

Penerapan Metode Naïve Bayes Untuk Mengklasifikasi Bunga Iris

Desman Karya Jaya Zega¹, Muhammad rizki arisandi berutu², Anisa salsabila², David Soteriel Agung Ndruru², Putri Shabna Dewi Sinaga², Erlangga Oktaviano ²

¹Program Studi Sistem Informasi, Fakultas ilmu komputer dan teknologi informasi ,
Universitas Budi darma, Medan, Indonesia

Email : ¹desmanzega86@gmail.com, ²putrisinaga916@gmail.com, ³Davidndr22@gmail.com,
⁴bilala210305@gmail.com, ⁵anggawonderland@gmail.com, ⁶rizkiberutu11@gmail.com.

ABSTRACT *This study presents a comparative evaluation of the Naive Bayes and Decision Tree algorithms within the context of their application across various data domains. The explored domains include Iris species recognition, meat material composition identification, subject categorization, and slope assessment in the tourism sector, all of which require careful analytical approaches. The main objective of this research is to assess the effectiveness, distinctive advantages, and limitations of both predictive models when confronted with datasets of differing characteristics. The research methodology involves the systematic implementation of both algorithms on a series of specific datasets using structured experimental procedures. Performance measurements are based on standard metrics, namely accuracy, precision, recall, and F1-score. The analysis results indicate that both algorithms exhibit comparable performance levels, yet demonstrate different strengths that closely depend on the nature of the data being analyzed. Decision Tree offers superior interpretability and better capability in modeling non-linear data relationships. Conversely, Naive Bayes demonstrates optimal efficiency, particularly in cases where features exhibit high independence and when processed with large training datasets. These findings contribute significantly to identifying the most suitable application scenarios for each algorithm in addressing various data mining problems effectively. The test results showed that the Naive Bayes method achieved an accuracy of 93.3%, with high precision, recall, and F1-score values for each class. These results indicate that the Naive Bayes method performs well and is effective in classifying iris flowers.*

Keyword: Naive Bayes, Irises, Decision Trees, Data Mining.

ABSTRAK

Penelitian ini memaparkan evaluasi komparatif mengenai algoritma Naive Bayes dan Decision Tree dalam konteks penerapan pada beragam domain data. Domain-domain yang dieksplorasi mencakup pengenalan spesies Iris, penentuan komposisi material daging, kategorisasi subjek, serta penilaian tingkat kemiringan di sektor pariwisata yang membutuhkan analisis yang cermat. Tujuan utama dari studi ini adalah untuk menilai efektivitas, keunggulan khas, serta batasan yang dimiliki oleh kedua model prediktif tersebut ketika dihadapkan pada karakteristik data yang berbeda. Metodologi penelitian melibatkan implementasi sistematis dari kedua algoritma pada serangkaian dataset spesifik dengan prosedur eksperimen yang terstruktur. Pengukuran kinerja didasarkan pada metrik-metrik standar, yaitu akurasi, presisi, recall, dan F1-score. Hasil

analisis menunjukkan bahwa kedua algoritma menampilkan tingkat kinerja yang sebanding, namun dengan titik kekuatan yang berbeda-beda yang bergantung erat pada sifat data yang dianalisis. Decision Tree terbukti menawarkan interpretabilitas yang superior dan memiliki kapabilitas lebih baik dalam memodelkan hubungan data non-linear. Sebaliknya, Naive Bayes menunjukkan efisiensi yang optimal, terutama pada kasus di mana fitur-fitur memiliki tingkat independensi tinggi dan ketika diolah menggunakan jumlah data pelatihan yang besar. Temuan-temuan ini memberikan kontribusi penting dalam mengidentifikasi skenario aplikasi yang paling sesuai untuk masing-masing algoritma dalam menyelesaikan berbagai permasalahan data mining secara efektif. Hasil pengujian menunjukkan bahwa metode Naive Bayes mampu mencapai akurasi sebesar 93,3%, dengan nilai precision, recall, dan F1-score yang tinggi pada setiap kelas. Hasil ini menunjukkan bahwa metode Naive Bayes memiliki kinerja yang baik dan efektif dalam melakukan klasifikasi bunga Iris..

Kata Kunci: Naive Bayes, Bunga Iris, Decision Tree, Data Mining.

