
PEMILIHAN KUALITAS PRODUK KELAPA SAWIT MENGGUNAKAN METODE NAÏVE BAYES DI LABUHANBATU SELATAN

Dinda Nurhasanah^{1*}, Dinda Ayu Lestari², Syamsiah Simatupang³

^{1,2,3}Sistem Informasi, STMIK Royal Kisaran

email: ^{*}dindanur11@gmail.com

Abstract: Indonesia is one of the highest palm oil producers in the world. Palm oil is one of the well-known natural resources in the world today because it has high economic value, such as being able to produce palm oil, fuel and others. In Indonesia, one of the largest palm oil producers is the province of Riau, according to the Central Statistics Agency (BPS) noted that Riau will be the largest palm oil producing region in Indonesia in 2021 with as many as 8,629 thousand tons. The quality of CPO is one of the main factors for customers to make purchasing decisions, especially the standards contained in the Indonesian National Standard (SNI) 01-2901-2006. The final quality of this palm oil product is determined by the combined results of CPO quality and kernel quality. Good quality raw materials will affect the selling price of raw materials for a good quality final product. Selection of the quality of palm oil products still uses old methods such as laboratory testing to check water content, free fatty acid levels, and others in the manufacture of CPO, which takes several days to get the results. The purpose of this study is to classify the quality of palm oil products with a method. Having an application/web can help the laboratory produce good CPO. The method used in this study is the Naïve Bayes method to obtain data quickly and accurately. The accuracy of the Naïve Bayes method in determining the quality of palm oil products is 82.05%.

Keywords: classification, quality, palm oil, naïve bayes

Abstrak: Indonesia merupakan salah satu penghasil minyak kelapa sawit tertinggi di dunia. Kelapa sawit adalah salah satu sumber daya alam yang terkenal saat ini didunia karena memiliki nilai ekonomis yang tinggi seperti dapat menghasilkan minyak sawit, bahan bakar dan lainnya. Di indonesia salah satu penghasil kelapa sawit terbesar adalah provinsi Riau, menurut Badan Pusat Statistik (BPS) mencatat Riau menjadi daerah penghasil kelapa sawit terbesar Di Indonesia pada tahun 2021 sebanyak 8.629 ribu ton. Kualitas CPO adalah salah satu faktor utama bagi pelanggan untuk membuat keputusan pembelian khususnya standar yang terdapat pada Standar Nasional Indonesia (SNI) 01-2901-2006. Kualitas akhir dari produk kelapa sawit ini ditentukan dari gabungan hasil kualitas CPO dan kualitas kernel. Bahan mentah yang berkualitas baik akan mempengaruhi harga jual bahan mentah untuk kualitas akhir produk yang baik. Pemilihan kualitas produk kelapa sawitnya masih memakai cara yang lama seperti pengujian laboratorium untuk pengecekan kadar air, kadar asam lemak bebas, dan lainnya dalam pembuatan CPO, yang membutuhkan beberapa hari untuk mendapatkan hasilnya. Tujuan dari penelitian ini adalah mengklasifikasikan kualitas produk kelapa sawit dengan sebuah metode. Dengan adanya aplikasi/web dapat membantu pihak laboratorium dalam menghasilkan CPO yang baik. Metode yang digunakan dalam penelitian ini adalah metode *Naïve Bayes* untuk memperoleh data dengan cepat dan akurat. Tingkat akurasi metode *Naïve Bayes* dalam menentukan kualitas produk kelapa sawit adalah sebesar 82,05%.

Kata Kunci: klasifikasi, kualitas, kelapa sawit, naïve bayes

PENDAHULUAN

Pada zaman sekarang ini kemajuan teknologi sangat pesat, apalagi ditambah dengan berbagai alat atau pun sistem yang semakin canggih dan mempermudah pekerjaan manusia. Teknologi memiliki peranan penting untuk membuat dan menunjang pekerjaan manusia semakin cepat dan mudah. Sudah banyak manusia yang membuat atau pun menciptakan teknologi dengan tujuan yang mereka inginkan sendiri, agar tidak perlu repot-repot untuk melakukan suatu kegiatan yang melelahkan karena pekerjaan yang tidak memakai teknologi. Kemajuan teknologi ini sangat berdampak positif khususnya dalam bidang perkebunan, salah satunya perkebunan kelapa sawit.

