

Klasifikasi Jenis Bunga Iris Berdasarkan Fitur Morfologi Menggunakan Algoritma Naive Bayes

Ely Novita Sari, Deci Irmayani*, Budianto Bangun

Fakultas Sains Dan Teknologi, Program Studi Sistem Informasi, Universitas Labuhanbatu, Rantauprapat, Indonesia

Email: ¹elinovitasari886@gmail.com, ^{2*}deacyirmayani@gmail.com, ³budiantobangun44@gmail.com

Email Penulis Korespondensi: deacyirmayani@gmail.com

Submitted: 20/05/2025; Accepted: 22/05/2025; Published: 23/05/2025

Abstrak—Penelitian ini bertujuan untuk mengklasifikasikan jenis bunga Iris berdasarkan fitur morfologi menggunakan algoritma Naive Bayes. Bunga Iris terdiri dari tiga jenis, yaitu Iris-Setosa, Iris-Versicolor, dan Iris-Virginica, yang dapat dibedakan berdasarkan panjang dan lebar kelopak serta panjang dan lebar daun. Dataset yang digunakan dalam penelitian ini adalah dataset Iris yang berisi informasi mengenai empat fitur morfologi dari ketiga jenis bunga tersebut. Algoritma Naive Bayes dipilih karena memiliki keunggulan dalam melakukan klasifikasi berbasis probabilitas secara sederhana, cepat, dan efektif, terutama untuk data dengan fitur-fitur independen. Tahapan penelitian ini meliputi pengumpulan data, ekstraksi fitur, pembagian data menjadi data latih dan data uji, pelatihan model menggunakan algoritma Naive Bayes, serta pengujian model untuk mengevaluasi akurasi klasifikasi. Hasil penelitian menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan data uji dengan baik, di mana nilai probabilitas tertinggi diperoleh pada kelas Iris-Versicolor, dengan nilai $P(\text{Versicolor} | X)=1$. Hal ini membuktikan bahwa data uji memiliki kemiripan paling tinggi terhadap spesies tersebut jika dibandingkan dengan dua spesies lainnya. Dengan demikian, algoritma Naive Bayes efektif diterapkan untuk klasifikasi jenis bunga Iris berdasarkan fitur morfologi yang dimiliki.

Kata Kunci: Klasifikasi; Bunga Iris; Fitur Morfologi; Naive Bayes; Data Mining

Abstract—This study aims to classify the types of Iris flowers based on morphological features using the Naive Bayes algorithm. Iris flowers consist of three types, namely Iris-Setosa, Iris-Versicolor, and Iris-Virginica, which can be distinguished based on the length and width of the petals as well as the length and width of the sepals. The dataset used in this research is the Iris dataset, which contains information on four morphological features from these three types of flowers. The Naive Bayes algorithm was chosen because of its advantages in performing probability-based classification in a simple, fast, and effective manner, especially for data with independent features. The research stages include data collection, feature extraction, splitting the data into training and testing sets, training the model using the Naive Bayes algorithm, and testing the model to evaluate classification accuracy. The results of the study show that the Naive Bayes model is able to classify the test data accurately, with the highest probability value obtained in the Iris-Versicolor class, with a value of $P(\text{Versicolor} | X)=1$. This indicates that the test data has the highest similarity to that species compared to the other two species. Thus, the Naive Bayes algorithm proves effective for classifying types of Iris flowers based on their morphological features.

Keywords: Classification; Iris Flowers; Morphological Features; Naive Bayes; Data Mining

1. PENDAHULUAN

Indonesia merupakan negara yang memiliki keanekaragaman flora yang sangat luas, yang mencakup berbagai jenis tanaman, baik tanaman hias, tanaman obat, maupun tanaman yang memiliki nilai ekonomi penting. Salah satu jenis tanaman hias yang cukup populer adalah bunga iris. Bunga iris dikenal dengan keindahan dan keberagaman warna serta bentuk kelopaknya, yang membuatnya sangat diminati di berbagai negara, termasuk Indonesia. Bunga iris sendiri memiliki berbagai spesies dengan karakteristik morfologi yang bervariasi, seperti ukuran kelopak, panjang tangkai, lebar daun, dan tinggi tanaman[1]. Setiap jenis bunga iris, seperti Iris Setosa, Iris Versicolor, dan Iris Virginica, dapat dibedakan melalui ciri-ciri morfologi tersebut, yang menjadikannya objek yang menarik untuk dipelajari dalam bidang taksonomi[2].

Proses identifikasi jenis bunga iris berdasarkan ciri-ciri morfologi memang dapat dilakukan secara manual. Namun, proses ini sering kali memakan waktu yang cukup lama, memerlukan keahlian khusus, dan rawan menghasilkan kesalahan, terutama ketika terdapat banyak spesimen yang harus diklasifikasikan. Dalam dunia botani dan pertanian, ketepatan dalam mengidentifikasi jenis tanaman menjadi sangat penting, karena kesalahan klasifikasi dapat berpengaruh pada penelitian lanjutan maupun penggunaan tanaman tersebut dalam industri tertentu. Oleh karena itu, dibutuhkan solusi yang lebih efisien dan akurat untuk melakukan klasifikasi jenis bunga iris dengan mempertimbangkan data morfologi.

Dalam beberapa tahun terakhir, teknologi data mining (penambangan data) telah berkembang pesat dan menawarkan solusi praktis untuk mengatasi masalah klasifikasi dalam berbagai bidang, termasuk dalam identifikasi tanaman. Salah satu algoritma machine learning yang dapat diterapkan untuk klasifikasi tanaman berdasarkan data morfologi adalah Naive Bayes[3]. Algoritma ini merupakan metode probabilistik yang sederhana namun efektif, yang dapat mengklasifikasikan data dengan banyak fitur dengan cara menghitung probabilitas untuk setiap kategori berdasarkan data historis yang ada[4]. Naive Bayes beroperasi dengan asumsi bahwa setiap fitur bersifat independen satu sama lain, meskipun dalam kenyataannya, beberapa fitur bisa saling terkait[5]. Dengan menggunakan prinsip ini, algoritma Naive Bayes dapat menghitung probabilitas setiap jenis bunga iris berdasarkan fitur morfologinya, seperti panjang kelopak, lebar daun, dan panjang tangkai, dan menghasilkan prediksi yang akurat.

Penerapan Naive Bayes dalam klasifikasi bunga iris menawarkan banyak keuntungan, di antaranya adalah kemampuannya untuk mengolah dataset besar dengan efisien dan memberikan hasil yang cepat dengan tingkat akurasi yang tinggi. Selain itu, Naive Bayes dapat menangani data yang bersifat kategorikal maupun numerik, menjadikannya sangat fleksibel untuk diterapkan pada data morfologi bunga iris yang memiliki berbagai jenis atribut[6]. Seiring dengan perkembangan teknologi, penggunaan algoritma ini menjadi solusi yang semakin praktis dalam sistem klasifikasi bunga iris, yang dapat digunakan oleh peneliti atau ahli botani untuk mengidentifikasi jenis bunga iris dengan lebih cepat dan tepat.