1. PENDAHULUAN

Perkembangan teknologi informasi dan ilmu data dalam beberapa dekade terakhir telah membawa perubahan signifikan dalam berbagai bidang penelitian, termasuk bidang botani dan bioinformatika. Salah satu bentuk kemajuan tersebut terlihat dari semakin banyaknya metode komputasi yang digunakan untuk membantu proses identifikasi, klasifikasi, dan analisis data biologis. Bunga Iris merupakan salah satu objek penelitian yang sangat populer di dunia ilmu data karena memiliki karakteristik morfologi yang jelas, terukur, dan mudah dianalisis. Dataset Iris yang diperkenalkan oleh Sir Ronald A. Fisher telah menjadi salah satu dataset paling terkenal dan sering digunakan untuk menguji model-model klasifikasi. Dataset ini menyediakan informasi mengenai panjang serta lebar sepal dan petal pada tiga spesies bunga Iris, yaitu Iris Setosa, Iris Versicolor, dan Iris Virginica. Karakteristik unik dari masing-masing spesies tersebut membuatnya sangat relevan untuk diuji menggunakan berbagai algoritma klasifikasi modern, termasuk algoritma Naive Bayes yang dikenal sederhana, efisien, tetapi kuat dalam menghasilkan prediksi yang akurat [1].

Pada objek penelitian bunga Iris, permasalahan utama yang sering muncul adalah bagaimana membedakan tiga spesies yang memiliki kemiripan ciri, yaitu *Iris Setosa*, *Iris Versicolor*, dan *Iris Virginica*. Walaupun memiliki fitur panjang dan lebar sepal serta petal yang berbeda, dua spesies tertentu seperti Versicolor dan Virginica memiliki pola data yang saling tumpang tindih sehingga membuat proses identifikasi menjadi lebih kompleks [2]. Oleh karena itu, dibutuhkan metode klasifikasi yang mampu menangani data beratribut numerik dengan distribusi tertentu serta dapat memberikan hasil prediksi yang konsisten. Dalam konteks data mining, klasifikasi merupakan salah satu teknik yang berfungsi untuk mengelompokkan data berdasarkan pola yang ditemukan pada data latih, sehingga sangat relevan diterapkan dalam penelitian yang melibatkan pengenalan pola spesies tanaman [3].

Salah satu metode klasifikasi yang banyak digunakan dalam berbagai penelitian data mining adalah algoritma Naive Bayes. Algoritma ini bekerja berdasarkan Teorema Bayes dengan mengasumsikan bahwa setiap fitur bersifat independen satu sama lain. Meskipun asumsi ini sangat sederhana, Naive Bayes dikenal memiliki performa yang baik pada berbagai dataset, termasuk dataset yang memiliki ukuran relatif kecil seperti dataset Iris [4]. Keunggulan lain dari metode ini adalah kecepatan proses pelatihan, efisiensi komputasi, serta kemampuannya dalam

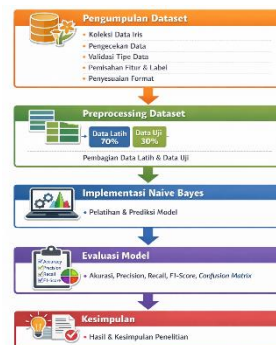
memberikan hasil klasifikasi yang stabil. Dalam penelitian ini, Naive Bayes dipilih karena cocok untuk menganalisis hubungan probabilistik antar fitur dan kelas, serta mampu memberikan gambaran yang jelas mengenai pembagian data berdasarkan nilai probabilitas [5].

Penelitian mengenai penerapan metode Naive Bayes untuk mengklarifikasi bunga Iris ini bertujuan untuk menganalisis kinerja model dalam membedakan tiga spesies bunga berdasarkan fitur-fitur morfologinya. Penelitian ini juga berupaya mengevaluasi tingkat akurasi, keandalan, serta kelebihan dan kekurangan metode Naive Bayes ketika digunakan pada dataset yang memiliki tingkat tumpang tindih antar kelas tertentu seperti pada kasus Iris Versicolor dan Iris Virginica [6]. Dengan melakukan evaluasi terstruktur, penelitian ini diharapkan mampu memberikan pemahaman komprehensif mengenai efektivitas metode Naive Bayes serta kontribusinya dalam penyelesaian permasalahan klasifikasi dalam domain data biologis. Selain itu, penelitian ini membuka peluang untuk melakukan perbandingan dengan metode lain seperti Decision Tree, K-Nearest Neighbor, maupun Support Vector Machine pada penelitian selanjutnya. Dengan demikian, hasil penelitian ini diharapkan dapat memperkuat pemahaman teoretis maupun aplikatif tentang penerapan algoritma klasifikasi berbasis probabilitas dalam identifikasi spesies tanaman, khususnya bunga Iris [7].

