
Klasifikasi Kualitas Tembakau Layak Panen dengan Naïve Bayes Classifier

Dwi Sri Anggun

Program Studi Teknik Informatika, Universitas Muhammadiyah Jember, d.srianggun@gmail.com

Keywords:

Quality,
Production,
Classification,
Naïve Bayes,
Tobacco,

ABSTRACT

Tobacco quality has an important role in determining the quality of cigarette production. One of the stages in postharvest tobacco that needs attention is sorting, where this process is carried out by graders. Sometimes a grader is less accurate in determining the quality of tobacco, the human error factor is usually the main cause. In this research, a classification model using the Naïve Bayes method will be made to help determine the quality of tobacco. From the results of research that has been carried out the level of accuracy between the information requested by the user and the answer given by the system (Precision) is 86%, the success rate of the system in retrieving an information (recall) is 68% and the level of closeness between the predicted value and the actual value. reach (accuracy) 67%.

Kata Kunci

Kualitas,
Produksi,
Klasifikasi,
Naïve Bayes,
Tembakau,

ABSTRAK

Kualitas tembakau memiliki peranan penting dalam menentukan kualitas produksi rokok. Salah satu tahapan dalam pascapanen tembakau yang perlu mendapat perhatian adalah sortasi, dimana proses ini dilakukan oleh grader. Adakalanya seorang grader kurang akurat dalam menentukan kualitas tembakau, factor human error biasanya yang menjadi penyebab utamanya. Dalam penelitian ini akan dibuat model klasifikasi dengan metode *Naïve Bayes* untuk membantu menentukan kualitas tembakau. Dari hasil penelitian yang telah dilakukan tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem (*Precision*) adalah 86%, tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi (*recall*) adalah 68% dan tingkat kedekatan antara nilai prediksi dengan nilai aktual mencapai (*accuracy*) 67%.

Korespondensi Penulis:

Dwi Sri Anggun,
Universitas Muhammadiyah Jember
Jl. Karimata No 49 Jember
Email: dwisria@gmail.com

1. PENDAHULUAN

Tembakau merupakan tumbuhan dari genus *Nicotiana* yang berasal dari Amerika Utara dan Amerika Selatan. Dari studi arkeolog yang telah dilakukan menunjukkan bahawa penggunaan tembakau dimulai sekitar abad pertama sebelum masehi, yaitu ketika orang Maya di Amerika menggunakan daun tembakau tersebut untuk merokok dalam upacara keagamaan mereka. Kemudian pemakaian tembakau mulai menyebar ke berbagai daerah sekitar amerika, dan pada akhirnya menyebar ke berbagai benua [1]. Proses penanaman tembakau biasanya di tanam pada bulan April. Proses penanaman tembakau sampai tanaman tersebut panen membutuhkan waktu sekitar enam bulan. Proses pemanenan tembakau dimulai dengan pemetikan daun tembakau oleh petani dan dilanjutkan dengan proses pengeringan. Proses pengeringan ini dilakukan dibawah terik sinar matahari dengan jangka waktu dua hari. Setelah dikeringkan daun tembakau akan dipilah berdasarkan kualitasnya untuk selanjutnya dijadikan bahan pembuatan rokok. Didalam pabrik daun tembakau ada yang langsung dipakai dan ada juga yang disimpan bertahun-tahun, hal ini dilakukan untuk produk-produk tertentu.

Kualitas tembakau akan berpengaruh langsung terhadap kualitas rokok. Untuk menjaga kualitas rokok diperlukan suatu tahapan sortasi. Proses sortasi ini merupakan sebuah tahapan untuk memilih kualitas daun tembakau yang didasarkan pada ukuran, berat, warna, dan sebagainya[2]. Proses sortasi dilakukan oleh seorang grader. Seorang

grader ini dapat menyortir daun tembakau berdasarkan kualitasnya. Proses penyortiran ini dilakukan dengan mengelompokkan daun tembakau berdasarkan tingkatan-tingkatan tertentu.[3]. Seorang grader harus mempunyai kemampuan yang baik dalam menilai kualitas daun tembakau. Penilaian kualitas daun tembakau ini didasarkan berdasarkan pada warna, aroma, dan pengalaman. Akan tetapi proses penilaian ini membutuhkan tenaga yang lebih, karena rentan terjadinya *human error*. *Human error* ini terjadi karena faktor kelelahan, penglihatan, pencahayaan, keadaan emosi, perbedaan persepsi fisik. Untuk mengurangi *human error* ini maka dibutuhkan sebuah sistem yang dapat membantu mengetahui kualitas tembakau layak panen atau sebaliknya.