Perkebunan kelapa sawit adalah salah satu sumber daya alam yang terkenal saat ini didunia karena memiliki nilai ekonomis yang tinggi seperti dapat menghasilkan minyak sawit, bahan bakar dan lainnya[1]. Diindonesia penghasil kelapa sawit terbesar adalah provinsi Riau, menurut Badan Pusat Statistik (BPS) mencatat Riau menjadi daerah penghasil kelapa sawit terbesar diindonesia pada tahun 2021 sebanyak 8.629 ribu ton. Sementara untuk provinsi Sumatra Utara yang dimana termasuk daerah Labuhanbatu Selatan menghasilkan sebanyak 5.310,90 ribu ton. Dari penghasilan yang banyak ini kelapa sawit akan diolah menjadi bahan mentah seperti CPO (*Crude Palm Oil*). Kualitas CPO adalah salah satu faktor utama bagi pelanggan untuk membuat keputusan pembelian khususnya standar yang terdapat pada Standar Nasional Indonesia (SNI) 01-2901-2006. Persyaratan mutu maksimum CPO yang dinyatakan adalah jumlah kotoran, asam lemak bebas, dan kadar air 0,5 %. Oleh

karena itu penerapan teknologi disini sangat penting dalam klasifikasi penentuan CPO disetiap perusahaan.

Selama ini dalam pemilihan kualitas produk kelapa sawit khususnya pada CPO masih dalam kategori yang repot seperti harus diuji laboratorium, dipertimbangkan secara manual seperti melihat warna yang dihasilkan apakah pucat, sedang, atau terang, melihat jenis bibit dari pohon kelapa sawit, dan kadar air dalam kelapa sawit tersebut. Sehingga para pelanggan yang ingin membuat keputusan untuk membeli berpikir dua kali atau masih dipertimbangkan lagi, karena proses yang lama dan masih kurang efektif dan efisien.

Sama halnya pada lokasi penelitian di perusahaan sawit di Labuhanbatu Selatan yang dalam pemilihan kualitas produk kelapa sawitnya masih memakai cara yang lama seperti pengujian laboratorium untuk pengecekan kadar air, kadar asam lemak bebas, dan lainnya dalam pembuatan CPO, yang membutuhkan beberapa hari untuk mendapatkan hasilnya. Dan dari hasil pekerjaan yang memakan waktu dapat berdampak pada pelanggan yang ingin membuat keputusan dalam membeli produk CPO dalam tempo yang cepat.

Dari permasalahan di atas peneliti tertarik dalam mencari solusi agar dapat mempermudah, mempercepat pekerjaan, serta pengklasifikasian data secara efisien dan akurat, dan dapat meningkatkan nilai beli dari pelanggan dengan adanya keterlibatan teknologi sistem informasi. Metode yang digunakan dalam penelitian ini adalah metode *Naive Bayes* untuk memperoleh data dengan cepat dan akurat. Metode *naive bayes* ini termasuk kedalam *data mining* yang dimana klasifikasi

didukung oleh konsep pengenalan pola, klasifikasi kualitas produk kelapa sawit sebagai kategori yang dihasilkan berdasarkan kadar kotoran pada CPO, kadar air CPO, kadar lemak bebas, indeks daya pemucatan, beta karoten, dan kadar pada kotoran kernel. Sehingga dapat menghasilkan sebuah informasi tentang bahan mentah CPO yang baik. Dengan adanya aplikasi/web dapat membantu pihak laboratorium dalam menghasilkan CPO.

Metode *Naive Bayes* merupakan sebuah metode klasifikasi yang berasal pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes [5].

Pada penerapannya dalam penelitian dibidang perkebunan, metode *Naive Bayes* juga telah banyak digunakan salah satunya oleh [4] dengan judul “Klasifikasi Kualitas Produk Kelapa Sawit Menggunakan Metode *Naive Bayes*” tujuan dari penelitian ini yaitu, dengan metode ini dapat mengklasifikasikan kualitas produk kelapa sawit.

METODE RISET

Metode riset memuat hal-hal penting terkait pelaksanaan penelitian dan tahapan-tahapan yang akan ditempuh dalam penelitian ini untuk mencapai suatu hasil atau output yang diinginkan [2].

1. Atribut dalam Riset

Dalam *machine learning*, penelitian ini dikategorikan sebagai penelitian supervised learning yaitu sebuah teknik model dimana data yang akan diproses memiliki label/target/class dengan tujuan untuk mengetahui klasifikasi data-data yang didapat, dengan menggunakan atribut akan diketahui klasifikasi yang penulis teliti [3]. Oleh sebab itu, atribut yang terdapat dalam penelitian ini terdiri dari Untuk menentukan klasifikasi kualitas produk mentah kelapa sawit ini berdasarkan 8 (delapan) atribut yaitu nilai Dirt CPO, Moisture CPO, FFA CPO, DOBI CPO, carotin CPO, dirt kernel, moisture kernel, dan broken kernel.