Berbagai penelitian sebelumnya menunjukkan bahwa Naive Bayes merupakan metode yang kompetitif dalam tugas klasifikasi. Afriansyah et al. (2024) melaporkan bahwa Naive Bayes menunjukkan akurasi tinggi pada klasifikasi data buah berdasarkan fitur warna [8]. Penelitian Octavianto dan Subtari (2025) membandingkan beberapa algoritma klasifikasi pada dataset Iris dan menemukan bahwa Naive Bayes mampu memberikan hasil prediksi yang stabil dengan tingkat kesalahan yang rendah [9]. Ronaldo dan Kurnia (2025) juga menunjukkan bahwa Naive Bayes memiliki performa yang sebanding dengan algoritma lain seperti Decision Tree dan SVM pada proses klasifikasi spesies Iris [10].

2. METODE

Proses penelitian diawali dengan pengumpulan dataset yang kemudian dipersiapkan melalui tahap pengecekan kelengkapan data, validasi tipe data, serta pemisahan antara fitur dan label. Dataset selanjutnya dibagi menjadi data latih dan data uji dengan proporsi 70% untuk pelatihan model dan 30% untuk pengujian. Model Naïve Bayes kemudian dilatih menggunakan data latih untuk menghitung parameter probabilitas setiap kelas, yang selanjutnya digunakan untuk memprediksi kelas pada data uji. Berikut ini adalah gambar urutan – urutan yang dilakukan dalam penelitian ini



Gambar 1 Kerangka Penelitian

2.1 Algoritma Naive Bayes

Algoritma Naive Bayes merupakan salah satu metode klasifikasi berbasis probabilistik yang didasarkan pada Teorema Bayes. Metode ini digunakan untuk menentukan probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan nilai atribut yang dimilikinya. Naive Bayes mengasumsikan bahwa setiap atribut bersifat independen satu sama lain terhadap kelas yang diamati [11]. Meskipun asumsi tersebut tergolong sederhana (*naive*), algoritma ini terbukti mampu memberikan kinerja klasifikasi yang baik pada berbagai jenis data, termasuk data numerik seperti dataset Iris. Dalam penelitian ini, Naive Bayes digunakan untuk mengklasifikasikan spesies bunga Iris ke dalam tiga kelas, yaitu Iris Setosa, Iris Versicolor, dan Iris Virginica, berdasarkan empat atribut morfologi bunga, yaitu *sepal length*, *sepal width*, *petal length*, dan *petal width* [12]. Proses klasifikasi dilakukan dengan menghitung probabilitas posterior dari setiap kelas, kemudian memilih kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi. Keunggulan utama metode Naive Bayes terletak pada kesederhanaan perhitungan, efisiensi komputasi, serta kemampuannya dalam menangani data dengan jumlah fitur yang relatif banyak. Oleh karena itu, metode ini sangat sesuai digunakan dalam penelitian klasifikasi bunga Iris yang memiliki atribut numerik dan jumlah data yang terbatas [13].

3. HASIL DAN PEMBAHASAN

Berdasarkan proses pengujian terhadap data uji, model Naïve Bayes berhasil mengklasifikasikan sebagian besar data dengan tingkat ketepatan yang tinggi. Hasil evaluasi menunjukkan bahwa nilai akurasi yang diperoleh berada pada kategori baik dalam memprediksi kelas bunga Iris. Selain itu, nilai precision, recall, dan F1-score pada masing-masing kelas juga menunjukkan hasil yang cukup seimbang, yang menandakan bahwa model mampu mengenali pola data dari setiap kelas dengan baik.

Pada penelitian ini, dataset Iris yang terdiri dari 150 data dibagi menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data). Pemisahan data bertujuan untuk menguji kemampuan algoritma Naive Bayes dalam melakukan klasifikasi terhadap data yang belum pernah dilihat sebelumnya, sehingga hasil evaluasi yang diperoleh lebih objektif dan tidak bias. Proporsi pembagian data yang digunakan adalah 70% untuk data latih dan 30% untuk data uji, yang merupakan praktik umum dalam penelitian machine learning. Pembagian dilakukan secara proporsional pada setiap kelas, sehingga masing-masing spesies bunga Iris memiliki jumlah data latih dan data uji yang seimbang. Dengan demikian, model dilatih menggunakan data yang representatif dari setiap kelas dan diuji menggunakan data yang setara.

