

Implementasi Metode Naïve Bayes Untuk Memprediksi Hasil Panen Padi

Dinda Meliana Hidayah
Teknik Informatika, Program Pascasarjana, Universitas Pamulang
e-mail: dmeliana123@gmail.com

Abstrak—Penelitian ini bertujuan untuk mendiskusikan prediksi hasil panen padi di kecamatan geberbitung kabupaten sukabumi dengan menggunakan algoritma naive bayes. Kekurangan prediksi hasil panen mengakibatkan kurangnya informasi untuk meningkatkan yang dibutuhkan peningkatan produksi saat panen. Kekurangan tersebut dapat diantisipasi dengan adanya sistem untuk memprediksi hasil panen. Salah satunya dengan sistem klasifikasi Naive bayes. Hasil prediksi dalam perhitungan pemodelan naive bayes diatas diketahui bahwa nilai Meningkatkan lebih besar dari pada nilai Menurun. Sehingga dapat diambil kesimpulan bahwa prediksi dalam penelitian ini Meningkatkan dan tidak termasuk dalam prediksi menurun. Prediksi salah satu proses memperkirakan secara sistematis berupa suatu yang paling mungkin terjadi dimasa depan berdasarkan informasi di masa sekarang ataupun dimasa lalu yang dimiliki, agar setiap kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat disesuaikan atau diperkecil. Prediksi tidak perlu memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan mencoba berusaha mendapatkan jawaban sesingkat atau sedekat mungkin yang akan terjadi.

Kata Kunci— Panen; Klasifikasi; Naive Bayes.

I. PENDAHULUAN

Panen di sebuah bidang pertanian sangat tidak asing sering disebut tahap akhir dari budidaya tanaman (cocok tanam). Panen merupakan kegiatan pemungutan atau bisa disebut juga pemetikan hasil bumi oleh setiap petani, bisa memakai alat apa saja yang terpenting dapat mengarah pada pencapaian hasil bumi (Darwis, 2019). Kegiatan akhir panen ini adalah adanya pemindahan tanaman dari tempat tumbuh ke lokasi yang lebih aman, kemudian diproses, dikonsumsi, atau ada juga yang disimpan. Beberapa tanaman sebagian besar mencapai periode kualitas maksimum, misalnya tanaman umbi-umbian atau pohon dapat dibiarkan di ladang atau dipanen sesuai kebutuhan. Namun periode kualitas maksimum matang ini apabila dibiarkan terbuka akan membusuk.

Kegiatan pengumpulan hasil panen ini dilakukan dari berbagai jenis lahan pertanian. Kegiatan pemanenan ini meliputi menuai, menumpuk, menangani merontokan, membersihkan, dan mengangkut. Sehingga panen ini dapat dilakukan secara individu atau dibantu dengan alat dan mesin. Penting untuk dapat memaksimalkan hasil gabah, dan meminimalkan kerusakan gabah dan penurunan kualitas. Berbagai tindakan atau perlakuan yang diberikan pada hasil pertanian setelah panen sampai komoditas berada pada tangan konsumen dalam bidang pertanian istilah ini diartikan pasca panen. Secara keilmuan istilah tersebut lebih tepatnya disebut pasca produksi (post production) yang bisa dibagi dalam dua bagian atau tahapan. Yaitu pasca panen (postharvest) dan pengolahan (processing). Penanganan pasca panen juga sering disebut pengolahan primer (primary processing) istilah yang digunakan merupakan perlakuan dari mulai panen sampai komoditas dapat dikonsumsi “segar” atau untuk persiapan pengolahan berikutnya (Mahmud & Purnomo, 2014).

Sebagai upaya mengatasi hasil panen di setiap lahan petani per tahunnya mengalami naik turun yang tidak stabil, agar tidak terjadinya kenaikan harga yang ekstrim memerlukan adanya prediksi yang tepat dan akurat untuk hasil panennya. Kekurangan prediksi hasil panen mengakibatkan kurangnya informasi untuk meningkatkan yang dibutuhkan peningkatan produksi saat panen. Kekurangan tersebut dapat diantisipasi dengan adanya sistem untuk memprediksi hasil panen. Salah satunya dengan sistem klasifikasi. Sistem klasifikasi dalam memprediksi hasil panen adalah sistem yang dapat memprediksi kapan panen dapat dilakukan. Implementasi yang diterapkan untuk memprediksi data yaitu Data mining salah satu bentuk implementasi untuk mencari sebuah model dan pola yang mampu memprediksi pada suatu data berdasarkan data sebelumnya di periode waktu tertentu.

