

**SKRIPSI**

**IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK  
MEMPREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS PT.  
PLASMEN SISUMUT)**

**DISUSUN OLEH KHAIRANI LISA NASUTION NPM.2009010033**



**UMSU**  
Unggul | Cerdas | Terpercaya

**PROGRAM STUDI SISTEM INFORMASI FAKULTAS ILMU  
KOMPUTER DAN TEKNOLOGI INFORMASI UNIVERSITAS  
MUHAMMADIYAH SUMATERA UTARA MEDAN**

**2025**

**IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK  
MEMPREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS PT.  
PLASMEN SISUMUT)**

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana  
Komputer (S.Kom) dalam Program Studi Sistem Informasi pada Fakultas  
Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah  
Sumatera Utara**

**KHAIRANI LISA NASUTION**

**NPM. 2009010033**

**PROGRAM STUDI SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI**

**UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

**MEDAN**

**2025**

## LEMBARAN PENGESAHAN

**Judul Skripsi** : IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK  
MEMPREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS  
PT. PLASMEN SISUMUT)

**Nama Mahasiswa** : KHAIRANI LISA NASUTION

**NPM** : 2009010033

**Program Studi** : SISTEM INFORMASI

Menyetujui

Dosen Pembimbing



Fatma Sari Hualung, S.Kom, M.Kom

NIDN. 0117019301

Ketua Program Studi



(Martiano, S.Pd, S.Kom., M.Kom)

NIDN. 0128029302

Dekan



(Dr. Al-Khewarizmi, S.Kom., M.Kom.)

NIDN. 0127099201

## PERNYATAAN ORISINALITAS

### IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MEMPREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS PT. PLASMEN SISUMUT)

#### SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 19 April 2025

Yang membuat pernyataan



KHAIRANI LISA NASUTION  
NPM. 2009010033

**PERNYATAAN PERSETUJUAN PUBLIKASI  
KARYA ILMIAH UNTUK KEPENTINGAN  
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : KHAIRANI LISA NASUTION  
NPM : 2009010033  
Program Studi : Sistem Informasi  
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

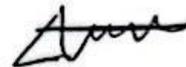
**IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK  
MEMREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS PT.  
PLASMEN SISUMUT)**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 19 April 2025

Yang membuat pernyataan



KHAIRANI LISA NASUTION

NPM. 2009010033

## **RIWAYAT HIDUP**

### DATA PRIBADI

Nama Lengkap : Khairani Lisa Nasution

Tempat dan Tanggal Lahir : Sei Rumbia kotapinang, labuhan batu selatan, sumatra utara.

Alamat Rumah : Dusun pekan sumut labuhan batu selatan kotapinang labuhan batu selatan.

Telepon/Faks/HP : 082185719859

E-mail : khaiaira493@gmail.com

Instansi Tempat Kerja : Belum Bekerja

Alamat Kantor : -

### DATA PENDIDIKAN

SD : SDN NEGRI 112241 LABUHAN BATU SELATAN

TAMAT : 2013

SMP : SMP NEGRI 2 KOTAPINANG LABUHAN BATU SELATAN

TAMAT : 2016

SMA : SMAN 1 KOTAPINANG LABUHAB BATU SELATAN

TAMAT : 2019

## ABSTRAK

Penyakit pada bibit kelapa sawit dapat berdampak signifikan terhadap produktivitas perkebunan kelapa sawit. Metode konvensional dalam mendeteksi penyakit sering kali membutuhkan waktu yang lama dan kurang efisien. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan algoritma Naïve Bayes dalam memprediksi penyakit pada bibit kelapa sawit berdasarkan karakteristik fisiknya. Data yang digunakan dalam penelitian ini meliputi atribut warna daun, struktur batang, kondisi tanah, serta beberapa gejala umum seperti bercak coklat, garis kuning, kelayuan, busuk batang, dan nekrosis. Model Naïve Bayes dikembangkan dan diuji menggunakan dataset yang diperoleh dari PT. Plasmen Sisumut, dengan evaluasi menggunakan confusion matrix untuk mengukur akurasi, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi yang cukup baik dalam mengklasifikasikan penyakit bibit sawit. Dengan adanya sistem ini, diharapkan petani dan pengelola perkebunan dapat mendeteksi penyakit lebih cepat dan mengambil tindakan pencegahan yang lebih efektif.

**Kata Kunci:** Naïve Bayes, Prediksi Penyakit, Bibit Kelapa Sawit, Data Mining, Confusion Matrix.

## **ABSTRACT**

*Diseases in oil palm seedlings can significantly impact plantation productivity. Conventional methods for detecting diseases are often time-consuming and inefficient. Therefore, this study aims to implement the Naïve Bayes algorithm to predict diseases in oil palm seedlings based on their physical characteristics. The dataset used in this research includes attributes such as leaf color, stem structure, soil conditions, and common symptoms like brown spots, yellow streaks, wilting, stem rot, and necrosis. The Naïve Bayes model was developed and tested using data obtained from PT. Plasmen Sisumut, with evaluation conducted using a confusion matrix to measure accuracy, precision, recall, and F1-score. The test results indicate that the developed model achieves a high level of accuracy in classifying oil palm seedling diseases. With this system, it is expected that farmers and plantation managers can detect diseases more quickly and take more effective preventive measures.*

**Keywords:** *Naïve Bayes, Disease Prediction, Oil Palm Seedlings, Data Mining, Confusion Matrix.*

## KATA PENGANTAR



Alhamdulillah, puji syukur kepada Allah SWT atas segala rahmat dan hidayah nya, sehingga penulis dapat menyelesaikan skripsi berjudul **“IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MEMREDIKSI PENYAKIT BIBIT SAWIT (STUDI KASUS PT. PLASMEN SISUMUT)”**. Skripsi ini disusun untuk memenuhi syarat mendapatkan gelar sarjana dalam program studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Universitas Muhammadiyah Sumatra Utara. Dalam proses penelitian dan penyusunan skripsi ini, banyak pelajaran dan tantangan yang dihadapi, yang semuanya memberikan manfaat di masa depan. Semua pencapaian ini tidak lepas dari dukungan dan motivasi dari banyak pihak. Oleh karena itu, penulis menyampaikan terima kasih sebesar-besarnya dan penghargaan setinggi tingginya kepada :

1. Bapak Prof. Dr. Agussani, M.AP., Selaku Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Bapak Martiano S.pd, S.Kom., M. Kom Selaku Ketua Program Studi Sistem Informasi yang selalu memberikan dukungan.
4. Bapak Halim Maulana, S.T, M.Kom., MTA Wakil Dekan 1 Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Ibu Fatma Sari Hutagalung, S.Kom, M.Kom selaku Dosen Pembimbing saya yang telah membimbing saya sampai ke tahap ini.
6. Teruntuk para Bapak/Ibu Dosen FIKTI (UMSU) atas ilmu pengetahuan yang telah diberikan kepada saya selama perkuliahan.
7. Seluruh Staff dan Keanggotaan Biro Kemahasiswaan yang mendukung dalam proses pengerjaan penelitian ini.
8. Teruntuk orang tua dan orang-orang tersayang saya, Khairul Heri Nasution, Masniar Harahap, Suci Lestari Nasution dan Ravindra Singh Dhillon yang selalu mendukung dan menyemangati perjalanan saya.

9. Teruntuk Teman-teman saya Angkatan Sistem Informasi 2020 yang sudah banyak membantu saya dalam segala hal dan memberikan dukungan yang lebih.
10. Teruntuk Diri sendiri. Saya ucapkan banyak terima kasih kepada diri saya sendiri yang telah berjuang dari semester awal hingga semester akhir yang telah memaksimalkan segala hal dengan maksimal dan baik, yang telah sabar dan mengikuti segala alur hidup yang rumit.

Dalam menyelesaikan tugas akhir ini saya sebagai peneliti tentu mempunyai banyak kekurangan dan kesalahan baik disengaja maupun tidak disengaja. Maka dari itu, saya akan sangat menghargai setiap masukan dan kritik untuk membangun diri saya agar lebih baik kedepannya. Akhir kata, saya memohon maaf sebesarbesarnya dan berharap tugas akhir ini dapat berguna dan memberikan manfaat.

Terimakasih Saya Ucapkan

Medan, 19 Mei 2025

Penyusun



Khairani Lisa Nasution

Npm : 2009010033

## DAFTAR ISI

DAFTAR ISI.....	xi
DAFTAR GAMBAR .....	xiv
DAFTAR TABEL .....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian .....	5
BAB II LANDASAN TEORI.....	7
2.1 Kelapa Sawit .....	7
2.2 Penyakit Tanaman Kelapa Sawit.....	7
2.3 Pengertian Sistem Informasi .....	8
2.4 Konsep <i>Data Mining</i> .....	9
2.5 Algoritma <i>Naïve Bayes</i> .....	9
2.6 Evaluasi Model dengan <i>Confussion Matrix</i> .....	11
2.6 Pra-Pemrosesan Data .....	12
2.7 <i>Website</i> .....	14
2.8 <i>Hypertext Markup Language (HTML)</i> .....	14
2.9 Hypertext Preprocessor (PHP) .....	14
2.10 XAMPP .....	14

2.11 <i>Database</i> .....	15
2.12 MySQL.....	15
2.13 <i>Flask API</i> .....	15
2.14 <i>Google Colaboratory</i> .....	16
2.15 <i>Pandas</i> .....	16
2.16 <i>Scikit-Learn</i> .....	16
2.17 Unified Modelling Language (UML).....	17
<b>BAB III METODE PENELITIAN</b> .....	<b>19</b>
3.1. Lokasi Penelitian.....	19
3.2. Objek Penelitian.....	19
3.3. Studi Literatur dan Pustaka .....	19
3.4. Pengumpulan Data .....	21
3.5. Bagan Alir ( <i>Flowchart</i> ).....	22
3.6. Pemeriksaan Data.....	24
3.7. Pembersihan dan Transformasi Data.....	25
3.8. Perhitungan dan Implementasi <i>Naïve Bayes</i> .....	27
3.9. Implementasi <i>Naïve Bayes</i> menggunakan <i>Python</i> .....	41
3.10. Evaluasi Algoritma Klasifikasi .....	41
Perhitungan presisi dari masing masing penyakit adalah sebagai berikut : .....	43
Perhitungan <i>recall</i> dari masing masing penyakit adalah sebagai berikut : .....	43
3.11. Desain Sistem.....	44
3.12. Perancangan <i>Interface</i> .....	54

3.13. Penulisan Kode Program Sistem .....	56
3.14. Pengujian Sistem .....	56
3.15. Integrasi Model Naïve Bayes .....	57
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>60</b>
4.1. Hasil Pelatihan Model <i>Naïve Bayes</i> .....	60
4.1.1. Dataset yang Digunakan .....	60
4.1.2. Perhitungan Probabilitas Prior .....	61
4.1.3. Perhitungan Probabilitas kondisional.....	62
4.1.4. Perhitungan Probabilitas Posterior .....	64
4.1.5. Evaluasi Model.....	67
4.2. Hasil Tampilan Aplikasi .....	79
4.2.1. Tampilan Menu <i>Login</i> .....	79
4.2.2. Tampilan Menu Dashboard .....	79
4.2.3. Tampilan Menu Bibit .....	80
4.2.4. Tampilan Menu Data Penyakit.....	81
4.2.5. Tampilan Menu Gejala .....	82
4.2.6. Tampilan Menu Dataset .....	85
4.2.7. Tampilan Menu Latih Model & Prediksi .....	87
4.2.8. Tampilan Lihat Hasil Prediksi.....	89
4.2.9. Tampilan Laporan Prediksi .....	90
4.3. Pengujian Program .....	90
4.3.1. Pengujian Login .....	91

4.3.2. Tabel Pengujian <i>Dashboard</i> .....	91
4.3.3. Tabel Pengujian Data Bibit .....	91
4.3.4. Pengujian Data Penyakit .....	92
4.3.5. Tabel Pengujian Data Prediksi .....	92
BAB V KESIMPULAN DAN SARAN .....	94
DAFTAR PUSTAKA.....	95

## DAFTAR GAMBAR

Gambar 3. 1 Flowchart Alur Penelitian .....	22
Gambar 3. 2 Kode Import Library Pandas .....	26
Gambar 3. 3 Kode memasukkan data kedalam Dataframe .....	26
Gambar 3. 4 Kode identifikasi dan menghapus nilai data yang hilang .....	27
Gambar 3. 5 Rancangan Use Case Diagram .....	44
Gambar 3. 6 Activity Diagram Login .....	46
Gambar 3. 7 Activity Diagram Data Bibit .....	47
Gambar 3. 8 Activity Diagram Prediksi Bibit .....	48
Gambar 3. 9 Activity Diagram Laporan .....	49
Gambar 3. 10 Activity Diagram Logout .....	50
Gambar 3. 11 Sequence Diagram.....	51
Gambar 3. 12 Class Diagram .....	52
Gambar 3. 13 Rancangan Form Login .....	53
Gambar 3. 14 Rancangan Form Dashboard .....	53
Gambar 3. 15 Rancangan Form Olah Data Bibit .....	54
Gambar 3. 16 Rancangan Form Prediksi .....	54

Gambar 3. 17 Rancangan Form Laporan .....	55
Gambar 3. 18 Cara Install Flask .....	56
Gambar 4. 1 Kode Perhitungan Probabilitas Kondisional .....	62
Gambar 4. 2 Kode untuk memprediksi Naive Bayes .....	65
Gambar 4. 3 Hasil Prediksi Dalam Confussion Matrix.....	73
Gambar 4. 4 Tampilan Menu Login .....	77
Gambar 4. 5 Tampilan Menu Tambah Data Bibit.....	79
Gambar 4. 6 Tampilan Menu Kelola Data Bibit .....	79
Gambar 4. 7 Tampilan Menu Tambah Penyakit .....	80
Gambar 4. 8 Tampilan Menu Kelola Penyakit.....	80
Gambar 4. 9 Tampilan Menu Tambah Gejala .....	81
Gambar 4. 10 Tampilan Daftar Gejala .....	81
Gambar 4. 11 Tampilan Tambah Atribut Gejala .....	82
Gambar 4. 12 Tampilan Menu Kelola Atribut Gejala.....	82
Gambar 4. 13 Tampilan Tambah Gejala Bibit .....	83
Gambar 4. 14 Tampilan Kelola Gejala Bibit .....	83
Gambar 4. 15 Tampilan Menu Tambah Dataset .....	84
Gambar 4. 16 Tampilan Daftar Dataset .....	84
Gambar 4. 17 Tampilan Lihat Dataset Detail .....	85
Gambar 4. 18 Dataset setelah difilter tanggal .....	86
Gambar 4. 19 Tampilan Probabilitas Prior .....	86
Gambar 4. 20 Tampilan Probabilitas Kondisional .....	86
Gambar 4. 21 Hasil Evaluasi Model .....	87
Gambar 4. 22 Tampilan Prediksi Model .....	87
 <b>DAFTAR TABEL</b>	
Tabel 3. 1 Studi Literatur dan Pustaka .....	20

Tabel 3. 2 Tabel Contoh Dataset .....	28
Tabel 3. 3 Tabel Contoh Confussion Matrix .....	41
Tabel 4. 1 10 Baris dataset yang digunakan .....	59
Tabel 4. 2 Tabel Hasil Perhitungan Probabilitas Kondisional .....	62
Tabel 4. 3 Tabel Data Sampel Bibit Sawit .....	63
Tabel 4. 4 Tabel Hasil Pengujian Login.....	89
Tabel 4. 5 Tabel Pengujian Dashboard .....	90
Tabel 4. 6 Tampilan Pengujian Data Bibit .....	90
Tabel 4. 7 Tabel Pengujian Data Penyakit .....	91
Tabel 4. 8 Tabel Pengujian Data Prediksi .....	91

## **BAB I PENDAHULUAN**

### **1.1 Latar Belakang**

Kelapa sawit merupakan sumber utama minyak nabati yang dapat dikonsumsi. Saat ini, kelapa sawit menjadi fokus utama dalam industri perkebunan, dengan luas penanaman mencapai hampir 15 juta hektar secara global. Karena keunggulannya sebagai sumber minyak nabati dan bahan baku agroindustri, kelapa sawit terus menjadi fokus utama dalam sektor pertanian. Menurut data terbaru dari Kementerian Pertanian, Indonesia menduduki posisi puncak sebagai negara penghasil kelapa sawit terluas di dunia. Hal ini menjadikan kelapa sawit sebagai komoditas penting dalam perekonomian Indonesia, terutama dalam bidang ekspor. Oleh karena itu, menjaga kualitas bibit kelapa sawit menjadi prioritas, terutama untuk memastikan hasil produksi yang optimal.

Tanaman kelapa sawit saat ini merupakan salah satu jenis tanaman Perkebunan yang menduduki posisi penting disektor pertanian umumnya, dan sektor perkebunan khususnya, hal ini disebabkan karena dari sekian banyak tanaman yang menghasilkan minyak atau lemak. Perkebunan kelapa sawit per satu hektarnya biaya produksi yang dibutuhkan sebesar 9,7 juta/hektar/tahun dengan nilai produksi mencapai Rp 17 juta/hektar/tahun. Dengan begitu buah kelapa sawit menjadi komoditas perkebunan yang memegang peranan penting bagi perekonomian di Indonesia sebagai salah satu penyumbang devisa non-migas yang cukup besar (Andrianto & Irawan, 2023).

Kualitas bibit kelapa sawit sangat menentukan produktivitas jangka panjang dari tanaman kelapa sawit itu sendiri. Faktor-faktor seperti ukuran daun, warna daun, bentuk batang, buah, akar, pucuk dan tandan buah adalah beberapa aspek penting yang perlu diperhatikan dalam proses seleksi bibit. Namun, salah satu tantangan terbesar dalam menjaga kualitas bibit adalah mengidentifikasi dan mengendalikan penyakit yang dapat menyerang bibit sejak dini. Penyakit pada bibit sawit tidak hanya menghambat pertumbuhan, tetapi juga berpotensi menurunkan hasil panen secara signifikan ketika tanaman sudah dewasa.

Hal serupa dihadapi oleh PT. Plasmen Sisumut, dalam proses pemilihan bibit, penelitian tradisional atau manual memerlukan waktu yang cukup lama, terutama karena keterbatasan metode tersebut dalam menghadapi ketidakpastian dan beberapa faktor yang mempengaruhi kualitas bibit sawit. Akibatnya, hal ini memperpanjang siklus waktu dan menyebabkan PT. Plasmen Sisumut cukup lama dalam mengetahui bibit sawitnya.

Oleh karena itu, diperlukan metode penilaian kualitas bibit kelapa sawit yang cukup efektif, yang dapat dilakukan secara cepat dan efisien. Salah satu langkah yang dapat ditempuh adalah pembuatan aplikasi prediksi menggunakan algoritma *Data mining* yang kemudian diintegrasikan dengan sistem informasi.

*Data mining* adalah proses menemukan informasi yang berguna dari sebuah basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi dari sekumpulan data besar untuk membantu dalam pengambilan Keputusan. *Data mining* dapat digunakan untuk menggali nilai tambah berupa informasi yang tidak diketahui secara manual dari suatu basis data. (Harahap & Sulindawaty, 2020)

Terdapat beberapa jenis algoritma yang dipakai dalam melakukan *Data Mining*, salah satunya adalah metode Naïve Bayes, klasifikasi dengan naïve bayes dilakukan berdasarkan teorema bayes dengan asumsi antar variabel penjelas independent. Cara ini dapat diasumsikan sebagai keadaan atau ketiadaan dari sebuah kejadian tertentu dari suatu grup yang tidak berhubungan dengan keadaan atau ketiadaan kejadian lain (Haba & Pelangi, 2020).

Algoritma Naïve bayes telah digunakan dalam beberapa penelitian, salah satunya penelitian yang dilakukan Sinaga et al. (2022) dengan judul penerapan algoritma Naïve Bayes untuk klasifikasi prediksi penerimaan siswa baru. Hasil pengujian menggunakan 30 data dapat diperoleh jika algoritma ini dapat memperoleh Tingkat akurasi sebesar 86,6%, sehingga dapat disimpulkan bahwa proses prediksi klasifikasi yang cepat dengan akurasi tinggi, prosesnya melibatkan perhitungan probabilitas posterior untuk menentukan kemungkinan suatu bibit terinfeksi penyakit tertentu berdasarkan fitur-fitur yang diamati.

Menurut (Rahman et al., 2022), Sistem Informasi merupakan pengintegrasian unsur sistem dan unsur informasi, Sistem Informasi merupakan kumpulan dari elemen-elemen yang berinteraksi untuk menghasilkan informasi bagi pengguna yang menggunakannya. Dengan menggunakan algoritma naïve bayes yang diintegrasikan kedalam sistem, sistem dapat memprediksi kualitas bibit kelapa sawit dengan cepat berdasarkan karakteristik fisik bibit, seperti ukuran dan warna daun, bentuk batang, dan kondisi buah serta akar. Dalam konteks pemilihan bibit kelapa sawit, Naïve bayes mampu mempelajari pola dari data pelatihan yang telah diberi label, sehingga sistem dapat mendeteksi bibit yang berpotensi berpenyakit sejak dini. Keunggulan ini dapat memungkinkan PT. Plasmen Sisumut untuk memperoleh hasil klasifikasi bibit yang efisien. Oleh karena itu, berdasarkan

latar belakang diatas, penelitian ini berjudul, “**Implementasi Algoritma Naïve Bayes untuk memprediksi penyakit bibit sawit (Studi Kasus PT.Plasmem Sisumut)**”

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang penulis jabarkan sebelumnya, penulis merumuskan beberapa masalah sebagai berikut :

1. Bagaimana pemanfaatan algoritma Naïve Bayes dapat membantu memprediksi penyakit pada bibit kelapa sawit berdasarkan karakteristik fisik bibit secara cepat dan akurat?
2. Bagaimana Tingkat akurasi algoritma Naïve Bayes dalam memprediksi penyakit bibit sawit?

## **1.3 Batasan Masalah**

Untuk menghindari pembahasan diluar permasalahan, maka dibutuhkan batasan masalah yang diuraikan sebagai berikut:

1. Penelitian ini hanya akan menggunakan karakteristik bibit kelapa sawit sebagai input untuk identifikasi penyakit, yaitu berdasarkan ukuran daun, warna daun, bentuk daun, jumlah daun, struktur batang, kondisi buah, dan sistem akar.
2. Metode klasifikasi yang digunakan dalam penelitian ini adalah metode *Naïve Bayes*
3. Penelitian ini akan diimplementasikan dalam bentuk aplikasi berbasis web

4. Dataset yang digunakan dalam penelitian ini terbatas pada beberapa jenis penyakit umum yang menyerang bibit kelapa sawit.
5. Aplikasi yang dikembangkan akan diuji dengan menggunakan data yang telah dikumpulkan dari PT. Plasmen Sisumut dari tahun 2023-2024, dan tidak mencakup semua kondisi lapangan yang mungkin terjadi di berbagai lokasi lainnya.