Oleh karena itu, penelitian ini bertujuan untuk mengembangkan sistem klasifikasi bunga iris yang menggunakan algoritma Naive Bayes. Sistem ini diharapkan dapat memberikan hasil yang cepat, akurat, dan mudah digunakan dalam mengidentifikasi jenis bunga iris berdasarkan fitur morfologi. Dengan penerapan algoritma Naive Bayes, proses identifikasi bunga iris dapat dilakukan secara otomatis, mengurangi potensi kesalahan manusia, serta mempercepat proses pengolahan data dalam bidang taksonomi tanaman.

Beberapa penelitian terdahulu yang berkaitan dengan kesesuaian metode, sehingga dapat dijadikan acuan dalam penelitian ini. Penelitian oleh Anita Desiani dkk, pada tahun 2021 menghasilkan penerapan metode Support Vector Machine (SVM) dalam klasifikasi bunga Iris. Hasilnya, metode percentage split dengan akurasi 96,7% dan metode k-fold cross validation dengan akurasi 92,6% menunjukkan bahwa SVM efektif digunakan untuk klasifikasi bunga Iris[7]. Penelitian oleh Muhammad Ridho Pradana dkk, pada tahun 2022 mengaplikasikan algoritma Support Vector Machine (SVM) untuk klasifikasi data bunga Iris. Hasil penelitian menunjukkan bahwa metode SVM berhasil mencapai tingkat akurasi sebesar 95% dalam mengklasifikasikan data bunga Iris, dengan pengolahan data yang melibatkan 150 data dan 3 kelas. Penelitian ini membuktikan bahwa SVM efektif digunakan untuk memecahkan masalah klasifikasi, meskipun dengan dataset yang relatif kecil[8].

Penelitian oleh Endang Etriyantri dkk, pada tahun 2022, menerapkan algoritma Naive Bayes Classifier dan C4.5 untuk memprediksi kelulusan mahasiswa di STMIK Bina Nusantara Jaya Lubuklinggau. Hasilnya, algoritma C4.5 mencapai akurasi 79,08%, lebih tinggi dari Naive Bayes, dan efektif untuk prediksi kelulusan[9]. Penelitian oleh Fitriana Harahap dkk, pada tahun 2021, menerapkan algoritma Naive Bayes Classifier untuk memprediksi pembelian cat. Hasil penelitian menunjukkan bahwa metode Naive Bayes berhasil mencapai akurasi prediksi sebesar 80% dengan mengklasifikasikan 48 dari 60 data pembelian cat dengan benar. Penerapan algoritma ini membantu perusahaan dalam memprediksi minat pembelian pelanggan dan menentukan pasokan produk secara lebih efektif[10].

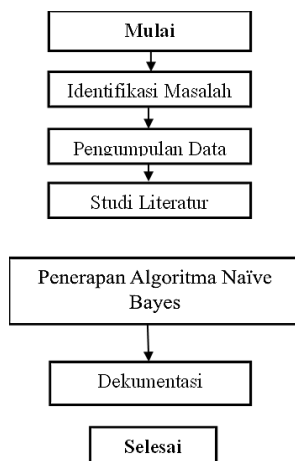
Penelitian sebelumnya banyak menggunakan algoritma SVM untuk klasifikasi bunga Iris dan menunjukkan akurasi tinggi. Sementara itu, algoritma Naive Bayes lebih banyak diterapkan pada kasus prediksi non-morfologis seperti kelulusan mahasiswa dan prediksi pembelian. Belum banyak penelitian yang secara khusus menerapkan algoritma Naive Bayes untuk klasifikasi bunga Iris berdasarkan fitur morfologi. Oleh karena itu, penelitian ini mengisi celah tersebut dengan mengkaji efektivitas Naive Bayes dalam klasifikasi bunga Iris berdasarkan panjang dan lebar kelopak serta mahkota bunga.

Berdasarkan hasil yang diperoleh dari beberapa penelitian terdahulu, penelitian ini menggunakan algoritma Naive Bayes untuk klasifikasi jenis bunga Iris berdasarkan fitur morfologi. Diharapkan metode ini dapat memberikan hasil yang lebih akurat dalam mengidentifikasi jenis bunga Iris berdasarkan ciri-ciri morfologi yang dimilikinya.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Kerangka penelitian menggambarkan urutan langkah-langkah yang diperlukan dalam proses penelitian. Setiap tahap yang ada saling terhubung secara terstruktur dan sistematis. Penyusunan tahap-tahap ini bertujuan untuk memudahkan pelaksanaan penelitian secara efektif dan efisien.



Gambar 1. Kerangka Penelitian

Berdasarkan struktur penelitian yang telah dijelaskan sebelumnya, penulis akan menjelaskan langkah-langkah yang diambil dalam penelitian ini sebagai berikut:

a. Identifikasi Masalah

Langkah pertama dalam penelitian ini adalah mengidentifikasi permasalahan dalam mengklasifikasikan jenis bunga Iris berdasarkan fitur morfologi. Fokus utamanya adalah menemukan metode yang dapat memberikan hasil klasifikasi yang lebih akurat, objektif, dan efisien dengan menggunakan algoritma Naive Bayes.

b. Pengumpulan Data

Pada tahap ini, dilakukan pengumpulan data citra bunga Iris yang mencakup fitur morfologi seperti panjang dan lebar kelopak serta panjang dan lebar daun. Data ini digunakan sebagai input dalam proses klasifikasi untuk menentukan jenis bunga Iris.

c. Studi Literatur

Penulis melakukan kajian literatur dengan membaca berbagai referensi seperti jurnal, buku, dan artikel ilmiah yang relevan. Studi ini mencakup teori-teori terkait dengan data mining, metode Naive Bayes, serta teknik pengolahan citra yang digunakan untuk klasifikasi jenis bunga Iris.

d. Penerapan Metode Naive Bayes

Pada tahap ini, algoritma Naive Bayes diterapkan untuk mengklasifikasikan jenis bunga Iris berdasarkan fitur morfologi yang telah diekstraksi. Data pelatihan digunakan untuk membangun model klasifikasi, dan data uji digunakan untuk mengukur akurasi model tersebut.

e. Dokumentasi

Tahap akhir mencakup pencatatan seluruh proses penelitian, termasuk tahapan pengumpulan data, penerapan algoritma, serta analisis hasil klasifikasi. Dokumentasi disusun secara sistematis agar dapat memberikan pemahaman yang jelas bagi pembaca atau peneliti lain yang ingin mengembangkan penelitian lebih lanjut.

2.2 Bunga Iris

Bunga Iris adalah tanaman berbunga yang berasal dari keluarga Iridaceae, dengan lebih dari 300 spesies yang tersebar luas di berbagai belahan dunia. Nama "Iris" diambil dari bahasa Yunani yang berarti "pelangi," mengacu pada ragam warna bunga iris yang sangat beragam, mulai dari ungu, biru, kuning, putih, hingga merah muda. Bunga ini mudah dikenali berkat bentuk kelopakannya yang elegan, simetris, serta panjang dan sempit[11].