Data mining merupakan bentuk panggilan data yang digunakan untuk menggali sebuah pengetahuan dari jumlah data yang benar dan akurat. Penggunaan data mining ini biasanya dengan menggunakan metode atau algoritma tertentu. Dalam klasifikasi algoritma naive bayes ada juga beberapa algoritma yang digunakan untuk klasifikasi hasil panen salah satunya algoritma backpropagation, algoritma backpropagation juga salah satu algoritma yang populer dan sangat efektif. Backpropagation bersifat adaptive yaitu dapat menyesuaikan pada dataset dan fault tolerance (kesalahan error kecil) dalam menyelesaikan masalah pada sistem (Tanaman et al., 2018). Namun algoritma naive bayes merupakan salah satunya juga algoritma yang memprediksi atau klasifikasi data secara akurat dan banyak digunakan. Teknik Data mining salah satu algoritma yang digunakan untuk klasifikasi adalah naive bayes. Dengan ketidak tergantungan yang kuat teorema bayes merupakan teknik prediksi berdasarkan kemungkinan yang sederhana pada penerapan aturan bayes dengan ketidak tergantungan. Proses klasifikasi naive bayes banyak digunakan dan lebih diminati karena kecepatan dan kesederhanaannya (Basit, 2020).

Prediksi merupakan salah satu proses memperkirakan secara sistematis berupa suatu yang paling mungkin terjadi dimasa depan berdasarkan informasi di masa sekarang maupun dimasa lalu yang dimiliki, agar setiap kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat disesuaikan atau diperkecil. Prediksi tidak perlu memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan mencoba berusaha mendapatkan jawaban sesingkat atau sedekat mungkin yang akan terjadi. Berdasarkan latar belakang di atas maka penelitian ini akan membangun sebuah sistem atau aplikasi klasifikasi dalam memprediksi hasil panen agar data lebih akurat kedepannya dengan judul “Prediksi Hasil Panen Padi Di Kecamatan Gegerbitung Kabupaten Sukabumi Dengan Menggunakan Algoritma Naive Bayes”.

II. METODE PENELITIAN

A. Pengertian Prediksi

Prediksi adalah suatu proses yang memperkirakan tentang suatu hal yang akan terjadi pada waktu yang akan datang berdasarkan informasi dan data yang ada. Untuk mengurangi ketidakpastian dan membuat perkiraan lebih baik dari apa yang akan terjadi dimasa yang akan datang, dan ini adalah tujuan dari melakukan prediksi data (Petrus Katemba & Koro, 2015). Panen merupakan kegiatan untuk menentukan kualitas dan kuantitas produksi. Dalam kegiatan ini penanganan panen dan pasca panen dapat mengakibatkan kerugian yang sangat besar. Ketahanan pangan yang diperlukan, oleh sebab itu penganan panen dan pasca panen ini secara benar perlu mendapatkan prioritas dalam proses memprediksi proses usahatani. (Darwis, 2019).

Adapun prediksi panen merupakan ramalan atau perkiraan untuk mengetahui tempat dan waktu panen yang tepat, misalnya memprediksi panen padi lebih baik dilakukan pada musim tanam kemarau terjadi pada bulan agustus sampai oktober dibanding musim hujan, karena waktu akan lebih efektif waktu panen yang dilakukan berdasarkan tujuannya untuk kemasakan yang diinginkan. Berdasarkan deskripsi umur panen padi yang tepat 30 sampai 35 hari setelah berbunga merata atau antara 135 sampai 145 hari setelah tanam. Kadar air pada umur panen optimum dicapai setelah kadar air gabah mencapai 22 -23% pada musim kemarau, dan antara 24 -26% pada musim hujan. (Tanaman et al., 2018)

B. Algoritma Naive Bayes

Naive bayes adalah sebuah algoritma klasifikasi dan prediksi Data mining yang mempertimbangkan setiap atribut tidak saling berhubungan karena itu algoritma genetika digunakan untuk membantu naive bayes dalam menentukan atribut-atribut yang harus digunakan. Dengan cara ini sehingga dapat meningkatkan akurasi. (Oramas et al., 2016) Algoritma naive bayes adalah sebuah bentuk klasifikasi data menggunakan probabilitas dan metode statik. Metode ini pertama kali diperkenalkan oleh seorang ilmuwan inggris thomas bayes, yang digunakan untuk memprediksi peluang yang akan muncul dimasa yang akan datang berdasarkan pengalaman masa lalu. (Pratiwi et al., 2021) Diasumsikan bahwa algoritma naive bayes dalam mengklasifikasi bahwa ada atau tidak ciri tertentu dari sebuah kelas yang tidak ada hubungannya dengan ciri dari kelas lainnya. Rumus persamaan dari naive bayes adalah:

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

Keterangan:

- X : data dengan class yang belum diketahui
- H : hipotesis data merupakan sebuah class spesifik
- $P(H|X)$: probabilitas hipotesis H berdasarkan kondisi X (*Posteriori Probability*)
- $P(H)$: probabilitas hipotesis H (*Prior Probability*)
- $P(X|H)$: probabilitas X berdasarkan kondisi hipotesis H
- $P(H)$: probabilitas dari X

Alur dalam algoritma naive bayes adalah sebagai berikut :

1. Membaca data training
2. Menghitung jumlah data dan probabilitas. Apabila data numerik maka harus cari nilai mean atau standar deviasi dari masing-masing parameter dan mencari nilai probabilitas yang sesuai dengan kategori yang sama dengan cara menghitung jumlah data agar dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan jumlah nilai dalam tabel mean, standart deviasi dan probabilitas (Aripin, 2019)

III. HASIL DAN PEMBAHASAN

A. Analisis Data

Analisis data dalam penelitian ini merupakan penelitian yang dilakukan Data mining menggunakan algoritma naive bayes untuk memprediksi atau klasifikasi hasil panen dari seluruh data set dari data masa lalu. Berikut data hasil panen yang meningkat dan menurun diambil dari masa 8 tahun kebelakang.

Tabel 1.
 Data Penelitian

Tahun	Musim	Lahan	Cuaca	Air	Daun	Batang	Isi Buah Padi	Harga Padi	Hasil Panen
1	pertama	Sedang	rendah	kering	aga coklat	sebagian roboh	buah dipinggir	meningkat	panen
2	Kedua	Besar	rendah	sedang	Kuning	berdiri	buah ditengah	meningkat	panen
3	ketiga	Sedang	tinggi	banjir	Kuning	berdiri	buah ditengah	menurun	gagal
4	pertama	Kecil	tinggi	sedang	Coklat	roboh	buah di pinggir	meningkat	gagal
5	pertama	Sedang	rendah	sedang	aga coklat	sebagian roboh	buah di pinggir	menurun	panen
6	kedua	Sedang	tinggi	banjir	Kuning	sebagian roboh	buah di pinggir	meningkat	panen
7	ketiga	Kecil	tinggi	sedang	Kuning	berdiri	buah di pinggir	meningkat	panen
8	ketiga	Besar	rendah	kering	Coklat	roboh	buah di pinggir	menurun	gagal

Transformasi data ini pengubahan data untuk mengelompokkan data yang beragam menjadi beberapa kategori. Berikut transformasi data dalam penelitian ini adalah sebagai berikut :

Tabel 2.
 Variabel Yang Telah Mendapatkan Data Training

Atribut/Variable	Total	Jumlah		Harga Padi		
		8	Panen	Gagal	Meningkat	Menurun
Lahan	Besar	2	1	1	1	1
	Sedang	4	3	1	2	2
	Kecil	2	1	1	2	0
Cuaca	Rendah	4	3	1	2	2
	Tinggi	4	2	2	3	1
	Kering	2	1	1	1	1
Air	Sedang	4	3	1	3	1
	Banjir	2	1	1	1	1
	Kuning	4	3	1	3	1
Daun	aga coklat	2	2	0	1	1
	Coklat	2	0	2	1	1
	Berdiri	3	2	1	2	1
Batang	sebagian roboh	3	2	1	2	1
	roboh	2	0	2	1	1
	buah dipinggir	6	4	2	4	2
Isi Buah Padi	buah ditengah	2	1	1	1	1
	Besar	2	1	1	1	1

B. Perhitungan Prediksi

Menghitung jumlah atribut / variabel yang telah mendapatkan data training dan data testing, maka selanjutnya perhitungan probabilitas. Berikut perhitungan probabilitas :