#### **1.4 Tujuan Penelitian**

Penelitian ini memiliki tujuan sebagai berikut :

1. Mengembangkan model prediksi penyakit bibit kelapa sawit menggunakan metode Naïve Bayes berdasarkan karakteristik fisik bibit.
2. Mengimplementasikan model Naïve Bayes ke dalam aplikasi berbasis web yang dapat digunakan untuk memprediksi penyakit bibit kelapa sawit secara cepat dan efisien.
3. Mengevaluasi tingkat akurasi dan efektivitas algoritma Naïve Bayes dalam memprediksi penyakit pada bibit kelapa sawit.

#### **1.5 Manfaat Penelitian**

Dengan berjalannya penelitian ini, penulis mengharapkan manfaat sebagai berikut :

1. Mempermudah petani dan pengelola perkebunan dalam memprediksi penyakit bibit kelapa sawit dengan cepat, sehingga pengelolaan bibit dapat dilakukan lebih optimal.
2. Mendorong penerapan teknologi berbasis web dalam bidang pertanian untuk meningkatkan produktivitas dan efisiensi dalam pengelolaan kesehatan bibit kelapa sawit.

## BAB II LANDASAN TEORI

### 2.1 Kelapa Sawit

Kelapa sawit merupakan tanaman yang termasuk dalam keluarga *Palmae*. Awalnya, minyak nabati ini pertama kali tumbuh di Brasil sebelum kemudian menyebar ke wilayah Amerika Ekuatorial, Afrika, Pasifik Selatan, dan Asia Tenggara. Benih kelapa sawit pertama kali ditanam di Indonesia pada tahun 1984, yang didatangkan dari Mauritius, Afrika. Pada Tahun 1911, Perkebunan kelapa sawit pertama di Indonesia oleh Schadt, seorang Jerman, di daerah Tanahitam, Hulu Sumatera Utara (Marcelina et al., 2022).

### 2.2 Penyakit Tanaman Kelapa Sawit

Terdapat beberapa jenis penyakit tanaman kelapa sawit, diantaranya adalah sebagai Berikut :

#### a. Penyakit Daun Bibit Muda (*Antracnose*)

Penyakit ini disebabkan oleh jamur *Colletotrichum gloeosporioides*. Gejalanya berupa bercak coklat kehitaman pada daun muda bibit sawit. Bercak ini dapat menyebar, menyebabkan kematian daun dan mempengaruhi pertumbuhan tanaman secara keseluruhan. Penyakit ini sering muncul di kondisi kelembapan tinggi.

#### b. Penyakit Garis Kuning (*Patch Yellow*)

Penyakit ini ditandai dengan munculnya garis-garis kuning pada daun bibit sawit. Penyebabnya bisa berupa kekurangan hara, terutama unsur kalium, atau adanya infeksi patogen. Garis kuning dapat mengurangi efisiensi fotosintesis, sehingga memperlambat pertumbuhan tanaman.

7

#### c. Busuk Batang Atas (*Upper Stem Rot*)

Penyakit ini disebabkan oleh jamur patogen yang menyerang bagian atas batang tanaman sawit. Gejalanya meliputi busuk pada batang atas, biasanya dimulai dari ujung batang hingga ke bawah, yang dapat menyebabkan tanaman mati. Penyakit ini umumnya muncul pada tanaman dewasa tetapi dapat juga menyerang bibit.

d. Penyakit Akar (*Blas Disease*)

Penyakit ini menyerang akar bibit sawit dan dapat disebabkan oleh beberapa patogen, termasuk *Rhizoctonia solani*. Gejalanya termasuk pembusukan pada akar yang menyebabkan tanaman layu dan mati. Penyakit ini sering ditemukan di tanah yang lembab dan tergenang air.

e. Penyakit Pangkal Batang Busuk (*Basal Stem Rot* atau *Genoderma*)

Penyakit ini disebabkan oleh jamur *Ganoderma boninense*. Penyakit ini menyerang pangkal batang, menyebabkan pembusukan dan pelapukan pada jaringan kayu. Gejalanya termasuk adanya jamur yang muncul di permukaan batang dan tanaman yang mati perlahan-lahan. Ini adalah salah satu penyakit paling serius yang menyerang tanaman sawit, termasuk bibit.

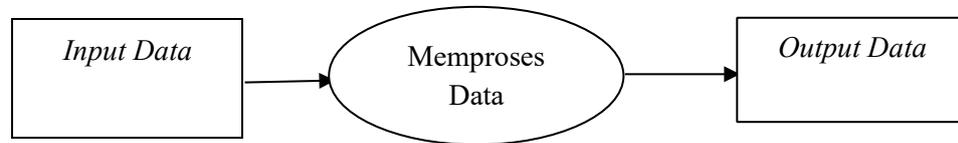
f. Penyakit Tandan Busuk (*Bunch Rot*)

Penyakit ini menyerang tandan buah sawit, menyebabkan pembusukan pada buah yang sedang berkembang. Penyebab utamanya adalah infeksi jamur *Marasmius palmivorus* atau bakteri tertentu. Tandan yang terinfeksi akan mengalami kerusakan parah sehingga mengurangi hasil produksi.

### 2.3 Pengertian Sistem Informasi

Menurut Ludwig Von Bartalanfy, Sistem adalah seperangkan unsur yang saling terkait dalam suatu hubungan antara unsur unsur tertentu dengan lingkungan sekitarnya. Selain itu, Sistem informasi diartikan sebagai sistem dalam organisasi

atau perusahaan yang memenuhi kebutuhan untuk mengolah transaksi harian, mendukung operasi operasi manajerial, serta mendukung kegiatan dengan tujuan strategis dari organisasi tersebut. (Soufitri, 2023: 5-6).



**Gambar 2. 1** Gambaran jalannya Suatu Sistem

#### 2.4 Konsep *Data Mining*

*Data Mining* merupakan bentuk penggalian data yang digunakan untuk menggali pengetahuan dari jumlah data yang besar (Putra et al., 2023). *Data Mining* diperlukan dalam melakukan prediksi untuk hubungan ditemukan yang memiliki arti, pola, dan kecenderungan dengan diperiksa sekumpulan besar data yang disimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola statistika ataupun matematis (Iriane & Nurfaizah, 2023). Dalam *Data Mining*, proses pencarian pola atau informasi yang berguna dari data yang telah dipilih atau diolah dinamakan sebagai *Knowledge Data Discovery* (KDD).

#### 2.5 Algoritma *Naïve Bayes*

Algoritma *Naïve Bayes* merupakan salah satu algoritma klasifikasi dalam *Machine Learning* yang sederhana namun tergolong dalam klasifikasi yang kuat. Algoritma *naïve bayes* adalah pengklasifikasi probabilistic berdasarkan penerapan teorema *bayes* untuk probabilitas bersyarat. Asumsina adalah bahwa semua fitur independen dan tidak terkait satu sama lain yang menjadi latar belakang mengapa *classifier* ini disebut *naif*. (Marpaung et al., 2022). *Naïve bayes* menggunakan ciri klasifikasi, asumsi yang sangat kuat akan independensi dari masing masing kondisi

atau kejadian. Adapun cara kerja algoritma *naïve bayes* adalah menggunakan pengklasifikasi probabilitas sederhana untuk menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari kumpulan data yang diberikan (Furqan et al., 2022). Dalam algoritma Naïve Bayes, beberapa parameter utama perlu diatur untuk mengoptimalkan performa model. Beberapa diantaranya adalah :

### 1. *Laplace Smoothing*

*Laplace Smoothing*, juga disebut sebagai *add-one smoothing*, merupakan teknik yang digunakan untuk mengatasi masalah Ketika fitur tertentu tidak muncul dalam data pelatihan untuk kelas tertentu. Metode ini disebut *smoothing* karena menambahkan nilai kecil dalam perhitungan probabilitas guna mencegah hasil probabilitas bernilai nol

### 2. Probabilitas Prior

Probabilitas prior merupakan probabilitas awal dari setiap kelas sebelum memperhitungkan fitur-fitur yang tersedia. Ini mencerminkan seberapa sering suatu kelas muncul dalam data pelatihan, memberikan Gambaran tentang frekuensi relative suatu kategori tanpa melihat fitur yang terkait.

### 3. Probabilitas *Likelihood*

Probabilitas Likelihood adalah probabilitas kemunculan fitur tertentu dalam data pelatihan untuk setiap kelas. Hal ini menggambarkan seberapa sering fitur tersebut muncul dalam data yang sudah diberi label sesuai dengan kelas tertentu.

Rumus dari naïve bayes dapat dilihat seperti dibawah ini :

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Keterangan :

- a)  $P(A|B)$  adalah probabilitas posterior, yaitu probabilitas bahwa peristiwa A terjadi mengingat bahwa peristiwa B telah terjadi.
- b)  $P(B|A)$  adalah probabilitas *likelihood*, yaitu probabilitas bahwa peristiwa B terjadi jika peristiwa A sudah terjadi
- c)  $P(A)$  adalah probabilitas prior, yaitu probabilitas awal dari peristiwa A sebelum mempertimbangkan peristiwa B
- d)  $P(B)$  adalah probabilitas total peristiwa B yang sering dianggap konstan dapat membandingkan probabilitas posterior untuk berbagai nilai A.

## 2.6 Evaluasi Model dengan *Confussion Matrix*

*Confussion Matrix* adalah sebuah table yang menunjukkan hasil klasifikasi, baik yang benar maupun yang salah (Normawati & Prayogi, 2021). Tabel ini memiliki empat istilah utama

1. *True Negative* (TN) yang berarti model klasifikasi memprediksi data ada di kelas *negative* namun yang sebenarnya data memang ada di kelas *negative*.
2. *True Positive* (TP) yang berarti model klasifikasi memprediksi data ada di kelas positif namun sebenarnya memang data berada di kelas positif.
3. *False Negative* (FN) yang berarti model klasifikasi memprediksi data ada di kelas *negative* namun yang sebenarnya data ada di kelas positif.

4. *False Positive* (FP) yang berarti model klasifikasi memprediksi data ada di kelas positif namun yang sebenarnya data ada di kelas *negative*.

Dari defenisi *confusion matrix*, beberapa poin poin dalam *confusion matrix* digunakan untuk menghitung *precision*, *recall*, dan *f1 score*. *Precision* adalah perbandingan antara *True Positive* (TP) dengan banyaknya data yang diprediksi positif, secara matematis dapat dilihat dibawah ini

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots (2.2)$$

Untuk *recall* sendiri adalah perbandingan antara *true positive* (TP) dengan banyaknya data yang sebenarnya positif. Dapat dinyatakan secara matematis seperti dibawah ini

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (2.3)$$

Sedangkan *F1 Score* adalah nilai tengah dari *precision* dan *recall*. Nilai terbaik dari *F1 Score* adalah 1 dan nilai terburuknya adalah 0, secara matematis dapat dituliskan seperti dibawah ini

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \dots\dots\dots (2.4)$$

Nilai *F1 Score* yang baik menandakan bahwa model klasifikasi kita punya *precision* dan *recall* yang baik (Setiawan, 2020).

## 2.6 Pra-Pemrosesan Data

Pra-pemrosesan data adalah tahapan untuk mengolah data mentah dengan menghilangkan beberapa permasalahan yang mengganggu saat pemrosesan data. Hal tersebut dikarenakan oleh data yang formatnya tidak konsisten. Melalui proses

ini, pemodelan algoritma KNN akan berjalan lebih efektif dan efisien. Adapun tahapan tahapan dalam pra-pemrosesan data adalah :

1. Pertama-tama dalam tahap awal *preprocessing* data, langkah yang esensial adalah melakukan data cleaning. Proses ini melibatkan pemilihan kembali data mentah untuk menghilangkan entri yang tidak lengkap, tidak relevan, atau tidak akurat. Dengan melakukan ini, kita dapat menghindari kesalahpahaman saat melakukan analisis terhadap data tersebut.
2. Langkah berikutnya adalah *data integration*, yang diperlukan karena preprocessing data melibatkan penggabungan data dari berbagai sumber ke dalam satu dataset. Penting untuk memastikan bahwa data dari berbagai sumber memiliki format yang seragam.
3. Setelah itu, kita melanjutkan ke tahap transformasi data. Sebagaimana dijelaskan sebelumnya, data yang berasal dari sumber yang berbeda mungkin memiliki format yang beragam. Oleh karena itu, perlu dilakukan penyesuaian format agar seluruh data yang terkumpul memiliki struktur yang seragam, memudahkan proses analisis data.
4. Tahap terakhir dalam *preprocessing* data adalah mengurangi jumlah data, yang dikenal sebagai *data reduction*. Tujuan utamanya adalah mengurangi sampel data tanpa mengubah hasil analisis. Terdapat tiga teknik yang dapat diterapkan pada tahap ini, yaitu pengurangan dimensi (*dimensionality reduction*), pengurangan jumlah (*numerosity reduction*), dan kompresi data (Binus Student Corner, 2022).

## 2.7 *Website*

*Website* adalah sekumpulan halaman yang terdiri atas beberapa laman yang berisi informasi dalam bentuk data digital, baik berupa teks, gambar, video, audio, dan animasi lainnya yang di sediakan melalui jalur koneksi internet (Tumini & Fitria, 2021).

## 2.8 *Hypertext Markup Language (HTML)*

HTML atau *HyperText Markup Language* adalah Bahasa *markup* sebagai dasar untuk menciptakan halaman web. Fungsinya untuk mengatur struktur dari sebuah situs web. HTML juga digunakan untuk menandai bagian mana yang berperan sebagai judul, isi, daftar tabel, dan lainnya. (Pratama, 2020).

## 2.9 *Hypertext Preprocessor (PHP)*

PHP adalah bahasa pemrograman skrip yang terintegrasi dengan HTML dan dieksekusi di sisi server, sementara yang dikirim ke browser hanya hasil akhirnya. Saat pengguna internet mengakses situs yang menggunakan PHP, server akan memproses semua perintah PHP terlebih dahulu dan kemudian mengirimkan hasilnya dalam format HTML ke browser pengguna. Dengan demikian, kode PHP asli tidak terlihat di browser. PHP dapat digunakan untuk mengambil informasi dari formulir web dan menggunakannya untuk berbagai fungsi (Aniqsa, 2019).

## 2.10 XAMPP

XAMPP adalah sebuah *Software Web Apache* yang didalamnya sudah tersedia *database server* MySQL dan mendukung pemrograman php. XAMPP merupakan software yang mudah digunakan dan gratis dan mendukung instalasi di *linux* dan *windows* (Sari et al., 2022).

### **2.11 Database**

*Database* adalah kumpulan data yang saling berhubungan, disimpan bersama dalam suatu media dengan cara yang meminimalisir duplikasi data. Data ini disimpan agar mudah diakses atau ditampilkan kembali, dan dapat digunakan oleh satu atau lebih aplikasi tanpa ketergantungan pada program tertentu. Data disusun sedemikian rupa sehingga proses penambahan, pengambilan, dan perubahan data dapat dilakukan dengan mudah dan tertatur (Tumini & Fitria, 2021).

### **2.12 MySQL**

MySQL merupakan perangkat lunak basis data relasi atau *Relational Database Management System* (RDBMS) yang diterapkan menggunakan lisensi GPL (*General Public License*). MySQL memiliki kinerja, kecepatan proses yang tidak kalah dengan database-database besar lainnya (Aniqsa, 2019). Beberapa pertimbangan dalam memilih MySQL antara lain :

1. Kecepatan
2. Mudah Digunakan
3. Terbuka
4. Kapabilitas
5. Konektifitas dan Keamanan

### **2.13 Flask API**

*Flask* adalah sebuah framework web yang ditulis menggunakan bahasa pemrograman *Python*, yang termasuk dalam kategori *microframework*. Meskipun *Flask* tergolong ringan, framework ini menawarkan fleksibilitas yang tinggi bagi pengembang untuk membangun aplikasi web yang terstruktur. Flask memudahkan pengelolaan tampilan dan logika aplikasi web, sehingga cocok digunakan untuk proyek web sederhana hingga aplikasi yang cukup kompleks. Salah satu

keunggulannya adalah kemudahan dalam penggunaannya, terutama bagi pemula dalam pengembangan aplikasi berbasis web. Flask juga memungkinkan integrasi yang baik dengan berbagai pustaka dan alat bantu Python lainnya, seperti ORM (Object Relational Mapping) untuk database dan RESTful API untuk komunikasi antar server (Budianto et al., 2023).

#### **2.14 Google Colaboratory**

*Google Colaboratory* adalah dokumen yang dapat dieksekusi yang memungkinkan pengguna menyimpan, membuat, dan berbagi program melalui *Google Drive*. Keunikan dari *Google Colaboratory* adalah mudah digunakan dan tidak perlu konfigurasi apapun untuk menjalankan kode *Python* (Raza, 2024).

#### **2.15 Pandas**

*Pandas* adalah *library* pada *Python* yang paling sering dipakai untuk menganalisis data. Cara kerja *pandas* cukup menarik, Data berupa format CSV atau SQL diubah menjadi objek dalam *Python* dengan bentuk baris dan kolom yang disebut sebagai *Dataframe*. Objek *Dataframe* ini akan terlihat sangat mirip dengan tabel yang terdapat dalam aplikasi pengolahan statistika pada umumnya. Penggunaan *Library Pandas* juga dapat membantu dalam verifikasi data, pengolahan data, normalisasi data, penggabungan dan penyatuan data, inspeksi data, serta membuat dan menyimpan data .

#### **2.16 Scikit-Learn**

*Scikit-Learn* adalah *library* dalam bahasa pemrograman *python* yang membantu melakukan proses pada data ataupun melakukan pelatihan pada data untuk kebutuhan *data science* seperti klasifikasi, *clustering*, dan lain-lain. *Library* ini sangat populer dikarenakan banyak sekali model pembelajaran mesin yang dapat dipanggil menggunakan *scikit-learn* (Bisa AI Academy, 2023).

## 2.17 Unified Modelling Language (UML)

*Unified Modelling Language (UML)* adalah bahasa pemodelan visual yang digunakan untuk menentukan, memvisualisasikan, membangun berbagai aspek dari sebuah sistem perangkat lunak. UML sendiri memiliki fungsi sebagai alat untuk menangkap pemahaman mengenai sistem yang perlu dibangun.

Bayangkan sistem sebagai kelompok objek yang berbeda yang bekerja sama untuk melakukan pekerjaan yang berguna untuk pengguna. Bagian struktur statis mendefinisikan jenis objek yang penting untuk sistem dan bagaimana mereka terkait satu sama lain. Bagian perilaku dinamis menggambarkan bagaimana objek berubah seiring waktu dan berkomunikasi satu sama lain untuk mencapai tujuan tertentu. Dengan memodelkan sistem dari berbagai sudut pandang yang terkait, kita dapat memahami sistem tersebut untuk berbagai keperluan (Rumbaugh et al., 2021).

Alat bantu yang digunakan dalam perancangan sistem menggunakan Unified Modelling Language (UML) adalah sebagai berikut :

### 1. *Use Case Diagram*

*Use Case Diagram* menunjukkan bagaimana perilaku suatu sistem ketika digunakan oleh orang lain, *Use Case Diagram* memecah fungsionalitas sistem menjadi tindakan yang bermakna bagi pengguna (*Actor*) yang menggunakan sistem. Pengguna (*Actor*) mencakup manusia, serta sistem dan proses komputer lainnya (Rumbaugh et al., 2021)..

### 2. Diagram Aktivitas (*Activity Diagram*)

Diagram aktivitas (*Activity Diagram*) menggambarkan aliran kerja dari sebuah sistem. Diagram aktivitas dapat mencakup cabang dan bercabangnya control dalam sebuah sistem yang berjalan secara bersamaan.

Cabang cabang ini mewakili aktivitas yang dapat dilakukan secara bersamaan (Rumbaugh et al., 2021).

### 3. *Sequence Diagram*

Sequence Diagram menjelaskan perilaku objek objek yang ada didalam *use case* dengan menjelaskan siklus hidup dari objek serta pesan pesan atau data yang dikirim oleh objek tersebut.

### 4. *Class Diagram* (Diagram Kelas)

Class diagram adalah salah satu jenis diagram struktur dalam UML yang dengan jelas menggambarkan struktur serta deskripsi kelas, atribut, metode, dan hubungan antara setiap objek. Class diagram bersifat statis, artinya diagram ini tidak menunjukkan apa yang terjadi saat kelas-kelas tersebut berinteraksi, melainkan menjelaskan jenis hubungan yang ada..

## **BAB III METODE PENELITIAN**

Metode Penelitian dapat diartikan sebagai Langkah-langkah ilmiah untuk mendapatkan data dengan tujuan dan manfaat tertentu (Ramdhan, 2021). Penentuan bibit sawit dalam penelitian ini akan melalui beberapa proses penelitian.

### **3.1. Lokasi Penelitian**

Lokasi penelitian di lakukan di PT. Plasmen Sisumut yang terletak di Desa Sisumut, kecamatan Kota Pinang, Kabupaten Labuhan Batu Selatan, Provinsi Sumatera Utara.

### **3.2. Objek Penelitian**

Penelitian akan berfokus pada karakteristik bibit kelapa sawit seperti ukuran dan warna daun, bentuk batang dan hasil identifikasi lainnya yang berpotensi mendeteksi penyakit.

### **3.3. Studi Literatur dan Pustaka**

Tahapan ini dilakukan guna mengetahui kebutuhan-kebutuhan yang digunakan dalam menyelesaikan masalah yang diteliti dan memahami lebih dalam metode yang akan digunakan. Berbagai referensi kemudian dipelajari sebagai fondasi dasar yang kuat dalam mengerjakan penelitian dengan mempelajari jurnal dan artikel terkait dengan penelitian. Berikut studi literatur dan Pustaka terkait.

Tabel 3. 1 Studi Literatur dan Pustaka

No	Nama Peneliti	Judul Penelitian	Hasil Penelitian
1	(Raza, 2024)	Sistem Deteksi Berita Hoax Menggunakan Algoritma Naïve Bayes dan Random Forest pada Machine Learning	Berhasil dikembangkan sistem deteksi berita palsu menggunakan <i>Naïve Bayes</i> dengan tingkat akurasi mencapai 96,68%
2	Sinaga et al. (2022)	Penerapan Algoritma <i>Naïve Bayes</i> untuk Klasifikasi Prediksi Penerimaan Siswa Baru	Berdasarkan pengujian sebanyak 30 data testing yang diolah, diperoleh tingkat akurasi sebesar 86,6% yaitu 26 siswa diterima dan sebanyak 4 tidak diterima
3	Jennie et al. (2021)	Sistem Pengambilan Keputusan Beban Kinerja Menggunakan <i>Naïve Bayes</i> Studi Kasus PDAM Bandarmasih	Dalam pembuatan aplikasi ini, menggunakan sebuah metode waterfall dalam pengembangan aplikasinya sedangkan untuk sistem pengambilan Keputusan menggunakan metode Naive Bayes Classifier dimana peneliti akan membandingkan data latih dan data uji yang akhirnya diperoleh hasil probabilitas tertinggi.
4	Maryana et al. (2023)	Sistem Pendukung Keputusan Kenaikan Jabatan Menggunakan Metode Naive Bayes (Studi Kasus : PT Buana Mulia Indonesia)	Sistem memiliki tingkat keakuratan menggunakan Naïve bayes sebesar 80%

### 3.4. Pengumpulan Data

Data yang diperlukan dalam penelitian ini adalah data karakteristik fisik dan kondisi bibit kelapa sawit yang tersedia di tempat penelitian. Beberapa metode yang digunakan dalam pengumpulan data adalah sebagai berikut :

#### 1. Observasi / Pengamatan Langsung

Penulis melakukan pengamatan langsung ke PT. Plasmen Sisumut untuk mendapatkan data-data yang diperlukan, terutama data mengenai karakteristik fisik bibit kelapa sawit yang telah dikategorikan sehat atau berpenyakit.

