

BAB II

TINJAUAN PUSTAKA

Dalam bab ini dijelaskan tentang penelitian sistem prediksi yang menggunakan metode algoritma *Naïve Bayes*. Bab ini juga menjelaskan tentang teori penunjang yang digunakan dalam pembuatan sistem.

2.1 Studi Sebelumnya

Berdasarkan penelitian terdahulu Harianto (2019), yang berjudul “Klasifikasi Hasil Panen Padi Berdasarkan Fisiologis Menggunakan Metode *Naïve Bayes Classification*” dari hasil penelitian tersebut dapat disimpulkan bahwa sistem prediksi dapat membantu para petani guna memprediksi tingkat keberhasilan panen pada musim tersebut.

Penelitian kedua yang dilakukan oleh Zulfikar (2019), dengan judul “Penerapan Sistem Pendukung Keputusan Dengan Metode *Naïve Bayes* Dalam Menentukan Kualitas Bibit Padi Unggul Pada Balai Pertanian Pasar Miring” dari hasil penelitian tersebut dapat disimpulkan bahwa metode *Naïve Bayes* dapat membantu dalam pengambilan sebuah keputusan dengan mengolah data yang telah dikumpulkan dengan tujuan untuk memprediksi hasil dan langkah yang harus diambil kedepannya.

Penelitian ketiga yang dilakukan oleh Basit (2020), yang berjudul “Implementasi Algoritma *Naïve Bayes* Untuk Memprediksi Hasil Panen Padi” dari hasil penelitian tersebut dapat disimpulkan bahwa metode *Naïve Bayes* banyak digunakan untuk proses klasifikasi karena *Naïve Bayes* lebih disukai disebabkan kecepatan dan kesederhanaannya. Penelitian keempat yang relevan dilakukan oleh Maesaroh (2017), yang berjudul “Sistem Prediksi Produktifitas Pertanian Padi Menggunakan Data Mining” dari hasil penelitian tersebut dapat disimpulkan bahwa metode *Naïve Bayes* memiliki presentasi tingkat kesalahan terkecil dari pada berapa metode lainnya.

Dari empat penelitian terkait yang telah dijelaskan, maka dapat diambil kesimpulan bahwa sistem prediksi dibutuhkan untuk membantu mengambil keputusan dalam menganalisa faktor penyebab keberhasilan dalam panen sehingga dapat meningkatkan tingkat keberhasilan panen pada musim selanjutnya, maka penulis ingin membuat sistem prediksi hasil panen tanaman tembakau berbasis *website* menggunakan metode naïve bayes.

2.2 Dasar Teori

berikut adalah teori-teori relevan yang menjadi acuan penulis dalam melakukan penelitian dan pembuatan sistem.

2.2.1 Pengertian Prediksi

Prediksi dalam Metode ilmiah atau proses ilmiah merupakan proses keilmuan untuk memperoleh pengetahuan secara sistematis berdasarkan bukti fisis. Ilmuwan melakukan pengamatan serta membentuk hipotesis dalam usahanya untuk menjelaskan fenomena alam. Prediksi yang dibuat berdasarkan hipotesis tersebut diuji dengan melakukan eksperimen. Jika suatu hipotesis lolos uji berkali-kali, hipotesis tersebut dapat menjadi suatu teori ilmiah.

2.2.2 Pengertian Tanaman Tembakau

Tanaman tembakau merupakan tanaman perkebunan sebagai bahan baku industri untuk produk Rokok yang memiliki nilai ekonomis sangat menjanjikan. Sistem budidaya tanaman tembakau sama dengan sistem budidaya untuk tanaman perkebunan lainnya yaitu terdiri dari pembibitan, penyiapan lahan, penanaman, pemeliharaan, panen dan pasca panen.

2.2.3 Pengertian Data Mining

Data mining adalah salah satu bentuk implementasi yang diterapkan untuk mencari sebuah model dan pola yang mampu melakukan prediksi pada suatu data berdasarkan data sebelumnya di periode waktu tertentu.

2.2.4 Pengertian Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa

lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi (Herdianto, 2016).

2.2.5 Naïve Bayes

Naïve Bayes Classifier merupakan, pengklasifikasi probabilitas sederhana berdasarkan pada Teorema Bayes. Teorema Bayes akan dikombinasikan dengan “Naïve” yang artinya pada setiap atribut/*variable* bersifat bebas (*independent*). Naïve Bayes Classifier dapat dilatih dengan efisien dalam pembelajaran terawasi (*supervised learning*).

Keuntungan penggunaan adalah bahwa metoda ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yg diperlukan dalam proses pengklasifikasian. Karena yg diasumsikan sebagai *variabel independent*, maka hanya varian dari suatu variabel dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari *matriks kovarians*.

Kelebihan dan Kekurangan Naive Bayes adalah :

1. Kelebihan
 - a. Mudah untuk dibuat
 - b. Hasil bagus
2. Kekurangan
 - a. Asumsi independence antar atribut membuat akurasi berkurang (karena biasanya ada keterkaitan).