Naïve Bayes merupakan sebuah algoritma data mining yang bertujuan untuk mengklasifikasi sebuah permasalahan. Dimana algoritma ini didasarkan pada perhitungan statistika yaitu probabilitas. Algoritma menggunakan teori *Bayes* dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [4]. Algoritma *Naïve Bayes* cukup baik untuk menangani permasalahan klasifikasi terkait dengan permasalahan bidang pertanian diantaranya adalah untuk diagnosa penyakit tanaman blimbing[5] dan penentuan kualitas bibit unggul padi [6]. Pada penelitian ini akan diterapkan solusi untuk mengatasi permasalahan yang ada pada pemilihan kualitas daun pada tembakau. Dimana akan dibangun sebuah sistem prediksi untuk menentukan kualitas tembakau layak panen atau tidak. Metode yang dipergunakan adalah *Naïve Bayes*, metode ini dipakai karena metode ini terbukti cukup baik digunakan untuk proses klasifikasi dan mempunyai tingkat akurasi yang tinggi [7].

2. METODE PENELITIAN

Proses penelitian ini didasarkan pada kerangka konsep yang dibuat oleh penulis. Adapun didalam kerangka penelitian akan dijelaskan prose penyelesaian penelitian mulai dari awal sampai terakhir penelitian. Kerangka konsep penelitian ini dapat dijelaskan pada gambar 1 dibawah ini



Gambar 1. Kerangka Kerja Penelitian

Adapun keterangan dari gambar 1 dapat dijelaskan sebagai berikut

1. Studi Literatur
Pada tahapan ini akan dilakukan pencarian teori-teori yang berhubungan dengan topik penelitian yang dibahas. Studi literatur didapatkan dari proses pencarian referensi dari buku, jurnal, internet dan sebagainya..
2. Pengumpulan Data
Proses pengumpulan data dilakukan dengan mencari data penelitian. Data yang digunakan adalah data tembakau yang layak panen atau sebaliknya
3. Analisis sistem
Pada tahapan ini akan dilakukan identifikasi kebutuhan dari system yang dibuat. Sistem yang dibangun menggunakan aplikasi berbasis web yang didalamnya terdapat proses klasifikasi dengan metode *Naïve bayes*.
4. Pengembangan Sistem
Tahapan pengembangan system merupakan tahapan setelah proses analisis system telah dilakukan. Proses ini berisi tahapan pembuatan *code*.
5. Pembuatan Laporan
Pada tahapan ini dilakukan pembuatan laporan yang berisi dokumentasi dari kegiatan penelitian, mulai tahapan awal penelitian sampai dengan akhir penelitian.

Naive Bayes merupakan algoritma *machine learning* yang diperuntukkan untuk proses klasifikasi. Algoritma ini didasarkan pada teori probabilitas *Bayes*. Algoritma *Naive Bayes* adalah teknik klasifikasi berdasarkan penerapan teori *Bayes*. Pada teori *Bayes* semua *predictor* diasumsikan independen satu sama lain. Karena sifatnya yang

independent maka tidak adanya keterkaitan antara fitur satu dengan yang lainnya dalam sebuah data. Dalam proses pengolahan data, data akan dibagi menjadi dua yaitu data latih dan data uji.

Tabel 1. Sampel Data Latih

No	Tinggi	Lebar Daun	Panjang Daun	Jumlah Daun	Status
1	Kos	42	63	4	Panen
2	Kak	75	83	4	Panen
3	Tng	71	74	4	Panen
4	Kos	28	39	2	Tidak Panen
5	Kak	68	79	2	Panen
6	Tng	69	80	4	Panen
7	Kos	49	70	4	Panen
8	Kak	59	67	4	Tidak Panen
9	Tng	64	73	4	Panen
10	Kos	50	62	4	Panen

Keterangan:

Kos : Setinggi Kaos Kaki

Kak : Setinggi Pinggang

Tng : Setinggi Bahu

Dari data training di atas akan diuji dengan data uji dengan penyelesaian sebagai berikut:

Tabel 2. Sampel Data Uji/Testing

No	Tinggi	Lebar Daun	Panjang Daun	Jumlah Daun	Status
1	Kos	50	63	3	?

Sebelum menentukan probabilitas setiap kriteria, terlebih dahulu dilakukan pemisahan antara Panen dan Tidak Panen.