#### 2. Wawancara

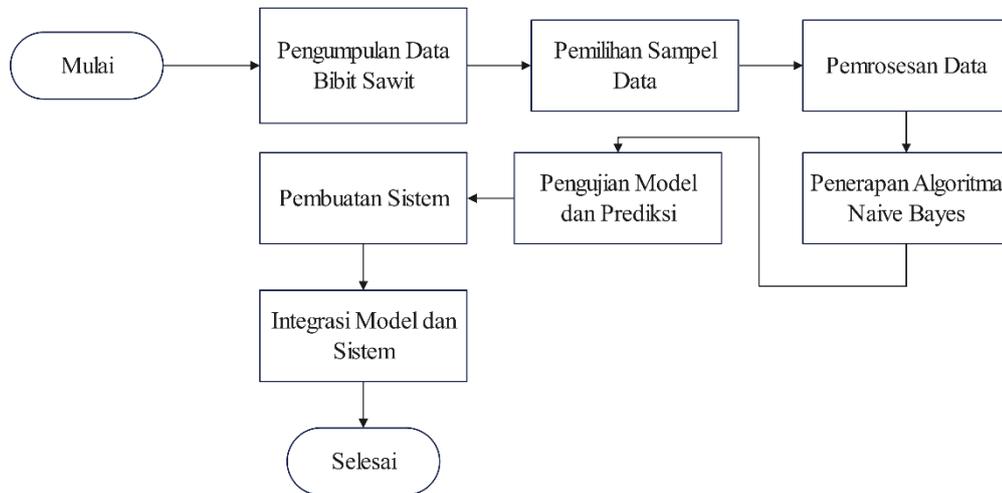
Penulis melakukan wawancara dan diskusi tanya jawab kepada pihak ahli di PT. Plasmen Sisumut untuk memahami jenis-jenis penyakit yang biasanya menyerang bibit kelapa sawit, faktor-faktor penyebabnya, dan langkah-langkah yang biasanya dilakukan. Beberapa pertanyaan yang diajukan dalam wawancara meliputi :

- a) Apa saja penyakit umum yang sering menyerang bibit kelapa sawit?
- b) Apa saja gejala fisik yang menunjukkan bahwa bibit mungkin terserang penyakit tertentu?
- c) Bagaimana cara membedakan gejala penyakit yang mirip satu sama lain?
- d) Bagaimana prosedur identifikasi penyakit pada bibit yang diterapkan di perusahaan?
- e) Seberapa akurat metode identifikasi penyakit yang saat ini digunakan?
- f) Berapa lama waktu yang dibutuhkan untuk mengidentifikasi penyakit secara manual?

- g) Apa tantangan utama dalam mendeteksi penyakit pada tahap awal pertumbuhan bibit?

### 3.5. Bagan Alir (*Flowchart*)

Berikut bagan alir dari penelitian terkait :



**Gambar 3. 1 Flowchart Alur Penelitian** Berikut penjelasan setiap tahapan penelitian diatas :

#### 1. Pengumpulan Data Bibit Sawit

Tahap ini melibatkan proses pengumpulan data terkait bibit sawit dari PT. Plasmen Sisumut. Data yang dikumpulkan mencakup berbagai karakteristik bibit, seperti ukuran daun, warna daun, bentuk batang, kondisi akar, serta indikasi gejala penyakit yang mungkin muncul.

#### 2. Pemilihan sampel data

Setelah data dikumpulkan, dilakukan pemilihan sampel dari data acak yang diberikan oleh tempat penelitian. Sampel yang dipilih harus mewakili berbagai kondisi bibit sawit, baik yang sehat, maupun yang terkena penyakit, agar model yang akan dibangun dapat mempelajari pola klasifikasi dengan baik.

### 3. Pemrosesan data

Data yang telah dipilih kemudian diproses agar sesuai dengan format yang digunakan dalam algoritma Naïve Bayes. Proses ini meliputi pembersihan data, transformasi data ke dalam bentuk numerik atau kategori, serta pembagian data menjadi set pelatihan dan pengujian.

### 4. Penerapan algoritma naïve bayes

Pada tahap ini, algoritma Naïve Bayes diterapkan untuk mempelajari pola dalam data. Algoritma ini bekerja dengan menghitung probabilitas setiap fitur bibit sawit terhadap kemungkinan terkena penyakit tertentu. Model dilatih menggunakan data latih yang telah diproses sebelumnya, sehingga dapat mengenali karakteristik yang berhubungan dengan penyakit bibit sawit.

### 5. Pengujian model dan prediksi

Setelah model naïve bayes dilatih, dilakukan pengujian dengan menggunakan data uji untuk mengevaluasi performanya dalam memprediksi penyakit bibit sawit. Pengujian dilakukan dengan menghitung akurasi, precision, recall, dan F1-score guna memastikan bahwa model mampu memberikan hasil prediksi yang valid.

### 6. Pembuatan Sistem

Pada tahap ini, sistem berbasis web dikembangkan untuk memungkinkan pengguna, seperti petani atau manajer perkebunan, menggunakan model prediksi secara mudah. Sistem ini dirancang agar pengguna dapat memasukkan data bibit sawit, menjalankan prediksi, serta melihat hasil klasifikasi yang diberikan oleh model Naïve Bayes.

## 7. Integrasi Model dan Sistem

Tahap akhir dalam penelitian ini adalah menghubungkan model yang dibuat dengan sistem web. Model prediksi yang dikembangkan menggunakan Python diintegrasikan dengan sistem menggunakan API Flask, sehingga pengguna dapat mengakses model secara online dan mendapatkan hasil prediksi langsung dari sistem yang telah dikembangkan.

### 3.6. Pemeriksaan Data

Tahap ini dilakukan untuk memeriksa dan identifikasi masalah yang terdapat dalam data dan memastikan data tersebut memiliki kualitas yang baik. Pemeriksaan data akan dilakukan dengan bantuan *library pandas*. Menurut Dicoding Team (2023) , adapun beberapa masalah umum yang biasanya dijumpai dalam sebuah data adalah sebagai berikut :

#### a. *Missing Value*

Masalah ini muncul ketika adanya nilai yang hilang dari sebuah data dan biasanya direpresentasikan dengan nilai NaN dalam *library pandas*. *Library pandas* sendiri menyediakan metode bernama `isnull()` atau `isna()` untuk mengidentifikasi masing masing nilai yang hilang dalam sebuah data yang kemudian dipadukan dengan metode `sum()` untuk menghitung jumlah *missing value* dalam data.

#### b. *Invalid Value*

Masalah ini muncul Ketika terdapat nilai yang tidak masuk akal dan tidak sesuai dengan ketentuan.

#### c. *Duplicate Value*

Masalah ini terjadi Ketika terdapat data yang memiliki nilai yang sama persis pada setiap kolomnya. Library *pandas* menyediakan metode *duplicated()* untuk mengidentifikasi apakah terdapat duplikasi terhadap data

d. *Inaccurate Value*

Masalah ini muncul ketika nilai pada sebuah data tidak sesuai dengan hasil observasi. Masalah ini umumnya muncul karena adanya *human error* dalam pencatatan transaksi.

e. *Inconsistent Value*

Masalah ini muncul Ketika sebuah data memiliki nilai yang tidak konsisten baik dari segi satuan maupun ketentuan penilaian.

### 3.7. Pembersihan dan Transformasi Data

Pembersihan data dilakukan berdasarkan temuan-temuan masalah yang ditemui pada proses pemeriksaan data, masalah-masalah yang ditemukan ditangani dengan teknik-teknik tertentu yang akan dibantu menggunakan *library pandas*.

Sedangkan tahapan transformasi data merupakan tahapan dimana data yang sudah melewati tahap pemeriksaan dan pembersihan diubah dengan membuat ringkasan (agregasi) sehingga data tersebut dapat digunakan dalam proses *data mining* menggunakan metode *k-nearest neighbor*.

Berikut beberapa tahapan penerapan pengolahan data tersebut menggunakan *library pandas* :

#### 1. Mengimpor *Library Pandas*

Langkah ini pertama kali dilakukan agar dapat menggunakan fitur-fitur yang disediakan *pandas* untuk pemrosesan data, adapun potongan kode yang digunakan dalam mengimpor *Library Pandas* dalam lingkungan pengembangan *Python* dapat dilihat pada gambar dibawah ini :

```
import pandas as pd
```

**Gambar 3. 2 Kode Import Library Pandas**

Potongan kode di atas adalah contoh penggunaan pernyataan import untuk mengimpor library *Pandas* dengan alias *pd*. Alias digunakan untuk mempermudah pemanggilan saat *pandas* digunakan dalam pemrosesan data.

## 2. Memasukkan data dalam format csv

*Pandas* memberikan kemudahan untuk memasukkan data kedalam lingkungan pengembangan *python*, adapun memasukkan data ke dalam lingkungan pengembangan model menggunakan bahasa *python* dan *library pandas* dapat dilihat dalam potongan kode dibawah ini :

```
# Memuat file CSV ke dalam DataFrame  
data = pd.read_csv('nama_file.csv')
```

**Gambar 3. 3 Kode memasukkan data kedalam DataFrame**

Setelah kode ini dieksekusi, data dari file CSV akan dimuat ke dalam objek DataFrame yang disebut *data*, yang dapat Anda gunakan untuk melakukan berbagai operasi pemrosesan data menggunakan *Pandas*.

## 3. Verifikasi data

Pemeriksaan apakah ada data yang hilang atau data tidak valid dapat dipermudah dengan menggunakan *library pandas*. Pengecekan nilai yang hilang

dalam suatu data dapat dilakukan dengan potongan kode yang terdapat di gambar 3.3 dibawah ini. Kode tersebut juga secara otomatis menghapus nilai yang hilang tersebut.

```
# Misalnya, identifikasi nilai yang hilang
data.dropna(inplace=True)
```

**Gambar 3. 4 Kode identifikasi dan menghapus nilai data yang hilang**

### 3.8. Perhitungan dan Implementasi *Naïve Bayes*

Setelah data dalam keadaan siap untuk dianalisa, langkah berikutnya adalah melakukan perhitungan *Naïve Bayes*. Adapun rumus yang digunakan dalam *Naïve Bayes* adalah sebagai Berikut :

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Keterangan :

- a.  $P(A|B)$  adalah probabilitas posterior, yaitu probabilitas bahwa peristiwa A terjadi mengingat bahwa peristiwa B telah terjadi.
- b.  $P(B|A)$  adalah probabilitas *likelihood*, yaitu probabilitas bahwa peristiwa B terjadi jika peristiwa A sudah terjadi
- c.  $P(A)$  adalah probabilitas prior, yaitu probabilitas awal dari peristiwa A sebelum mempertimbangkan peristiwa B
- d.  $P(B)$  adalah probabilitas total peristiwa B yang sering dianggap konstan dapat membandingkan probabilitas posterior untuk berbagai nilai A.

Berikut adalah contoh studi kasus Naïve Bayes dalam menentukan penyakit pada bibit sawit. Diketahui bahwa dataset berikut menggambarkan ciri-ciri yang dimiliki oleh satu bibit sawit.

**Tabel 3. 2 Tabel Contoh Dataset**

<b>Bercak Coklat</b>	<b>Garis Kuning</b>	<b>Daun Layu</b>	<b>Busuk Batang</b>	<b>Nekrosis</b>	<b>Penyakit</b>
Tidak	Tidak	Ya	Tidak	Tidak	Bias Disease
Ya	Ya	Tidak	Tidak	Ya	Antraknosa
Ya	Tidak	Tidak	Ya	Tidak	Antraknosa
Ya	Ya	Tidak	Ya	Tidak	Antraknosa
Ya	Tidak	Ya	Tidak	Ya	Antraknosa
Tidak	Tidak	Tidak	Tidak	Ya	Bunch Rot
Tidak	Tidak	Ya	Tidak	Tidak	Bias Disease
Ya	Ya	Ya	Ya	Tidak	Antraknosa
Tidak	Ya	Tidak	Tidak	Ya	Patch Yellow
Tidak	Tidak	Ya	Ya	Ya	Upper Stem Rot
Ya	Ya	Ya	Ya	Tidak	Antraknosa
Tidak	Tidak	Ya	Tidak	Ya	Bunch Rot
Tidak	Tidak	Ya	Ya	Ya	Upper Stem Rot
Tidak	Tidak	Ya	Tidak	Ya	Bunch Rot
Ya	Tidak	Ya	Ya	Tidak	Antraknosa
Tidak	Tidak	Tidak	Ya	Tidak	Ganoderma
Tidak	Tidak	Ya	Tidak	Tidak	Bias Disease
Ya	Tidak	Ya	Ya	Tidak	Antraknosa
Ya	Ya	Ya	Ya	Ya	Antraknosa
Tidak	Ya	Tidak	Ya	Ya	Patch Yellow

Dalam dataset diatas, berisi beberapa contoh fitur gejala penyakit pada bibit sawit serta label penyakit yang terdektesi, berikut adalah penjelasan struktur dataset :

- a. **Gejala Bercak Coklat** : Indikator apakah bibit sawit memiliki bercak coklat (ya / tidak)

- b. **Gejala Garis Kuning** : Indikator apakah terdapat garis kuning pada daun bibit sawit (ya / tidak)
- c. **Gejala Layu** : Indikator apakah bibit mengalami kelayuan (ya / tidak)
- d. **Gejala Busuk Batang** : Indikator apakah bibit menunjukkan gejala busuk batang
- e. **Gejala Nekrosis** : Indikator apakah terjadi nekrosis (kematian jaringan pada tanaman)
- f. **Penyakit** : Label target yang menunjukkan jenis penyakit pada bibit sawit.

Selanjutnya, berdasarkan data diatas, dilakukan pencarian penyakit bibit sawit berdasarkan gejala dibawah ini :

- a. Gejala Bercak Coklat = Ya
- b. Gejala Garis Kuning = Tidak
- c. Gejala Layu = Ya
- d. Gejala Busuk Batang = Tidak
- e. Gejala Nekrosis = Tidak

Adapun Langkah-langkah perhitungan Naïve Bayes secara manual adalah sebagai berikut :

#### 1. Menghitung probabilitas prior

Probabilitas prior adalah kemungkinan awal bahwa suatu bibit sawit memiliki penyakit tertentu sebelum mempertimbangkan gejala. Yang didapatkan dengan rumus :

$$P(C) = \frac{\text{Jumlah sampel dengan penyakit } C}{\text{Total sampel dalam dataset}}$$

Dalam contoh dataset diatas, dataset memiliki total 20 sampel penyakit sawit, dimana distribusi penyakitnya adalah sebagai berikut :

a. Penyakit Antraknosa

Dalam dataset, terdapat 9 sampel kasus yang memiliki penyakit antraknosa, adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Antraknosa}) = \frac{9}{20}$$

$$P(\text{Antraknosa}) = 0,45$$

b. Penyakit Blas Disease

Dalam dataset, terdapat 3 sampel kasus yang memiliki penyakit blas disease, adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Blas Disease}) = \frac{3}{20}$$

$$P(\text{Blas Disease}) = 0,15$$

c. Penyakit *Bunch Rot*

Dalam dataset, terdapat 3 sampel kasus yang memiliki penyakit *bunch rot* adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Blas Disease}) = \frac{3}{20}$$

$$P(\text{Blas Disease}) = 0,15$$

d. Penyakit *Patch Yellow*

Dalam dataset, terdapat 2 sampel kasus yang memiliki penyakit *patch yellow* adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Patch Yellow}) = \frac{2}{20}$$

$$P(\text{Patch Yellow}) = 0,10$$

e. Penyakit *Upper Stem Rot*

Dalam dataset, terdapat 2 sampel kasus yang memiliki penyakit *Upper Stem Rot* adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Upper Stem Rot}) = \frac{2}{20}$$

$$P(\text{Patch Yellow}) = 0,10$$

f. Penyakit *Ganoderma*

Dalam dataset, terdapat 2 sampel kasus yang memiliki penyakit *Upper Stem Rot* adapun perhitungan probabilitas priornya adalah sebagai berikut :

$$P(\text{Ganoderma}) = \frac{1}{20}$$

$$P(\text{Ganoderma}) = 0,05$$

## 2. Menghitung probabilitas kondisional

Probabilitas kondisional  $P(X|C)$  adalah peluang suatu gejala  $X$  muncul, dengan syarat bahwa bibit sawit memiliki penyakit tertentu  $C$ . Probabilitas ini dihitung dengan rumus :

$$P(X | C) = \frac{\text{Jumlah sampel dengan gejala } X \text{ dalam kelas } C}{\text{Jumlah total sampel dalam kelas } C}$$

Berikut contoh perhitungan probabilitas kondisional

a. Penyakit Antraknosa

Dalam dataset terkait, terdapat total 9 kasus terkait, berikut adalah perhitungan probabilitas kondisional setiap gejala yang terdapat dalam penyakit ini

$$P(\text{Bercak Coklat} = \text{Ya} | \text{Antraknosa}) = \frac{9}{9}$$

$$P(\text{Bercak Coklat} = \text{Ya} | \text{Antraknosa}) = 1.0$$

Berikut merupakan probabilitas kondisional untuk gejala lainnya :

$$P(\text{Garis Kuning} = \text{Tidak} | \text{Antraknosa}) = \frac{4}{9}$$

$$P(\text{Garis Kuning} = \text{Tidak} | \text{Antraknosa}) = 0,44$$

$$P(\text{Layu} = \text{Ya} | \text{Antraknosa}) = \frac{6}{9}$$

$$P(\text{Layu} = \text{Ya} | \text{Antraknosa}) = 0,667$$

$$P(\text{Busuk Batang} = \text{Tidak} | \text{Antraknosa}) = \frac{2}{9}$$

$$P(\text{Busuk Batang} = \text{Tidak} | \text{Antraknosa}) = 0,22$$

$$P(\text{Nekrosis} = \text{Tidak} | \text{Antraknosa}) = \frac{6}{9}$$

$$P(\text{Nekrosis} = \text{Tidak} | \text{Antraknosa}) = 0,667$$

Berdasarkan perhitungan probabilitas kondisional, gejala bercak coklat memiliki probabilitas sebesar 1.0 dalam kasus Antraknosa, yang berarti setiap bibit sawit yang terinfeksi Antraknosa pasti memiliki bercak coklat. Selain itu, kelayuan juga merupakan gejala yang cukup sering terjadi dengan probabilitas 0.667, menunjukkan bahwa dua pertiga dari bibit sawit dengan Antraknosa mengalami kelayuan. Sebaliknya, gejala garis kuning memiliki probabilitas 0.444 untuk tidak muncul, yang berarti garis kuning bukan gejala dominan dalam kasus Antraknosa.

Busuk batang juga jarang terjadi dengan probabilitas 0.222, sementara nekrosis tidak selalu hadir, dengan kemungkinan 66.7% kasus Antraknosa tidak menunjukkan nekrosis. Ini menunjukkan bahwa bercak coklat dan kelayuan adalah gejala utama Antraknosa, sedangkan garis kuning, busuk batang, dan nekrosis kurang umum terjadi.

Oleh karena itu, probabilitas kondisional untuk sebuah gejala X memiliki penyakit Antraknosa adalah :

$$P(X | Atraknosa) = 1.0 \times 0.444 \times 0.667 \times 0.22 \times 0.667$$

$$P(X | Atraknosa) = 0.044$$

#### b. Penyakit Blas Disease

Dalam dataset terkait, terdapat total 3 kasus terkait, berikut adalah perhitungan probabilitas kondisional setiap gejala yang terdapat dalam penyakit ini

$$P(\text{Bercak Coklat} = \text{Ya} | \text{Blas Disease}) = \frac{0}{3} = 0.0$$

$$P(\text{Garis Kuning} = \text{Tidak} | \text{Blas Disease}) = \frac{3}{3} = 1.0$$

$$P(\text{Layu} = \text{Ya} | \text{Blas Disease}) = \frac{3}{3} = 1.0$$

$$P(\text{Busuk Batang} = \text{Tidak} | \text{Blas Disease}) = \frac{3}{3} = 1.0$$

$$P(\text{Nekrosis} = \text{Tidak} | \text{Blas Disease}) = \frac{3}{3} = 1.0$$

Untuk penyakit Blas Disease, tidak ada kasus yang memiliki bercak coklat, sehingga jika bercak coklat muncul, kemungkinan bibit sawit mengalami Blas

Disease sangat kecil. Selain itu, 100% kasus Blas Disease tidak memiliki garis kuning, yang berarti gejala ini tidak pernah muncul dalam penyakit ini. Gejala layu dan busuk batang memiliki probabilitas 1.0, yang menunjukkan bahwa kedua gejala ini selalu muncul pada bibit sawit dengan Blas Disease. Sementara itu, nekrosis juga tidak pernah muncul dalam kasus Blas Disease. Dari hasil ini, dapat disimpulkan bahwa bibit sawit yang terkena Blas Disease akan selalu mengalami kelayuan dan busuk batang, tetapi tidak akan memiliki bercak coklat, garis kuning, atau nekrosis.

Oleh karena itu, probabilitas kondisional untuk sebuah gejala X memiliki penyakit *Blas Disease* adalah :

$$P(X | Blas Disease) = 0$$

c. Penyakit *Bunch Rot*

Dalam dataset terkait, terdapat total 3 kasus terkait, berikut adalah perhitungan probabilitas kondisional setiap gejala yang terdapat dalam penyakit ini :

$$P(Bercak Coklat = Ya | Bunch Rot) = \frac{0}{3} = 0.0$$

$$P(Garis Kuning = Tidak | Bunch Rot) = \frac{3}{3} = 1.0$$

$$P(Layu = Ya | Bunch Rot) = \frac{2}{3} = 0.667$$

$$P(Busuk Batang = Tidak | Bunch Rot) = \frac{3}{3} = 1.0$$

$$P(\text{Nekrosis} = \text{Tidak} | \text{Bunch Rot}) = \frac{0}{3} = 0.0$$

Dalam kasus Bunch Rot, bibit sawit tidak memiliki bercak coklat dengan probabilitas 1.0, sehingga bercak coklat bukanlah indikator utama penyakit ini. Sebaliknya, garis kuning tidak pernah muncul dalam kasus Bunch Rot, yang berarti jika garis kuning terlihat, penyakit ini sangat kecil kemungkinannya terjadi. Kelayuan memiliki probabilitas sebesar 0.667, yang menunjukkan bahwa dua pertiga dari bibit sawit dengan penyakit ini akan mengalami layu. Busuk batang terjadi di semua kasus Bunch Rot (1.0), sehingga ini adalah gejala utama penyakit ini. Sementara itu, nekrosis tidak muncul dalam kasus Bunch Rot, yang berarti gejala ini bukan bagian dari penyakit tersebut. Dengan demikian, bibit sawit yang mengalami busuk batang dan kelayuan tanpa adanya bercak coklat, garis kuning, atau nekrosis kemungkinan besar menderita Bunch Rot.