Gambar 2. Bunga IRIS

Tanaman iris termasuk dalam kategori herba, dengan daun-daun yang panjang dan sempit serta batang tegak. Beberapa spesies iris tumbuh di lingkungan yang lebih kering, sedangkan yang lainnya dapat ditemukan di daerah lembap seperti tepi sungai atau danau. Iris tumbuh subur di berbagai iklim, terutama di wilayah yang memiliki cuaca sedang, dan dapat ditemukan hampir di seluruh dunia[12]. Bunga iris memiliki dua jenis kelopak yang berbeda, yaitu kelopak luar yang disebut "falls" dan biasanya lebih lebar serta menggantung, sementara kelopak dalam yang disebut "standards" lebih tegak dan sempit. Bagian bunga yang paling mencolok adalah kelopak luar dengan warna cerah, seringkali dihiasi dengan pola atau garis-garis menarik yang menambah keindahannya[13]. Iris tidak hanya dikenal sebagai tanaman hias yang memperindah taman atau lanskap, tetapi beberapa spesies iris juga memiliki manfaat dalam pengobatan tradisional, berkat sifat antiinflamasi yang dimilikinya. Selain itu, iris digunakan dalam industri kosmetik. Keanekaragaman morfologi bunga iris, seperti ukuran dan bentuknya, membuatnya menjadi objek penelitian yang menarik. Fitur-fitur fisik ini sering digunakan dalam pengklasifikasian spesies iris, seperti dalam studi menggunakan data mining untuk membedakan berbagai jenis iris berdasarkan karakteristiknya[14].

2.3 Fitur Morfologi

Fitur Morfologi merujuk pada karakteristik fisik yang dapat diamati pada organisme, khususnya pada tanaman, yang dapat digunakan untuk mengidentifikasi dan membedakan spesies. Dalam konteks penelitian botani, fitur morfologi pada tanaman mencakup berbagai aspek seperti ukuran, bentuk, dan struktur bagian-bagian tubuh tanaman, seperti daun, batang, bunga, dan akar. Fitur-fitur ini memiliki peran penting dalam klasifikasi dan identifikasi spesies tanaman, serta memberikan informasi yang berguna untuk studi taksonomi dan ekologi. Pada tanaman berbunga, salah satu fitur morfologi yang sering digunakan untuk identifikasi adalah morfologi bunga itu sendiri. Ciri-ciri seperti panjang dan lebar kelopak bunga, panjang dan lebar mahkota bunga, serta panjang tangkai bunga adalah beberapa fitur yang dapat digunakan untuk membedakan jenis tanaman satu dengan yang lainnya. Misalnya, pada klasifikasi bunga iris, panjang kelopak (sepal length), lebar kelopak (sepal width), panjang mahkota (petal length), dan lebar

mahkota (petal width) merupakan fitur utama yang digunakan untuk membedakan tiga spesies bunga iris, yakni *Iris-setosa*, *Iris-versicolor*, dan *Iris-virginica*. Fitur morfologi juga memiliki hubungan yang erat dengan adaptasi tanaman terhadap lingkungan tempat mereka tumbuh. Misalnya, tanaman yang tumbuh di daerah kering cenderung memiliki daun yang lebih kecil dan tebal untuk mengurangi kehilangan air, sementara tanaman yang tumbuh di daerah basah mungkin memiliki daun yang lebih besar dan tipis. Oleh karena itu, analisis fitur morfologi tidak hanya berguna untuk identifikasi spesies, tetapi juga untuk memahami bagaimana tanaman beradaptasi dengan kondisi lingkungan tertentu. Secara keseluruhan, fitur morfologi memainkan peran yang sangat penting dalam sistem klasifikasi tanaman, baik untuk tujuan ilmiah, konservasi, maupun aplikasi praktis dalam bidang pertanian dan hortikultura[15].

2.4 Data Mining

Data Mining adalah proses eksplorasi data dalam skala besar untuk menemukan pola, hubungan, atau informasi berharga yang mendukung pengambilan keputusan. Data mining memanfaatkan teknik seperti statistik, pembelajaran mesin, dan analisis data untuk menyederhanakan kompleksitas data yang besar dan mengungkap wawasan tersembunyi[16]. Dalam konteks klasifikasi jenis bunga iris berdasarkan fitur morfologi, data mining digunakan untuk menganalisis data yang mencakup panjang kelopak, lebar kelopak, panjang mahkota, dan lebar mahkota bunga. Proses ini bertujuan untuk mengidentifikasi hubungan antara fitur morfologi tersebut dan jenis bunga iris, yaitu *Iris-setosa*, *Iris-versicolor*, atau *Iris-virginica*[17]. Tahapan data mining dalam penelitian ini dimulai dengan pengumpulan dataset yang relevan, seperti dataset bunga iris yang sering digunakan dalam penelitian analisis data. Langkah berikutnya adalah pembersihan dan pra-proses data untuk memastikan data siap dianalisis. Setelah data diproses, dilakukan eksplorasi untuk memahami karakteristik dasar dan pola awal dalam dataset. Melalui penerapan algoritma klasifikasi yang sesuai, data mining membantu dalam menetapkan pola dan hubungan yang ada antara atribut morfologi bunga dan kategori jenis bunga. Proses ini memberikan hasil yang dapat digunakan untuk membangun model prediksi berbasis data yang akurat[18]. Aplikasi data mining dalam klasifikasi bunga iris menunjukkan pentingnya teknik ini dalam dunia penelitian dan industri. Dengan data mining, pengolahan data biologis menjadi lebih efisien, memungkinkan peneliti untuk menghasilkan wawasan mendalam yang sebelumnya sulit diakses. Proses ini juga berkontribusi pada pengembangan metode analisis yang lebih canggih untuk data yang serupa dalam berbagai bidang.

2.5 Algoritma Naive Bayes

Algoritma Naive Bayes adalah metode klasifikasi berbasis probabilitas yang digunakan untuk memprediksi kelas suatu data berdasarkan atribut-atribut yang ada. Algoritma ini didasarkan pada Teorema Bayes, yang menghubungkan probabilitas terjadinya suatu peristiwa dengan informasi yang sudah diketahui. Naive Bayes menggunakan prinsip *conditional probability*, di mana setiap atribut dianggap saling independen satu sama lain, meskipun dalam kenyataannya, hubungan antar atribut mungkin ada. Asumsi independensi ini menjadikan nama algoritma ini menjadi "naive" atau "naif"[19]. Dalam penggunaannya, Naive Bayes bekerja dengan menghitung probabilitas dari setiap kelas berdasarkan fitur-fitur yang ada pada data. Secara umum, langkah pertama adalah menghitung probabilitas prior untuk setiap kelas. Probabilitas ini menggambarkan kemungkinan suatu kelas muncul berdasarkan data yang ada. Kemudian, Naive Bayes menghitung probabilitas kondisional untuk setiap atribut yang ada dalam data. Setelah itu, algoritma akan mengalikan probabilitas prior dan probabilitas kondisional untuk setiap kelas yang mungkin. Hasil akhirnya adalah kelas dengan probabilitas tertinggi yang dipilih sebagai prediksi untuk data baru[20]. Naive Bayes memiliki beberapa keunggulan, seperti kesederhanaan, kecepatan, dan kemampuannya untuk bekerja dengan baik meskipun ada ketergantungan antar fitur. Meskipun asumsi independensi ini seringkali tidak realistis, Naive Bayes tetap memberikan hasil yang cukup akurat dalam banyak kasus, terutama dalam aplikasi teks seperti klasifikasi email atau analisis sentimen. Selain itu, algoritma ini juga dapat menangani data dengan jumlah atribut yang besar secara efisien. Naive Bayes adalah metode probabilistik berdasarkan Teorema Bayes, yang menghitung probabilitas suatu kelas berdasarkan fitur yang ada. Formula dasarnya adalah:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

