machine learning, di mana pengujian fungsional menjadi tahap krusial untuk memastikan bahwa seluruh fitur yang diimplementasikan dapat memberikan hasil yang konsisten dan dapat diandalkan dalam skenario penggunaan nyata. Hasil pengujian ini memperkuat temuan bahwa sistem klasifikasi berbasis algoritma Naïve Bayes yang dikembangkan tidak hanya unggul dari sisi akurasi model, tetapi juga dari aspek keandalan operasional dan kemudahan penggunaan. Dengan demikian, sistem ini layak digunakan sebagai alat bantu dalam proses seleksi dan klasifikasi bibit unggul kelapa sawit, serta dapat diandalkan untuk meningkatkan efisiensi dan akurasi dalam pengambilan keputusan di bidang perkebunan kelapa sawit.

Dalam penerapan algoritma Naïve Bayes untuk klasifikasi bibit unggul kelapa sawit ini, tantangan utama yang dihadapi adalah ketergantungan pada kualitas data input.

Beberapa data bibit unggul yang digunakan dalam aplikasi mengandung fitur yang hilang atau tidak lengkap. Hal ini memengaruhi ketepatan prediksi dan memerlukan penanganan khusus, seperti pengisian data yang hilang atau penggunaan teknik imputasi untuk mengatasi masalah tersebut. Selain itu, meskipun aplikasi memberikan hasil klasifikasi dengan cepat, kecepatan proses menjadi tantangan dalam skala besar. Pengujian menunjukkan bahwa aplikasi bekerja optimal pada data dengan jumlah kecil hingga menengah, namun pada dataset yang lebih besar, waktu respons sedikit meningkat, yang memerlukan peningkatan efisiensi algoritma atau optimasi sistem.

Selain masalah kualitas data, kecepatan proses juga menjadi tantangan, terutama ketika aplikasi digunakan pada skala data yang besar. Hasil pengujian menunjukkan bahwa aplikasi bekerja sangat optimal pada dataset berukuran kecil hingga menengah, dengan waktu respons rata-rata 0,5 hingga 1 detik per data. Namun, ketika jumlah data yang diproses meningkat secara signifikan, waktu respons aplikasi juga cenderung meningkat, yang dapat mengurangi efisiensi sistem secara keseluruhan. Untuk mengatasi hal ini, diperlukan peningkatan efisiensi algoritma, misalnya dengan optimasi kode atau pemanfaatan teknik komputasi paralel agar sistem tetap responsif meskipun menangani volume data yang besar.

Kedua tantangan ini, yaitu kualitas data input dan efisiensi pemrosesan, merupakan aspek krusial dalam pengembangan sistem klasifikasi berbasis Naïve Bayes untuk pemilihan bibit unggul kelapa sawit. Penelitian ini menegaskan pentingnya tahapan pra-pemrosesan data, seperti pembersihan data dari nilai kosong dan duplikat, serta normalisasi atribut, untuk memastikan data yang digunakan sudah layak dianalisis. Di sisi lain, pengembangan sistem juga perlu memperhatikan aspek skalabilitas dan optimasi agar aplikasi tetap dapat memberikan hasil klasifikasi yang cepat dan akurat pada berbagai skala penggunaan. Dengan demikian, keberhasilan implementasi Naïve Bayes dalam klasifikasi bibit unggul kelapa sawit sangat ditentukan oleh upaya menjaga kualitas data dan efisiensi proses secara berkelanjutan.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, penerapan algoritma Naïve Bayes dalam sistem klasifikasi pemilihan bibit unggul kelapa sawit terbukti mampu meningkatkan objektivitas dan efisiensi proses seleksi, yang sebelumnya sangat bergantung pada metode manual yang rentan subjektivitas dan inkonsistensi. Sistem yang dikembangkan berhasil mengklasifikasikan bibit unggul dengan tingkat akurasi tinggi, yakni mencapai 95%, serta

recall sempurna (1,00) untuk kelas bibit unggul, menandakan seluruh bibit unggul dalam data uji dapat dikenali dengan baik oleh model. Hal ini menunjukkan bahwa pendekatan data mining berbasis Naïve Bayes efektif dalam menangani dataset berukuran besar dan variabel agronomis yang kompleks, serta mampu mempercepat proses klasifikasi menjadi hanya 0,5–1 detik per data, sehingga sangat mendukung kebutuhan industri kelapa sawit yang menuntut efisiensi dan ketepatan dalam seleksi bibit.

Namun, penelitian ini juga menemukan beberapa keterbatasan signifikan. Salah satu kendala utama adalah ketidakseimbangan kelas dalam dataset, di mana proporsi bibit tidak unggul jauh lebih sedikit dibandingkan bibit unggul. Hal ini menyebabkan model cenderung bias terhadap kelas mayoritas dan kurang mampu mengenali bibit tidak unggul, yang tercermin dari nilai precision, recall, dan f1-score yang rendah pada kelas minoritas. Selain itu, kualitas data input sangat memengaruhi performa model, terutama jika terdapat data yang hilang atau tidak lengkap. Tantangan lain adalah penurunan efisiensi sistem pada skala data yang sangat besar, di mana waktu respons aplikasi cenderung meningkat seiring bertambahnya volume data yang diproses.