Tahapan dari proses algoritma Naive Bayes adalah :

1. Siapkan Dataset
2. Hitung jumlah kelas pada data training
3. Hitung jumlah kasus yang sama dengan kelas yang sama
4. Kalikan semua hasil sesuai dengan data testing yang akan dicari kelasnya

5. Bandingkan hasil tiap kelas, nilai tertinggi ditetapkan sebagai kelas terbaru

Berikut Formula umum Naive Bayes: Keterangan :

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

X : Data testing yang kelasnya belum di ketahui

H : Hipotesis data X yang merupakan satu klas spesifik

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

$P(H)$: Probabilitas hipotesis (prior probability)

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

$P(X)$: Probabilitas X

Data prediksi hasil panen tanaman tembakau berdasarkan variabel yang telah ditentukan.

1. prediksi hasil panen tanaman tembakau berdasarkan variabel yang telah ditentukan.

Tabel 2.1 Dataset Untuk Contoh Perhitungan

No	PUPUK	CUACA	Serangan Hama	kondisi tanah	KELAS
1	Cukup	Baik	Sedang	Kering	Meningkat
2	Cukup	Panas	Sedang	Kering	Meningkat
3	Kurang	Panas	Rendah	Basah	Meningkat
4	Cukup	Panas	Rendah	Kering	Meningkat
5	Sedang	Panas	Tinggi	Lembab	Meningkat
6	Sedang	Panas	Rendah	Lembab	Meningkat
7	Kurang	Panas	Tinggi	Basah	Menurun

Tabel 2.1 Lanjutan Dataset Untuk Contoh Perhitungan

8	Kurang	Baik	Rendah	Basah	Meningkat
9	Cukup	Hujan	Tinggi	Kering	Menurun
10	Cukup	Baik	Rendah	Kering	meningkat

1. menentukan probabilitas variabel terhadap kategori, dari data training ditentukan nilai probabilitas dari setiap variabel:
 - a. probabilitas penggunaan pupuk pada setiap kategori hasil panen

Tabel 2.2 Probabilitas Penggunaan Pupuk

Himpunan	Jumlah Kategori Pupuk		Probabilitas Kategori Pupuk	
	Meningkat	Menurun	Meningkat	Menurun
Cukup	4	1	0,5	0,5
Sedang	2	0	0,25	0
Kurang	2	1	0,25	0,5

- b. Probabilitas kultur cuaca pada setiap kategori hasil panen

Tabel 2.3 Probabilitas Kultur cuaca

Himpunan	Jumlah Kategori Cuaca		Probabilitas Kategori Cuaca	
	Meningkat	Menurun	Meningkat	Menurun
Panas	5	1	0,625	0,5
Baik	3	0	0,375	0
Hujan	0	1	0	0,5

c. Probabilitas serangan hama pada setiap kategori hasil panen

Tabel 2.4 Probabilitas Serangan Hama

Himpunan	Jumlah Kategori Serangan Hama		Probabilitas Kategori Serangan Hama	
	Meningkat	Menurun	Meningkat	Menurun
Tinggi	1	2	0,125	1
Sedang	2	0	0,25	0
Rendah	5	0	0,625	0

d. Probabilitas kondisi tanah pada setiap kategori hasil panen

Tabel 2.4 Probabilitas Kondisi Tanah

Himpunan	Jumlah Kategori Kondisi Tanah		Probabilitas Kategori Kondisi Tanah	
	Meningkat	Menurun	Meningkat	Menurun
Kering	4	1	0,5	0,5
Lembab	2	0	0,25	0
Basah	2	1	0,25	0,5

e. Probabilitas untuk setiap kategori hasil panen

Tabel 2.5 Probabilitas Hasil Panen

Himpunan	Meningkat	Menurun
Jumlah	8\10	2\10

Dalam Pengguna menginputkan data kriteria hasil panen tembakau dengan kriteria:

1. Penggunaan pupuk : Cukup
2. Kultur Cuaca : Baik
3. Serangan hama : Sedang
4. Kondisi tanah : Kering

$$5. \text{Meningkat} = 0,5 * 0,375 * 0,25 * 0,5 * 0,8 \\ = 0,01875$$

$$6. \text{Menurun} = 0,2 * 0,5 * 0 * 0 * 0,5 * 0,2 \\ = 0$$

Tabel 2.6 Perhitungan Naïve Bayes

Kategori	Himpunan	Meningkat	Menurun
		0,8	0,2
Pupuk	Cukup	0,5	0,5
Cuaca	Baik	0,375	0
Hama	Sedang	0,25	0
Tanah	Kering	0,5	0,5
Akhir		0,01875	0

Karena nilai akhir kelas “Meningkat” lebih besar dari kelas “Menurun” maka hasil prediksi data uji X masuk dalam kategori kelas “meningkat”.

2.2.6 Pengertian Web

Pemrograman Web terdiri dari dua kata yaitu pemrograman (proses/cara/pembuatan) dan web yang artinya jaringan komputer yang terdiri dari kumpulan situs internet yang menampilkan informasi seperti teks, gambar, suara, animasi. Dapat kita simpulkan pemrograman web adalah pemrograman yang menampilkan sebuah informasi melalui internet "tanpa internet program web tersebut tidak akan bisa berjalan".

