



KLASIFIKASI KUALITAS TEMBAKAU LAYAK PANEN DENGAN NAIVE BAYES CLASSIFIER DI DUSUN KAMPONG SITUBONDO

Muhammad Yasin^{a*}, Zaohol Fatah^b

^a Sain & Teknologi / Teknologi Informasi, muhammadaliyasin474@gmail.com, Universitas Ibrahimy Situbondo,
Jawa Timur

^b Sain & Teknologi / Sistem Informasi, Zaeholfatah@gmail.com, Universitas Ibrahimy Situbondo, Jawa Timur

* Korespondensi

ABSTRACT

To classify the quality of tobacco that is suitable for harvest in Dusun Kampong using the Naive Bayes Classifier algorithm. Tobacco as a plantation commodity has high economic value, both in the domestic and international markets. This study uses a qualitative method with data collection through direct observation and interviews with local tobacco farmers and traders. The Naive Bayes algorithm is applied to predict harvest eligibility based on the attributes of height, leaf width, leaf length, and number of leaves. Testing is done manually and using the RapidMiner tool to produce accurate predictions. The results of the study show the effectiveness of the Naive Bayes algorithm in classifying tobacco quality.

Keywords: classification, tobacco, naive bayes, data mining.

Abstrak

Untuk mengklasifikasikan kualitas tembakau yang layak panen di Dusun Kampong menggunakan algoritma Naive Bayes Classifier. Tembakau sebagai komoditas perkebunan memiliki nilai ekonomi tinggi, baik di pasar domestik maupun internasional. Penelitian ini menggunakan metode kualitatif dengan pengumpulan data melalui observasi langsung dan wawancara dengan petani serta pedagang tembakau setempat. Algoritma Naive Bayes diterapkan untuk memprediksi kelayakan panen berdasarkan atribut tinggi, lebar daun, panjang daun, dan jumlah daun. Pengujian dilakukan secara manual dan menggunakan alat RapidMiner untuk menghasilkan prediksi yang akurat. Hasil penelitian menunjukkan efektivitas algoritma Naive Bayes dalam klasifikasi kualitas tembakau.

Kata Kunci: klasifikasi, tembakau, nave bayes, data mining.

1. PENDAHULUAN

Tanaman tembakau merupakan tanaman hasil perkebunan tetapi bukan tanaman pangan, bagian yang sering dimanfaatkan adalah daunnya, daun biasanya digunakan sebagai bahan utama pembuatan rokok. Industri di Indonesia menggunakan 80% tembakau lokal untuk racikan rokok keretek. Tipe tembakau lokal begitu beragam, karena daya adaptasi dan berkembang tipe tembakau terbatas hanya pada suatu wilayah membuat tipe tembakau dinamakan sesuai dengan wilayah tembakau di kembangkan.[1] Tembakau merupakan salah satu komoditi perkebunan yang memiliki daya jual tinggi baik di pasar domestik maupun internasional. Tembakau merupakan fancy product dari salah satu komoditi perkebunan yang memiliki daya jual tinggi baik di pasar domestik maupun internasional. Indonesia menjadi salah satu negara penghasil tembakau terbesar kelima setelah Amerika Serikat dengan jumlah produksi mencapai 196.300 ton, sementara peringkat pertama adalah Cina dengan produksi 2.995.400 ton dan diikuti Brazil dan India (FAO, 2017). Berdasarkan data statistik, negara tujuan ekspor utama tembakau dari Indonesia adalah Amerika Serikat yang mencapai 40 persen diikuti Srilanka 15 persen dan Belanda 12 persen.[2] Data mining didefinisikan sebagai satu set teknik yang digunakan secara otomatis untuk mengeksplorasi secara menyeluruh dan membawa ke permukaan relasi-relasi yang kompleks pada set data yang sangat besar. Set data yang dimaksud di sini adalah set data

yang berbentuk tabulasi, seperti yang banyak diimplementasikan dalam teknologi manajemen basis data relasional. Akan tetapi, teknik-teknik data mining dapat juga diaplikasikan pada representasi data yang lain, seperti domain data spasial, berbasis teks, dan multimedia (citra).[3] Dalam dunia pertanian, klasifikasi dapat membantu dalam identifikasi penyakit tanaman, prediksi hasil panen, dan pengelolaan sumber daya alam.[4]