➤ Kelayakan Panen:

Berdasarkan pada tabel 3.1, didapat nilai probabilitas untuk kriteria sebagai berikut:

- Tinggi : Kos (Setinggi Kaos Kaki)

Tinggi = Kos dengan Kelas Panen adalah 3/8.

$$P(\text{Tinggi} = \text{Kos} \mid \text{Kelas} = \text{Panen}) = 3 / 8 = 0,375$$

- Lebar Daun = 50

Untuk menghitung probabilitas kriteria Lebar Daun Panen akan digunakan rumus pada persamaan (2) dikarenakan data pada Lebar Daun merupakan data yang bernilai kontinyu sebagai berikut :

$$P(\text{Lebar Daun} = 50 \mid \text{Kelas} = \text{Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 14,961}} e^{-\frac{(50-57,5)^2}{2 \times (14,961)^2}}$$

$$= \frac{1}{9,693} e^{-\frac{56,25}{447,663}} = 0,103 \times 0,882 = 0,091$$

- Panjang Daun = 63

$$P(\text{Panjang Daun} = 63 \mid \text{Kelas} = \text{Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 12,684}} e^{-\frac{(63-69)^2}{2 \times (12,684)^2}}$$

$$= \frac{1}{8,925} e^{-\frac{36}{160,884}} = 0,112 \times 0,800 = 0,090$$

- Jumlah Daun = 3

$$P(\text{Jumlah Daun} = 3 \mid \text{Kelas} = \text{Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 0,843}} e^{-\frac{(3-3,6)^2}{2 \times (0,843)^2}}$$

$$= \frac{1}{2,301} e^{-\frac{0,36}{1,422}} = 0,435 \times 0,253 = 0,110$$

➤ Kelayakan Tidak Panen:

- Tinggi : Kos (Setinggi Kaos Kaki)

Tinggi = Kos dengan Kelas Tidak Panen adalah 1/2.

$$P(\text{Tinggi} = \text{Kos} \mid \text{Kelas} = \text{Tidak Panen}) = 1 / 2 = 0,500$$

- Lebar Daun = 50

$$P(\text{Lebar Daun} = 50 \mid \text{Kelas} = \text{Tidak Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 16,920}} e^{-\frac{(50-54,6)^2}{2 \times (16,920)^2}}$$

$$= \frac{1}{10,308} e^{-\frac{21,16}{572,6}} = 0,097 \times 0,037 = 0,004$$

- Panjang Daun = 63

$$P(\text{Panjang Daun} = 63 \mid \text{Kelas} = \text{Tidak Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 16,628}} e^{-\frac{(63-67)^2}{2 \times (16,628)^2}}$$

$$= \frac{1}{10,219} e^{-\frac{16}{553}} = 0,098 \times 0,029 = 0,003$$

- Jumlah Daun = 3

$$P(\text{Jumlah Daun} = 3 \mid \text{Kelas} = \text{Panen}) = \frac{1}{\sqrt{2 \times 3,14 \times 1,095}} e^{-\frac{(3-3,2)^2}{2 \times (1,095)^2}}$$

$$= \frac{1}{2,623} e^{-\frac{0,04}{2,4}} = 0,381 \times 0,017 = 0,006$$

Setelah diperoleh nilai probabilitas setiap kriteria, kemudian dilakukan perkalian nilai kriteria sesuai dengan kelas panen atau tidak sebagai berikut :

$$P(\text{Kelas} = \text{Panen}) = \frac{8}{10} \times 0,375 \times 0,091 \times 0,090 \times 0,110 = 0,0002702700$$

$$P(\text{Kelas} = \text{Tidak Panen}) = \frac{2}{10} \times 0,500 \times 0,004 \times 0,003 \times 0,006 = 0,0000000072$$

Untuk menghasilkan nilai Probabilitas maka dilakukan normalisasi terhadap kelas panen atau tidak panen tersebut sehingga jumlah nilai yang diperoleh sama dengan 1.

$$\text{Panen} = \frac{0,0002702700}{0,0002702700 + 0,0000000072} = 0,999973361$$

$$\text{Tidak Panen} = \frac{0,0000000072}{0,0002702700 + 0,0000000072} = 0,0000266393$$

Dari nilai probabilitas yang diperoleh maka dapat disimpulkan dari data tembakau yang diuji di atas diketahui probabilitas tembakau tersebut masuk ke kelas Panen lebih besar dari pada probabilitas Tidak Panen. Adapun nilai persentase yang dihasilkan dari nilai probabilitas kedua kelas tersebut dapat dilihat pada tabel 3.3 di bawah ini :