Contoh Bahasa Program Bahasa Script dan Design dalam Pemrograman Web :

1. HTML adalah singkatan dari HyperText Markup Language sebuah bahasa markup yang biasa digunakan membuat sebuah halaman web dan dapat menampilkan informasi melalui web browser. Html sendiri memiliki struktur dasar (berisi kode html) yang sudah ditetapkan. Contoh : Seperti html, tittle, head, body.

2. CSS adalah singkatan dari Cascading Style Sheets adalah sebuah bahasa yang memiliki kode - kode perintah untuk membangun suatu style (gaya) dari dokumen yang ditulis dari bahasa markup. kode-kode css tersebut hanya berfungsi sebagai pengatur, penata dll.
3. PHP adalah singkatan dari Hypertext Preprocessor adalah kumpulan kode/script yang digunakan sebagai pengolah data dari program yang akan dibuat atau dibangun.
4. JAVASCRIPT adalah bahasa scripting yang berjalan di sisi client, bahasa pemrograman web ini bersifat Client Side Programming Language dimana pemrosesannya dilakukan oleh client. salah satu contoh hasil yang dibuat atau ditulis dari bahasa ini adalah sebuah animasi yang sering kita lihat pada website/blog dll.

Dari ke-4 bahasa pemrograman yang dapat membangun dan melakukan pemrograman web diatas masih banyak lagi bahasa-bahasa pemrograman web lain. Namun cukup diketahui ke empat bahasa pemrograman dan script diatas yang banyak dan sering digunakan saat ini.

2.2.7 PHP

PHP adalah bahasa pemrograman web yang digunakan untuk membuat halaman web dinamis. Walaupun pada perkembangannya, PHP saat ini juga dapat digunakan untuk membuat aplikasi selain web, seperti aplikasi desktop. PHP pertama kali di kembangkan pada tahun 1995 oleh Rasmus Lerdorf, namun sekarang di ambil oleh oleh The PHP Group. Pada awalnya PHP adalah singkatan dari Personal Home Page, namun dalam perkembangannya, di ubah menjadi PHP: *Hypertext Preprocessor*, sebuah kepanjangan rekursif. PHP dirilis dalam lisensi PHP License, yang sedikit berbeda dengan lisensi GNU *General Public License* (GPL) yang biasa digunakan untuk proyek Open Source. Namun penggunaan PHP tetap tidak dikenakan biaya (gratis). Kemudahan dan kepopuleran PHP sudah menjadi standar bagi programmer web di seluruh dunia. Dan menurut wikipedia, PHP telah terinstall pada lebih dari 244 juta website dan 2,1 web server hingga saat ini. Fungsi PHP Dalam Pemrograman Web adalah Dalam membuat halaman

web, PHP sebenarnya bukanlah bahasa yang wajib digunakan. Sebuah website sederhana dapat dibuat tanpa menggunakan PHP sama sekali. Anda bisa membuat sebuah website, murni dengan menghubungkan beberapa halaman HTML saja. Namun jika anda ingin membuat web yang dinamis, bisa menyimpan ke dalam database, membuat halaman yang berubah-ubah sesuai input dari user, maka pada saat itulah PHP dibutuhkan. Untuk pembuatan web, kode PHP biasanya di sisipkan kedalam dokumen HTML. Karena hal inilah PHP disebut juga sebagai *scripting language* atau bahasa pemrograman script (Adhi 2014).

2.2.8 Penegrtian Data Flow Diagram(DFD)

Data Flow Diagram(DFD) merupakan suatu cara atau metode untuk membuat rancangan sebuah sistem yang mana berorientasi pada alur data yang bergerak pada sebuah sistem nantinya. Dalam pembuatan Sistem Informasi, DFD sering digunakan. DFD dibuat oleh para analis untuk membuat. Berikut ini adalah simbol-simbol yang di gunakan untuk menggambarkan *Data Flow Diagram(DFD)*.

2.2.9 Basis Data

Sistem manajemen basis data atau kadang disingkat SMDB, adalah suatu sistem atau perangkat lunak yang dirancang untuk mengelola suatu basis data dan menjalankan operasi terhadap data yang diminta banyak pengguna. Contoh tipikal SMDB adalah akuntansi, sumber daya manusia, dan sistem pendukung pelanggan, SMDB telah berkembang menjadi bagian standar di bagian pendukung (*back office*) suatu perusahaan. Contoh SMDB adalah Oracle, SQL server 2000/2003, MS Access, MySQL dan sebagainya. DBMS merupakan perangkat lunak yang dirancang untuk dapat melakukan utilisasi dan mengelola koleksi data dalam jumlah yang besar. DBMS juga dirancang untuk dapat melakukan manipulasi data secara lebih mudah. Sebelum adanya DBMS, data pada umumnya disimpan dalam bentuk *flat file*, yaitu *file teks* yang ada pada sistem operasi. Sampai sekarangpun masih ada aplikasi yang menyimpan data dalam bentuk flat secara langsung.