Keterangan:

$P(C|X)$: Probabilitas kelas C diberikan fitur X (klasifikasi yang ingin ditentukan)

$P(X|C)$: Probabilitas mendapatkan fitur X dalam kelas C

$P(C)$: Probabilitas awal dari kelas C (probabilitas prior)

$P(X)$: Probabilitas mendapatkan fitur X secara umum

3. HASIL DAN PEMBAHASAN

3.1 Analisa

Pada tahap analisis sistem, langkah ini dapat dianggap sebagai tahap awal dalam menyelesaikan masalah klasifikasi jenis bunga Iris berdasarkan fitur morfologi. Analisis sistem sangat penting untuk memastikan bahwa pengembangan model klasifikasi berjalan dengan lancar dan terstruktur, sehingga dapat menghasilkan sistem yang akurat dalam menentukan spesies bunga Iris. Bagian ini berfokus pada perhitungan dan perancangan yang diperlukan dalam proses

klasifikasi bunga Iris menggunakan algoritma Naive Bayes. Dalam penelitian ini, digunakan dataset Iris yang terdiri dari tiga jenis bunga, yaitu Iris-setosa, Iris-versicolor, dan Iris-virginica, dengan empat fitur morfologi utama: panjang mahkota bunga, lebar mahkota bunga, panjang kelopak bunga, dan lebar kelopak bunga. Data ini diolah untuk menentukan pola klasifikasi berdasarkan pendekatan probabilistik yang diterapkan dalam algoritma Naive Bayes. Setelah semua data yang diperlukan berhasil dikumpulkan, langkah selanjutnya adalah mengidentifikasi pola dalam dataset untuk membangun model klasifikasi yang efektif. Tahap berikutnya melibatkan perhitungan probabilitas untuk setiap kelas menggunakan teorema Bayes, yang mencakup perhitungan probabilitas prior dan likelihood berdasarkan distribusi data. Selain itu, dilakukan studi kepustakaan untuk mendukung perancangan dan validasi model klasifikasi agar dapat diterapkan secara optimal dalam mengklasifikasikan jenis bunga Iris dengan tingkat akurasi yang tinggi.

3.2 Penerapan Algoritma Naïve Bayes

Tabel 1 menampilkan sampel data bunga Iris yang terdiri dari empat fitur morfologi: panjang mahkota bunga, lebar mahkota bunga, panjang kelopak bunga, dan lebar kelopak bunga. Setiap sampel diklasifikasikan ke dalam tiga spesies, yaitu Iris-setosa, Iris-versicolor, dan Iris-virginica. Iris-setosa memiliki kelopak lebih pendek, Iris-versicolor berukuran sedang, sedangkan Iris-virginica memiliki kelopak terbesar. Data ini digunakan dalam klasifikasi dengan algoritma Naive Bayes, yang menganalisis probabilitas setiap fitur untuk menentukan spesies bunga. Proses ini memungkinkan sistem mengidentifikasi spesies baru berdasarkan karakteristik morfologinya.

Tabel 1. Dataset Bunga Iris

No	Panjang Mahkota Bunga (cm)	Lebar Mahkota Bunga (cm)	Panjang Kelopak Bunga (cm)	Lebar Kelopak Bunga (cm)	Spesies (Klasifikasi)
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	3.6	1.4	0.2	Iris-setosa
6	7	3.2	4.7	1.4	Iris-versicolor
7	6.4	3.2	4.5	1.5	Iris-versicolor
8	6.9	3.1	4.9	1.5	Iris-versicolor
9	5.5	2.3	4	1.3	Iris-versicolor
10	6.5	2.8	4.6	1.5	Iris-versicolor
11	6.3	3.3	6	2.5	Iris-virginica
12	5.8	2.7	5.1	1.9	Iris-virginica
13	7.1	3	5.9	2.1	Iris-virginica
14	6.3	2.9	5.6	1.8	Iris-virginica
15	6.5	3	5.8	2.2	Iris-virginica

Tabel 1 menampilkan data uji bunga Iris yang digunakan untuk menguji keakuratan model klasifikasi berbasis algoritma Naive Bayes. Data uji ini terdiri dari empat fitur morfologi utama, yaitu panjang mahkota bunga, lebar mahkota bunga, panjang kelopak bunga, dan lebar kelopak bunga. Pada tabel ini, spesies bunga belum ditentukan (ditandai dengan "?"), yang berarti model akan memprediksi spesies berdasarkan pola data latih sebelumnya. Dengan menggunakan perhitungan probabilitas, model akan menentukan kemungkinan spesies yang paling sesuai dengan fitur morfologi yang diberikan, membantu dalam identifikasi bunga secara otomatis.

Tabel 2. Data Uji Bunga Iris

No	Panjang Mahkota Bunga (cm)	Lebar Mahkota Bunga (cm)	Panjang Kelopak Bunga (cm)	Lebar Kelopak Bunga (cm)	Spesies (Klasifikasi)
1	3.6	1.5	1.6	0.5	?

Berdasarkan data sebelumnya yang telah dianalisis, data uji ini dapat diklasifikasikan ke dalam salah satu spesies bunga Iris yang tersedia, yaitu Iris-setosa, Iris-versicolor, atau Iris-virginica. Proses klasifikasi dilakukan dengan membandingkan nilai fitur morfologi dari data uji dengan data latih yang telah dikumpulkan sebelumnya. Dengan menggunakan algoritma Naive Bayes, perhitungan probabilitas untuk setiap spesies dilakukan berdasarkan panjang dan lebar mahkota serta panjang dan lebar kelopak bunga. Hasil perhitungan menunjukkan spesies dengan probabilitas tertinggi, sehingga data uji ini dapat dikategorikan ke dalam spesies bunga yang paling sesuai berdasarkan karakteristik morfologinya.

Langkah pertama dalam menyelesaikan klasifikasi menggunakan algoritma Naive Bayes adalah menghitung Prior Probability $P(C)$, yaitu probabilitas awal dari masing-masing kelas sebelum mempertimbangkan fitur yang ada. Prior Probability dihitung dengan membagi jumlah sampel dalam setiap kelas dengan total jumlah sampel dalam dataset. Nilai ini menggambarkan kemungkinan kemunculan suatu kelas tanpa memperhitungkan atribut lainnya. Dengan menentukan $P(C)$ terlebih dahulu, algoritma dapat membangun dasar perhitungan probabilitas bersyarat pada

tahap berikutnya. Langkah ini penting karena memberikan pemahaman awal tentang distribusi kelas dalam dataset sebelum fitur digunakan dalam proses klasifikasi.

Iris-Setosa memiliki 5 sampel

Iris-Versicolor memiliki 5 sampel

Iris-Virginica memiliki 5 sampel

Maka:

$$P(\text{Setosa}) = \frac{5}{15} = 0.33$$

$$P(\text{Versicolor}) = \frac{5}{15} = 0.33$$

$$P(\text{Virginica}) = \frac{5}{15} = 0.33$$

Langkah selanjutnya dalam penerapan algoritma Naive Bayes adalah menghitung parameter Gaussian, yaitu mean (rata-rata, μ) dan standar deviasi (σ) untuk setiap fitur dalam setiap kelas. Parameter ini diperlukan untuk menghitung probabilitas likelihood menggunakan distribusi Gaussian.