Naive Bayes adalah salah satu metode klasifikasi yang populer dan sering digunakan dalam machine learning dan data mining.[5]

2. METODOLOGI PENELITIAN

Penelitian lapangan atau *field research* yakni penelitian yang dilaksanakan secara sistematis untuk mengambil data di lapangan. Dengan pendekatan menggunakan penelitian kualitatif memaparkan yaitu penelitian kualitatif adalah penelitian yang dilaksanakan dalam menemukan dan mendeskripsikan suatu kegiatan yang dilakukan.[6] Selain itu penelitian ini juga menggunakan metode Algoritma Naive Bayes. Yaitu pengklasifikasian, Klasifikasi adalah tindakan untuk memberikan kelompok pada setiap keadaan dan setiap kelompok berisi atribut, salah satunya adalah kelas attribute. [7] Penerapan metode algoritma Naive Bayes diharapkan mampu memprediksi pendataan kualifikasi kualitas tembakau layak panen dan tidak layak panen.

Pengumpulan data penulis lakukan dengan mencari data atau informasi yang relevan dengan masalah yang diteliti, yaitu :

a. Pengamatan (*Observation*)

Observasi atau pengamatan adalah suatu teknik pengumpulan data dengan cara melakukan pengamatan secara langsung ke lapangan terhadap objek yang diteliti. Khususnya data tembakau di dusun kampung.[8]

b. Wawancara (*Interview*)

Wawancara merupakan Teknik pengumpulan data yang dilakukan melalui tatap muka dan Tanya jawab langsung antara pengumpul data terhadap narasumber/sumber data. Pengumpulan data dengan melakukan tanya jawab dengan pihak-pihak pedagang atau petani tembakau di dusun kampung.[9]

Pada perhitungan Algoritma Naive Bayes klasifikasi tanaman tembakau layak panen ini dalam menentukan kemungkinan atau probabilitas terhadap suatu penentuan klasifikasi maka menggunakan rumus berikut :

$$P(X|H)$$

$$P(X|H)P(H)$$

$$P(X)$$

Rumus 1 Probabilitas Naive Bayes

Dimana :

X merupakan data sampel dengan kelas yang tidak diketahui

H ialah hipotesa yaitu X adalah data dengan kelas (label) C.

P(H) adalah peluang dari hipotesa H

P(X) adalah peluang data sampel yang diamati

P(X|H) merupakan peluang dari data sampel X, bila hipotesa benar (valid).[10]

3. HASIL DAN PEMBAHASAN

Tabel 1. Data Trening

Nama	Tinggi (cm)	Lebar Daun (cm)	Panjang Daun (cm)	Jumlah Daun	Status
H. Mindon	145	26	38	15	Tidak Layak Panen
H. Ibrohim	150	25	40	17	Tidak Layak Panen
H. Abrori	153	28	41	17	Tidak Layak Panen
Toto	155	27	42	16	Tidak Layak Panen
Busairi	158	29	46	19	Tidak Layak Panen
Kori	160	28	45	18	Tidak Layak Panen

H. Yayan	162	30	44	18	Tidak Layak Panen
H. Hamdani	165	29	48	19	Layak Panen
H. Nurul	168	31	49	20	Layak Panen
H. Norhaton	170	30	50	20	Layak Panen
wildan	172	33	53	21	Layak Panen
H. Umar	175	32	52	22	Layak Panen
H. Akbar	176	36	56	25	Layak Panen
H. Sakli	178	34	54	23	Layak Panen
H. Heri	180	35	55	24	Layak Panen

Tabel 2. Data Testing

Nama	Tinggi (cm)	Lebar Daun (cm)	Panjang Daun (cm)	Jumlah Daun	Status (Prediksi)
H. Subhan	145	24	39	16	?
Hosnia	175	33	53	23	?