Oleh karena itu, probabilitas kondisional untuk sebuah gejala X memiliki penyakit *Bunch Rot* adalah :

$$P(X | \text{Bunch Rot}) = 0$$

#### d. Penyakit *Patch Yellow*

Dalam dataset terkait, terdapat total 2 kasus terkait, berikut adalah perhitungan probabilitas kondisional setiap gejala yang terdapat dalam penyakit ini :

$$P(\text{Bercak Coklat} = \text{Ya} | \text{Patch Yellow}) = \frac{0}{2} = 0.0$$

$P$

$$P(\text{Garis Kuning} = \text{Tidak} \mid \text{Patch Yellow}) = \frac{0}{2} = 0.0$$

$$P(\text{Layu} = \text{Ya} \mid \text{Patch Yellow}) = \frac{2}{2} = 1.0$$

$$P(\text{Busuk Batang} = \text{Tidak} \mid \text{Patch Yellow}) = \frac{1}{2} = 0.5$$

$$P(\text{Nekrosis} = \text{Tidak} \mid \text{Patch Yellow}) = \frac{0}{2} = 0.0$$

Bibit sawit yang terkena Patch Yellow tidak pernah memiliki bercak coklat atau garis kuning, yang berarti kedua gejala ini tidak berhubungan dengan penyakit ini. Gejala layu muncul dalam semua kasus Patch Yellow (1.0), sehingga kelayuan bisa menjadi indikator utama penyakit ini. Busuk batang memiliki probabilitas 0.5, yang berarti penyakit ini bisa terjadi baik dengan atau tanpa busuk batang. Selain itu, nekrosis tidak pernah muncul dalam kasus Patch Yellow, sehingga gejala ini bukan bagian dari penyakit ini. Dengan kata lain, kelayuan adalah gejala yang paling kuat dalam mendeteksi Patch Yellow, sementara bercak coklat, garis kuning, dan nekrosis tidak relevan.

Oleh karena itu, probabilitas kondisional untuk sebuah gejala  $X$  memiliki penyakit *Bunch Rot* adalah :

$$P(X \mid \text{Patch Yellow}) = 0$$

e. Penyakit *Upper Stem Rot*

Dalam dataset terkait, terdapat total 2 kasus terkait, berikut adalah perhitungan probabilitas kondisional setiap gejala yang terdapat dalam penyakit ini :

$P$

$$P(\text{Bercak Coklat} = \text{Ya} \mid \text{Upper Stem Rot}) = \frac{1}{2} = 0.5$$

$$P(\text{Garis Kuning} = \text{Tidak} \mid \text{Upper Stem Rot}) = \frac{2}{2} = 1.0$$

$$P(\text{Layu} = \text{Ya} \mid \text{Upper Stem Rot}) = \frac{1}{2} = 0.5$$

$$P(\text{Busuk Batang} = \text{Tidak} \mid \text{Upper Stem Rot}) = \frac{2}{2} = 1.0$$

$$P(\text{Nekrosis} = \text{Tidak} \mid \text{Upper Stem Rot}) = \frac{1}{2} = 0.5$$

Dalam kasus *Upper Stem Rot*, bercak coklat muncul pada 50% kasus, yang berarti ini bukan indikator yang sangat kuat tetapi juga tidak bisa diabaikan. Garis kuning tidak pernah muncul dalam kasus penyakit ini (1.0), sehingga gejala ini bukan bagian dari *Upper Stem Rot*. Gejala layu memiliki probabilitas 0.5, yang berarti penyakit ini bisa terjadi dengan atau tanpa layu. Sementara itu, busuk batang selalu muncul (1.0), sehingga ini adalah gejala utama dari penyakit ini. Terakhir, nekrosis muncul pada 50% kasus, yang berarti tidak semua bibit sawit yang terkena penyakit ini mengalami nekrosis. Dengan demikian, busuk batang adalah indikator yang paling dominan untuk penyakit *Upper Stem Rot*, sementara bercak coklat, layu, dan nekrosis bisa muncul atau tidak

Oleh karena itu, probabilitas kondisional untuk sebuah gejala X memiliki penyakit *Bunch Rot* adalah :

$$P(X \mid \text{Upper Stem Rot}) = 0.5 \times 1.0 \times 0.5 \times 1.0 \times 0.5$$

$$P(X \mid \text{Upper Stem Rot}) = 0.0125$$

*P*

### 3. Probabilitas Posterior

Probabilitas posterior adalah kemungkinan bibit sawit memiliki penyakit tertentu setelah mempertimbangkan semua gejala yang ada. Dihitung dengan rumus teorima bayes.

Karena  $P(X)$  sama untuk semua kelas penyakit, kita cukup membandingkan nilai pembilang untuk menentukan penyakit yang paling mungkin terjadi:

$$P(C|X) \propto P(C) \prod_i P(X_i|C)$$

Artinya, naïve bayes akan mengalikan probabilitas prior  $P(C)$  dengan semua probabilitas kondisional untuk setiap gejala yang kita observasi.

Untuk studi kasus ini, kita memiliki gejala yang diamati:

- **Gejala Bercak Coklat = Ya**
- **Gejala Garis Kuning = Tidak**
- **Gejala Layu = Ya**
- **Gejala Busuk Batang = Tidak**
- **Gejala Nekrosis = Tidak**

Kita akan menghitung probabilitas penyakit:

1. Antraknosa

$P(\text{Antraknosa} | X) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$P(\text{Antraknosa} | X) = 0.45 \times 0.444$

$P(\text{Antraknosa} | X) = 0.0198$

2. Blas Disease

$P(\text{Blas Disease} | X) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Blas Disease} | X) = 0.15 \times 0.0$$

$$P(\text{Blas Disease} | X) = 0.0$$

### 3. Bunch Rot

$P(\text{Bunch Rot} | X) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Bunch Rot} | X) = 0.15 \times 0.0$$

$$P(\text{Bunch Rot} | X) = 0.0$$

### 4. Patch Yellow

$P(\text{Patch Yellow} | X) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Patch Yellow} | X) = 0.10 \times 0.0$$

$$P(\text{Patch Yellow} | X) = 0.0$$

### 5. Upper Stem Rot

$P(\text{Upper Stem Rot} | X) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Upper Stem Rot} | X) = 0.10 \times 0.125$$

$$P(\text{Upper Stem Rot} | X) = 0.0125$$

Berdasarkan hasil perhitungan, penyakit dengan probabilitas tertinggi adalah Antraknosa (0.0198), diikuti oleh Upper Stem Rot (0.0125). Penyakit lainnya memiliki probabilitas 0 karena ada satu atau lebih gejala dengan probabilitas nol dalam penyakit tersebut, yang menyebabkan hasil akhirnya nol.

Dari contoh simulasi perhitungan ini, kita dapat menyimpulkan bahwa bibit sawit dengan gejala yang diamati kemungkinan besar terkena Antraknosa. Namun, ada kemungkinan lebih kecil bahwa bibit tersebut memiliki penyakit Upper Stem Rot.

### **3.9. Implementasi Naïve Bayes menggunakan Python**

Dalam Pembangunan model *Naïve Bayes* untuk klasifikasi penyakit bibit sawit, penggunaan *library Scikit-Learn* memungkinkan proses yang lebih cepat dan efisien. Model ini bekerja dengan menggunakan data numerik yang mewakili karakteristik penyakit. Setiap data tersebut dikategorikan ke dalam beberapa label penyakit sesuai dengan data yang tersedia. Data yang telah dikumpulkan dibagi menjadi set pelatihan dan set pengujian, yang bertujuan untuk melatih model agar mampu mengenali pola dari setiap kategori penyakit.

### **3.10. Evaluasi Algoritma Klasifikasi**

Model klasifikasi yang dibuat kemudian diuji menggunakan data uji yang dipisahkan dari awal untuk memastikan bahwa model dapat memprediksi data baru dengan baik. Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1 Score*, yang berfungsi untuk mengukur sejauh mana model mampu memprediksi penyakit bibit sawit dengan benar. Akurasi dihitung sebagai rasio jumlah prediksi yang benar dibandingkan dengan total prediksi yang dilakukan, sementara *precision* mengukur proporsi prediksi yang benar terhadap keseluruhan prediksi positif yang diberikan model.

*Recall* menunjukkan seberapa baik model dalam menangkap semua kasus positif yang sebenarnya ada, sedangkan *F1 Score* adalah metrik yang menggabungkan *precision* dan *recall* untuk memberikan Gambaran keseimbangan antara keduanya. Sebagai contoh, misalkan hasil pengujian model terhadap 20

sampel uji menghasilkan 13 prediksi benar dan 7 prediksi salah, maka akurasi model dihitung sebagai berikut :

$$Akurasi = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Data}} \times 100\%$$

$$Akurasi = \frac{13}{20} \times 100\%$$

$$Akurasi = 65\%$$

Selain akurasi, *precision*, *recall*, dan *F1 Score* berdasarkan hasil prediksi model menggunakan *Confussion Matrix*, misalkan terdapat data sebagai berikut :

**Tabel 3. 3 Tabel Contoh Confussion Matrix**

Aktual/ Prediksi	Bias Disease	Antraknosa	Bunch Rot	Patch Yellow	Upper Stem Rot	Ganoderma	Total Actual
Bias Disease	3 (TP)	1 (FN)	0	0	0	0	4
Antraknosa	0	4 (TP)	0	0	0	0	4
Bunch Rot	0	0	2	0	0	0	2
Patch Yellow	0	0	0	1 (TP)	0	0	1
Upper Stem Rot	0	0	0	0	2 (TP)	0	2
Ganoderma	0	0	0	0	0	1 (TP)	1

<b>Total</b>	<b>3</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>20</b>
<b>Predicted</b>							

Perhitungan presisi dari masing masing penyakit adalah sebagai berikut :

$$\text{Presisi Bias Disease} = \frac{3}{3+0} = \frac{3}{3} = 1$$

$$\text{Presisi Antraknosa} = \frac{4}{4+1} = \frac{4}{5} = 0,8$$

$$\text{Presisi Bunch Rot} = \frac{2}{2+0} = \frac{2}{2} = 1$$

$$\text{Presisi Patch Yellow} = \frac{1}{1+0} = \frac{1}{1} = 1$$

$$\text{Presisi Upper Stem Rot} = \frac{2}{2+0} = \frac{2}{2} = 1$$

$$\text{Presisi Ganoderma} = \frac{1}{1+0} = \frac{1}{1} = 1$$

Perhitungan *recall* dari masing masing penyakit adalah sebagai berikut :

$$\text{Recall Bias Disease} = \frac{3}{3+1} = \frac{3}{4} = 0,75$$

$$\text{Recall Antraknosa} = \frac{4}{4+0} = \frac{4}{4} = 1,0$$

$$\text{Recall Bunch Rot} = \frac{2}{2+0} = \frac{2}{2} = 1$$

$$\text{Recall Patch Yellow} = \frac{2}{2+0} = \frac{2}{2} = 1$$

$$\text{Recall Upper Stem Rot} = \frac{2}{2+0} = \frac{2}{2} = 1$$

$$\text{Recall Ganoderma} = \frac{1}{1+0} = \frac{1}{1} = 1$$

Perhitungan *F1 Score* untuk masing masing adalah

$$\text{F1 Bias Disease} = 2 \times \frac{1.0 \times 0.75}{1.0 + 0.75} = 2 \times \frac{0.75}{1.75} = 0.8571$$

$$F1 \text{ Antraknosa} = 2 \times \frac{0.8 \times 1.00}{0.8 + 1.00} = 2 \times \frac{0.80}{1.00} = 0.8889$$

$$F1 \text{ Bunch Rot} = 2 \times \frac{1.0 \times 1.0}{1.0 + 1.0} = 2 \times \frac{1.00}{1.00} = 1.000$$

$$F1 \text{ Patch Yellow} = 2 \times \frac{1.0 \times 1.0}{1.0 + 1.0} = 2 \times \frac{1.00}{1.00} = 1.000$$

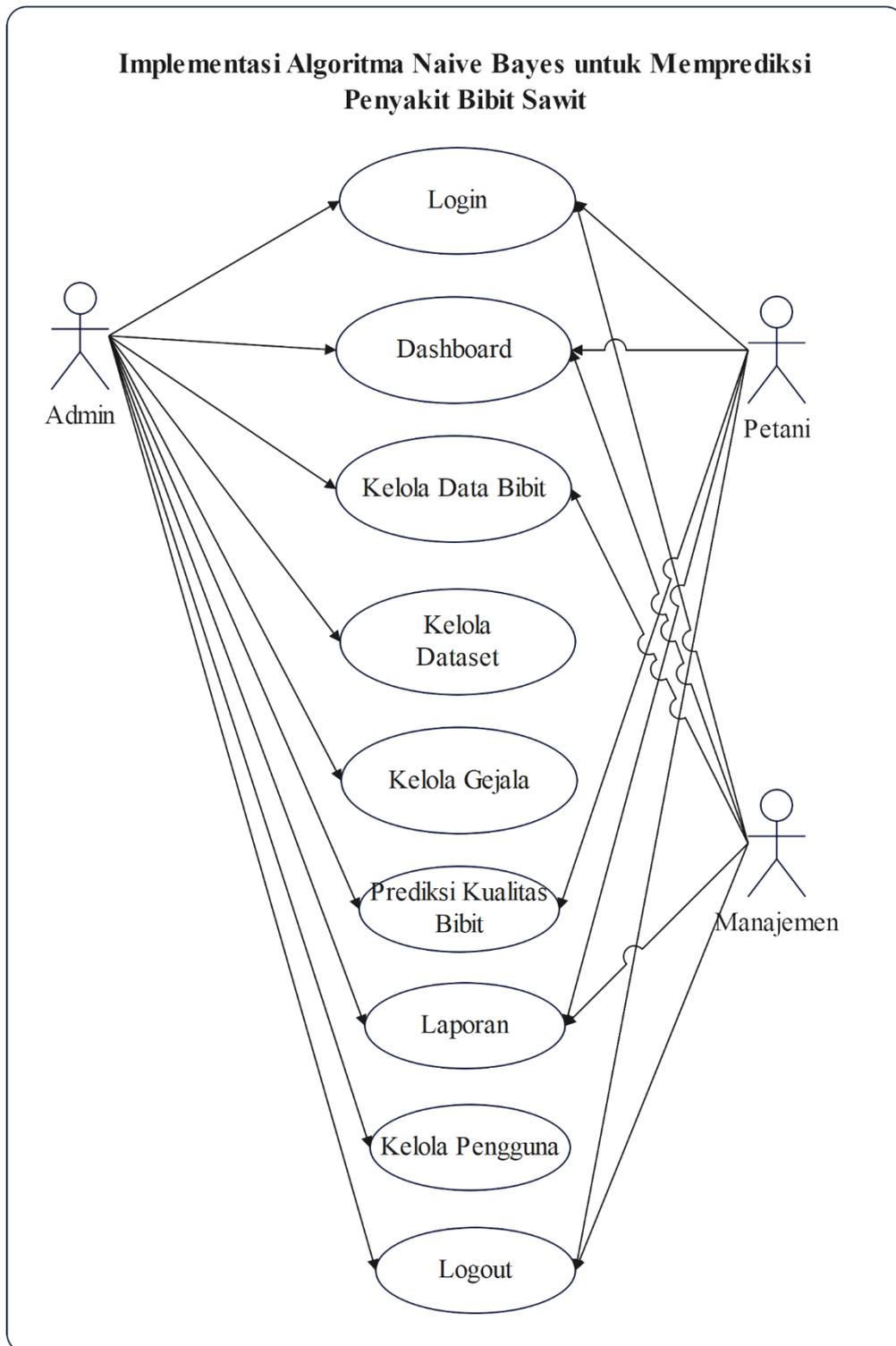
$$F1 \text{ Upper Stem Rot} = 2 \times \frac{1.0 \times 1.0}{1.0 + 1.0} = 2 \times \frac{1.00}{1.00} = 1.000$$

### 3.11. Desain Sistem

Dalam proses ini, sistem akan didesain sedemikian rupa sesuai dengan kebutuhan-kebutuhan fitur yang harus tersedia dalam yang akan dibuat. Proses ini berfokus pada beberapa perancangan seperti struktur data dan arsitektur perangkat lunak yang dibuat dengan pemodelan UML seperti *use case diagram*, *class diagram*, *activity diagram*, dan *diagram* serta rancangan tampilan sistem yang akan dibuat dan rancangan basis data beserta masing masing relasi antar tabelnya.

#### 1. Rancangan *Use Case Diagram*

Secara garis besar, proses sistem yang akan dirancang di tampilkan dalam sebuah diagram *use case*. Yang mana dalam *use case* berisi informasi mengenai aktor dalam sistem dan fitur yang dapat diakses olehnya. Tampilan *use case diagram* dalam terlihat pada gambar dibawah ini.

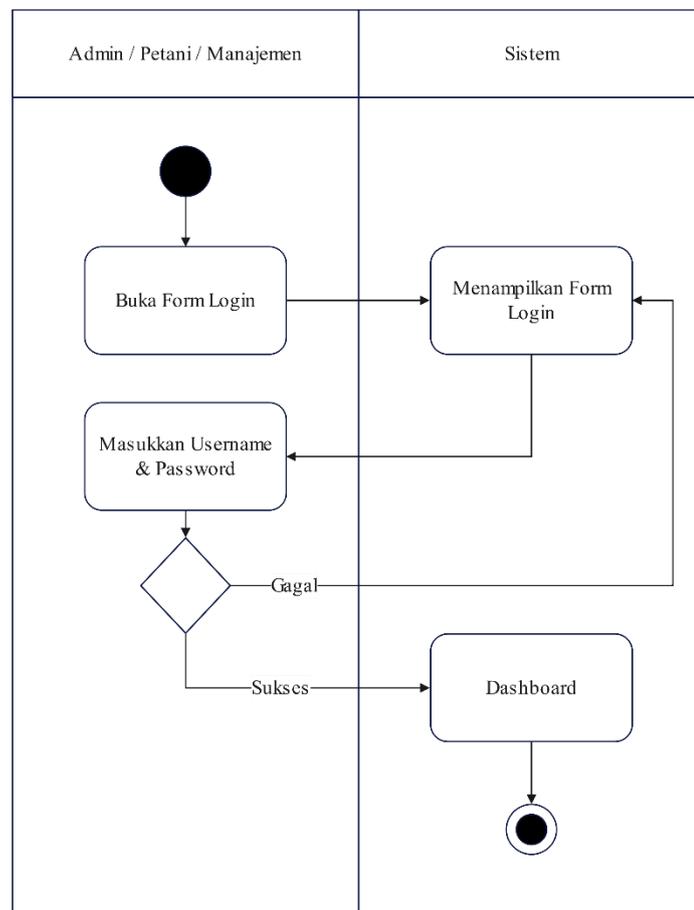


**Gambar 3. 5 Rancangan *Use Case Diagram***

*Use Case Diagram* yang diperbarui ini menggambarkan alur kerja dalam sistem prediksi penyakit bibit sawit menggunakan algoritma Naïve Bayes, dengan tiga jenis pengguna utama: Admin, Petani, dan Manajemen. Admin memiliki akses penuh untuk mengelola data bibit, dataset, gejala, pengguna, serta laporan dan melakukan prediksi kualitas bibit. Petani memiliki akses untuk melakukan prediksi kualitas bibit, melihat laporan hasil prediksi, serta mengakses *dashboard* dan *logout*. Manajemen hanya memiliki akses untuk melihat laporan hasil prediksi dan mengakses *dashboard*, tanpa dapat mengedit atau menambahkan data baru. Semua pengguna harus melakukan *login* terlebih dahulu sebelum mengakses sistem dan dapat melakukan *logout* setelah selesai.

## 2. *Activity Diagram Login*

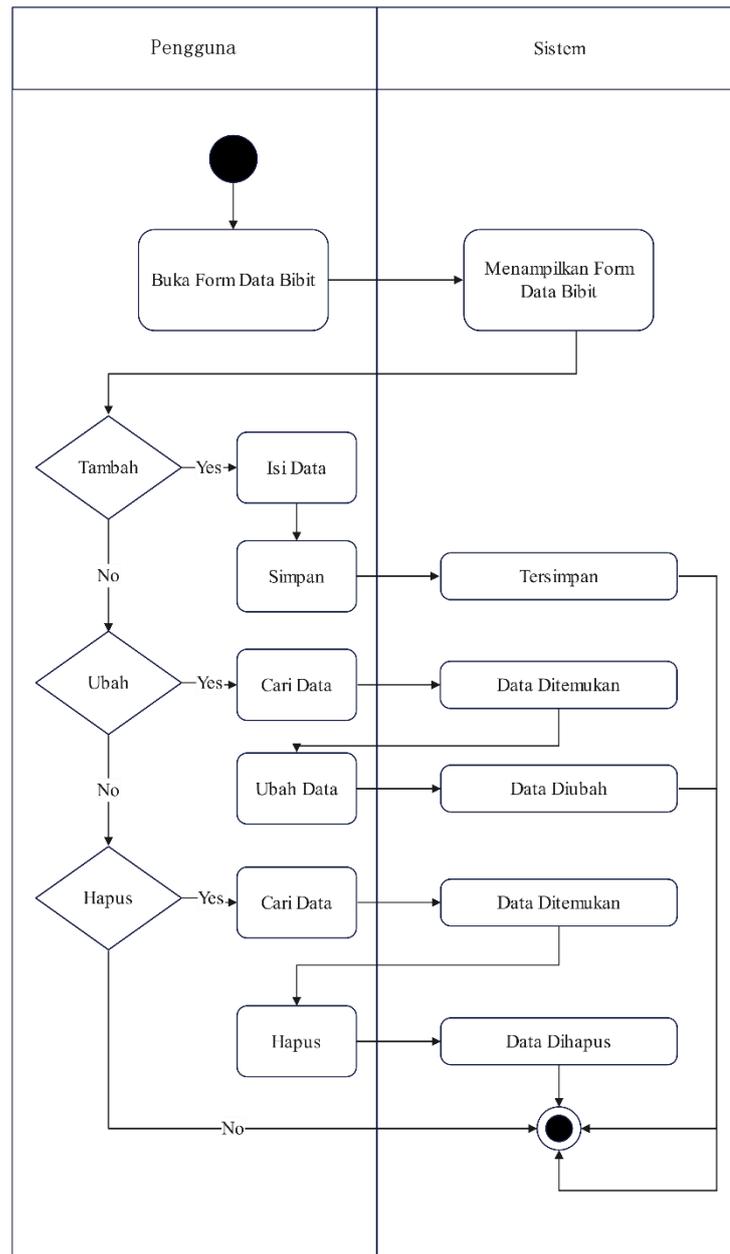
Aktivitas yang dilakukan saat aktor (Admin, Petani, dan Manajemen) melakukan *login* ke sistem, dalam diagram ini menggambarkan proses autentikasi pengguna dalam sistem. Proses dimulai ketika Admin, Petani, atau Manajemen membuka form login, yang kemudian ditampilkan oleh sistem. Pengguna memasukkan username dan password, lalu sistem melakukan verifikasi. Jika kredensial yang dimasukkan tidak valid, sistem akan menampilkan kembali form login untuk mencoba kembali. Jika kredensial benar, pengguna diarahkan ke Dashboard, yang menandakan bahwa login berhasil dan sistem siap digunakan. dapat dilihat secara rinci pada gambar dibawah ini.