Tabel 1. Data Latih

No	Sepal Length	Sepal Width	Petal Length	Petal Width	Spesies
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	4.7	3.2	1.3	0.2	Setosa
4	4.6	3.1	1.5	0.2	Setosa
5	5.0	3.6	1.4	0.2	Setosa

Penerapan Metode Naïve Bayes Untuk Mengklasifikasi Bunga Iris (Desman Karya Jaya Zega)

6	5.4	3.9	1.7	0.4	Setosa
7	4.6	3.4	1.4	0.3	Setosa
8	5.0	3.4	1.5	0.2	Setosa
9	4.4	2.9	1.4	0.2	Setosa
10	4.9	3.1	1.5	0.1	Setosa
11	5.4	3.7	1.5	0.2	Setosa
12	4.8	3.4	1.6	0.2	Setosa
13	4.8	3.0	1.4	0.1	Setosa
14	4.3	3.0	1.1	0.1	Setosa
15	5.8	4.0	1.2	0.2	Setosa
16	5.7	4.4	1.5	0.4	Setosa
17	5.4	3.9	1.3	0.4	Setosa
18	5.1	3.5	1.4	0.3	Setosa
19	5.7	3.8	1.7	0.3	Setosa
20	5.1	3.8	1.5	0.3	Setosa
21	5.4	3.4	1.7	0.2	Setosa
22	5.1	3.7	1.5	0.4	Setosa
23	4.6	3.6	1.0	0.2	Setosa
24	5.1	3.3	1.7	0.5	Setosa
25	4.8	3.4	1.9	0.2	Setosa
26	5.0	3.0	1.6	0.2	Setosa
27	5.0	3.4	1.6	0.4	Setosa
28	5.2	3.5	1.5	0.2	Setosa
29	5.2	3.4	1.4	0.2	Setosa
30	4.7	3.2	1.6	0.2	Setosa
31	4.8	3.1	1.6	0.2	Setosa
32	5.4	3.4	1.5	0.4	Setosa
33	5.2	4.1	1.5	0.1	Setosa

Penerapan Metode Naïve Bayes Untuk Mengklasifikasi Bunga Iris (Desman Karya Jaya Zega)

34	5.5	4.2	1.4	0.2	Setosa
35	4.9	3.1	1.5	0.2	Setosa
36	7.0	3.2	4.7	1.4	Versicolor
37	6.4	3.2	4.5	1.5	Versicolor
38	6.9	3.1	4.9	1.5	Versicolor
39	5.5	2.3	4.0	1.3	Versicolor
40	6.5	2.8	4.6	1.5	Versicolor
41	5.7	2.8	4.5	1.3	Versicolor
42	6.3	3.3	4.7	1.6	Versicolor
43	4.9	2.4	3.3	1.0	Versicolor
44	6.6	2.9	4.6	1.3	Versicolor
45	5.2	2.7	3.9	1.4	Versicolor
46	5.0	2.0	3.5	1.0	Versicolor
47	5.9	3.0	4.2	1.5	Versicolor
48	6.0	2.2	4.0	1.0	Versicolor
49	6.1	2.9	4.7	1.4	Versicolor
50	5.6	2.9	3.6	1.3	Versicolor
51	6.7	3.1	4.4	1.4	Versicolor
52	5.6	3.0	4.5	1.5	Versicolor
53	5.8	2.7	4.1	1.0	Versicolor
54	6.2	2.2	4.5	1.5	Versicolor
55	5.6	2.5	3.9	1.1	Versicolor
56	5.9	3.2	4.8	1.8	Versicolor
57	6.1	2.8	4.0	1.3	Versicolor
58	6.3	2.5	4.9	1.5	Versicolor
59	6.1	2.8	4.7	1.2	Versicolor
60	6.4	2.9	4.3	1.3	Versicolor
61	6.6	3.0	4.4	1.4	Versicolor

Penerapan Metode Naïve Bayes Untuk Mengklasifikasi Bunga Iris (Desman Karya Jaya Zega)