- 1) Menghitung jumlah probabilitas hasil panen gagal panen dan harga meningkat dan menurunnya padi
 - $P(Y = \text{Panen}) = 5/8 : 62.5\%$ “jadi hasil jumlah data hasil panen pada data training/dataset dibagi dengan jumlah keseluruhan data”.
 - $P(Y = \text{Gagal}) = 3/8 : 37.5\%$ “hasil jumlah data gagal panen pada data training dibagi dengan jumlah data keseluruhan data”
 - $P(Y = \text{Meningkat}) = 4/8 : 50\%$ “jumlah data harga padi yang meningkat pada data training dibagi dengan jumlah data keseluruhan”
 - $P(Y = \text{Menurun}) = 4/8 : 50\%$ “jumlah data harga padi yang menurun pada data training yang dibagi dengan jumlah data keseluruhan”
- 2) Menghitung probabilitas nilai atribut
 - Total data training = 8
 - Data panen = 5
 - Data gagal= 3
 - Data meningkat= 4
 - Data menurun = 4

Tabel 3.
 Transformasi Data

Atribut/Variable	Jumlah	Hasil Panen		Harga Padi			Persentase			
		8	Panen	Gagal	Meningkat	Menurun	Panen	Gagal	Meningkat	Menurun
Lahan	Besar	1	1	1	1	1	0.2	0.3	0.2	0.2
	Sedang	3	1	2	2	3	0.6	0.3	0.5	0.5
	Kecil	1	1	2	0	1	0.2	0.3	0.5	0
Cuaca	Rendah	3	1	2	2	3	0.6	0.3	0.5	0.5
	Tinggi	2	2	3	1	2	0.4	0.6	0.7	0.2
	Kering	1	1	1	1	1	0.2	0.3	0.2	0.2

Atribut/Variable	Jumlah	Hasil Panen		Harga Padi		Persentase				
		Panen	Gagal	Meningkat	Menurun	Panen	Gagal	Meningkat	Menurun	
Air	Sedang	3	1	3	1	3	0.6	0.3	0.7	0.2
	Banjir	1	1	1	1	1	0.2	0.3	0.2	0.2
	Kuning	3	1	3	1	3	0.6	0.3	0.7	0.2
Daun	aga coklat	2	0	1	1	2	0.4	0	0.2	0.2
	Coklat	0	2	1	1	0	0	0.6	0.2	0.2
	Berdiri	2	1	2	1	2	0.4	0.2	0.5	0.2
Batang	sebagian roboh	2	1	2	1	2	0.4	0.2	0.5	0.2
	roboh	0	2	1	1	0	0	0.5	0.2	0.2
	buah dipinggir	4	2	4	2	4	0.8	0.6	1	0.5
Isi Buah Padi	buah ditengah	1	1	1	1	1	0.2	0.3	0.2	0.2
	Besar	1	1	1	1	1	0.2	0.3	0.2	0.2

Semua hasil variabel dihitung dari setiap klasifikasi setelah mengetahui nilai dari setiap atribut/variabel terhadap setiap kelas probabilitas dengan rumus persentase. semisal menghitung data testing yaitu:

Tabel 4.
 Hasil Penelitian

Tahun	Lahan	Cuaca	Air	Daun	Batang	Isi Buah Padi
9	Sedang	Tinggi	Sedang	Kuning	sebagian roboh	buah di pinggir

Hasilnya sebagai berikut :

- Panen : $0.8 \times 0.8 \times 0.8 \times 0.8 \times 0.8 \times 1.2 = 0.393216$
- Gagal : $1.33 \times 1.33 \times 1.33 \times 1.33 \times 1.33 \times 2 = 8.4279$
- Meningkatkan: $1 \times 1 \times 1 \times 1 \times 1.5 = 1.5$
- Menurun : $1 \times 1 \times 1 \times 1 \times 1.5 = 1.5$

Dari hasil perhitungan tersebut maka menunjukkan persentase hasil panen 0.393216 dan gagal panen 8.4279, maka hasil yang dicari adalah gagal dengan harga meningkat.

C. Hasil Prediksi

Hasil prediksi dalam perhitungan pemodelan naive bayes diatas diketahui bahwa nilai $P(Y|Panen)$ lebih kecil dari pada nilai $P(Y|Gagal)$. Dan harga padi $P(Y|Meningkat)$ lebih besar dari $P(Y|Menurun)$ Sehingga dapat diambil kesimpulan bahwa prediksi dalam penelitian ini gagal panen dengan harga padi yang Meningkatkan dan tidak termasuk dalam prediksi Hasil panen dan Harga padi Menurun.

IV. KESIMPULAN

Berdasarkan pemaparan di atas, ada beberapa kesimpulan yang dapat ditarik dalam pembahasan diatas adalah sebagai berikut. Pertama, dengan adanya aplikasi prediksi maka dapat mempermudah petani dalam memprediksi hasil panen. Kedua, adanya website ini menjadi lebih bisa mengontrol panen dan gagal dalam bertani dengan prediksi hasil dari bagaimana musimnya.