Berdasarkan temuan tersebut, rekomendasi untuk penelitian selanjutnya adalah melakukan penyeimbangan kelas pada dataset, misalnya dengan teknik oversampling atau undersampling, agar model dapat mengenali kedua kelas secara lebih proporsional. Selain itu, penelitian lanjutan perlu mengintegrasikan teknik pra-pemrosesan data yang lebih canggih, seperti imputasi data hilang dan normalisasi fitur, untuk meningkatkan kualitas data input. Pengembangan sistem juga dapat diarahkan pada optimasi algoritma dan pemanfaatan komputasi paralel guna menjaga efisiensi pemrosesan pada skala data yang lebih besar. Penelitian mendatang disarankan untuk mengeksplorasi kombinasi algoritma atau pendekatan ensemble untuk meningkatkan akurasi dan generalisasi model, serta memperluas variabel agronomis yang dianalisis agar sistem klasifikasi semakin adaptif terhadap dinamika lingkungan dan kebutuhan industri kelapa sawit.

DAFTAR REFERENSI

- Agustina, N., Citra, D. H., Purnama, W., Nisa, C., & Kurnia, A. R. (2022). Implementasi algoritma Naive Bayes untuk analisis sentimen ulasan Shopee pada Google Play Store: The implementation of Naïve Bayes algorithm for sentiment analysis of Shopee reviews on Google Play Store. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(1), 47–54.

- Alfina, O., & Harahap, F. (2019). Pemodelan UML sistem pendukung keputusan dalam penentuan kelas siswa siswa tunagrahita. *Methomika: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, 3(2), 143–150.
- Amelia, N., Dilla, S. F., Azizah, S., Fahira, Z., & Darlis, A. (2023). Efektivitas peran guru dalam kurikulum Merdeka Belajar. *Jurnal Ilmiah Wahana Pendidikan*, 9(2), 421–426.
- Asrin, F. (2023). Pengujian fungsionalitas sistem inventaris barang pada Sekolah Menengah Kejuruan Citra Borneo menggunakan black box testing. *Jurnal Ilmiah ILKOMINFO-Ilmu Komputer & Informatika*, 6(2), 131–143.
- Br Ginting, D. Y., br Ginting, R., & Sembiring, D. J. (2020). Sistem pendukung keputusan dengan menggunakan metode Analytic Hierarchy Process (AHP). Yogyakarta: Penerbit Andi.
- Elvera, S. E., & Yesita Astarina, S. E. (2021). *Metodologi penelitian*. Yogyakarta: Penerbit Andi.
- Fatima, F., Setiawan, E., Renata, R., & Ramadhani, A. (2024, Oktober). Strategi pengelolaan berkelanjutan kelapa sawit di Indonesia. *Forum Ekonomi: Jurnal Ekonomi, Manajemen dan Akuntansi*, 26(4), 803–807.
- Firmansyah, H., Saputra, S. A., & Falah, M. (2025). Pengaruh dimensionality reduction menggunakan Self-Organizing Maps terhadap klasifikasi Naive Bayes pada data kampanye pemasaran. *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research*, 2(1b), 1853–1862.
- Kushariyadi, K., Apriyanto, H., Herdiana, Y., Asy'ari, F. H., Judijanto, L., Pasrun, Y. P., & Mardikawati, B. (2024). *Artificial Intelligence: Dinamika perkembangan AI beserta penerapannya*. PT. Sonpedia Publishing Indonesia.
- Mahendra, G. S., Tampubolon, L. P. D., Arni, S., Kharisma, L. P. I., Resmi, M. G., Sudipa, I. G. I., ... & Syam, S. (2023). *Sistem pendukung keputusan: Teori dan penerapannya dalam berbagai metode*. PT. Sonpedia Publishing Indonesia.
- Mauliana, P., Wiguna, W., & Permana, A. Y. (2020). Pengembangan e-helpdesk support system berbasis web di PT Akur Pratama. *Jurnal Responsif: Riset Sains dan Informatika*, 2(1), 19–29.
- Prastyo, E. H. A., Suhartono, S., Faisal, M., Yaqin, M. A., & Firdaus, R. A. J. (2024). Naive Bayes classification untuk prediksi cacat perangkat lunak. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 9(2), 782–791.
- Punkastyo, D. A., Septian, F., & Syaripudin, A. (2024). Implementasi data mining menggunakan algoritma Naïve Bayes untuk prediksi kelulusan siswa. *Journal of System and Computer Engineering*, 5(1), 24–35.
- Purba, B., & Syahputra, R. (2021). Implementasi metode Naive Bayes Classifier pada evaluasi kepuasan mahasiswa terhadap pembelajaran daring. *Infotekjar: Jurnal Nasional Informatika dan Teknologi Jaringan*, 6(1), 85–91.

- Rahmadden, S. K. M. K., Wulandari, D., Renova, M., & Sari, R. (2024). *Machine Learning*. Serasi Media Teknologi.
- Senika, A., Rasiban, R., & Iskandar, D. (2022). Implementasi metode Naïve Bayes dalam penilaian kinerja sales marketing pada PT. Pachira Distrinusa. *Jurnal Media Informatika Budidarma*, 6(1), 701.
- Setiawan, Z., Fajar, M., Priyatno, A. M., Putri, A. Y. P., Aryuni, M., Yuliyanti, S., ... & Wijaya, A. (2023). *Buku ajar data mining*. PT. Sonpedia Publishing Indonesia.
- Siregar, I. M., Pratama, D., & Himawan, C. (2024). Penggunaan Jaccard Similarity Coefficient dalam optimasi proses rekrutmen karyawan berbasis profil dan kompetensi. *SINTECH (Science and Information Technology) Journal*, 7(2), 101–111.
- Suriani, U. (2023). Penerapan data mining untuk memprediksi tingkat kelulusan mahasiswa menggunakan algoritma decision tree C4.5. *Journal of Computer and Information Systems Ampera*, 4(2), 55–65.