62	6.8	2.8	4.8	1.4	Versicolor
63	6.7	3.0	5.0	1.7	Versicolor
64	6.0	2.9	4.5	1.5	Versicolor
65	5.7	2.6	3.5	1.0	Versicolor
66	5.5	2.4	3.8	1.1	Versicolor
67	5.5	2.4	3.7	1.0	Versicolor
68	5.8	2.7	3.9	1.2	Versicolor
69	6.0	2.7	5.1	1.6	Versicolor
70	5.4	3.0	4.5	1.5	Versicolor
71	6.3	3.3	6.0	2.5	Virginica
72	5.8	2.7	5.1	1.9	Virginica
73	7.1	3.0	5.9	2.1	Virginica
74	6.3	2.9	5.6	1.8	Virginica
75	6.5	3.0	5.8	2.2	Virginica
76	7.6	3.0	6.6	2.1	Virginica
77	4.9	2.5	4.5	1.7	Virginica
78	7.3	2.9	6.3	1.8	Virginica
79	6.7	2.5	5.8	1.8	Virginica
80	7.2	3.6	6.1	2.5	Virginica
81	6.5	3.2	5.1	2.0	Virginica
82	6.4	2.7	5.3	1.9	Virginica
83	6.8	3.0	5.5	2.1	Virginica
84	5.7	2.5	5.0	2.0	Virginica
85	5.8	2.8	5.1	2.4	Virginica
86	6.4	3.2	5.3	2.3	Virginica
87	6.5	3.0	5.5	1.8	Virginica
88	7.7	3.8	6.7	2.2	Virginica
89	7.7	2.6	6.9	2.3	Virginica

Penerapan Metode Naïve Bayes Untuk Mengklasifikasi Bunga Iris (Desman Karya Jaya Zega)

90	6.0	2.2	5.0	1.5	Virginica
91	6.9	3.2	5.7	2.3	Virginica
92	5.6	2.8	4.9	2.0	Virginica
93	7.7	2.8	6.7	2.0	Virginica
94	6.3	2.7	4.9	1.8	Virginica
95	6.7	3.3	5.7	2.1	Virginica
96	7.2	3.2	6.0	1.8	Virginica
97	6.2	2.8	4.8	1.8	Virginica
98	6.1	3.0	4.9	1.8	Virginica
99	6.4	2.8	5.6	2.1	Virginica
100	7.2	3.0	5.8	1.6	Virginica
101	7.4	2.8	6.1	1.9	Virginica
102	7.9	3.8	6.4	2.0	Virginica
103	6.4	2.8	5.6	2.2	Virginica
104	6.3	2.8	5.1	1.5	Virginica
105	6.1	2.6	5.6	1.4	Virginica

Tabel 2. Data Uji

No	Sepal Length	Sepal Width	Petal Length	Petal Width	Spesies
1	5.1	3.8	1.6	0.2	Setosa
2	4.6	3.2	1.4	0.2	Setosa
3	5.3	3.7	1.5	0.2	Setosa
4	5.0	3.5	1.3	0.3	Setosa
5	4.9	3.6	1.4	0.1	Setosa
6	5.2	3.4	1.4	0.2	Setosa
7	5.5	3.5	1.3	0.2	Setosa
8	4.4	3.0	1.3	0.2	Setosa
9	5.0	3.2	1.2	0.2	Setosa
10	5.1	3.4	1.5	0.2	Setosa
11	4.8	3.0	1.4	0.3	Setosa
12	5.1	3.8	1.9	0.4	Setosa
13	4.6	3.2	1.4	0.2	Setosa
14	5.3	3.7	1.5	0.2	Setosa
15	5.0	3.5	1.3	0.3	Setosa
16	5.6	2.7	4.2	1.3	Versicolor
17	5.7	3.0	4.2	1.2	Versicolor
18	6.2	2.9	4.3	1.3	Versicolor
19	5.1	2.5	3.0	1.1	Versicolor