**Gambar 3. 6 Activity Diagram Login**

### 3. Activity Diagram Data Bibit

Aktivitas yang dilakukan saat admin mengelola data bibit, mulai dari menambahkan data bibit baru, mengubah data bibit yang sudah ada dapat dilihat pada dibawah ini.

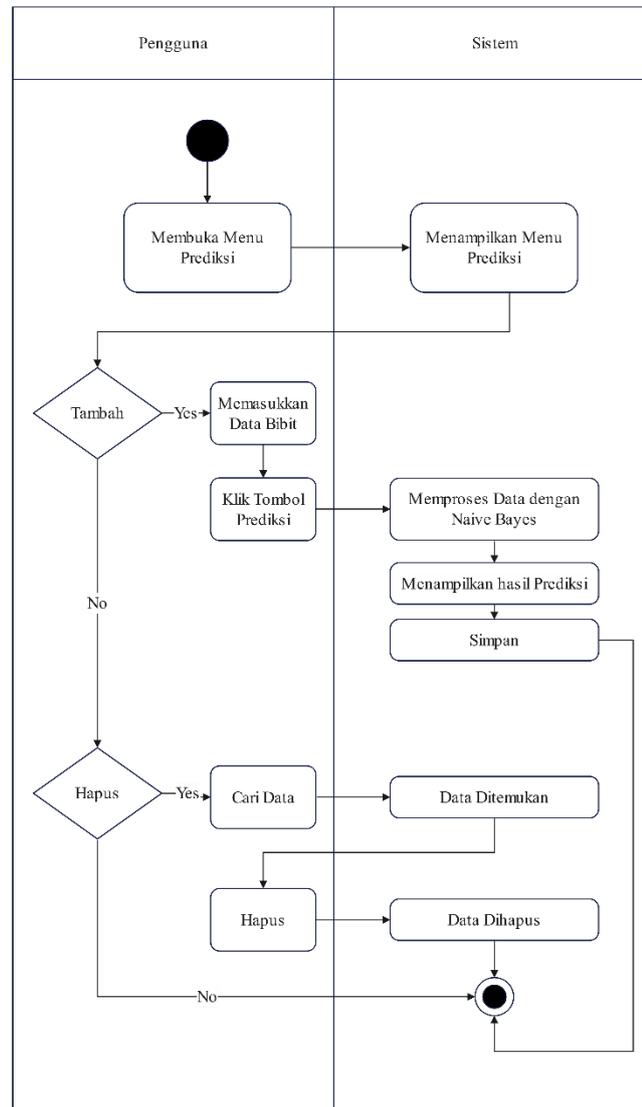


**Gambar 3. 7 Activity Diagram Data Bibit**

#### 4. Activity Diagram Data Prediksi Kualitas Bibit

Aktivitas yang dilakukan dibawah ini menggambarkan proses dalam Menu Prediksi Kualitas Bibit, yang memungkinkan pengguna (Admin, Petani, atau Manajemen) memprediksi kualitas bibit. Proses dimulai dengan pengguna membuka menu prediksi, kemudian mereka dapat memilih data bibit dan

melakukan prediksi dengan mengklik tombol prediksi. Sistem akan memproses data menggunakan algoritma Naïve Bayes dan menampilkan hasil prediksi.

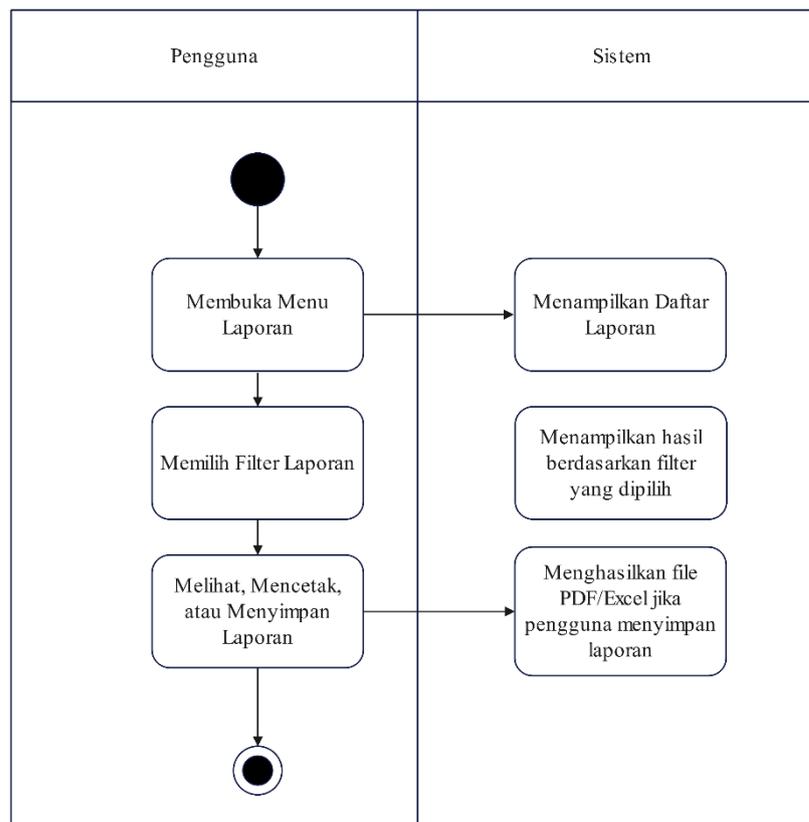


**Gambar 3. 8 Activity Diagram Prediksi Bibit**

### 5. Activity Diagram Laporan

*Activity Diagram* ini menggambarkan alur proses pengelolaan laporan prediksi kualitas bibit sawit dalam sistem. Pengguna (Admin, Petani, atau Manajemen) dapat membuka menu laporan untuk melihat daftar hasil prediksi yang telah dilakukan sebelumnya. Setelah itu, pengguna dapat memilih filter laporan

berdasarkan kriteria tertentu seperti tanggal, jenis penyakit, atau status bibit. Sistem kemudian menampilkan laporan sesuai filter yang dipilih, dan pengguna dapat melakukan tindakan seperti melihat detail laporan, mencetak laporan, atau menyimpannya dalam format PDF atau Excel. Jika selesai, pengguna dapat kembali ke menu utama atau keluar dari menu laporan

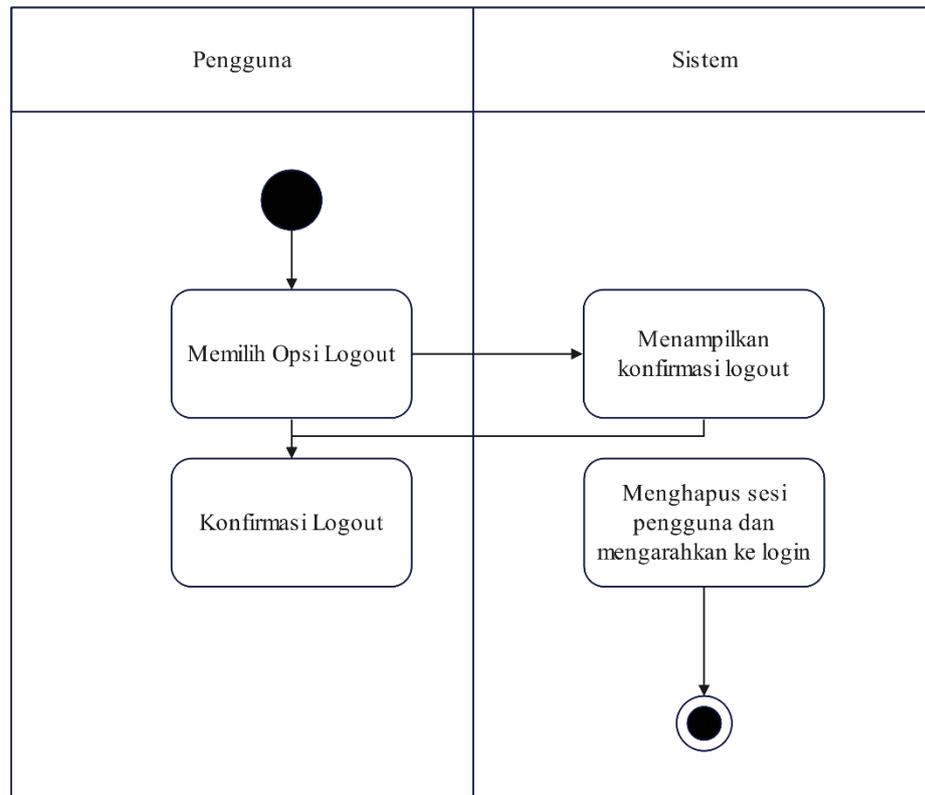


**Gambar 3.9 Activity Diagram Laporan**

#### 6. Activity Diagram Logout

*Activity Diagram Logout* menggambarkan proses pengguna keluar dari sistem. Alur prosesnya dimulai dari pengguna memilih opsi *logout* dari sistem, lalu sistem menampilkan konfirmasi *logout* untuk memastikan pengguna ingin keluar. Setelah pengguna mengkonfirmasi *logout*, sistem akan menghapus sesi pengguna dan

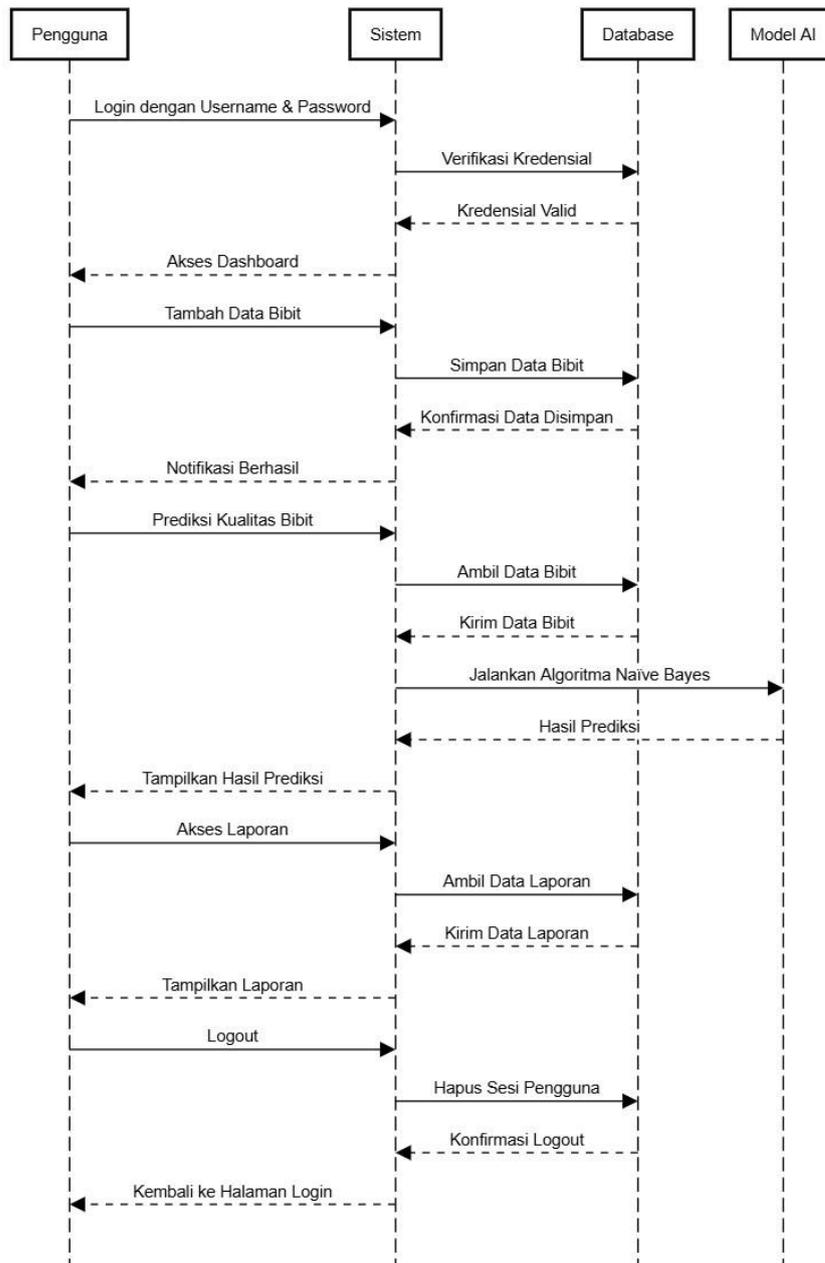
mengarahkan pengguna ke halaman login. *Activity Diagram* terkait dalam dilihat pada gambar dibawah ini.



**Gambar 3. 10 Activity Diagram Logout**

### 7. Sequence Diagram

*Sequence Diagram* dibawah ini menggambarkan alur interaksi antara pengguna (*Admin, Petani, atau Manajemen*) dengan sistem dalam berbagai proses seperti *login*, penambahan data bibit, prediksi kualitas bibit, pembuatan laporan, dan *logout*.

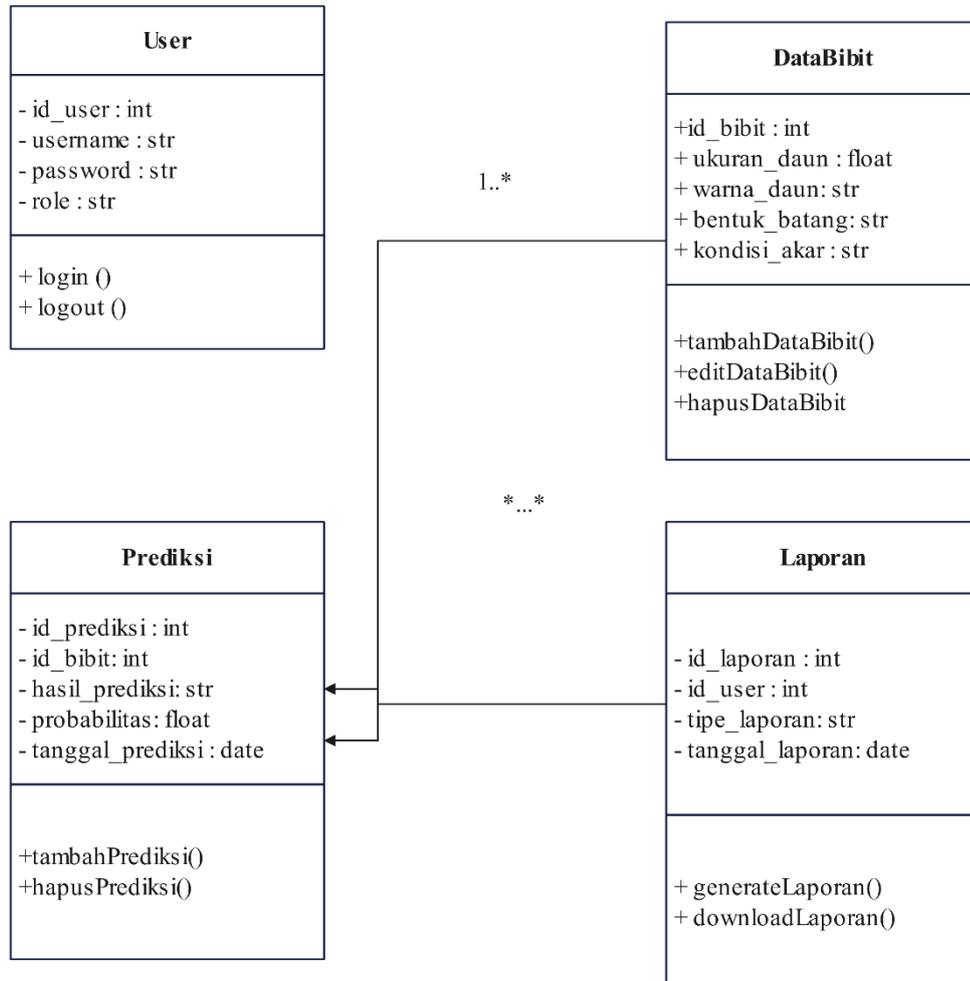


**Gambar 3. 11 Sequence Diagram**

### 8. Class Diagram

*Class Diagram* sistem ini terdiri dari empat kelas utama: User, DataBibit, Prediksi, dan Laporan. User memiliki atribut login dan peran (Admin, Petani, Manajemen). DataBibit menyimpan informasi bibit sawit yang dimasukkan oleh pengguna. Prediksi menggunakan algoritma Naïve Bayes untuk menentukan

kualitas bibit berdasarkan data yang diberikan. Laporan memungkinkan pengguna mengakses dan menyimpan hasil prediksi dalam berbagai format.



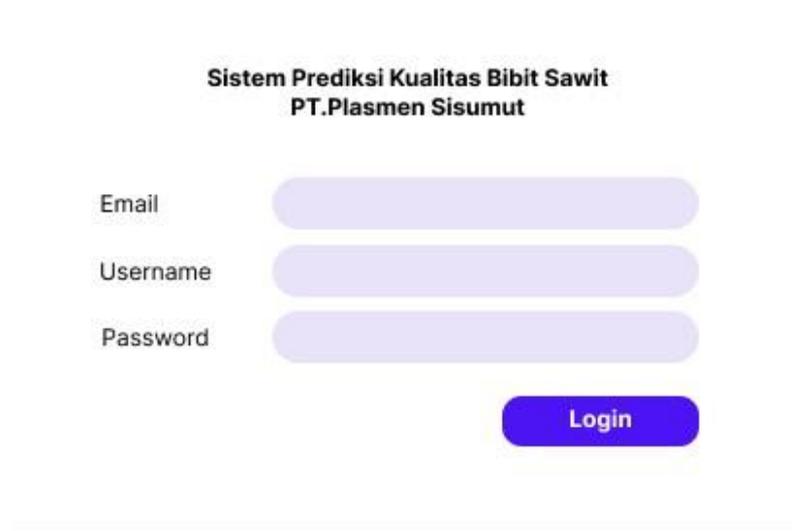
**Gambar 3. 12 Class Diagram**

### 3.12. Perancangan *Interface*

Tahapan ini merancang tampilan interaksi antar muka pengguna untuk digunakan dalam sistem aplikasi, rancangan terkait adalah sebagai berikut :

#### 1. Rancangan *form login*

Rancangan *form* menu *login* dapat dilihat pada gambar dibawah ini.

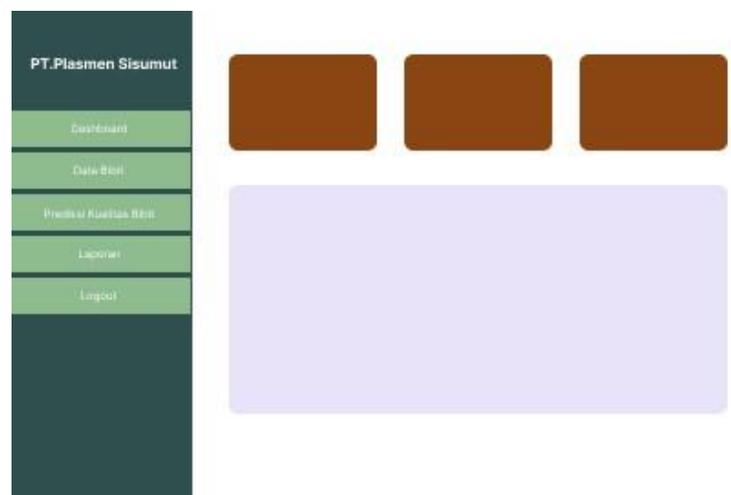


The image shows a login form titled "Sistem Prediksi Kualitas Bibit Sawit PT.Plasmen Sisumut". It features three input fields: "Email", "Username", and "Password", each with a light purple rounded rectangular placeholder. Below these fields is a blue rounded rectangular button labeled "Login".

**Gambar 3. 13 Rancangan Form *Login***

#### 2. Rancangan *form dashboard*

Rancangan *form* menu *dashboard* dapat dilihat pada gambar dibawah ini.



The image shows a dashboard layout. On the left is a dark green vertical sidebar menu with the following items: "PT.Plasmen Sisumut", "Dashboard", "Data Bibit", "Prediksi Kualitas Bibit", "Laporan", and "Logout". To the right of the sidebar, there are three brown square placeholders at the top and a large light purple rectangular placeholder below them.

**Gambar 3. 14 Rancangan Form Dashboard**

3. Rancangan *form* olah data bibit

Rancangan *form* menu olah data bibit dapat dilihat pada gambar dibawah ini.

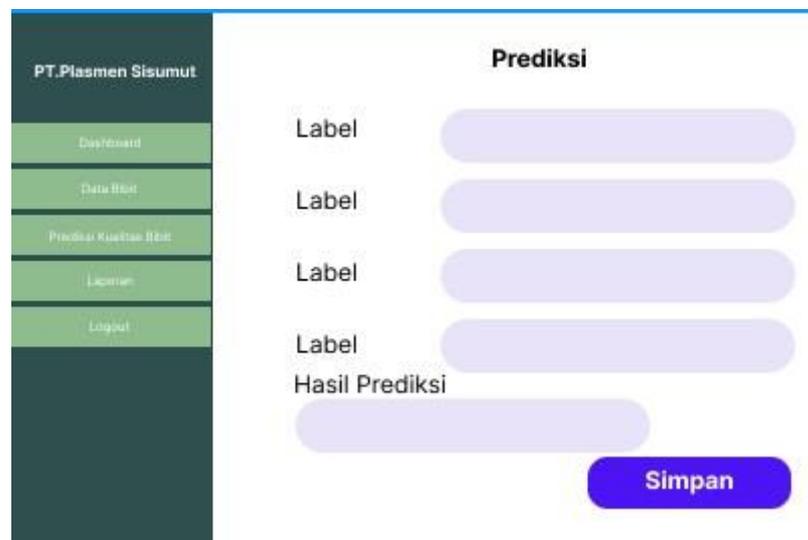


The image shows a mobile application interface for 'PT. Plasmen Sisumut'. On the left is a vertical sidebar menu with the following items: 'Dashboard', 'Data Bibit', 'Prediksi Kualitas Bibit', 'Laporan', and 'Logout'. The main content area is titled 'Tambah Data Bibit' and contains five input fields, each preceded by the text 'Label'. At the bottom right of this section is a blue button labeled 'Tambah Data'.