DAFTAR PUSTAKA

- [1] Agustian, S., & Wibowo, H. (2019). Perbandingan Metode Moving Average untuk Prediksi Hasil Produksi Kelapa Sawit. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri (SNTIKI), November*, 156–162.
- [2] Aripin, J. J. (2019). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi pada BPR Pantura. <https://repository.nusamandiri.ac.id/index.php/repo/viewitem/13890>
- [3] Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknik*, 21(1), 60–64. <https://doi.org/10.18196/st.211211>
- [4] Basit, A. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Hasil Panen Padi. *Jurnal Teknik Informatika Kaputama (JTIK) 2020*, 4(2), 208–213.
- [5] Darwis, V. (2019). Potensi Kehilangan Hasil Panen Dan Pasca Panen Jagung Di Kabupaten Lampung Selatan. *Journal of Food System & Agribusiness*, 2(1), 55–66. <https://doi.org/10.25181/jofsa.v2i1.1110>
- [6] Hasan, M. (2017). Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Menggunakan Algoritma Naive Bayes Berbasis Forward Selection. *ILKOM Jurnal Ilmiah*, 9(3), 317–324. <https://doi.org/10.33096/ilkom.v9i3.163.317-324>
- [7] Mahmud, Y., & Purnomo, S. S. (2014). Keragaman Agronomis beberapa Varietas Unggul Baru Tanaman Padi (Oryza sativa L.) pada Model Pengelolaan Tanaman Terpadu. *Jurnal Ilmiah Solusi*, 1(1), 1–10.
- [8] Maudi, M. F., Nugraha, A. L., & Sasmito, B. (2014). Desain Aplikasi Sistem Informasi Pelanggan PDAM Berbasis Webgis (Studi Kasus : Kota Demak). *Jurnal Geodesi Undip*, 3(3), 98–110.
- [9] Mustika, W. P., Mardian, M., & Rinawati, R. (2018). Analytical Hierarchy Process Untuk Menganalisa Faktor Pemilihan Web Browser Pada Desktop. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 2(1), 83. <https://doi.org/10.30645/j-sakti.v2i1.57>
- [10] Novendri. (2019). Pengertian Web. *Lentera Dumai*, 10(2), 46–57.
- [11] Nurudin, A. F. (2015). Aplikasi Prediksi Hasil Panen Padi Dengan Metode Least Square (Studi Kasus : RT.001 RW.006 Ds.Warujayeng Kab.Nganjuk).

Artikel Skripsi Universitas Nusantara PGRI Kediri, 6.

- [12] Petrus Katemba, & Koro, R. (2015). Menggunakan Regresi Linear. *Jurnal Ilmiah Flash*, 3, 42–51.
- [13] Pratiwi, T. A., Irsyad, M., Kurniawan, R., Agustian, S., & Negara, B. S. (2021). Klasifikasi Kebakaran Hutan Dan Lahan Menggunakan Algoritma Naïve Bayes Di Kabupaten Pelalawan. *CESS (Journal of Computer Engineering, System and Science)*, 6(1), 139. <https://doi.org/10.24114/cess.v6i1.22555>
- [14] Riyanti, R., Ulinnuha Latifa, & Yuliarman Saragih. (2020). Pengembangan Learning Management System (LMS) Untuk Bahasa Pemrograman PHP. *Jurnal Ilmiah Core IT: Community Research Information Technology*, 5(1), 20–29.
- [15] Speed, J., & Engineering, S. P. (2013). *speed.web.id*. 5(2), 2–6.
- [16] Syafarina, G. A., & Rusdina, R. (2020). Prototype Aplikasi Pengelolaan Kegiatan Prakerin (Praktek Kerja Industri) Dan Pkl (Praktek Kerja Lapangan) Berbasis Web. *Technologia: Jurnal Ilmiah*, 11(3), 126. <https://doi.org/10.31602/tji.v11i3.3281>
- [17] Tanaman, P., Sawah, P., Kabupaten, M., Di, K., & Utara, S. (2018). *Padi Sawah Menurut Kabupaten / Kota*. July.
- [18] Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada DataSet Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437. <https://doi.org/10.30865/mib.v4i2.2080>
- [19] Yusuf, E., Riza, T. A., Ariefianto, T., & Elektro, F. (2013). *Implementasi Teknologi Load Balancer Dengan*. 11–16.