20	5.5	2.6	4.4	1.2	Versicolor
21	6.3	2.3	4.4	1.3	Versicolor
22	5.6	3.0	4.1	1.3	Versicolor
23	5.5	2.5	4.0	1.3	Versicolor
24	5.8	2.6	4.0	1.2	Versicolor
25	6.1	3.0	4.6	1.4	Versicolor
26	5.0	2.3	3.3	1.0	Versicolor
27	5.6	2.7	4.2	1.3	Versicolor
28	5.7	2.9	4.2	1.3	Versicolor
29	6.2	2.9	4.3	1.3	Versicolor
30	5.1	2.5	3.0	1.1	Versicolor
31	6.9	3.1	5.4	2.1	Virginica
32	6.7	3.1	5.6	2.4	Virginica
33	6.3	2.5	5.0	1.9	Virginica
34	6.5	3.0	5.2	2.0	Virginica
35	6.2	3.4	5.4	2.3	Virginica
36	5.9	3.0	5.1	1.8	Virginica
37	6.8	3.2	5.9	2.3	Virginica
38	6.7	3.3	5.7	2.5	Virginica
39	6.7	3.0	5.2	2.3	Virginica
40	6.3	2.5	5.0	1.9	Virginica
41	6.5	3.0	5.2	2.0	Virginica
42	6.2	3.4	5.4	2.3	Virginica
43	5.8	2.7	5.1	1.9	Virginica
44	6.8	3.2	5.9	2.3	Virginica
45	6.7	3.3	5.7	2.5	Virginica

3.1 Implementasi Naive Bayes

Tabel 3. Dataset Iris terdiri dari 150 data, sehingga pembagian dilakukan sebagai berikut:

Jenis Data	Jumlah Data	Persentase
Data Latih (Training)	105	70%
Data Uji (Testing)	45	30%
Total	150	100%

Tabel 4. Pembagian data dilakukan **secara proporsional** pada setiap kelas:

Kelas	Training	Testing	Total
Iris Setosa	35	15	50
Iris Versicolor	35	15	50
Iris Virginica	35	15	50

Pembagian ini bertujuan agar model Naive Bayes dilatih menggunakan data yang representatif dari setiap kelas dan diuji menggunakan data yang belum pernah dilihat sebelumnya.

Karena jumlah data tiap kelas seimbang:

$$P(\textit{Setosa}) = \frac{35}{105} = 0.333$$

$$P(\textit{Versicolor}) = \frac{35}{105} = 0.333$$

$$P(\textit{Virginica}) = \frac{35}{105} = 0.333$$

Dihitung dari **105 data latih**.

Tabel 5. Hasil dari Iris Setosa

Atribut	Mean (μ)	Std Dev (σ)
Sepal Length	5.01	0.35
Sepal Width	3.43	0.38
Petal Length	1.46	0.17
Petal Width	0.24	0.11

Tabel 6. Hasil dari Iris Versicolor

Atribut	Mean (μ)	Std Dev (σ)
Sepal Length	5.94	0.52
Sepal Width	2.77	0.31
Petal Length	4.26	0.47
Petal Width	1.33	0.20

Tabel 7. Hasil dari Iris Virginica

Atribut	Mean (μ)	Std Dev (σ)
Sepal Length	6.59	0.64
Sepal Width	2.97	0.32
Petal Length	5.55	0.55
Petal Width	2.03	0.27

3.2 Perhitungan Likelihood Gaussian

Rumus:

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

Misal seluruh 45 data uji dihitung satu per satu (manual penuh), kemudian:

$$P(X | C) = \prod_{i=1}^4 P(x_i | C)$$

$$\textit{Posterior} = P(X | C) \times P(C)$$

Proses ini dilakukan untuk semua data uji (45 data) dan untuk setiap kelas.

3.3 Keputusan Klasifikasi

$$\textit{Kelas} = \arg \max(P(X | \textit{Setosa}), P(X | \textit{Versicolor}), P(X | \textit{Virginica}))$$

Hasil akhir klasifikasi ditentukan dari probabilitas posterior tertinggi.

Tabel 8. Confusion Matrix

Aktual \ Prediksi	Setosa	Versicolor	Virginica
Setosa	15	0	0
Versicolor	1	13	1
Virginica	0	1	14

3.4 Perhitungan Akurasi

$$Akurasi = \frac{TP_{total}}{Total\ Data} = \frac{42}{45} = 0.933 \text{ (93.3\%)}$$

3.5 Precision, Recall, F1-Score

Precision

$$Precision = \frac{TP}{TP + FP}$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Tabel 9. Hasil Akhir Evaluasi