Menghitung nilai Mean (rata-rata, μ)

a. Iris-Setosa

Panjang Mahkota Bunga – Iris-Setosa:

$$\mu \text{ Setosa} = \frac{(5.1+4.9+4.7+4.6+5)}{5} = 4.86$$

Lebar Mahkota Bunga – Iris-Setosa:

$$\mu \text{ Setosa} = \frac{(3.5+3+3.2+3.1+3.6)}{5} = 3.28$$

Panjang Kelopak Bunga – Iris-Setosa:

$$\mu \text{ Setosa} = \frac{(1.4+1.4+1.3+1.5+1.4)}{5} = 1.4$$

Lebar Kelopak Bunga – Iris-Setosa:

$$\mu \text{ Setosa} = \frac{(0.2+0.2+0.2+0.2+0.2)}{5} = 0.2$$

b. Iris-versicolor

Panjang Mahkota Bunga – Iris-versicolor:

$$\mu \text{ versicolor} = \frac{(7.0+6.4+6.9+5.5+6.5)}{5} = 6.46$$

Lebar Mahkota Bunga – Iris-versicolor:

$$\mu \text{ versicolor} = \frac{(3.2+3.2+3.1+2.3+2.8)}{5} = 2.92$$

Panjang Kelopak Bunga – Iris-versicolor:

$$\mu \text{ versicolor} = \frac{(4.7+4.5+4.9+4.0+4.6)}{5} = 4.54$$

Lebar Kelopak Bunga – Iris-versicolor:

$$\mu \text{ Setosa} = \frac{(1.4+1.5+1.5+1.3+1.5)}{5} = 1.44$$

c. Iris-virginica

Panjang Mahkota Bunga – Iris-virginica:

$$\mu \text{ virginica} = \frac{(6.3+5.8+7.1+6.3+6.5)}{5} = 6.4$$

Lebar Mahkota Bunga – Iris-virginica:

$$\mu \text{ virginica} = \frac{(3.3+2.7+3.0+2.9+3.0)}{5} = 2.98$$

Panjang Kelopak Bunga – Iris-virginica:

$$\mu \text{ virginica} = \frac{(6.0+5.1+5.9+5.6+5.8)}{5} = 5.68$$

Lebar Kelopak Bunga – Iris-virginica:

$$\mu \text{ virginica} = \frac{(2.5+1.9+2.1+1.8+2.2)}{5} = 2.1$$

Berikut adalah nilai rata-rata (μ) dari setiap fitur dalam dataset Iris untuk tiga kelas bunga: Setosa, Versicolor, dan Virginica. Panjang dan lebar mahkota serta panjang dan lebar kelopak mencerminkan karakteristik khas dari masing-masing spesies. Setosa memiliki mahkota yang lebih pendek namun lebih lebar dibandingkan dua spesies lainnya. Versicolor dan Virginica memiliki panjang mahkota yang lebih besar, dengan Virginica menunjukkan kelopak yang paling panjang dan lebar. Perbedaan ini membantu dalam proses klasifikasi berdasarkan fitur bunga, yang digunakan dalam analisis probabilitas pada metode Naïve Bayes. Berikut adalah Tabel 3 hasil perhitungannya:

Tabel 3. Nilai rata-rata (μ)

Fitur Morfologi	μ Setosa	μ Versicolor	μ Virginica
Panjang Mahkota (cm)	4.86	6.46	6.4
Lebar Mahkota (cm)	3.28	2.92	2.98
Panjang Kelopak (cm)	1.4	4.54	5.68
Lebar Kelopak (cm)	0.2	1.44	2.1

Menghitung Standar Deviasi (σ)

a. Iris-Setosa

Panjang Mahkota Bunga – Iris-Setosa:

$$(\sigma)\text{Setosa} = \sqrt{\frac{(5.1 - 4.86)^2 + (4.9 - 4.86)^2 + (4.7 - 4.86)^2 + (4.6 - 4.86)^2 + (5.0 - 4.86)^2}{5}} = 0.185$$

Lebar Mahkota Bunga – Iris-Setosa:

$$(\sigma)\text{Setosa} = \sqrt{\frac{(5.1 - 4.86)^2 + (4.9 - 4.86)^2 + (4.7 - 4.86)^2 + (4.6 - 4.86)^2 + (5.0 - 4.86)^2}{5}} = 0.232$$

Panjang Kelopak Bunga – Iris-Setosa:

$$(\sigma)\text{Setosa} = \sqrt{\frac{(1.4 - 1.4)^2 + (1.4 - 1.4)^2 + (1.3 - 1.4)^2 + (1.5 - 1.4)^2 + (1.4 - 1.4)^2}{5}} = 0.063$$

Lebar Kelopak Bunga – Iris-Setosa:

$$(\sigma)\text{Setosa} = \sqrt{\frac{(0.2 - 0.2)^2 + (0.2 - 0.2)^2 + (0.2 - 0.2)^2 + (0.2 - 0.2)^2 + (0.2 - 0.2)^2}{5}} = 0$$

b. Iris-versicolor

Panjang Mahkota Bunga – Iris-versicolor:

$$(\sigma)\text{versicolor} = \sqrt{\frac{(7.0 - 6.46)^2 + (6.4 - 6.46)^2 + (6.9 - 6.46)^2 + (5.5 - 6.46)^2 + (6.5 - 6.46)^2}{5}} = 0.532$$

Lebar Mahkota Bunga – Iris-versicolor:

$$(\sigma)\text{versicolor} = \sqrt{\frac{(3.2 - 2.92)^2 + (3.2 - 2.92)^2 + (3.1 - 2.92)^2 + (2.3 - 2.92)^2 + (2.8 - 2.92)^2}{5}} = 0.343$$

Panjang Kelopak Bunga – Iris-versicolor:

$$(\sigma)\text{versicolor} = \sqrt{\frac{(4.7 - 4.54)^2 + (4.5 - 4.54)^2 + (4.9 - 4.54)^2 + (4.0 - 4.54)^2 + (4.6 - 4.54)^2}{5}} = 0.301$$

Lebar Kelopak Bunga – Iris-versicolor:

$$(\sigma)\text{versicolor} = \sqrt{\frac{(1.4 - 1.44)^2 + (1.5 - 1.44)^2 + (1.5 - 1.44)^2 + (1.3 - 1.44)^2 + (1.5 - 1.44)^2}{5}} = 0.080$$

c. Iris-virginica

Panjang Mahkota Bunga – Iris-virginica:

$$(\sigma)\text{virginica} = \sqrt{\frac{(6.3 - 6.4)^2 + (5.8 - 6.4)^2 + (7.1 - 6.4)^2 + (6.3 - 6.4)^2 + (6.5 - 6.4)^2}{5}} = 0.417$$

Lebar Mahkota Bunga – Iris-virginica:

$$(\sigma)\text{virginica} = \sqrt{\frac{(3.3 - 2.98)^2 + (2.7 - 2.98)^2 + (3.0 - 2.98)^2 + (2.9 - 2.98)^2 + (3.0 - 2.98)^2}{5}} = 0.194$$