2. Tahapan Riset

Tahapan yang akan ditempuh dalam penelitian ini digambarkan tahapan pertama yang akan dilakukan adalah *Data Preprocessing* dimana dalam proses tersebut terdapat *data cleaning* untuk melakukan *handling missing* pada data dan diproses, kemudian tahapan kedua setelah *data processing*, data akan di bagi menjadi dua yaitu data *Training* dan data *testing*, untuk pembagian data dan pengujian data, setelah itu memasuki testimasi proses metode *Naive Bayes*, dimana akan melalui proses kualifikasi, dan probabilitas data dengan tujuan mendapatkan hasil yang baik, dengan metode hasil yang baik pula.

3. Data Riset

Data riset yang akan digunakan adalah dataset Kualitas CPO yang berasal dari Pabrik Labuhanbatu Selatan. Pada tahun 2021. Data tersebut berjumlah sebanyak

200 baris data. Terdapat 8 kolom (*Dirt CPO*, *Moisture CPO*, *FFA CPO*, *DOBI CPO*, *Carotin CPO*, *Dirt Kernel*, *Moisture Kernel* dan *Broken Kernel*) digunakan untuk pengujian dan 1 kolom

(Kualitas) sebagai kolom target/label. Sampel dataset yang akan digambarkan dalam penelitian ini menggunakan 5 data teratas seperti yang terdapat dalam Gambar 1 berikut ini :

No	DirtCPO	MoistCPO	FFACPO	DOBICPO	CarotinCPO	DirtKernel	MoistKernel	BrokenKernel	Kualitas
0	1	0.022	0.17	2.31	2.54	447	7.47	7.40	14.91 Menengah
1	2	0.021	0.16	4.17	2.56	460	7.51	7.32	15.21 Menengah
2	3	0.026	0.17	2.76	2.54	502	6.81	8.55	17.41 Baik
3	4	0.027	0.25	0.49	2.31	448	7.41	8.05	19.25 Buruk
4	5	0.027	0.19	3.56	3.08	501	7.12	7.69	20.92 Menengah
...
96	97	0.542	4.76	8.32	2.42	432	7.76	9.75	12.87 Buruk
97	98	0.754	6.54	3.43	5.12	542	5.33	5.21	32.76 Baik
98	99	0.535	8.65	8.67	7.64	874	9.54	9.75	23.43 Menengah
99	100	0.654	9.65	9.65	8.54	829	8.64	2.54	12.34 Menengah
100	101	0.398	0.65	8.53	5.33	452	1.64	6.44	12.43 Menengah

Gambar 1. Sampel Dataset

5. Pembagian Data

Pembagian data dalam penelitian ini menggunakan *data training* dan *data testing* dengan rasio 90:10. Dimana *data training* memiliki kisaran data sebanyak 10 dari dataset. Algoritma yang menggunakan rasio 90:10 memiliki perfoma yang lebih baik dibandingkan dengan algoritma lainnya dimana tingkat akurasinya dapat diperoleh sebesar 99,40%. Adapun rasio 80:20, 70:30, dan 60:40 yang digunakan dalam penelitian ini sehingga rasio pembagian data lebih bervariatif dimana tujuan yang diharapkan adalah agar memperoleh model naive bayes manakah yang memiliki hasil estimasi terbaik.

6. Klasifikasi dengan Naive Bayes

Naive Bayes Classifier merupakan suatu metode klasifikasi yang berakar

pada Teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistic yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu prediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama Naive Bayes adalah asumsi yang sangat kuat akan indepenensi dari masing-masing kondisi/kejadian.

Metode Naive Bayes ini hanya membutuhkan jumlah *data training* yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai variable independent, maka hanya varian dari suatu variable dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan matriks.

Adapun rumus yang lain dalam naive bayes adalah sebagai berikut :

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

Keterangan :

X : Data dengan class yang belum diketahui

H : hipotesis data X merupakan suatu class spesifik

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

P(H) : Probabilitas hipotesis H

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

P(X) : Probabilitas X

7. Model Pengukuran Akurasi

Akurasi dalam pengukuran merupakan tingkat ke dekatkan pengukuran kuantitas terhadap nilai yang sebenarnya, sedangkan Kepresision dari suatu system pengukuran diartikan sejauh mana pengulangan pengukuran dalam kondisi yang tidak berubah mendapatkan hasil yang sama. Akurasi merupakan ketepatan dari sebuah model dalam memprediksi/mengklasifikasi sebuah studi kasus. Berikut ini merupakan rumus yang digunakan untuk menilai akurasi dari metode *Naïve Bayes* yang digunakan dalam penelitian ini.