3.1 Pengujian Manual

Pengujian perhitungan manual untuk membuktikan hasil klasifikasi dapat di lakukan perhitungan manual sebagai berikut :

- Menghitung kemungkinan tingkat layak panen :
 $P(\text{Status} = \text{Layak Panen}) = 8/15 = 0,533333333$
- Total data menghitung kemungkinan tingkat layak panen :
 $P(\text{Status} = \text{Tidak Layak Panen}) = 7/15 = 0,466666667$

3.1.1 Tinggi CM

$P(X_1 = 140 - 159 \mid \text{Klasifikasi} = \text{layak panen})$
 $P(X_1 = 0/8) = 0$
 $P(X_1 = 140 - 159 \mid \text{Klasifikasi} = \text{tidak layak panen})$
 $P(X_1 = 5/7) = 0,714285714$

$P(X_1 = 160 - 179 \mid \text{Klasifikasi} = \text{layak panen})$
 $P(X_1 = 7/8) = 0,875$
 $P(X_1 = 160 - 179 \mid \text{Klasifikasi} = \text{tidak layak panen})$
 $P(X_1 = 2/7) = 0,285714286$

$P(X_1 = 180 - 199 \mid \text{Klasifikasi} = \text{layak panen})$
 $P(X_1 = 1/8) = 0,125$
 $P(X_1 = 180 - 199 \mid \text{Klasifikasi} = \text{tidak layak panen})$
 $P(X_1 = 0/7) = 0$

3.1.2 Lebar Daun CM

$P(X_1 = 20 - 29 \mid \text{Klasifikasi} = \text{layak panen})$
 $P(X_1 = 1/8) = 0,125$
 $P(X_1 = 20 - 29 \mid \text{Klasifikasi} = \text{tidak layak panen})$
 $P(X_1 = 7/7) = 1$

$P(X_1 = 30 - 39 \mid \text{Klasifikasi} = \text{layak panen})$
 $P(X_1 = 7/8) = 1$
 $P(X_1 = 30 - 39 \mid \text{Klasifikasi} = \text{tidak layak panen})$
 $P(X_1 = 1/7) = 0,125$