Panjang Kelopak Bunga – Iris-virginica:

$$(\sigma)_{\text{virginica}} = \sqrt{\frac{(6.0 - 5.68)^2 + (5.1 - 5.68)^2 + (5.9 - 5.68)^2 + (5.6 - 5.68)^2 + (5.8 - 5.68)^2}{5}} = 0.319$$

Lebar Kelopak Bunga – Iris-virginica:

$$(\sigma)_{\text{virginica}} = \sqrt{\frac{(2.5 - 2.1)^2 + (1.9 - 2.1)^2 + (2.1 - 2.1)^2 + (1.8 - 2.1)^2 + (2.2 - 2.1)^2}{5}} = 0.245$$

Berikut adalah nilai standar deviasi (σ) dari setiap fitur dalam dataset Iris untuk tiga kelas bunga: Setosa, Versicolor, dan Virginica. Standar deviasi menunjukkan sebaran data di sekitar rata-rata, memberikan gambaran tentang variasi ukuran morfologi bunga dalam masing-masing spesies. Setosa memiliki standar deviasi yang relatif kecil, terutama pada panjang dan lebar kelopak, yang menunjukkan ukuran bunga yang lebih seragam. Versicolor menunjukkan variasi yang lebih tinggi dibandingkan Setosa, terutama pada panjang mahkota dan panjang kelopak. Sementara itu, Virginica memiliki variasi terbesar dalam lebar kelopak, yang mencerminkan perbedaan ukuran yang lebih signifikan dalam spesies ini. Perbedaan dalam standar deviasi ini membantu dalam analisis klasifikasi, seperti metode Naïve Bayes, yang mempertimbangkan sebaran data dalam menentukan probabilitas suatu spesies berdasarkan fitur morfologinya. Berikut adalah Tabel 4 hasil perhitungannya:

Tabel 4. Nilai standar deviasi (σ)

Fitur Morfologi	σ Setosa	σ Versicolor	σ Virginica
Panjang Mahkota (cm)	0.185	0.532	0.417
Lebar Mahkota (cm)	0.232	0.343	0.194
Panjang Kelopak (cm)	0.063	0.301	0.319
Lebar Kelopak (cm)	0	0.080	0.245

Langkah selanjutnya dalam penerapan algoritma Naïve Bayes adalah menghitung probabilitas likelihood menggunakan distribusi Gaussian. Probabilitas likelihood ini dihitung berdasarkan parameter Gaussian, yaitu mean (rata-rata, μ) dan standar deviasi (σ), untuk setiap fitur dalam setiap kelas. Probabilitas ini menunjukkan seberapa besar kemungkinan suatu nilai fitur muncul dalam suatu kelas tertentu.

a. Iris-Setosa

Panjang Mahkota Bunga (3.6 | Iris-Setosa):

$$\begin{aligned}
 P(3.6|\text{Setosa}) &= \frac{1}{\sqrt{2\pi(0.185)^2}} e^{-\frac{(3.6-4.86)^2}{2(0.185)^2}} \\
 &= \frac{1}{\sqrt{2\pi(0.0342)}} e^{-\frac{(3.6-4.86)^2}{2(0.0342)}} \\
 &= \frac{1}{\sqrt{0.2149}} e^{-\frac{1.6225}{0.0684}} \\
 &= \frac{1}{0.4637} e^{-23.73} \\
 &= 1.82 \times 10^{-10}
 \end{aligned}$$

Lebar Mahkota Bunga (1.5|Iris-Setosa):

$$\begin{aligned}
 P(1.5|\text{Setosa}) &= \frac{1}{\sqrt{2\pi(0.232)^2}} e^{-\frac{(1.5-3.28)^2}{2(0.232)^2}} \\
 &= \frac{1}{\sqrt{2\pi(0.0538)}} e^{-\frac{3.2184}{0.1076}} \\
 &= \frac{1}{0.5838} e^{-29.91} \\
 &= 2.83 \times 10^{-13}
 \end{aligned}$$

Panjang Kelopak Bunga (1.6| Iris-Setosa):

$$\begin{aligned}
 P(1.6|\text{Setosa}) &= \frac{1}{\sqrt{2\pi(0.063)^2}} e^{-\frac{(1.6-1.4)^2}{2(0.063)^2}} \\
 &= \frac{1}{\sqrt{2\pi(0.004)}} e^{-\frac{0.04}{0.008}} \\
 &= \frac{1}{0.158} e^{-5} \\
 &= 0.041
 \end{aligned}$$

Lebar Kelopak Bunga (0.5|Iris-Setosa):

$$P(0.5|\text{Setosa}) = \frac{1}{\sqrt{2\pi(0)^2}} e^{-\frac{(1.5-0)^2}{2(0)^2}}$$

Karena standar deviasi =0, hasilnya adalah 0.

b. Iris-versicolor

Panjang Mahkota Bunga (3.6 | Iris-versicolor):

$$\begin{aligned} P(3.6|\text{Versicolor}) &= \frac{1}{\sqrt{2\pi(0.532)^2}} e^{-\frac{(3.6-6.46)^2}{2(0.532)^2}} \\ &= \frac{1}{\sqrt{2\pi(0.283)}} e^{-\frac{8.1796}{0.566}} \\ &= \frac{1}{1.332} e^{-14.45} \\ &= 3.97 \times 10^{-7} \end{aligned}$$

Lebar Mahkota Bunga (1.5 | Iris-versicolor):

$$\begin{aligned} P(1.5|\text{Versicolor}) &= \frac{1}{\sqrt{2\pi(0.343)^2}} e^{-\frac{(1.5-2.92)^2}{2(0.343)^2}} \\ &= \frac{1}{\sqrt{2\pi(0.118)}} e^{-\frac{2.0176}{0.236}} \\ &= \frac{1}{0.857} e^{-8.55} \\ &= 2.21 \times 10^{-4} \end{aligned}$$

Panjang Kelopak Bunga (1.6 | Iris-versicolor):

$$\begin{aligned} P(1.6|\text{Versicolor}) &= \frac{1}{\sqrt{2\pi(0.301)^2}} e^{-\frac{(1.6-4.54)^2}{2(0.301)^2}} \\ &= \frac{1}{\sqrt{2\pi(0.0906)}} e^{-\frac{8.7076}{0.1812}} \\ &= \frac{1}{0.754} e^{-48.06} \\ &= 2.54 \times 10^{-21} \end{aligned}$$

Lebar Kelopak Bunga (0.5 | Iris-versicolor):

$$\begin{aligned} P(0.5|\text{Versicolor}) &= \frac{1}{\sqrt{2\pi(0.080)^2}} e^{-\frac{(0.5-1.44)^2}{2(0.080)^2}} \\ &= \frac{1}{\sqrt{2\pi(0.0064)}} e^{-\frac{0.8836}{0.0128}} \\ &= \frac{1}{0.201} e^{-96.03} \\ &= 5.22 \times 10^{-30} \end{aligned}$$

c. Iris-Virginica

Panjang Mahkota Bunga (3.6 | Iris-virginica):

$$\begin{aligned} P(3.6|\text{Virginica}) &= \frac{1}{\sqrt{2\pi(0.417)^2}} e^{-\frac{(3.6-6.4)^2}{2(0.417)^2}} \\ &= \frac{1}{\sqrt{2\pi(0.174)}} e^{-\frac{7.84}{0.348}} \\ &= \frac{1}{1.047} e^{-22.52} \\ &= 1.55 \times 10^{-10} \end{aligned}$$