Tabel 3 Hasil Persentase Probabilitas Panen dan Tidak Panen

Kelas	Probabilitas	Presentasi
Panen	0,99997336	99,997
Tidak Panen	0,000026639	0,003

3. HASIL DAN ANALISIS

Pengujian merupakan tahapan yang paling penting dalam pembuatan sebuah system. Sistem akan diuji untuk menilai seberapa layak aplikasi dapat digunakan. Pada penelitian ini akan dilakukan 2 pengujian utama yaitu: Pengujian *black box* dan pengujian akurasi dari metode *Naïve Bayes* yang dipakai. Pengujian *Black Box* ditujukan untuk menguji fungsionalitas dari sistem, Sedangkan pengujian dari akurasi ditujukan untuk menguji tingkat akurasi dari metode *Naïve Bayes*. Adapun detail dari pengujian didalam penelitian ini adalah sebagai berikut .

a. Pengujian *Black Box*

Adapun detail dari skenario uji coba *black box* adalah sebagai berikut

Tabel 4 Pengujian *Black Box*

Kode Uji	Skenario Uji	Hasil yang diharapkan	Kesimpulan	
			Berhasil	Tidak Berhasil
Menu utama	Memilih tombol dashboard	Menampilkan halaman utama	√	
	Memilih tombol data training	Menampilkan data training	√	
	Memilih tombol pengujian testing	Menampilkan form data pengujian	√	
	Memilih tombol tentang pembuat	Menampilkan biodata aplikasi	√	
Menu data training	Pencarian data trainig	Menampilkan data training berdasarkan data filter	√	
	Memilih tombol tambah	Menampilkan form data training	√	
	Memilih tombol edit	Menampilkan form edit data training	√	
	Memilih tombol hapus	Menghapus data training	√	
Menu Pengujian data testing	Memilih tombol hitung	Menampilkan Hasil Panen atau tidak Panen	√	
Menu tentang pembuat	Informasi pengembang	Menampilkan Informasi Pengembang	√	

Pada table 4.8 diatas adalah hasil pengujian fitur sistem yang telah dibangun, dari hasil diatas maka sistem dapat disimpulkan berhasil dibangun.

1. Menu Utama
 - a. Tombol Dashboard pada menu/tombol ini dinyatakan berhasil jika pada halaman dashnoard tidak terjadi *error* atau kesalahan halaman dashboard.
 - b. Tombol Data Training pada menu ini dinyatakan berhasil jika data yang ditampilkan adalah data training dan tidak bercampur dengan data uji.
 - c. Tombol Pengujian Data Testing pada halaman ini dinyatakan berhasil jika inputan user “data uji” berhasil menentukan kelayakan kualitas panen tembakau atau tidak di PTPN X Kertosari Jember.
 - d. Tombol Tentang Pembuat pada tombol ini dinyatakan berhasil jika pada halaman ini hanya menampilkan profil penyusun tugas akhir.
2. Menu Data Training
 - a. Pencarian data training dinyatakan berhasil jika user melakukan filter/pencarian data sesuai keyword yang dicari.
 - b. Tombol tambah dinyatakan berhasil jika user melakukan tambah data berhasil di simpan maka dinyatakan berhasil.
 - c. Tombol edit dinyatakan berhasil jika data yang di edit akan berhasil tersimpan sesuai perubahan dan tidak menambah data baru.
 - d. Tombol hapus dinyatakan berhasil jika data yang dihapus berhasil dan tidak menghapus data yang lain.
3. Menu Pengujian Data Testing
 - a. Tombol hitung dinyatakan berhasil jika user melakukan penghitungan data yang diuji dengan metode *naïve bayes* menampilkan hasil penentuan kelayakan kualitas panen tembakau dan sesuai dengan perhitungan manual.
4. Menu Tentang Pembuat
 - a. Informasi pengembang pada tombol ini dinyatakan berhasil jika pada halaman ini hanya menampilkan profil penyusun tugas akhir.

b. Pengujian Akurasi

Pengujian *naïve bayes* dilakukan untuk mengetahui tingkat akurasi sistem klasifikasi kualitas tembakau layak panen dengan menggunakan metode *naïve bayes* dengan dataset dan *data testing* yang berbeda-beda, hasil pengujian sebagai berikut:

Diketahui :

- Dataset Panen : 42
- Dataset Tidak Panen : 37
- Data Testing Panen : 24
- Data Testing Tidak Panen : 12
- Hasil Klasifikasi Panen : 22
- Hasil Klasifikasi Tidak Panen : 14