**Gambar 3. 15 Rancangan Form Olah Data Bibit**

4. Rancangan *form* prediksi

Rancangan *form* menu prediksi dapat dilihat pada gambar dibawah ini.



The image shows a mobile application interface for 'PT. Plasmen Sisumut'. On the left is a vertical sidebar menu with the following items: 'Dashboard', 'Data Bibit', 'Prediksi Kualitas Bibit', 'Laporan', and 'Logout'. The main content area is titled 'Prediksi' and contains four input fields, each preceded by the text 'Label'. Below these is a larger input field labeled 'Hasil Prediksi'. At the bottom right of this section is a blue button labeled 'Simpan'.

**Gambar 3. 16 Rancangan Form Prediksi**

### 5. Rancangan *form* laporan

Rancangan *form* laporan dapat dilihat pada gambar dibawah ini.

The image shows a web application interface. On the left is a dark green sidebar menu with the following items: 'PT.Plasmen Sisumut', 'Dashboard', 'Data Riak', 'Prediksi Kualitas Baku', 'Laporan', and 'Logout'. On the right is a white form titled 'Laporan'. It contains two input fields: 'Periode' and 'Filter Jenis', both with light purple rounded rectangular inputs. Below these fields is a blue button with the text 'Generate' in white.

**Gambar 3. 17 Rancangan Form Laporan**

### 3.13. Penulisan Kode Program Sistem

Pada tahap ini untuk pembuatan sistem diimplementasikan kedalam kode program, Pembangunan sistem aplikasi sendiri akan menggunakan Bahasa pemrograman PHP untuk *front-end*, bahasa pemrograman Python untuk *back-end* dan menjalankan klasifikasi menggunakan *k-nearest neighbor*, serta untuk basis data menggunakan MySQL. Alur sistem yang dibangun sesuai dengan diagram UML atau rancangan aplikasi yang dibuat sebelumnya pada saat desain sistem.

### 3.14. Pengujian Sistem

Setelah pembangunan sistem selesai, pengujian perlu dilakukan untuk menguji fungsionalitas dari setiap fitur dalam aplikasi. Pengujian dilakukan dengan menggunakan aplikasi web yang dijalankan secara lokal (*localhost*) untuk

mengajukan praktek. Tujuan utama dari pengujian ini adalah untuk mengidentifikasi bug dan memverifikasi bahwa aplikasi beroperasi sesuai dengan harapan.

Selama pengujian program, penulis juga memastikan bahwa aplikasi dapat menangani beban kerja yang diantisipasi tanpa mengalami penurunan kinerja atau kegagalan sistem. Hal ini penting untuk memastikan bahwa aplikasi siap digunakan.

### 3.15. Integrasi Model Naïve Bayes

Dengan menggunakan Flask, memungkinkan pembuatan REST API untuk model sehingga model Naïve Bayes dapat diakses dari aplikasi atau sistem berbasis web. Berikut langkah-langkah yang akan digunakan

#### a. Instalasi Flask

Pastikan *Flask* sudah terpasang, jika belum dapat dilakukan dengan kode dibawah ini

```
pip install Flask
```

**Gambar 3. 18 Cara Install Flask**

#### b. Membuat API dengan Flask

Berikut merupakan contoh pembuatan API dengan *Flask*

```

from flask import Flask, request,
jsonify import joblib import numpy as np

# Inisialisasi Flask app app
= Flask(__name__)

# Muat model yang disimpan
model = joblib.load('model_sawit_nb.pkl')

@app.route('/predict', methods=['POST']) def
predict():
    # Ambil data dari request POST
    data = request.get_json()

    # Ubah data ke numpy array untuk prediksi
    features = np.array(data['features']).reshape(1, -1)

    # Prediksi menggunakan model
    prediction = model.predict(features)

    # Buat respon dalam bentuk JSON
    response = {
        'prediction': int(prediction[0]) # Pastikan prediksi dalam
        format integer
    }
    return
    jsonify(response)
    if __name__ ==
    '__main__':
    app.run(debug=True)

```

### Penjelasan Kode:

- a) `app = Flask(__name__)`: Inisialisasi aplikasi Flask.
- b) Memuat Model: `model = joblib.load('model_sawit_nb.pkl')` untuk memuat model yang sudah disimpan dengan Joblib.
- c) Route `/predict`:

1. Method POST: Mengambil data input dari permintaan POST dalam bentuk JSON (harus berupa list fitur).
2. Prediksi: Mengubah data ke array numpy dan menggunakan `model.predict()` untuk prediksi.
3. Respon JSON: Hasil prediksi dikembalikan dalam format JSON.

Setelah Api telah dibuat, hal yang dilakukan selanjutnya adalah mengambil data dari sistem web yang sudah dibangun, dan kemudian mengirimkan data *form* tersebut ke API untuk kemudian diambil prediksinya menggunakan model yang sudah dibangun sebelumnya.

Setelah API dibuat, langkah selanjutnya adalah mengambil data dari form di sistem web yang sudah dibangun dan mengirimkan data tersebut ke API. API kemudian akan memproses data tersebut menggunakan model yang telah dilatih sebelumnya untuk menghasilkan prediksi. Hasil prediksi tersebut akan dikirim kembali ke sistem web dan ditampilkan kepada pengguna.

## BAB IV HASIL DAN PEMBAHASAN

### 4.1. Hasil Pelatihan Model *Naïve Bayes*

#### 4.1.1. Dataset yang Digunakan

Dataset yang digunakan terdiri dari 100 data unik bibit kelapa sawit. Setiap data memuat informasi mengenai atribut pada bibit, seperti struktur batang, dan kondisi tanah, serta gejala yang muncul, seperti bercak coklat, garis kuning, kelayuan, busuk batang, dan nekrosis, dengan label berupa jenis penyakit. 10 baris pertama Dataset terkait dapat dilihat pada tabel dibawah ini

**Tabel 4. 1 10 Baris dataset yang digunakan**

nama	warna daun	struktur batang	kondisi tanah	bercak coklat	garis kuning	kelayuan	busuk batang	nekrosis	penyakit
Bibit 1	Hijau Tua	Kuat	Kering	Tidak	Tidak	Tidak	Tidak	Tidak	Hawar Daun
Bibit 2	Hijau Muda	Kuat	Lempung	Tidak	Tidak	Tidak	Tidak	Ya	Hawar Daun
Bibit 3	Hijau Muda	Lemah	Basah	Tidak	Ya	Ya	Ya	Tidak	Busuk Akar
Bibit 4	Hijau Muda	Kuat	Lempung	Tidak	Ya	Ya	Ya	Tidak	Busuk Akar
Bibit 5	Pucat	Kuat	Basah	Ya	Ya	Tidak	Tidak	Tidak	Blas Disease
Bibit 6	Hijau Tua	Sedang	Lempung	Tidak	Ya	Ya	Ya	Ya	Busuk Akar
Bibit 7	Hijau Muda	Sedang	Lempung	Ya	Tidak	Tidak	Tidak	Tidak	Hawar Daun
Bibit 8	Hijau Muda	Sedang	Kering	Tidak	Ya	Tidak	Tidak	Tidak	Blas Disease

Bib it 9	Hija u Mud a	Sedan g	Lempu ng Ya			Tida k Ya	Ya	Ya	Ya	Antrakn osa
Bibi t 10	Hija u Mud a	Kuat	Lempu ng Ya	.....		Ya	Ya	Ya	Tidak	Busuk Akar

59

#### 4.1.2. Perhitungan Probabilitas Prior

Probabilitas awal dihitung berdasarkan frekuensi kemunculan masing masing penyakit dalam dataset, adapun label penyakit yang ada dalam dataset beserta perhitungan probabilitas prior nya adalah sebagai berikut

##### 1. Penyakit Antraknosa

Terdapat 30 baris data sawit yang memiliki penyakit antraknosa, maka probabilitas prior nya adalah sebagai berikut

$$P(\text{Antraknosa}) = \frac{30}{100} = 0,3$$

##### 2. Penyakit *Blas Disease*

Terdapat 20 baris data sawit yang memiliki penyakit *blas*, maka probabilitas prior nya adalah sebagai berikut

$$P(\text{Blas Disease}) = \frac{20}{100} = 0,2$$

##### 3. Penyakit *Bunch Rot* (Busuk Tandan)

Terdapat 6 baris data sawit yang memiliki penyakit *bunch rot*, maka probabilitas prior nya adalah sebagai berikut

$$P(\text{Bunch Rot}) = \frac{6}{100} = 0,06$$

#### 4. Penyakit Busuk Akar

Terdapat 28 baris data sawit yang memiliki penyakit busuk akar maka probabilitas prior nya adalah sebagai berikut

$$P(\text{Busuk Akar}) = \frac{28}{100} = 0,28$$

#### 5. Penyakit Hawar Daun

Terdapat 16 baris data sawit yang memiliki penyakit busuk akar maka probabilitas prior nya adalah sebagai berikut

$$P(\text{Busuk Akar}) = \frac{16}{100} = 0,16$$

### 4.1.3. Perhitungan Probabilitas kondisional

Probabilitas kondisional dihitung untuk setiap fitur atribut dan gejala terhadap setiap penyakit menggunakan rumus *Naïve Bayes*, yaitu struktur batang, kondisi tanah, dan beberapa gejala lainnya. berikut contoh perhitungan likelihoodnya untuk atribut struktur batang di seluruh label penyakit :

#### 1. Struktur batang dengan nilai kuat pada label penyakit Antraknosa

Perhitungan probabilitas kondisional terkait adalah sebagai berikut

$$P(\text{Fitur} = x \mid \text{kelas} = c) = \frac{\text{Jumlah Fitur pada Kelas} + 1}{\text{Total Kelas} + \text{Jumlah Nilai Fitur Unik}}$$

$$P(\text{Struktur Batang} = \text{Kuat} \mid \text{Antraknosa}) = \frac{10 + 1}{30 + 3} = 0.333$$

2. Struktur batang dengan nilai kuat pada label penyakit *Blas Disease*

Perhitungan probabilitas kondisional terkait adalah sebagai berikut

$$P(\text{Fitur} = x \mid \text{kelas} = c) = \frac{\text{Jumlah Fitur pada Kelas} + 1}{\text{Total Kelas} + \text{Jumlah Nilai Fitur Unik}}$$

$$P(\text{Struktur Batang} = \text{Kuat} \mid \text{Blas Disease}) = \frac{7 + 1}{20 + 3} = 0.348$$

3. Struktur batang dengan nilai kuat pada label penyakit Busuk Akar

Perhitungan probabilitas kondisional terkait adalah sebagai berikut

$$P(\text{Fitur} = x \mid \text{kelas} = c) = \frac{\text{Jumlah Fitur pada Kelas} + 1}{\text{Kelas} + \text{Jumlah Nilai Fitur Unik}} \text{Total}$$

$$P(\text{Struktur Batang} = \text{Kuat} \mid \text{Busuk Akar}) = \frac{12 + 1}{28 + 3} = 0.419$$

Perhitungan dilanjutkan kepada seluruh atribut dan gejala di keseluruhan penyakit, agar lebih mudah, perhitungan dilakukan menggunakan *python* menggunakan kode dibawah ini

```
# LIKELIHOOD dengan Laplace Smoothing
likelihood = defaultdict(lambda: defaultdict(lambda: defaultdict(int)))

print("\n Likelihoods (P(feature=value | class)):")
for c in classes:
    df_c = df[df[label] == c]
    total_c = len(df_c)
    print(f"\nClass '{c}' ({total_c} data):")
    for f in features:
        unique_vals = df[f].unique()
        for v in unique_vals:
            count = len(df_c[df_c[f] == v])
            prob = (count + 1) / (total_c + len(unique_vals))
            likelihood[c][f][v] = prob
            print(f"P({f}={v} | {c}) = {prob:.4f}")
```

**Gambar 4. 1 Kode Perhitungan Probabilitas Kondisional**

Berikut hasil perhitungan probabilitas kondisional setelah dihitung menggunakan *Python*

Tabel 4. 2 Tabel Hasil Perhitungan Probabilitas Kondisional

Fitur	Value	Hawar Daun	Busuk Akar	<i>Blas Disease</i>	<i>Antraknosa</i>	<i>Bunch Rot</i>
Struktur Batang	Kuat	0.4211	0.4194	0.3478	0.3333	0.3333
	Lemah	0.2632	0.2581	0.2174	0.3333	0.3333
	Sedang	0.3158	0.3226	0.4348	0.3333	0.3333
Kondisi Tanah	Kering	0.2632	0.2903	0.2174	0.3333	0.1111
	Lempung	0.4737	0.3871	0.3043	0.3636	0.2222
	Basah	0.2632	0.3226	0.4783	0.3030	0.6667
Bercak Coklat	0	0.7778	0.7333	0.6364	0.0938	0.6250
	1	0.2222	0.2667	0.3636	0.9062	0.3750
Garis Kuning	0	0.9444	0.4000	0.0909	0.4688	0.5000
	1	0.0556	0.6000	0.9091	0.5312	0.5000
kelayuan	0	0.9444	0.0333	0.5909	0.7188	0.6250
	1	0.0556	0.9667	0.4091	0.2812	0.3750
Busuk Batang	0	0.5556	0.0333	0.9545	0.3438	0.3750
	1	0.4444	0.9667	0.0455	0.6562	0.6250
Nekrosis	0	0.6667	0.5667	0.5909	0.0938	0.3750
	1	0.3333	0.4333	0.4091	0.9062	0.6250

#### 4.1.4. Perhitungan Probabilitas Posterior

Perhitungan ini dilakukan untuk mengetahui hasil prediksi dari setiap bibit yang ada dan menentukan label penyakit terkait, berikut contoh perhitungan di salah satu data bibit sawit

Tabel 4. 3 Tabel Data Sampel Bibit Sawit

Fitur	Nilai
Struktur Batang	Kuat
Kondisi Tanah	Basah
Bercak Coklat	Ya
Garis Kuning	Ya
Kelayuan	Tidak
Busuk Batang	Ya
Nekrosis	Ya

Perhitungan posterior dan prediksinya adalah sebagai berikut :

1. Hawar Daun

$P(\text{Hawar Daun}) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Hawar Daun}) = 0.16 \times 0.4211 \times 0.2632 \times 0.2222 \times 0.0556 \times 0.9444 \times 0.4444 \times 0.333$$

$$P(\text{Hawar Daun}) = 0,00003065$$

2. Busuk Akar

$P(\text{Busuk Akar}) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Busuk Akar}) = 0.28 \times 0.4194 \times 0.3226 \times 0.2667 \times 0.6 \times 0.0333 \times 0.9667 \times 0.4333$$

$$P(\text{Busuk Akar}) = 0,000084557$$

3. Blas Disease

$P(\text{Blas Disease}) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Blas Disease}) = 0.20 \times 0.3478 \times 0.4783 \times 0.3636 \times 0.9091 \times 0.5909 \times 0.0455 \times 0.4091$$

$$P(\text{Blas Disease}) = 0,0001209624$$

4. Antraknosa

$P(\text{Antraknosa}) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Antraknosa}) = 0.30 \times 0.3333 \times 0.3030 \times 0.9062 \times 0.5312 \times 0.7188 \times 0.6562 \times 0.9062$$

$$P(\text{Antraknosa}) = 0,00623375$$

5. Bunch Rot

$P(\text{Bunch Rot}) = \text{Probabilitas Prior} \times \text{Probabilitas Kondisional}$

$$P(\text{Bunch Rot}) = 0.06 \times 0.3333 \times 0.6667 \times 0.3750 \times 0.5 \times 0.625 \times 0.625 \times 0.625$$

$$P(\text{Bunch Rot}) = 0,0006113$$

Dengan menjumlahkan seluruh hasil probabilitas diatas, maka nilai pembaginya adalah

$$P(X) = 0.00003065 + 0.00008456 + 0.00012096 + 0.00623375 + 0.0006113$$

$$P(X) = 0.00708122$$

Maka perhitungan posterior akhir untuk prediksi kelas adalah sebagai berikut

$$P(\text{Hawar Daun}) = \frac{0.00003065}{0.00708122} = 0,43\% P$$

$$P(\text{Busuk Akar}) = \frac{0.00008456}{0.00708122} = 1,19\%$$

$$P(\text{Blas Disease}) = \frac{0.0012096}{0.00708122} = 1,43\%$$

$$P(\text{Antraknosa}) = \frac{0.00003065}{0.00708122} = 88,01\%$$

$$P(\text{Bunch Rot}) = \frac{0.0006113}{0.00708122} = 8,63\%$$

Makanya prediksi penyakit bibit terkait adalah Antraknosa.

Perhitungan diatas dilakukan ke dalam seluruh data bibit untuk melatih model yang bisa melatih dan mengingat pola data agar dapat diprediksi dengan data baru, proses prediksi dilakukan menggunakan kode dibawah ini

```

# =====
# 4. FUNGSI PREDIKSI
# =====
def predict(sample):
    posteriors = {}
    for c in classes:
        log_prob = math.log(prior_prob[c])
        for f in features:
            val = sample[f]
            prob = likelihood[c][f].get(val, 1e-6)
            log_prob += math.log(prob)
        posteriors[c] = log_prob

# Softmax Normalisasi ke bentuk probabilitas
max_log = max(posteriors.values())
exp_vals = {k: math.exp(v - max_log) for k, v in posteriors.items()}
total = sum(exp_vals.values())
probs = {k: v / total for k, v in exp_vals.items()}
return max(probs, key=probs.get), probs

```

**Gambar 4. 2 Kode untuk memprediksi Naive Bayes**

Fungsi diatas digunakan untuk memprediksi penyakit berdasarkan data gejala. Pertama, ia menghitung kemungkinan (probabilitas) untuk setiap jenis penyakit dengan cara menjumlahkan log dari peluang awal (prior) dan peluang tiap gejala yang muncul (disebut likelihood). Jika ada gejala yang tidak ditemukan, ia menggunakan angka kecil agar tetap bisa dihitung. Setelah semua kemungkinan dihitung, hasilnya diubah menjadi bentuk persentase agar bisa dibandingkan dengan mudah. Terakhir, fungsi ini akan mengembalikan jenis penyakit yang paling mungkin beserta semua nilai probabilitas dari setiap penyakit.

#### 4.1.5. Evaluasi Model

Berikut merupakan tabel perbandingan hasil prediksi model dengan data aslinya, dapat dilihat pada tabel dibawah ini

<b>id_data</b>	<b>actual</b>	<b>predicted</b>	<b>Hawar Daun</b>	<b>Busuk Akar</b>	<b>Blas Disease</b>	<b>Antraknosa</b>	<b>Bunch Rot</b>
1	Hawar Daun	Hawar Daun	0,920605	0,001272	0,058931	0,006857	0,012335
2	Hawar Daun	Hawar Daun	0,828224	0,001297	0,057095	0,072283	0,0411
3	Busuk Akar	Busuk Akar	0,001322	0,91072	0,022166	0,004379	0,061413
4	Busuk Akar	Busuk Akar	0,002082	0,971498	0,012346	0,002875	0,011198
5	Blas Disease	Blas Disease	0,017789	0,000887	0,851753	0,078518	0,051054

6	Busuk Akar	Busuk Akar	0,001241	0,907974	0,016976	0,044155	0,029654
7	Hawar Daun	Hawar Daun	0,707904	0,000946	0,117483	0,144158	0,029509
8	Blas Disease	Blas Disease	0,050844	0,001838	0,922149	0,009729	0,015441
9	Antraknosa	Antraknosa	0,009696	0,354158	0,001561	0,605958	0,028628
10	Busuk Akar	Busuk Akar	0,001504	0,893383	0,017842	0,070279	0,016992
11	Hawar Daun	Hawar Daun	0,910149	0,034644	0,004495	0,013073	0,037638
12	Antraknosa	Antraknosa	0,050208	0,005357	0,001165	0,849089	0,094181
13	Blas Disease	Blas Disease	0,041771	0,001711	0,815882	0,093651	0,046986
14	Antraknosa	Antraknosa	0,078734	0,0056	0,000646	0,88767	0,02735

15	Hawar Daun	Hawar Daun	0,5378 53	0,0012 96	0,1747 96	0,093882	0,1921 74
16	Busuk Akar	Busuk Akar	0,0364 36	0,8484 75	0,0025 42	0,08834	0,0242 06
17	Antraknosa	Antraknosa	0,0004 14	0,3699 9	0,0067 94	0,597879	0,0249 23
18	Hawar Daun	Hawar Daun	0,9343 93	0,0009 57	0,0465 21	0,004218	0,0139 11
19	Antraknosa	Antraknosa	0,1198 47	0,0086 58	0,0009 84	0,844492	0,0260 2
20	Busuk Akar	Busuk Akar	0,0269 36	0,8467 86	0,0024 16	0,107757	0,0161 06
21	Blas Disease	Blas Disease	0,0014 65	0,0354 46	0,8256 35	0,06883	0,0686 24

22	Hawar Daun	Hawar Daun	0,7369 07	0,0403 92	0,0061 77	0,019052	0,1974 72
23	Blas Disease	Blas Disease	0,0232 9	0,0009 35	0,9292 85	0,004051	0,0424 39
24	Bunch Rot	Blas Disease	0,1173 1	0,1130 38	0,3888 67	0,178778	0,2020 07
25	Antraknosa	Antraknosa	0,0042	0,0003 28	0,4640 53	0,477849	0,0535 7
26	Blas Disease	Blas Disease	0,0459 25	0,0024 37	0,6919 8	0,203867	0,0557 9
27	Hawar Daun	Hawar Daun	0,9343 93	0,0009 57	0,0465 21	0,004218	0,0139 11
28	Busuk Akar	Busuk Akar	0,0307 65	0,8476 46	0,0014 86	0,068196	0,0519 08
29	Antraknosa	Antraknosa	0,0027 37	0,0074 46	0,0108	0,891741	0,0872 75
30	Hawar Daun	Hawar Daun	0,8014 78	0,0446 18	0,0067 19	0,012951	0,1342 35

31	Blas Disease	Blas Disease	0,0020 67	0,0610 18	0,8062 19	0,034673	0,0960 24
32	Busuk Akar	Busuk Akar	0,0329 39	0,9040 01	0,0032 5	0,003541	0,0562 7
33	Antraknos a	Antraknos a	0,0054 36	0,0003 67	0,1638 28	0,8165	0,0138 69
34	Busuk Akar	Busuk Akar	0,0230 67	0,8057 37	0,0045 52	0,083891	0,0827 54
35	Antraknos a	Hawar Daun	0,2769 07	0,2580 26	0,1395 07	0,091956	0,2336 04
36	Busuk Akar	Busuk Akar	0,0329 39	0,9040 01	0,0032 5	0,003541	0,0562 7
37	Blas Disease	Blas Disease	0,0020 67	0,0610 18	0,8062 19	0,034673	0,0960 24