3.1.3 Lebar Daun CM

$P(X1 = 30 - 39 \mid \text{Klasifikasi} = \text{layak panen})$

$P(X1 = 0/8) = 0$

$P(X1 = 30 - 39 \mid \text{Klasifikasi} = \text{tidak layak panen})$

$P(X1 = 1/7) = 0,142857143$

$P(X1 = 40 - 49 \mid \text{Klasifikasi} = \text{layak panen})$

$P(X1 = 2/8) = 0,25$

$P(X1 = 40 - 49 \mid \text{Klasifikasi} = \text{tidak layak panen})$

$P(X1 = 6/7) = 0,857142857$

$P(X1 = 50 - 59 \mid \text{Klasifikasi} = \text{layak panen})$

$P(X1 = 6/8) = 0,75$

$P(X1 = 50 - 59 \mid \text{Klasifikasi} = \text{tidak layak panen})$

$P(X1 = 0/7) = 0$

3.1.4 Jumlah Daun

$P(X1 = 10 - 19 \mid \text{Klasifikasi} = \text{layak panen})$

$P(X1 = 1/8) = 0,125$

$P(X1 = 10 - 19 \mid \text{Klasifikasi} = \text{tidak layak panen})$

$P(X1 = 7/7) = 1$

$P(X1 = 20 - 29 \mid \text{Klasifikasi} = \text{layak panen})$

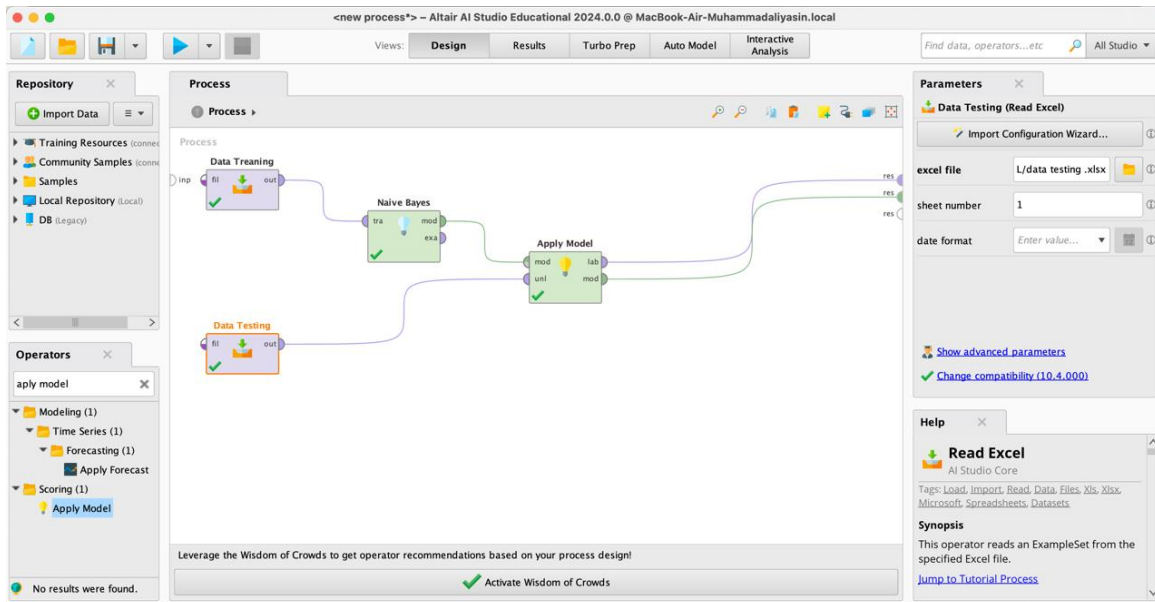
$P(X1 = 7/8) = 0,875$

$P(X1 = 20 - 29 \mid \text{Klasifikasi} = \text{tidak layak panen})$

$P(X1 = 0/7) = 0$

3.2 Pengujian Algoritma Naïve Bayes Menggunakan Tools Rapid Miner

Membuat desain proses yang akan menjalankan data training dan data testing kemudian menghubungkannya ke operator naive bayes dan *apply model*.



Gambar 1. Desain Proses RapidMiner

Hasil dari pemrosesan yang telah dilakukan sebelumnya dan telah mendapatkan hasil dari simple distribution.

Row No.	Status (Pre...	prediction(...	confidence(...	confidence(...	Nama	Tinggi (cm)	Lebar Daun...	Panjang Da...	Jumlah Da...
1	?	Tidak Layak...	1.000	0.000	H. Subhan	145	24	39	16
2	?	Layak Panen	0.000	1.000	Hosnia	175	33	53	23

Gambar 2. Simple Distribution RapidMiner

Dan terakhir adalah hasil prediksi dari tingkat kekuatan yang dimiliki oleh karakter *Shorekeeper* yang dihitung berdasarkan atribut pada data testing.