Maka nilai tingkat *precision*, *recall* & *accuracy* sebagai berikut:

TP : 19
 TN : 5
 FP : 3
 FN : 9

$$Precision = \frac{TP}{TP + FP} = \frac{19}{19 + 3} = 0,86 \times 100\% = 86\%$$

$$Recall = \frac{TP}{TP + FN} = \frac{19}{19 + 9} = 0,68 \times 100\% = 68\%$$

$$Akurasi = \frac{Jumlah\ data\ benar}{Jumlah\ Data\ Keseluruhan} = \frac{24}{36} \times 100\% = 67\%$$

Dari hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem (*Precision*) adalah 86%, tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi (*recall*) adalah 68% dan tingkat kedekatan antara nilai prediksi dengan nilai aktual mencapai (akurasi) 67%. Karena nilai nilai *precision*, *recall* dan akurasi melebihi dari 50% maka sistem layak atau efektif (Pao, Lee, 1989).

4. KESIMPULAN

Dari hasil analisis yang telah dilakukan pada pembahasan sebelumnya, maka penelitian ini dapat ditarik kesimpulan yaitu dari segi fungsionalitas aplikasi sudah berjalan dengan baik tidak terjadi kesalahan. Sementara itu untuk uji metodenya didapatkan hasil *precision* sebesar 86%, tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi (*recall*) adalah 68% dan tingkat kedekatan antara nilai prediksi dengan nilai aktual mencapai (*accuration*) 67%. Nilai tersebut tergolong cukup rendah, oleh karena itu didalam penelitian selanjutnya dapat dilakukan *preprocessing* data terlebih dahulu terhadap data yang telah didapat. dan dapat dibandingkan metode lain untuk menghasilkan tingkat akurasi yang lebih akurat. Kaitannya dengan data dapat dikembangkan dengan menambahkan kriteria atau faktor lain dalam penentuan kualitas kelayakan panen tembakau.

UCAPAN TERIMA KASIH

Ucapan terima kasih dapat diperuntukkan kepada pihak-pihak yang telah membantu dalam penelitian, khususnya PTP X Kertosari dan Universitas Muhammadiyah Jember yang telah memberikan kesempatan kepada penulis untuk melakukan penelitian

REFERENSI

- [1] S. Mishra and M. B. Mishra, "Tobacco: Its historical, cultural, oral, and periodontal health association," *J. Int. Soc. Prev. Community Dent.*, vol. 3, no. 1, pp. 12–18, 2013, doi: 10.4103/2231-0762.115708.
- [2] S. Titosastro and W. Musholaeni, "Penanganan Panen Dan Pasca Panen Tembakau Di Kabupaten Bojonegoro," *J. Buana Sains*, vol. 15, no. 2, pp. 155–164, 2015.
- [3] A. Rintiasti and I. Krisnadi, "GRADING WARNA DAUN TEMBAKAU BAWAH NAUNGAN MENGGUNAKAN JARINGAN SARAF TIRUAN. (Color Grading of Shaded Tobacco Leaves Using Artificial Neural Network)," *J. Ind. Has. Perkeb.*, vol. 12, no. 1, p. 43, 2017, doi: 10.33104/jihp.v12i1.2847.
- [4] G. Shobha and S. Rangaswamy, *Machine Learning*, 1st ed., vol. 38. Elsevier B.V., 2018. doi: 10.1016/bs.host.2018.07.004.

-
- [5] Y. V. Via, H. Maulana, and S. Miftakhoneki, "Penerapan Metode Naive Bayes Sebagai Diagnosa Hama Penyakit Tanaman Belimbing," *J. Ilm. Teknol. Inf. dan Robot.*, vol. 2, no. 2, pp. 27–32, 2020, doi: 10.33005/jifti.v2i2.35.
- [6] M. Zulfikar and H. Fahmi, "Penerapan Sistem Pendukung Keputusan Dengan Metode Naive Bayes Dalam Menentukan Kualitas Bibit Padi Unggul Pada Balai Pertanian Pasar Miring," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 159, 2019, doi: 10.32672/jnkti.v2i2.1566.
- [7] A. N. Sihananto and H. Maulana, "Studi Literatur Tentang Performa Naive Bayes Dalam Klasifikasi Data," *Pros. Semin. Nas. Inform. Bela Negara*, vol. 2, pp. 132–135, 2021, doi: 10.33005/santika.v2i0.134.