38	Antraknos a	Antraknos a	0,0776 78	0,0084 17	0,0018 03	0,821032	0,0910 69
39	Busuk Akar	Busuk Akar	0,0020 82	0,9714 98	0,0123 46	0,002875	0,0111 98
40	Hawar Daun	Hawar Daun	0,4702 78	0,0295 27	0,0035 84	0,452504	0,0441 08
41	Antraknos a	Antraknos a	0,0009 71	0,0328 31	0,3442 76	0,606136	0,0157 87
42	Antraknos a	Antraknos a	0,0928 29	0,0068 78	0,0012 7	0,872151	0,0268 72
43	Busuk Akar	Busuk Akar	0,0009 2	0,9090 68	0,0161 87	0,054032	0,0197 93
44	Blas Disease	Blas Disease	0,0027 87	0,0759 74	0,8238 6	0,068595	0,0287 83
45	Busuk Akar	Busuk Akar	0,0008 12	0,8692 02	0,0188 58	0,032515	0,0786 13
46	Busuk Akar	Busuk Akar	0,0382 43	0,9446 31	0,0017 15	0,004522	0,0108 89

47	Hawar Daun	Hawar Daun	0,7698 03	0,0012 36	0,0884 46	0,089579	0,0509 35
48	Busuk Akar	Busuk Akar	0,0012 69	0,9106 39	0,0354 63	0,003503	0,0491 26
49	Hawar Daun	Hawar Daun	0,5747 78	0,0267 32	0,0034 07	0,335184	0,0598 99
50	Busuk Akar	Busuk Akar	0,0014 14	0,8612 63	0,0279 5	0,088078	0,0212 95
51	Antraknosa	Antraknosa	0,0052 31	0,0098 84	0,0121 64	0,946975	0,0257 45
52	Blas Disease	Blas Disease	0,0295 17	0,0010 06	0,8327 53	0,115808	0,0209 17
53	Busuk Akar	Busuk Akar	0,0307 65	0,8476 46	0,0014 86	0,068196	0,0519 08

54	Antraknosa	Antraknosa	0,0052 31	0,0098 84	0,0121 64	0,946975	0,0257 45
55	Bunch Rot	Bunch Rot	0,0367 71	0,0818 82	0,1209 26	0,295103	0,4653 19
56	Bunch Rot	Blas Disease	0,0844 52	0,0834 63	0,4665 8	0,171605	0,1939 01
57	Blas Disease	Antraknosa	0,2024 19	0,1280 31	0,2372 21	0,32905	0,1032 79
58	Antraknosa	Busuk Akar	0,0003 37	0,4590 36	0,0156 5	0,456474	0,0685 02
59	Blas Disease	Blas Disease	0,0813 75	0,0028 68	0,8855 43	0,011678	0,0185 36
60	Hawar Daun	Hawar Daun	0,9243 3	0,0343 05	0,0027 39	0,009957	0,0286 69
61	Antraknosa	Antraknosa	0,1262 72	0,0003 48	0,0373 06	0,820282	0,0157 91
62	Blas Disease	Blas Disease	0,0177 89	0,0008 87	0,8517 53	0,078518	0,0510 54

63	Busuk Akar	Busuk Akar	0,0226 44	0,8773 42	0,0023 43	0,069017	0,0286 54
64	Antraknos a	Antraknos a	0,0027 37	0,0074 46	0,0108	0,891741	0,0872 75
65	Busuk Akar	Busuk Akar	0,0012 77	0,8970 86	0,0104 83	0,054531	0,0366 23
66	Antraknos a	Antraknos a	0,0052 31	0,0098 84	0,0121 64	0,946975	0,0257 45
67	Antraknos a	Antraknos a	0,0060 26	0,3301 91	0,0027 44	0,564953	0,0960 86
68	Hawar Daun	Hawar Daun	0,5448 04	0,0012 6	0,1062 33	0,114114	0,2335 89
69	Antraknos a	Antraknos a	0,0078 95	0,0005 42	0,2379 08	0,741068	0,0125 88

70	Antraknos a	Antraknos a	0,0004 14	0,3699 9	0,0067 94	0,597879	0,0249 23
71	Antraknos a	Antraknos a	0,1203 52	0,0003 69	0,0782 26	0,71075	0,0903 03
72	Busuk Akar	Busuk Akar	0,0230 23	0,8563 76	0,0014 3	0,08421	0,0349 61
73	Blas Disease	Blas Disease	0,0417 71	0,0017 11	0,8158 82	0,093651	0,0469 86
74	Blas Disease	Blas Disease	0,0219 62	0,0012 95	0,7280 29	0,088631	0,1600 82
75	Blas Disease	Blas Disease	0,0019 6	0,0388 14	0,9205 24	0,002268	0,0364 34
76	Hawar Daun	Hawar Daun	0,9094 21	0,0455 64	0,0034 65	0,016165	0,0253 86
77	Busuk Akar	Busuk Akar	0,0007 71	0,8125 19	0,0179 04	0,049391	0,1194 15
78	Hawar Daun	Hawar Daun	0,7135 5	0,0015 47	0,1054 07	0,137005	0,0424 92

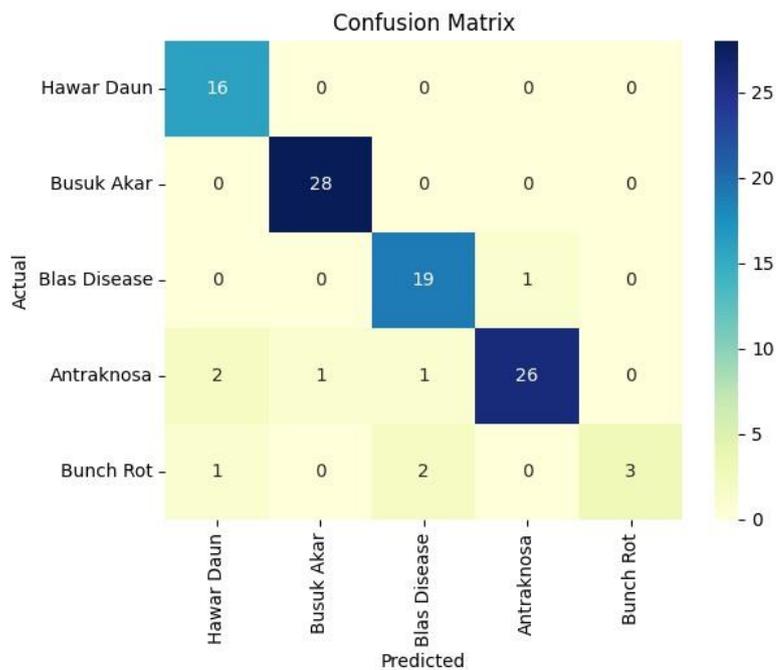
79	Bunch Rot	Hawar Daun	0,7285 88	0,0413 55	0,0029 9	0,151742	0,0753 25
80	Blas Disease	Blas Disease	0,0026 1	0,0438 47	0,8665 65	0,068114	0,0188 64
81	Antraknosa	Antraknosa	0,1095 85	0,0002 9	0,0194 26	0,854254	0,0164 45
82	Antraknosa	Antraknosa	0,0069 59	0,0128 2	0,0097 09	0,944825	0,0256 86
83	Antraknosa	Antraknosa	0,0499 45	0,0047 96	0,0005 27	0,929117	0,0156 15
84	Blas Disease	Blas Disease	0,0654 26	0,0026 12	0,7667 52	0,110014	0,0551 95
85	Antraknosa	Antraknosa	0,0032 41	0,0091 86	0,0213 19	0,880116	0,0861 37
86	Busuk Akar	Busuk Akar	0,0009 2	0,9090 68	0,0161 87	0,054032	0,0197 93
87	Busuk Akar	Busuk Akar	0,0020 96	0,9628 12	0,0124 27	0,00463	0,0180 35
88	Antraknosa	Antraknosa	0,0078 95	0,0005 42	0,2379 08	0,741068	0,0125 88
89	Bunch Rot	Bunch Rot	0,0334 15	0,0714 32	0,0659 33	0,321802	0,5074 18
90	Antraknosa	Blas Disease	0,2222 83	0,2316 48	0,2318 74	0,212933	0,1012 62
91	Busuk Akar	Busuk Akar	0,0012 41	0,9079 74	0,0169 76	0,044155	0,0296 54
92	Blas Disease	Blas Disease	0,0377 68	0,0014 79	0,9042 08	0,004927	0,0516 17
93	Busuk Akar	Busuk Akar	0,0014 94	0,9653 37	0,0189 86	0,004539	0,0096 44
94	Blas Disease	Blas Disease	0,0009 58	0,0241 49	0,8999 78	0,037514	0,0374 01

95	Busuk Akar	Busuk Akar	0,0009 47	0,8432 28	0,0317 55	0,066333	0,0577 36
96	Antraknos a	Antraknos a	0,0264 69	0,0351 08	0,0552 23	0,857878	0,0253 22
97	Busuk Akar	Busuk Akar	0,0196 53	0,8249 11	0,0026 85	0,040841	0,1119 11
98	Busuk Akar	Busuk Akar	0,0010 36	0,8520 8	0,0263 35	0,106503	0,0140 45
99	Antraknos a	Hawar Daun	0,5775 73	0,0368 3	0,0044 01	0,347339	0,0338 57
100	Bunch Rot	Bunch Rot	0,0334 15	0,0714 32	0,0659 33	0,321802	0,5074 18

Evaluasi kinerja model *Naïve Bayes* dilakukan menggunakan metrik Akurasi, *Precision*, *Recall*, dan *F1-Score*. Setelah dilakukan pelatihan dan prediksi terhadap dataset, berikut hasil evaluasi yang diperoleh

#### 1. *Confussion Matrix*

Gambar *Confussion Matrix* dapat dilihat pada gambar dibawah ini :



**Gambar 4. 3 Hasil Prediksi Dalam Confussion Matrix**

*Confusion matrix* di atas merepresentasikan performa model Naive Bayes dalam mengklasifikasikan lima jenis penyakit pada bibit sawit berdasarkan gejala yang diamati. Model menunjukkan performa sangat baik dalam mengklasifikasikan penyakit Hawar Daun dan Busuk Akar, dengan tingkat prediksi yang benar mencapai 100% (masing-masing 16 dan 28 data terklasifikasi tepat). Untuk Blas Disease, model berhasil mengklasifikasikan 19 dari 20 data dengan benar, hanya 1 data yang salah klasifikasi ke Antraknosa. Sementara itu, Antraknosa memiliki akurasi lebih rendah, dengan 4 dari 30 data salah diklasifikasikan ke kelas lain.

Bunch Rot menjadi kelas yang paling sering keliru diklasifikasikan, hanya 3 dari 6 data terdeteksi dengan benar.

## 2. Perhitungan Akurasi

Berdasarkan *Confussion Matrix* di penjabaran sebelumnya, jumlah prediksi yang benar untuk setiap kelas penyakit adalah sebagai berikut :

- a. Hawar Daun : 16
- b. Busuk Akar : 28
- c. *Blas Disease* : 19
- d. Antraknosa : 26
- e. *Bunch Rot* : 3

Maka perhitungan akurasi adalah

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data} \times 100\%$$

$$Akurasi = \frac{16 + 28 + 19 + 26 + 3}{100} \times 100\%$$

$$Akurasi = \frac{92}{100} \times 100\%$$

$$Akurasi = 92\%$$

Dengan demikian, model Naive Bayes menghasilkan akurasi sebesar 92%, yang menunjukkan bahwa model memiliki performa yang baik dalam mengklasifikasikan jenis penyakit pada bibit sawit berdasarkan fitur-fitur gejala yang diberikan.

## 3. Perhitungan *Precision*, *Recall*, dan *F1 Score*

Selain akurasi, evaluasi performa Naive Bayes juga dilakukan dengan menggunakan metrik *Precision*, *Recall*, dan *F1-Score* untuk masing-masing kelas penyakit bibit sawit.

a. Hawar Daun

**TP = 16**

**FP = 3** (2 dari Antraknosa, 1 dari Bunch Rot)

**FN = 0**

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{16}{16 + 3} = \frac{16}{19} \approx 0.8421$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{16}{16} = 1.00$$

$$\text{F1-score} = 2 \times \frac{0.8421 \times 1.00}{0.8421 + 1.00} \approx 0.9143$$

b. Busuk Akar

TP = 28

FP = 1 (dari Antraknosa)

FN = 0

$$\text{Precision} = \frac{28}{28 + 1} = \frac{28}{29} \approx 0.9655$$

$$\text{Recall} = \frac{28}{28} = 1.00$$

$$\text{F1-score} \approx 2 \times \frac{0.9655 \times 1.00}{0.9655 + 1.00} \approx 0.9824$$

c. Blas Disease

TP = 19

FP = 3 (1 dari Antraknosa, 2 dari Bunch Rot)

FN = 1 (1 salah ke Antraknosa)

$$\text{Precision} = \frac{19}{19 + 3} = \frac{19}{22} \approx 0.8636$$

$$\text{Recall} = \frac{19}{19 + 1} = \frac{19}{20} = 0.95$$

$$\text{F1-score} \approx 2 \times \frac{0.8636 \times 0.95}{0.8636 + 0.95} \approx 0.9047$$

d. Antraknosa

$$\text{TP} = 26$$

$$\text{FP} = 1 \text{ (dari Blas Disease)}$$

$$\text{FN} = 4 \text{ (2 salah ke Hawar Daun, 1 ke Busuk Akar, 1 ke Blas)}$$

$$\text{Precision} = \frac{26}{26 + 1} = \frac{26}{27} \approx 0.9630$$

$$\text{Recall} = \frac{26}{26 + 4} = \frac{26}{30} \approx 0.8667$$

$$\text{F1-score} \approx 2 \times \frac{0.9630 \times 0.8667}{0.9630 + 0.8667} \approx 0.9123$$

e. Bunch Rot

$$\text{TP} = 3$$

$$\text{FP} = 0$$

$$\text{FN} = 3 \text{ (1 ke Hawar Daun, 2 ke Blas Disease)}$$

$$\text{Precision} = \frac{3}{3 + 0} = 1.00$$

$$\text{Recall} = \frac{3}{3 + 3} = \frac{3}{6} = 0.50$$

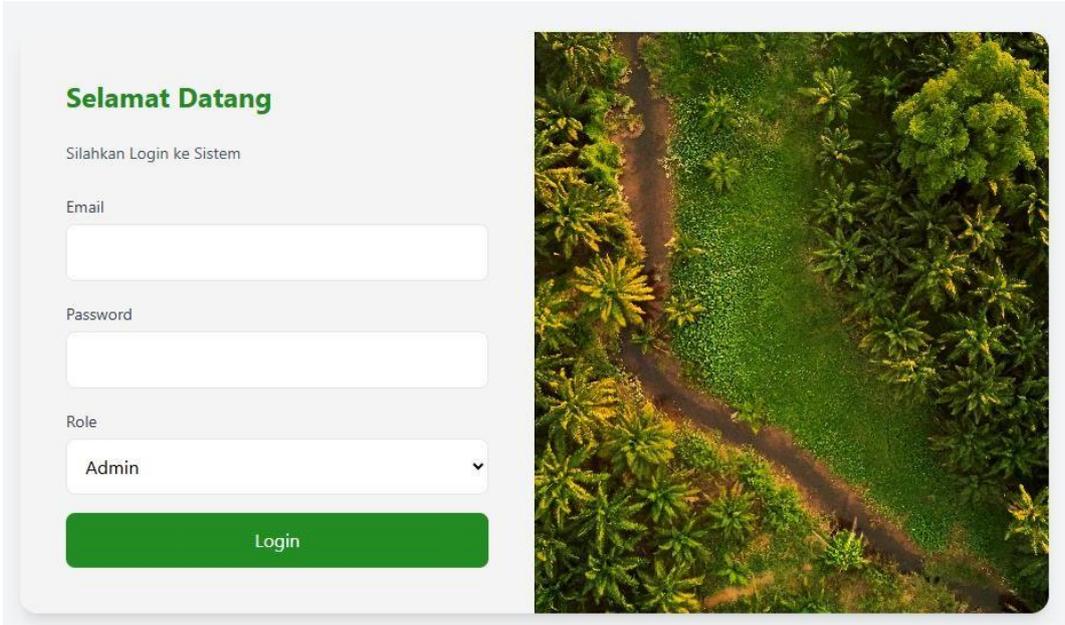
$$\text{F1-score} \approx 2 \times \frac{1.00 \times 0.50}{1.00 + 0.50} = 0.6667$$

## 4.2. Hasil Tampilan Aplikasi

Model Naïve Bayes yang dibangun diimplementasikan dalam aplikasi berbasis web agar mudah digunakan oleh para user, berikut merupakan hasil dari sistem web yang dikembangkan

### 4.2.1. Tampilan Menu *Login*

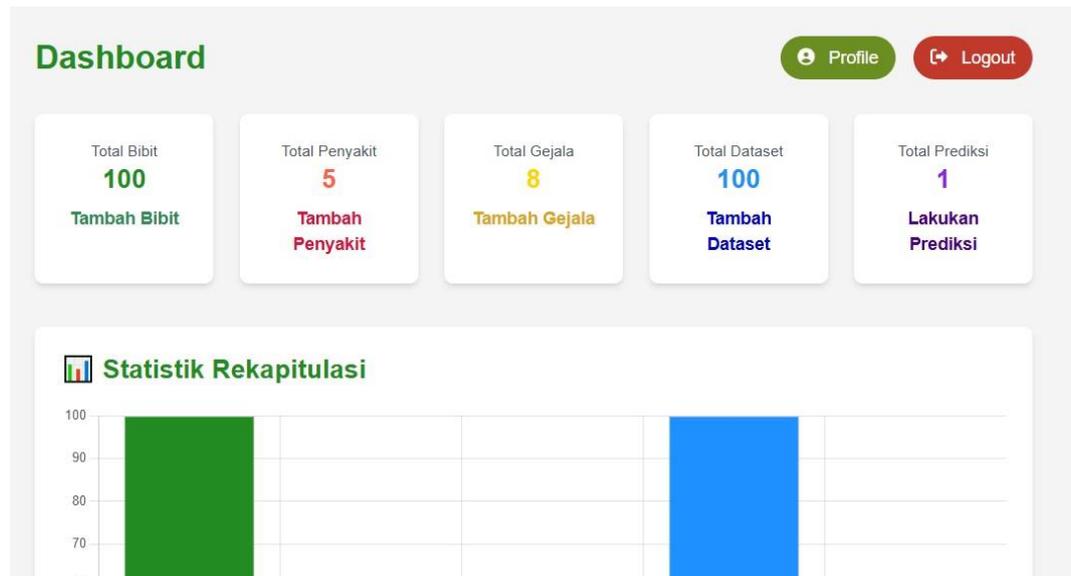
Tampilan menu *Login* dimana pengguna dapat masuk kedalam sistem dan memilih *role* nya dalam sistem dapat dilihat pada gambar dibawah ini



**Gambar 4. 4 Tampilan Menu Login**

### 4.2.2. Tampilan Menu Dashboard

Tampilan menu *Dashboard* dapat dilihat pada gambar dibawah ini



**Gambar 4. 5 Tampilan Menu *Dashboard***

### 4.2.3. Tampilan Menu Bibit

Tampilan menu untuk mengelola bibit terdiri atas menu tambah bibit dan Kelola data bibit.

#### 1. Menu Tambah Data Bibit

Tampilan menu data bibit dimana pengguna dapat menambah data berupa nama bibit, warna daun, struktur batang dan kondisi tanah dapat dilihat pada gambar dibawah ini

The "Tambah Bibit" form is displayed on a dashboard with "Profile" and "Logout" buttons. The form contains the following fields:

- Nama Bibit:** A text input field.
- Warna Daun:** A dropdown menu with "Hijau" selected.
- Struktur Batang:** A dropdown menu with "Tebal" selected.
- Kondisi Tanah:** A text area for input.

At the bottom of the form are two buttons: "Batal" (grey) and "Simpan Bibit" (green).