SimpleDistribution

Description
Distribution model for label attribute Status

Class Tidak Layak Panen (0.467)
5 distributions

Class Layak Panen (0.533)
5 distributions

Gambar 3. Hasil Prediksi Rapid Miner

4. KESIMPULAN DAN SARAN

Mengembangkan model klasifikasi kualitas tembakau layak panen menggunakan algoritma Naive Bayes di Dusun Kampong situbondo. Data yang dianalisis mencakup atribut tinggi tanaman, lebar dan panjang daun, serta jumlah daun. Penelitian menggunakan pendekatan field research dengan metode observasi dan wawancara. Hasil dari model menunjukkan kemampuan algoritma Naive Bayes untuk memprediksi status layak panen dengan data training dan testing melalui perhitungan manual dan tools RapidMiner. Akurasi klasifikasi didukung oleh pengolahan distribusi probabilitas dari data observasi, menunjukkan keberhasilan algoritma dalam mengelompokkan tanaman tembakau berdasarkan atribut yang signifikan. Hasil dari proses perhitungan manual Layak Panen 0,533333333 dan Tidak Layak Panen 0,466666667.

Ucapan Terima Kasih

Di balik pembuatan jurnal yang berjudul “Klasifikasi Kualitas Tembakau Layak Panen Dengan Naive Bayes Classifier di Dusun Kampong Situbondo” ini tidak luput dari bimbingan dan bantuan dari berbagai pihak. Saya ucapkan banyak-banyak terima kasih kepada:

- a. dosen pembimbing yang telah meluangkan waktunya untuk memberikan arahan, bimbingan dan serta wawasan yang menjadikan salah satu kunci keberhasilan penelitian ini.
- b. Kepada kedua orang tua yang selalu memotivasi saya untuk mencapai tujuan masa depan
- c. Yang terakhir saya ucapkan terimakasih kepada teman – teman yang selalu menemani saat penggarapan jurnal ini, Jaza kumullahu hairan

DAFTAR PUSTAKA

- [1] F. A. Fauzi, M. T. Furqon, and N. Yudistira, “Klasifikasi Jenis Tanaman Tembakau di Indonesia menggunakan Naïve Bayes dengan Seleksi Fitur Information Gain,” 2021. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [2] J. A. A. Y. I. Adhitya Wardhono, *Panduan Praktik Kelembagaan Usaha Budidaya Tembakau Besuki Na-Oogst*. 2018.
- [3] S. Kom. , M. Kom. D. A. P. S. Kom. , M. Kom. Amril Mutoi Siregar, *DATA MINING*.
- [4] S. Kom. , M. K. V. T. S. Kom. , M. K. Sri W. S. Kom. , M. K. S. E. Putri. Rahayu Mayang Sari, *Klasifikasi Forecasting Menggunakan Algoritma Naive Bayes*. 2024.
- [5] Rahayu Mayang Sari, *Klasifikasi Data Mining*. Serasi Media Teknologi, 2024.
- [6] R. Anisya Dwi Septiani and D. Wardana, “Implementasi Program Literasi Membaca 15 Menit Sebelum Belajar Sebagai Upaya Dalam Meningkatkan Minat Membaca.”
- [7] R. M. A. M. Y. H. S. Dinda Ayu Pratiwi, *SELEKSI CALON KELULUSAN TEPAT WAKTU MAHASISWA TEKNIK INFORMATIKA MENGGUNAKAN METODE NAÏVE BAYES*. kreatif, 2020.
- [8] “KUALITAS PELAYANAN KESEHATAN DI PUSAT KESEHATAN MASYARAKAT KEMBANG SERI KECAMATAN TALANG EMPAT KABUPATEN BENGKULU TENGAH,” *Jurnal Professional FIS UNIVED*, vol. 6, pp. 1–9, 2019.
- [9] E. Trivaika, M. Andri Senubekti, and A. Manajemen Informatika Dan Komputer HASS, “Perancangan Aplikasi Pengelola Keuangan Pribadi Berbasis Android,” vol. 16, no. 1, 2022, [Online]. Available: <https://journal.uniku.ac.id/index.php/ilkom>
- [10] W. I. R. M. H. K. S. Fadila, *Penerapan Metode Naive Bayes dan Skala Likert Pada Aplikasi Prediksi Kelulusan Mahasiswa*. kreatif, 2020.