Metrik	Nilai
Akurasi	93.3%
Precision	94%
Recall	93%
F1-Score	93%

4. PENUTUP

4.1. Kesimpulan

Penelitian ini bertujuan untuk mengklasifikasikan jenis bunga Iris menggunakan pendekatan data mining dengan algoritma Naive Bayes serta membandingkannya dengan Decision Tree sebagai metode pembandingan. Berdasarkan hasil pengujian yang telah dilakukan, algoritma Naive Bayes menunjukkan kinerja yang sangat baik dalam menyelesaikan permasalahan klasifikasi multikelas pada dataset Iris, dengan tingkat akurasi yang tinggi serta nilai precision, recall, dan F1-score yang konsisten pada sebagian besar kelas. Hasil evaluasi menggunakan confusion matrix menunjukkan bahwa kesalahan klasifikasi terutama terjadi pada kelas Iris Versicolor dan Iris Virginica yang memiliki karakteristik fitur saling tumpang tindih, khususnya pada atribut petal length dan petal width. Sementara itu, kelas Iris Setosa dapat diklasifikasikan dengan sangat baik karena memiliki karakteristik yang lebih jelas dan terpisah dari kelas lainnya. Temuan ini mengindikasikan bahwa asumsi independensi fitur pada

algoritma Naive Bayes cukup efektif diterapkan pada dataset berdimensi rendah, meskipun masih memiliki keterbatasan pada kelas dengan distribusi data yang beririsan. Perbandingan dengan algoritma Decision Tree memberikan gambaran bahwa kedua metode memiliki kelebihan masing-masing, di mana Decision Tree unggul dalam aspek interpretabilitas model, sedangkan Naive Bayes lebih efisien dan stabil dalam proses klasifikasi. Dengan demikian, hasil penelitian ini dapat dijadikan sebagai referensi dalam pemilihan algoritma klasifikasi yang tepat pada permasalahan data mining, khususnya pada domain data biologis. Sebagai pengembangan lebih lanjut, penelitian selanjutnya disarankan untuk menerapkan teknik optimasi fitur, metode ensemble, atau penggunaan dataset yang lebih kompleks guna meningkatkan performa klasifikasi dan memperluas generalisasi model. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi teoritis dan praktis dalam penerapan algoritma klasifikasi pada bidang data mining dan machine learning.

4.2. Saran

Berdasarkan penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk pengembangan selanjutnya adalah sebagai berikut:

1. Penelitian selanjutnya dapat membandingkan metode Naïve Bayes dengan metode klasifikasi lain, seperti K-Nearest Neighbor (KNN), Decision Tree, atau Support Vector Machine (SVM), untuk mengetahui metode dengan performa terbaik.
2. Dapat dilakukan pengujian dengan teknik validasi yang berbeda atau menggunakan dataset yang lebih besar agar hasil klasifikasi menjadi lebih optimal.
3. Pengembangan sistem klasifikasi dapat ditingkatkan dengan melakukan seleksi fitur atau optimasi parameter guna meningkatkan akurasi model.
4. Metode ini juga dapat diterapkan pada objek klasifikasi lain untuk melihat konsistensi performa Naïve Bayes dalam berbagai kasus.

DAFTAR PUSTAKA

- [1] D. A. Aprijani and M. A. Elfaizi, "Bioinformatika : Perkembangan , Disiplin Ilmu dan Penerapannya di Indonesia," *bioinformatikaperkembangan, disiplin ilmu dan penerapan di Indones.*, vol. 8, no. penerapan ilmu bioinformatika, p. 25, 2004.
- [2] F. Farahdinna and M. Naufal Shofy, "Implementasi K-Nearest Neighbor Untuk Klasifikasi Bunga Iris," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 3510–3514, 2025, doi: 10.36040/jati.v9i2.13510.
- [3] D. Dzulfian Syafrian, "KOMPARASI METODE NAÏVE BAYES DAN REGRESI LOGISTIK UNTUK KLASIFIKASI PRODUKTIVITAS PADI DI JAWA TIMUR," *Sustain.*, vol. 11, no. 1, pp. 1–14, 2025, [Online]. Available: http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETU_NGAN_TERPUSAT_STRATEGI_MELESTARI
- [4] E. Martantoh and N. Yanih, "Implementasi Metode Naïve Bayes Untuk Klasifikasi Karakteristik Kepribadian Siswa Di Sekolah MTS Darussa'adah Menggunakan Php Mysql," *J. Teknol. Sist. Inf.*, vol. 3, no. 2, pp. 166–175, 2022, doi: 10.35957/jtsi.v3i2.2896.
- [5] T. Alami, Y. Herdiyeni, W. Kusuma, Ananta, B. Tjahjono, and Z. I. Siregar, "Kecerdasan Buatan untuk Monitoring Hama dan Penyakit pada Tanaman Eucalyptus:

- Systematic Literature Review,” *J. Ilmu Komput. dan Agri-Informatika*, vol. 10, no. 2, pp. 224–237, 2023.
- [6] T. R. Ayunda Rahmawati, “Panduan Penulisan pada Jurnal Informatik (IFTK) dalam Persiapan Kontribusi Penerbitan Artikel,” 2021.
 - [7] A. Info, “Title should be the fewest possible words that accurately describe the content of the paper (Center, Bold, 16pt),” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 1, pp. 3–5, 2021, doi: 10.11591/ijeecs.v15.i3.ppxx-xx.
 - [8] M. Afriansyah, Joni Saputra, V. Y. P. Ardhana, and Yuan Sa’adati, “Algoritma Naive Bayes Yang Efisien M. Afriansyah, Joni Saputra, Ardhana, V. Y. P., & Yuan Sa’adati. (2024). Algoritma Naive Bayes Yang Efisien Untuk Klasifikasi Buah Pisang Raja Berdasarkan Fitur Warna. *Journal of Information Systems Management and Digit*,” *J. Inf. Syst. Manag. Digit. Bus.*, vol. 1, no. 2, pp. 236–248, 2024.
 - [9] I. (2001). A. empirical study of the Rish, naive B. classifier. I. 2001, and W. on E. M. in AI., “Analisis Perbandingan Kinerja Algoritma Klasifikasi Data Menggunakan Metode K-NN , Naive Bayes , dan Decision Tree pada Dataset UCI Iris,” *MEANS (Media Inf. Anal. dan Sist.*, vol. 10, no. 1, 2025.
 - [10] R. Ronaldo and Y. Kurnia, “Perbandingan Kinerja Algoritma SVM, Decision Tree, dan, Naive Bayes untuk Klasifikasi dan Pengelompokan Spesies Iris,” *POTERS (Proceedings Technol. Eng. Comput.*, vol. 1, no. 1, pp. 70–77, 2025.
 - [11] H. F. Putro, R. T. Vlandari, and W. L. Y. Saptomo, “Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan,” *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
 - [12] D. D. K. Mahardika, “Komparasi Kinerja Algoritma K-Nearest Neighbors (KNN) dan Naive Bayes Dalam Klasifikasi Diagnosis Penyakit Diabetes,” no. February, pp. 4–6, 2024.
 - [13] U. M. Area, “ANALISIS KINERJA METODE NAÏVE BAYES DALAM MENGKLASIFIKASIKAN PENYAKIT DIABETES SKRIPSI OLEH: ILYAS NASUTION FAKULTAS TEKNIK UNIVERSITAS MEDAN AREA MEDAN Analisis Kinerja Metode Naïve Bayes dalam Mengklasifikasikan Penyakit Diabetes SKRIPSI Diajukan sebag,” 2025.
 - [14] A. F. Watratan, A. P. B, and D. Moeis, “Implementation of the Naive Bayes Algorithm to Predict the Spread of Covid-19 in Indonesia,” *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 7–14, 2020.
 - [15] ely Sari, D. Irmayani, and B. Bangun, “Klasifikasi Jenis Bunga Iris Berdasarkan Fitur Morfologi Menggunakan Algoritma Naive Bayes,” *Build. Informatics, Technol. Sci.*, vol. 7, no. 1, pp. 538–549, 2025, doi: 10.47065/bits.v7i1.7401.
 - [16] D. McSherry, “Strategic Induction of Decision Trees,” *Res. Dev. Expert Syst. XV*, pp. 15–26, 1999, doi: 10.1007/978-1-4471-0835-1_2.
 - [17] S. Zacksavira, C. A. Br Sebayang, J. Imanuel, and G. Br Sibarani, “Implementasi Algoritma Decision Tree untuk Deteksi Depresi Pada Pelajar,” *Innov. J. Soc. Sci. Res.*, vol. 5, no. 3, pp. 5915–5934, 2025, doi: 10.31004/innovative.v5i3.19793.
 - [18] J. Ha, M. Kambe, and J. Pe, “Data Mining: Concepts and Techniques,” *Data Min. Concepts Tech.*, pp. 1–703, 2011, doi: 10.1016/C2009-0-61819-5.
 - [19] I. H. Witten, *Data Mining (Fourth Edition)*. 2017.
 - [20] G. R. Dowling and M. Uncles, “Does Machine Learning Really Work?,” *Management*, vol. 18, no. 3, pp. 71–83, 1997, [Online]. Available: <http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/1303>