Lebar Mahkota Bunga (1.5 | Iris-virginica):

$$P(1.5|\text{Virginica}) = \frac{1}{\sqrt{2\pi(0.194)^2}} e^{-\frac{(1.5-2.98)^2}{2(0.194)^2}}$$

$$\begin{aligned}
 &= \frac{1}{\sqrt{2\pi(0.0376)}} e^{-\frac{2.1984}{0.0752}} \\
 &= \frac{1}{0.487} e^{-29.25} \\
 &= 4.73 \times 10^{-13}
 \end{aligned}$$

Panjang Kelopak Bunga (1.6|Iris-virginica):

$$\begin{aligned}
 P(1.6|Virginica) &= \frac{1}{\sqrt{2\pi(0.319)^2}} e^{-\frac{(1.6-5.68)^2}{2(0.319)^2}} \\
 &= \frac{1}{\sqrt{2\pi(0.1018)}} e^{-\frac{16.64}{0.2036}} \\
 &= \frac{1}{0.800} e^{-81.72} \\
 &= 3.76 \times 10^{-36}
 \end{aligned}$$

Lebar Kelopak Bunga (0.5|Iris-virginica):

$$\begin{aligned}
 P(0.5|Virginica) &= \frac{1}{\sqrt{2\pi(0.245)^2}} e^{-\frac{(0.5-2.1)^2}{2(0.245)^2}} \\
 &= \frac{1}{\sqrt{2\pi(0.060)}} e^{-\frac{2.56}{0.12}} \\
 &= \frac{1}{0.615} e^{-21.33} \\
 &= 8.92 \times 10^{-10}
 \end{aligned}$$

Berikut adalah Tabel 5 yang merangkum hasil perhitungan probabilitas likelihood berdasarkan distribusi Gaussian untuk ketiga jenis Iris:

Tabel 5. Nilai probabilitas likelihood

Fitur Morfologi	σ Setosa	σ Versicolor	σ Virginica
Panjang Mahkota (cm)	1.82×10^{-10}	3.97×10^{-7}	1.55×10^{-10}
Lebar Mahkota (cm)	2.83×10^{-13}	2.21×10^{-4}	4.73×10^{-13}
Panjang Kelopak (cm)	0.041	2.54×10^{-21}	3.76×10^{-36}
Lebar Kelopak (cm)	0 (Karena standar deviasi =0)	5.22×10^{-30}	8.92×10^{-10}

Langkah terakhir dalam penerapan algoritma Naïve Bayes adalah menghitung probabilitas akhir untuk setiap kelas dengan menggabungkan probabilitas prior dan likelihood dari setiap fitur. Probabilitas prior menunjukkan seberapa sering suatu kelas muncul dalam data, sementara likelihood menunjukkan seberapa besar kemungkinan suatu fitur muncul dalam kelas tertentu. Semua probabilitas likelihood yang telah dihitung sebelumnya dikalikan dengan prior kelas untuk mendapatkan nilai akhir. Hasil ini dibandingkan di antara semua kelas, dan kelas dengan probabilitas tertinggi akan dipilih sebagai hasil klasifikasi. Dengan demikian, prediksi dilakukan berdasarkan probabilitas terbesar yang diperoleh.

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

Menghitung $P(X|C)P(C)$ untuk setiap kelas:

a. Iris-Setosa

$$\begin{aligned}
 (C|Setosa) &= P(3.6|Setosa) \times P(1.5|Setosa) \times P(1.6|Setosa) \times P(0.5|Setosa) \\
 &= (1.82 \times 10^{-10}) \times (2.83 \times 10^{-13}) \times (0.041) \times (0) \\
 &= 0
 \end{aligned}$$

Karena $P(0.5|Setosa) = 0$, maka hasilnya pasti 0.

$$P(X|Setosa)P(Setosa) = 0 \times 0.33 = 0$$

b. Iris-versicolor

$$\begin{aligned}
 P(C|versicolor) &= P(3.6|versicolor) \times P(1.5|versicolor) \times P(1.6|versicolor) \times P(0.5|versicolor) \\
 &= (3.97 \times 10^{-7}) \times (2.21 \times 10^{-4}) \times (2.54 \times 10^{-21}) \times (5.22 \times 10^{-30})
 \end{aligned}$$

$$= 1.17 \times 10^{-61}$$

$$P(X|\text{versicolor})P(\text{versicolor}) = 1.17 \times 10^{-61} \times 0.33 = 3.87 \times 10^{-62}$$

c. Iris-virginica

$$P(C|\text{virginica}) = P(3.6|\text{virginica}) \times P(1.5|\text{virginica}) \times P(1.6|\text{virginica}) \times P(0.5|\text{virginica})$$

$$= (1.55 \times 10^{-10}) \times (4.73 \times 10^{-13}) \times (3.76 \times 10^{-36}) \times (8.92 \times 10^{-10})$$

$$= 2.41 \times 10^{-68}$$

$$P(X|\text{virginica})P(\text{virginica}) = 2.41 \times 10^{-68} \times 0.33 = 3.95 \times 10^{-69}$$

Menghitung probabilitas posterior $P(C|X)$:

Dimana $P(X)$ adalah jumlah dari semua $P(X|C)P(C)$:

$$P(X) = (0) + (3.87 \times 10^{-62}) + (7.95 \times 10^{-69})$$

$$= 3.87 \times 10^{-62}$$

Sehingga,

$$P(\text{Setosa}|X) = \frac{0}{3.87 \times 10^{-62}} = 0$$

$$P(\text{Versicolor}|X) = \frac{3.87 \times 10^{-62}}{3.87 \times 10^{-62}} = 1$$

$$P(\text{Virginica}|X) = \frac{3.95 \times 10^{-69}}{3.87 \times 10^{-62}} = 2.05 \times 10^{-7}$$

Berdasarkan hasil perhitungan menggunakan algoritma Naïve Bayes pada fitur morfologi bunga Iris, diperoleh bahwa probabilitas tertinggi untuk data uji berada pada kelas Iris-Versicolor, dengan nilai $P(\text{Versicolor} | X) = 1$. Hal ini menunjukkan bahwa berdasarkan panjang dan lebar mahkota serta kelopak bunga, data uji memiliki kemiripan yang paling tinggi dengan spesies Iris-Versicolor dibandingkan dengan spesies lainnya (Iris-Setosa dan Iris-Virginica). Dengan demikian, algoritma Naïve Bayes terbukti dapat mengklasifikasikan jenis bunga Iris secara efektif berdasarkan fitur morfologi yang dimilikinya.