**Gambar 4. 6 Tampilan Menu Tambah Data Bibit**

## 2. Menu Kelola Data Bibit

Tampilan menu dimana pengguna dapat melihat data bibit yang tersimpan serta melakukan edit dan hapus data bibit terkait dapat dilihat pada gambar dibawah ini

**Daftar Bibit**

Tambah Bibit

Tampilkan 10 data per halaman

Cari:

ID	Nama Bibit	Warna Daun	Struktur Batang	Kondisi Tanah	Aksi
1	Bibit 1	Hijau Tua	Kuat	Kering	<a href="#">Edit</a> <a href="#">Hapus</a>
2	Bibit 2	Hijau Muda	Kuat	Lempung	<a href="#">Edit</a> <a href="#">Hapus</a>
3	Bibit 3	Hijau Muda	Lemah	Basah	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 4. 7 Tampilan Menu Kelola Data Bibit**

### 4.2.4. Tampilan Menu Data Penyakit

Tampilan menu untuk mengelola bibit terdiri atas menu tambah penyakit dan Kelola data penyakit

#### 1. Tampilan Menu Tambah Penyakit

Tampilan menu tambah penyakit dimana pengguna dapat menambah data berupa nama penyakit dan deskripsi penyakit dapat dilihat pada gambar dibawah ini

**Gambar 4. 8 Tampilan Menu Tambah Penyakit**

## 2. Tampilan Menu Kelola Penyakit

Tampilan menu dimana pengguna dapat melihat data penyakit yang tersimpan serta melakukan edit dan hapus data penyakit terkait dapat dilihat pada gambar dibawah ini

ID	Nama Penyakit	Deskripsi	Aksi
1	Hawar Daun	Hawar Daun	<a href="#">Edit</a> <a href="#">Hapus</a>
2	Busuk Akar	Busuk Akar	<a href="#">Edit</a> <a href="#">Hapus</a>
3	Bias Disease	Bias Disease	<a href="#">Edit</a> <a href="#">Hapus</a>
4	Antraknosa	Antraknosa	<a href="#">Edit</a> <a href="#">Hapus</a>
5	Bunch Rot	Bunch Rot	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 4. 9 Tampilan Menu Kelola Penyakit**

### 4.2.5. Tampilan Menu Gejala

Tampilan menu gejala terdiri atas beberapa kelompok tampilan seperti Tambah Gejala, Daftar Gejala, Tambah Atribut Gejala, Kelola Atribut Gejala, Tambah Gejala Bibit, dan Kelola Gejala Bibit

#### 1. Tampilan Tambah Gejala

Tampilan menu tambah gejala yang mana pengguna dapat menambah gejala atau mengimpor dari *excel* dapat dilihat pada gambar dibawah ini

**Gambar 4. 10 Tampilan Menu Tambah Gejala**

## 2. Tampilan Daftar Gejala

Tampilan menu data gejala yang mana pengguna dapat melihat daftar gejala yang ada dapat dilihat pada gambar dibawah ini

ID	Nama Gejala	Aksi
1	warna daun	<a href="#">Edit</a> <a href="#">Hapus</a>
2	struktur batang	<a href="#">Edit</a> <a href="#">Hapus</a>
3	kondisi tanah	<a href="#">Edit</a> <a href="#">Hapus</a>
4	bercak coklat	<a href="#">Edit</a> <a href="#">Hapus</a>
5	garis kuning	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 4. 11 Tampilan Daftar Gejala**

## 3. Tampilan Tambah Atribut Gejala

Tampilan menu tambah atribut gejala yang mana pengguna dapat menambah atribut gejala atau mengimpor dari *excel* dapat dilihat pada gambar dibawah ini

**Gambar 4. 12 Tampilan Tambah Atribut Gejala**

#### 4. Tampilan Kelola Atribut Gejala

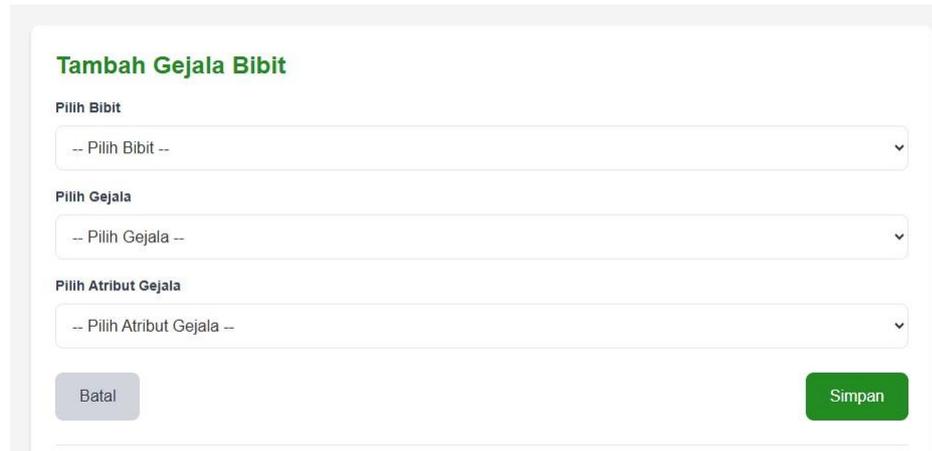
Tampilan menu Kelola atribut gejala yang mana pengguna dapat melihat daftar atribut gejala yang ada dapat dilihat pada gambar dibawah ini

ID	Nama Gejala	Nama Atribut	Aksi
1	warna daun	Hijau Tua	<a href="#">Edit</a> <a href="#">Hapus</a>
2	struktur batang	Kuat	<a href="#">Edit</a> <a href="#">Hapus</a>
3	kondisi tanah	Kering	<a href="#">Edit</a> <a href="#">Hapus</a>
4	bercak coklat	Tidak	<a href="#">Edit</a> <a href="#">Hapus</a>
5	garis kuning	Tidak	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 4. 13 Tampilan Menu Kelola Atribut Gejala**

#### 5. Tampilan Tambah Gejala Bibit

Tampilan menu tambah gejala bibit yang mana pengguna dapat menambah hubungan atau relasi antara gejala, atribut, dan bibit dapat dilihat pada gambar dibawah ini



**Tambah Gejala Bibit**

Pilih Bibit  
-- Pilih Bibit --

Pilih Gejala  
-- Pilih Gejala --

Pilih Atribut Gejala  
-- Pilih Atribut Gejala --

Batal Simpan

**Gambar 4. 14 Tampilan Tambah Gejala Bibit**

## 6. Tampilan Kelola Gejala Bibit

Tampilan menu Kelola gejala bibit yang mana pengguna dapat melihat daftar gejala bibit yang ada dapat dilihat pada gambar dibawah ini



**Kelola Gejala Bibit**

Tambah Gejala Bibit

Tampilkan 10 data per halaman Cari:

ID	Bibit	Gejala	Atribut Gejala	Aksi
1	Bibit 1	warna daun	Hijau Tua	<a href="#">Edit</a> <a href="#">Hapus</a>
2	Bibit 1	struktur batang	Kuat	<a href="#">Edit</a> <a href="#">Hapus</a>
3	Bibit 1	kondisi tanah	Kering	<a href="#">Edit</a> <a href="#">Hapus</a>
4	Bibit 1	bercak coklat	Tidak	<a href="#">Edit</a> <a href="#">Hapus</a>
5	Bibit 1	garis kuning	Tidak	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 4. 15 Tampilan Kelola Gejala Bibit**

### 4.2.6. Tampilan Menu Dataset

Tampilan dalam menu dataset terdiri dari tampilan tambah dataset baru, Kelola dataset, dan lihat dataset detail

#### 1. Tampilan Tambah Dataset

Tampilan pengguna untuk menambah dataset baru dapat dilihat pada gambar dibawah ini

**Tambah Dataset**

Nama Dataset

Pilih Bibit

Pilih Penyakit

Simpan Dataset

Import dari File Excel

Choose File No file chosen

Import Excel

**Gambar 4. 16 Tampilan Menu Tambah Dataset**

## 2. Tampilan Kelola Dataset

Tampilan pengguna untuk melihat dan mengelola dataset yang tersimpan dapat dilihat pada gambar dibawah ini

**Daftar Dataset**

Show 10 entries Search:

#	Nama Dataset	Bibit	Penyakit	Aksi
1	Impor: Bibit 100	Bibit 100	Bunch Rot	Lihat Detail Hapus
2	Impor: Bibit 99	Bibit 99	Antraknosa	Lihat Detail Hapus
3	Impor: Bibit 98	Bibit 98	Busuk Akar	Lihat Detail Hapus
4	Impor: Bibit 97	Bibit 97	Busuk Akar	Lihat Detail Hapus
5	Impor: Bibit 96	Bibit 96	Antraknosa	Lihat Detail Hapus

**Gambar 4. 17 Tampilan Daftar Dataset**

## 3. Tampilan Lihat Dataset Detail

Tampilan pengguna untuk melihat dan mengelola dataset detail dapat dilihat pada gambar dibawah ini

**Detail Dataset**

Nama Dataset: Impor: Bibit 99  
 Nama Bibit: Bibit 99  
 Penyakit: Antraknosa

#	Gejala	Atribut
1	warna daun	Hijau Tua
2	struktur batang	Kuat
3	kondisi tanah	Kering
4	bercak coklat	Ya

**Gambar 4. 18 Tampilan Lihat Dataset Detail**

#### 4.2.7. Tampilan Menu Latih Model & Prediksi

Tampilan ini terdiri atas tampilan latih model dan prediksi model, yang mana menu ini menjalankan algoritma *python* untuk melatih model naïve bayes, menyimpan model, dan melakukan prediksi terhadap data baru.

##### 1. Tampilan Menu Latih Model

Proses latih model dimulai dengan pengguna memilih rentang tanggal dataset yang ingin digunakan sebagai data untuk melatih model, terlihat seperti gambar dibawah ini

**Latih Model Naive Bayes**

Dari Tanggal: 23/03/2025

Sampai Tanggal: 23/03/2025

id_dataset	id_bibit	id_penyakit	nama_dataset	tanggal_buat	nama_bibit	nama_penyakit
98	98	2	Impor: Bibit 98	2025-03-23	Bibit 98	Busuk Akar
49	49	1	Impor: Bibit 49	2025-03-23	Bibit 49	Hawar Daun
42	42	4	Impor: Bibit 42	2025-03-23	Bibit 42	Antraknosa
13	13	3	Impor: Bibit 13	2025-03-23	Bibit 13	Blas Disease
76	76	1	Impor: Bibit 76	2025-03-23	Bibit 76	Hawar Daun
58	58	4	Impor: Bibit 58	2025-03-23	Bibit 58	Antraknosa

**Gambar 4. 19 Dataset setelah difilter tanggal**

Proses dilanjutkan dengan melatih model, proses kode akan dilakukan dalam *backend*, program kemudian menyimpan model yang ada dan menampilkan beberapa hasil Latihan seperti probabilitas prior seperti dibawah ini

#### Probabilitas Prior

- Antraknosa: 0.3
- Blas Disease: 0.2
- Bunch Rot: 0.06
- Busuk Akar: 0.28
- Hawar Daun: 0.16

**Gambar 4. 20 Tampilan Probabilitas Prior**

Probabilitas kondisional juga dapat terlihat seperti pada gambar dibawah ini

#### Probabilitas Kondisional (Likelihood)

##### ▼ Antraknosa

bercak coklat: 0 = 0.0938, 1 = 0.9062  
 busuk batang: 0 = 0.3438, 1 = 0.6562  
 garis kuning: 0 = 0.4688, 1 = 0.5312  
 kelayuan: 0 = 0.7188, 1 = 0.2812  
 kondisi\_tanah: 0 = 0.303, 1 = 0.3333, 2 = 0.3636  
 nekrosis: 0 = 0.0938, 1 = 0.9062  
 struktur\_batang: 0 = 0.3333, 1 = 0.3333, 2 = 0.3333

**Gambar 4. 21 Tampilan Probabilitas Kondisional**

Program juga menghasilkan hasil evaluasi model seperti dibawah ini

#### Evaluasi Model

- Akurasi: 0.92
- Presisi: 0.9268
- Recall: 0.8633
- F1-Score: 0.8761

**Gambar 4. 22 Hasil Evaluasi Model**

## 2. Tampilan Menu Prediksi Model

Tampilan menu prediksi model untuk sistem memprediksi data baru dapat dilihat pada gambar dibawah ini

**Gambar 4. 23 Tampilan Prediksi Model**

Bibit yang telah diproses kemudian menghasilkan prediksi seperti dibawah ini

**Gambar 4. 23 Tampilan Hasil Prediksi**

#### 4.2.8. Tampilan Lihat Hasil Prediksi

Tampilan menu lihat hasil prediksi dapat dilihat pada gambar dibawah ini

ID Prediksi	Bibit	Prediksi Penyakit	Probabilitas	Tanggal Prediksi	Aksi
7	Bibit 90	Blas Disease	0.2319	2025-04-14	Hapus

**Gambar 4. 24 Tampilan Lihat Hasil Prediksi**

#### 4.2.9. Tampilan Laporan Prediksi

Tampilan menu Laporan Prediksi dapat dilihat pada gambar dibawah ini.

No	Tanggal	Nama Bibit	Warna Daun	Struktur Batang	Kondisi Tanah	Hasil Prediksi	Probabilitas (%)
1	2025-04-14	Bibit 90	Hijau Tua	Kuat	Kering	Blas Disease	23.19%

**Gambar 4. 25 Tampilan Laporan Prediksi**

#### 4.3. Pengujian Program

Proses uji coba terhadap sistem wajib dilakukan dengan tujuan memastikan fitur fitur yang ada didalam sistem sudah siap untuk digunakan. Untuk melakukan pengujian, perangkat yang ada harus memiliki spesifikasi sebagai berikut :

1. Satu unit PC dengan spesifikasi *processor* minimum Intel Core I3, dengan RAM minimal 1 GB, dan menyediakan penyimpanan minimal 200 mb untuk menyimpan aplikasi
2. Telah terinstall perangkat lunak seperti XAMPP (Minimum versi 3.3.0 dengan PHP Versi minimum 8.1.2), MySQL server, Visual Studio Code/Sublime Text
3. Dapat terkoneksi dengan internet untuk dapat menjalankan beberapa framework yang ada dalam sistem seperti *Composer*, *Bootstrap*, dan *Jquery*

4. Terinstall *Python* dengan minimum versi 3.1.0, *Python* berfungsi sebagai *back end* server untuk menjalankan klasifikasi Pohon Keputusan

Pengujian sistem dilakukan di lingkungan lokal untuk mengidentifikasi dan memperbaiki masalah sebelum digunakan. Langkah langkah pengujian meliputi verifikasi fungsionalitas setiap fitur, seperti penambahan data obat secara otomatis dan manual, pengunggahan file excel, serta validasi data yang dimasukkan.

#### 4.3.1. Pengujian Login

Hasil pengujian *login* dapat dilihat pada gambar dibawah ini

**Tabel 4. 4 Tabel Hasil Pengujian *Login***

<b>Data Masukan</b>	<b>Spesifikasi</b>	<b>Hasil Pengujian</b>	<b>Kesimpulan</b>
Username dan Password valid	Redirect ke Tampilan Dashboard	Login Berhasil	[✓] Valid
Username dan Password salah	Login Gagal	Login Gagal	[✓] Valid
Username dan Password kosong	Login Gagal	Login Gagal	[✓] Valid

#### 4.3.2. Tabel Pengujian *Dashboard*

Hasil pengujian *dashboard* dapat dilihat pada gambar dibawah ini

**Tabel 4. 5 Tabel Pengujian *Dashboard***

<b>Data Masukan</b>	<b>Spesifikasi</b>	<b>Hasil Pengujian</b>	<b>Kesimpulan</b>
Dashboard	Menampilkan menu utama	Terdapat ringkasan data	[✓] Valid

#### 4.3.3. Tabel Pengujian Data Bibit

Hasil pengujian Data Bibit dapat dilihat pada gambar dibawah ini

**Tabel 4. 6 Tampilan Pengujian Data Bibit**

<b>Data Masukan</b>	<b>Spesifikasi</b>	<b>Hasil Pengujian</b>	<b>Kesimpulan</b>
Tambah Data Bibit	Memasukkan data bibit dan menampilkan data yang telah tersimpan	Data dapat ditambahkan dan muncul di data bibit	[✓] Valid
Edit Data Bibit	Menampilkan data bibit yang ingin diedit dan dapat mengedit data	Data berhasil diedit dan diperbarui	[✓] Valid
Hapus Data Bibit	Muncul pop-up konfirmasi penghapusan dan data berhasil dihapus	Data berhasil dihapus	[✓] Valid

#### 4.3.4. Pengujian Data Penyakit

Hasil pengujian data penyakit dapat dilihat pada gambar dibawah ini

**Tabel 4. 7 Tabel Pengujian Data Penyakit**

<b>Data Masukan</b>	<b>Spesifikasi</b>	<b>Hasil Pengujian</b>	<b>Kesimpulan</b>
Tambah Data Penyakit	Memasukkan data penyakit dan menampilkan data yang telah tersimpan	Data dapat ditambahkan dan muncul di data penyakit	[✓] Valid
Edit Data Penyakit	Menampilkan data penyakit yang ingin diedit dan dapat mengedit data	Data berhasil diedit dan diperbarui	[✓] Valid
Hapus Data Penyakit	Muncul pop-up konfirmasi penghapusan dan data berhasil dihapus	Data berhasil dihapus	[✓] Valid

#### 4.3.5. Tabel Pengujian Data Prediksi

Hasil pengujian data prediksi dapat dilihat pada gambar dibawah ini

**Tabel 4. 8 Tabel Pengujian Data Prediksi**

<b>Data Masukan</b>	<b>Spesifikasi</b>	<b>Hasil Pengujian</b>	<b>Kesimpulan</b>
Latih Model	Memilih dataset dan melatih model	Model Dilatih dan Hasil Pelatihan Model Muncul	[✓] Valid
Lakukan Prediksi	Melakukan Prediksi Berdasarkan Data Baru	Berhasil Menghasilkan Prediksi	[✓] Valid

## **BAB V KESIMPULAN DAN SARAN**

### **5.1. Kesimpulan**

Berdasarkan hasil penelitian dan pengujian program yang telah dilakukan, berikut Kesimpulan dari penelitian ini

1. Pemanfaatan algoritma Naïve Bayes dalam memprediksi penyakit pada bibit kelapa sawit berdasarkan karakteristik fisik terbukti dapat memberikan hasil yang cepat dan cukup akurat. Dengan menggunakan fitur seperti warna daun, struktur batang, kondisi tanah, dan gejala lainnya, algoritma ini mampu mengklasifikasikan jenis penyakit dengan pendekatan berbasis probabilitas. Keunggulan utama dari metode ini adalah kemampuannya dalam menangani dataset dengan jumlah terbatas serta efisiensi dalam proses komputasi. Selain itu, karena prinsip dasarnya yang sederhana, model dapat dengan mudah diimplementasikan dan diperbarui dengan data baru, sehingga meningkatkan fleksibilitasnya dalam sistem prediksi penyakit pada bibit sawit.
2. Tingkat akurasi algoritma Naïve Bayes dalam memprediksi penyakit bibit sawit bergantung pada kualitas dan jumlah data yang digunakan. Berdasarkan hasil pengujian, algoritma ini menunjukkan performa yang cukup baik dalam mengklasifikasikan penyakit (akurasi 92%), terutama jika dataset telah diproses dengan baik dan memiliki distribusi fitur yang sesuai.

### **5.2. Saran**

Adapun yang menjadi saran dalam penelitian ini adalah sebagai berikut :

93

1. Untuk meningkatkan akurasi prediksi, jumlah dan variasi data yang digunakan dalam pelatihan model perlu diperbanyak. Selain itu, seleksi fitur yang lebih relevan dapat dilakukan agar hanya atribut yang memiliki pengaruh signifikan

terhadap klasifikasi penyakit yang digunakan, sehingga mengurangi bias dalam model.

2. Mengombinasikan Naïve Bayes dengan algoritma lain seperti Decision Tree atau Random Forest dapat meningkatkan performa prediksi. Dengan pendekatan hybrid, model dapat menangani kelemahan asumsi independensi dalam Naïve Bayes dan menghasilkan klasifikasi yang lebih akurat.

## **DAFTAR PUSTAKA**

Andrianto, R., & Irawan, F. (2023). *Implementasi Metode Regresi Linear Berganda Pada Sistem Prediksi Jumlah Tonase Kelapa Sawit di PT. Paluta Inti Sawit.*

- Binus Student Corner. (2022). *Teknik pre-processing dan classification dalam data science – Master of Industrial Engineering*. Binus University.  
<https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classificationdalam-data-science/>
- Bisa AI Academy. (2023). *Bisa AI Academy - Course: Machine Learning Dengan Scikit Learn Python*. <https://bisa.ai/course/detail/MzU3/1>
- Budianto, A. J., Ocsa, P., & Saian, N. (2023). Pengembangan Modul Inventory Management pada Aplikasi Master Distribution Centre System Menggunakan Framework Flask di PT XYZ. *Jurnal Teknologi Informasi Dan Komunikasi*, 7(2), 2023. <https://doi.org/10.35870/jti>
- Dicoding Team. (n.d.). *Assessing Data | Belajar Analisis Data dengan Python | Dicoding Indonesia*. Retrieved February 11, 2024, from <https://www.dicoding.com/academies/555/tutorials/30985>
- Furqan, Mhd., Ramadhan Nasution, Y., & Fadillah, R. (2022). Klasifikasi Penyakit Kulit Menggunakan Algoritma Naïve Bayes Berdasarkan Tekstur Warna Berbasis Android. *Jurnal Sains Komputer & Informatika (J-SAKTI)*.
- Haba, R. K., & Pelangi, K. C. (2020). Pengelompokan Buah Jeruk menggunakan Naïve Bayes dan Gray Level Co-occurrence Matrix. *ILKOM Jurnal Ilmiah*, 12(1), 17–24. <https://doi.org/10.33096/ilkom.v12i1.494.17-24>
- Harahap, P. N., & Sulindawaty, S. (2020). Implementasi Data Mining Dalam Memprediksi Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus PT.Arma Anugerah Abadi Cabang Sei Rampah). *MATICS*, 11(2), 46. <https://doi.org/10.18860/mat.v11i2.7821>

- H.M., A. (2019). *Penerapan Metode C45 Dalam Memprediksi Pola Pembelian Bahan Campuran Olahan Karet (Studi Kasus : PT.Anugrah Sibolga Lestari)*. <http://repository.potensi-utama.ac.id/jspui/jspui/handle/123456789/3504>
- Iriane, R., & Nurfaizah. (2023). Penerapan Data Mining Untuk Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode K-Nearest Neighbor. *KLIK : Kajian Ilmiah Informatika Dan Komputer*, 3(5), 509–515. <https://djournals.com/klik>
- Jennie, G., Kasiliyani, A., Arifha Saputra, W., & Wardiah, I. (2021). *SISTEM PENGAMBILAN KEPUTUSAN BEBAN KINERJA MENGGUNAKAN NAIVE BAYES STUDI KASUS PDAM BANDARMASIH*.
- Marcelina, D., Yulianti, E., & Mair, Z. R. (2022). Penerapan Metode Forward Chaining Pada Sistem Pakar Identifikasi Penyakit Tanaman Kelapa Sawit. *Jurnal Ilmiah Informatika Global*, 13(2). <https://doi.org/10.36982/jiig.v13i2.2299>
- Marpaung, F., Aulia, F., Suryani SKom, N., & Cyra Nabila SKom, R. (2022). *COMPUTER VISION DAN PENGOLAHAN CITRA DIGITAL*. [www.pustakaaksara.co.id](http://www.pustakaaksara.co.id)
- Maryana, A., Putri, N., & Octaviano, A. (2023). *Sistem Pendukung Keputusan Kenaikan Jabatan Menggunakan Metode Naive Bayes (Studi Kasus : PT Buana Mulia Indonesia)*. <https://journal.mediapublikasi.id/index.php/logic>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).
- Pratama, A. (2020). *HTML Uncover - Panduan Belajar HTML Untuk Pemula*. [www.duniaikom.com](http://www.duniaikom.com)
- Putra, V., Tri Pranoto, G., & Eko Putra, F. (2023). Klasifikasi Kebutuhan Sparepart Dengan Algoritma K-Nearest Neighbor Untuk Meningkatkan Penjualan Sparepart.

- Bulletin of Information Technology (BIT)*, 4(2), 287–293.  
<https://doi.org/10.47065/bit.v3i1>
- Rahman, W., Saudin, L., & Sri Wahyuni, N. (2022). *BAHAN AJAR SISTEM INFORMASI MANAJEMEN*. [www.penerbitwidina.com](http://www.penerbitwidina.com)
- Ramdhan, M. (2021). *Metode Penelitian*.  
[https://books.google.co.id/books?id=Ntw\\_EAAAQBAJ&printsec=frontcover&hl=id&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](https://books.google.co.id/books?id=Ntw_EAAAQBAJ&printsec=frontcover&hl=id&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false)
- Raza, M. N. (2024). *SISTEM DETEKSI BERITA HOAX MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN RANDOM FOREST PADA MACHINE LEARNING*.
- Rumbaugh, James., Jacobson, Ivar., & Booch, Grady. (2021). *The unified modeling language reference manual*. Addison-Wesley.
- Sari, I. P., Jannah, A., Meuraxa, A. M., Syahfitri, A., & Omar, R. (2022). Perancangan Sistem Informasi Penginputan Database Mahasiswa Berbasis Web. *Hello World Jurnal Ilmu Komputer*, 1(2), 106–110. <https://doi.org/10.56211/helloworld.v1i2.57>
- Setiawan, S. (2020). *Membicarakan Precision, Recall, dan F1-Score*. Medium.Com.  
<https://stevkarta.medium.com/membicarakan-precision-recall-dan-f1-scoree96d81910354>
- Sinaga, S., Sembiring, R. W., & Sumarno, S. (2022). Penerapan Algoritma Naive Bayes untuk Klasifikasi Prediksi Penerimaan Siswa Baru. In *Journal of Machine Learning and Data Analytics (MALDA)* (Vol. 1, Issue 1).
- Tumini, & Fitria, M. (2021). PENERAPAN METODE SCRUM PADA E-LEARNING STMIK CIKARANG. *Jurnal Informatika SIMANTIK*, 6(1).  
<https://www.simantik.panca-sakti.ac.id>