$$\text{Akurasi} = \frac{\sum \text{Data Benar}}{100\%} \times n$$

Keterangan :

$\sum \text{Data Benar}$: Jumlah benar dalam sampel uji

n : Banyaknya jumlah sampel uji

HASIL DAN PEMBAHASAN

Dalam penggunaan metode *Naïve Bayes* penyederhanaan menggunakan nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dalam metode *Naïve Bayes* memakai klasifikasi berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel penjelas. Pada pembelajaran ini dipakai atau ditekankan pada hasil dan klasifikasi dari probabilitas. Keuntungan menggunakan metode *Naïve Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menemukan *estimasi parameter* yang diperlukan dalam proses pengklasifikasian.

a. Visualisasi Data

Visualisasi data bertujuan untuk memudahkan seseorang dalam mengamati dan memahami sebuah data. Terdapat beberapa perbandingan dan cara melihat data untuk mengungkapkan visualnya. Dengan menggunakan metode *Naïve Bayes* dalam resentase variabel x dan y untuk melihat hasil kandungan dari CPO yang digunakanapakah kandungan CPO terbaik, menengah, dan buruk. Visualisasi data disajikan dalam bentuk antar variabel x dengan variabel y dan penyajian dalam bentuk seperti diagram lingkaran, tiap variabel menggunakan warna yang berbeda agar lebih mudah dipahami dan mudah diingat. Untuk itu visualisasi dari data kelapa sawit yang digunakan dalam penelitian ini terdapat pada gambar diagram sebagaimana yang terdapat dibawah ini:

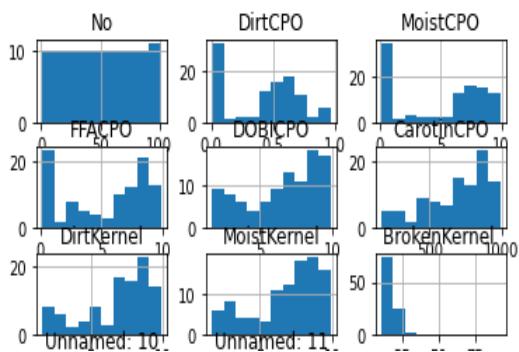


Gambar 2. Presentase Kandungan CPO Menurut Kualitas

Pada gambar diatas, hasil dari klasifikasi sebagai sumbu Y divisualisasikan dengan diagram lingkaran dimana persentase klasifikasi CPO sebanyak 100 data, klasifikasi CPO baik sebanyak 35,6%, klasifikasi CPO menengah sebanyak 46,5%, klasifikasi CPO buruk sebanyak 17,8%.

b. Perolehan grafik Kualitas CPO

Berikut ini adalah visualisasi data grafik pada kualitas CPO pada kelapa sawit sebagai berikut :



Gambar 3. Grafik Pada Kualitas CPO

Pada gambar diatas, terdapat 8 garfik dalam kualitas CPO yaitu : Dirt CPO, Moist CPO, FFA CPO, DOBI CPO, Carotin CPO, Dirt Kernel, Moist Kernel, Broken Kernel. Berdasarkan grafik diatas, dapat diketahui bahwa

kandungan CPO yang banyak terdapat pada DOBI CPO.

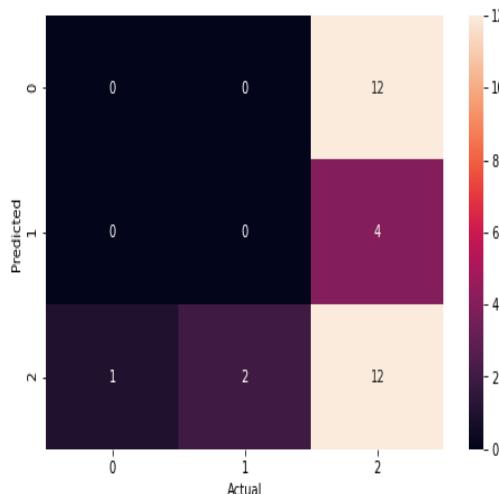
c. Tahap Pengujian Data

Pada penelitian ini menggunakan metode *naive bayes* dimana dilakukan perbandingan menggunakan *Naive Bayes* agar diperoleh hasil akurasi terbaik digunakan dalam estimasi kualitas CPO. Penelitian ini menggunakan pembagian data untuk menguji data sebanyak 3 bentuk diantaranya pembagian rasio pertama yaitu 90:10, kedua 80:20, dan rasio ketiga 70:30. Pengujian data dengan rasio untuk mencari model *Naive Bayes* dengan nilai tinggi. Sehingga, akan diketahui nilai *Naive Bayes* manakah yang terbaik terhadap dataset untuk mencari kualitas kandungan CPO. Berikut ini pembahasan penelitian untuk menentukan nilai *Naive Bayes* terbaik kedalam rincian sebagai berikut:

Tabel 1. Pengujian Data

No	Training Testing	Akurasi
1	90 : 10	36%
2	80 : 20	33%
3	70 : 30	38%

Berdasarkan tabel diatas, pada pengujian rasio 90:10, maka diperoleh akurasi 36%, pada pengujian rasio 80:20 diperoleh nilai akurasi 33% dan pengujian rasio 70:30 diperoleh nilai akurasi 38%. Sebuah *performance* pada naive bayes dapat dihitung nilai rata-rata sebesar 35,6%. Perbandingan dari pengujian metode *Naive Bayes* diatas, diperoleh rasio 70:30 adalah akurasi yang terbaik. Perolehan confusion matrix untuk metode *Naive Bayes* adalah :



Gambar 4. Visualisasi Confusion Matrix

Berdasarkan gambar diatas, maka didapat hasil nilai *Naïve Bayes* menghasilkan nilai *Precision* untuk kualitas CPO yang baik 0%, untuk yang kualitas CPO buruk 0%, dan untuk kualitas menengah 43%, nilai *recall* untuk yang kualitas baik 0%, yang kualitas CPO buruk 0%, dan untuk kualitas CPO menengah 80%, nilai *score* untuk yang kualitas CPO baik 0%, untuk kualitas CPO buruk 0% dan yang kualitas CPO menengah 56% untuk metode *Naïve Bayes*.

KESIMPULAN

Kualitas produk mentah seperti CPO (Crude Palm Oil) adalah salah satu faktor utama bagi pelanggan untuk membuat keputusan pembelian khususnya standar yang terdapat pada Standar Nasional Indonesia (SNI). Berdasarkan hasil penelitian yang telah dibahas, dapat diambil kesimpulan bahwa klasifikasi kualitas produk kelapa sawit dapat dilakukan melalui sebuah aplikasi agar lebih mempermudahkan dengan menerapkan algoritma *Naïve Bayes* berbasis web.

Dan tingkat akurasi metode *Naïve Bayes* dalam menentukan kualitas produk kelapa sawit adalah sebesar 36%. Untuk meningkatkan kinerja klasifikasi kualitas kelapa sawit disarankan dengan menambahkan jumlah data uji yang lebih banyak agar tingkat akurasi menjadi lebih akurat.

Berdasarkan hasil pengolahan data menggunakan metode *Naïve Bayes* menggunakan software *jupyter notebook*, *visual studio code* dan *python* sebagai bahasa pemrogramannya serta pengujinya menggunakan 3 rasio pengukuran nilai akurasi yaitu 90:10, 80:20, dan 70:30 maka diperoleh nilai akurasi paling bagus adalah rasio 70:30 dengan nilai akurasi 36%.

Dengan demikian metode *Naïve Bayes* digunakan sebagai metode klasifikasi yang akan diimplementasikan dalam pembuatan *deploy* aplikasi untuk mengestimasi kadar kualitas CPO yang dihasilkan kelapa sawit.

DAFTAR PUSTAKA

- [1] D. Suryani, A. Yulianti, E. L. Maghfiroh, and J. Alber, "Quality Classification of Palm Oil Products Using *Naïve Bayes* Method," *Sistemas*, vol. 11, no. 1, p. 251, 2022, doi: 10.32520/stmsi.v11i1.1713.
- [2] U. Ruswandi, M. Erihadiana, and A. Saepurahman, "Kajian Riset Pluralisme Dan Multikulturalisme," vol. 6, pp. 777–787, 2022.
- [3] Anggara, Bayu, and R M Nasrul

-
- Halim D, ‘Penerapan Data Mining Rumah Sakit Umum Prabumulih Menggunakan Algoritma Naive Bayes (Study Kasus : Penyakit Jantung)’, 209–20
- [4] Depari, Deo Haganta, Yuni Widiastiwi, Mayanda Mega Santoni, Fakultas Ilmu Komputer, Universitas Pembangunan, Nasional Veteran, and others, ‘Perbandingan Model Decision Tree , Naive Bayes Dan Random Forest Untuk Prediksi Klasifikasi Penyakit Jantung’, 4221 (2022), 239–48
- [5] Hafiz, Dimsyiar M Al, Khoirul Amaly, Javen Jonathan, and M Teranggono Rachmatullah, ‘Sistem Prediksi Penyakit Jantung Menggunakan Metode Naive Bayes’, 2.2, 151–57