4. KESIMPULAN

Berdasarkan rangkaian analisis dan eksperimen yang telah dilakukan pada bab-bab sebelumnya, dapat disimpulkan bahwa algoritma *Naïve Bayes* merupakan metode klasifikasi yang efektif dalam mengidentifikasi jenis bunga *Iris* berdasarkan fitur morfologis, yaitu panjang dan lebar kelopak (*sepal*) serta mahkota bunga (*petal*). Proses klasifikasi dilakukan dengan memanfaatkan pendekatan probabilistik yang mengandalkan distribusi data numerik untuk menghasilkan keputusan yang akurat dan efisien. Penelitian ini menunjukkan bahwa metode *Naïve Bayes* mampu melakukan klasifikasi secara cepat dengan tingkat ketepatan yang tinggi, terutama dalam kasus pengenalan spesies yang memiliki kemiripan karakteristik morfologi, seperti pada spesies *Iris-Versicolor*. Salah satu temuan penting dari penelitian ini adalah bahwa sistem klasifikasi dapat secara otomatis mengenali spesies bunga dengan tingkat kepercayaan tinggi. Hal ini dibuktikan melalui hasil perhitungan probabilitas tertinggi yang mengarah pada kelas *Iris-Versicolor*, dengan nilai $P(\text{Versicolor}|X) = 1$. Nilai tersebut mengindikasikan bahwa data uji sangat sesuai dengan ciri khas dari spesies tersebut. Keunggulan lain dari metode ini adalah kemampuannya dalam mengurangi potensi kesalahan identifikasi yang umum terjadi dalam proses taksonomi manual, sehingga dapat diadopsi sebagai solusi teknologi yang mendukung klasifikasi spesies berbasis morfologi. Dengan demikian, dapat disimpulkan bahwa algoritma *Naïve Bayes* tidak hanya unggul dalam hal kecepatan dan akurasi, tetapi juga sangat aplikatif dalam mendukung otomatisasi klasifikasi spesies tanaman secara efektif.

REFERENCES

- [1] A. Sabri, "Transfer Learning Model CNN Pralatih Untuk Klasifikasi Bunga Iris Berbasis Citra," *Prosiding Seminar SeNTIK*, vol. 7, no. 1, pp. 71–77, 2023, Accessed: Jun. 21, 2025. [Online]. Available: <https://ejournal.jakstik.ac.id/index.php/sentik/article/view/3416>.
- [2] P. Putra, A. M. H. Pardede, and S. Syahputra, "Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga," *JTIK (Jurnal Teknik Informatika Kaputama)*, vol. 6, no. 1, pp. 297–305, 2022, doi: 10.59697/jtik.v6i1.401.
- [3] Alvina Felicia Watratan, Arwini Puspita. B, and Dikwan Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," *Journal of Applied Computer Science and Technology*, vol. 1, no. 1, pp. 7–14, 2020, doi: 10.52158/jacost.v1i1.9.
- [4] M. V. Anand, B. Kiranbala, S. R. Srividhya, C. Kavitha, and M. Younus, "Gaussian Naïve Bayes Algorithm : A Reliable Technique Involved in the Assortment of the Segregation in Cancer," vol. 2022, 2022, doi: 10.1155/2022/2436946.
- [5] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.



- [6] I. W. Saputro and B. W. Sari, “Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa,” *Creative Information Technology Journal*, vol. 6, no. 1, p. 1, 2020, doi: 10.24076/citec.2019v6i1.178.
- [7] I. R. Anita Desiani, Irmeilyana Irmeilyana, Herlina Hanum, Yuli Andriani, Sri Indra Maiyanti, Clarita Margo Uteh, “Penerapan Metode Support Vector Machine Dalam Klasifikasi Bunga Iris,” *AJAI (Indonesia Jurnal of Applied Informatics)*, vol. 7 No. 1, 2022, doi: <https://doi.org/10.20961/ijai.v7i1.61486>.
- [8] Z. S. Muhammad Ridho Pradana, Rio Septian Hardinata, “Analisa Algoritma Support Vector Machine Pada Data Bunga Iris,” *Jurnal Dharma Agung*, vol. 30, no. April, pp. 477–487, 2022.
- [9] E. Etriyanti, D. Syamsuar, and N. Kunang, “Implementasi Data Mining Menggunakan Algoritma Naive Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa,” *Telematika*, vol. 13, no. 1, pp. 56–67, 2020, doi: 10.35671/telematika.v13i1.881.
- [10] F. Harahap, N. E. Saragih, E. T. Siregar, and H. Sariangsah, “Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat,” *Jurnal Ilmiah Informatika*, vol. 9, no. 01, pp. 19–23, 2021, doi: 10.33884/jif.v9i01.3702.
- [11] F. Farahdinna and M. N. Shofy, “Implementasi K-Nearest Neighbor Untuk Klasifikasi Bunga Iris,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 2, pp. 3510–3514, 2025, doi: 10.36040/jati.v9i2.13510.
- [12] F. D. Noviantha, I. F. Astuti, and A. H. Kridalaksana, “Sistem Pendukung Keputusan Untuk Penentuan Kategori Uang Kulliah Tunggal Dengan Metode Multifactor Evaluation Process (Studi Kasus : Universitas Mulawarman),” *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 13, no. 2, p. 88, 2019, doi: 10.30872/jim.v13i2.1552.
- [13] Z. Rahman, Z. Fatah, and J. D. Prasetyo, “Klasifikasi Spesies Bunga Iris Menggunakan Algoritma Klasifikasi Knn Di Rapidminer,” *Jurnal Ilmiah Multidisiplin Ilmu*, vol. 1, no. 6, pp. 60–66, 2024, doi: 10.69714/0syd5n74.
- [14] D. Rofianto, O. Arifin, and D. K. Widyawati, “Perbandingan Metode Klasifikasi Learning Vector Quantization Dengan Diskriminan Fisher Pada Data Bunga Iris,” *JSR: Jaringan Sistem Informasi Robotik*, vol. 7, no. 1, pp. 44–49, 2023, doi: 10.58486/jsr.v7i1.209.
- [15] N. E. Paulina, Z. E. Fitri, A. Madjid, and A. M. N. Imron, “Klasifikasi Kerusakan Mutu Tomat Berdasarkan Seleksi Fitur Menggunakan K-Nearest Neighbor,” *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 6, no. 2, pp. 144–154, 2021, doi: 10.26760/mindjournal.v6i2.144-154.
- [16] I. W. S. A. Nugraha, “Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means,” *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 2, pp. 234–244, 2023, doi: 10.5281/zenodo.7567622.
- [17] S. W. Harjono, N. W. Utami, and I. G. A. P. D. Putri, “Klasterisasi Tingkat Penjualan pada Startup Panak. id dengan Algoritma K-Means,” *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 17, no. 1, pp. 55–66, 2023, Accessed: Jun. 21, 2025. [Online]. Available: <https://pdfs.semanticscholar.org/4346/fa99cc56a0bea38f917923ccb6d0aca4f228.pdf>
- [18] M. Rosadi, D. A. Nurhasanah, and M. S. Hasibuan, “Clustering Panjang Ruas Jalan di BBPJN Sumut Menggunakan Algoritma K-Means,” *Journal of Computer Science and Informatics Engineering*, vol. 2, no. 1, pp. 29–38, 2023, doi: 10.55537/cosie.v2i1.567.
- [19] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *Jurnal Tekno Kompak*, vol. 15, no. 1, pp. 131–145, 2021, doi: 10.33365/jtk.v15i1.744.
- [20] M. K. K. Insan, U. Hayati, and O. Nurdiawan, “Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 478–483, 2023, doi: 10.36040/jati.v7i1.6373.