Jurnal Sains Teknologi dan Sistem Informasi
journal.yp3a.org/index.php/satesi
Vol. 5 No. 2 (2025) 183 – 189 ISSN Media Elektronik: 2807–8152

Klasifikasi Rambut Rontok Menggunakan Metode *Naive Bayes*

Arif Hidayat[1], Sutedi[2]
[1,2]Teknik Informatika, Ilmu Komputer, Institut Informatika dan Bisnis Darmajaya, Indonesia
Email: [1]arif.2421211030p@mail.darmajaya.ac.id, [2*]*sutedi@darmajaya.ac.id*

ABSTRACT
Hair loss is a common problem experienced by various groups, both men and women. In 2023, a survey found that 64.7% of hair problems were hair loss, followed by dandruff, dry hair, dull hair, limp hair, and oily hair. The causes of hair loss are diverse, ranging from genetic and hormonal factors to lifestyle. To accurately identify and classify the types of hair loss, a method is needed that can handle data efficiently and accurately. This study aims to classify types of hair loss using the Naive Bayes method, a classification algorithm in data mining based on probability and Bayes' theorem. The data used consists of various attributes such as staying up late, brain activity duration, stress levels, dandruff, and hormones. The Naive Bayes method was chosen because of its ability to handle data with a large number of attributes and produce fairly accurate predictions even with the assumption of independence between features. The results of the study show that the Naive Bayes method is capable of classifying types of hair loss with an accuracy rate of 76.67%. These findings indicate that this method can be used as a tool to aid in the rapid and efficient initial diagnosis of hair loss problems.

Keywords: Hair Loss, Data Mining, Classification, Naïve Bayes, Prediction.

ABSTRAK
Rambut rontok merupakan permasalahan umum yang dialami oleh berbagai kalangan, baik pria maupun wanita. Pada tahun 2023 lembaga survey mendapatkan 64,7% permasalahan pada rambut ialah rambut rontok disusul ketombe, rambut kering, kusam, lepek dan berminyak. Penyebab rambut rontok sangat beragam, mulai dari faktor genetik, hormonal, hingga gaya hidup. Untuk mengidentifikasi dan mengklasifikasikan jenis rambut rontok secara tepat, diperlukan metode yang mampu menangani data secara efisien dan akurat. Penelitian ini bertujuan untuk mengklasifikasikan jenis rambut rontok menggunakan metode *Naive Bayes*, sebuah algoritma klasifikasi dalam *data mining* yang berbasis pada *probabilitas* dan *Teorema Bayes*. Data yang digunakan terdiri dari berbagai atribut seperti begadang, durasi kerja otak, tingkat stres, ketombe, dan hormon. Metode *Naive Bayes* dipilih karena kemampuannya dalam menangani data dengan jumlah atribut yang banyak serta menghasilkan prediksi yang cukup akurat meskipun dengan asumsi independensi antar fitur. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* mampu mengklasifikasikan jenis rambut rontok dengan tingkat akurasi mencapai 76,67%. Temuan ini menunjukkan bahwa metode ini dapat digunakan sebagai alat bantu dalam diagnosis awal permasalahan rambut rontok secara cepat dan efisien.

Kata Kunci: Rambut Rontok, *Data Mining*, *Classification*, *Naïve Bayes*, Prediksi.

1. Pendahuluan

Setiap orang memiliki kebutuhan fisik dan kesehatan yang berbeda, salah satunya adalah rambut. Rambut pada kepala adalah salah satu bagian terpenting bagi seseorang khususnya secara estetika. Kesehatan kulit kepala dan rambut adalah suatu keadaan kulit kepala dan rambut dengan tidak adanya penyakit mengganggu seperti ketombe, rambut rontok, kering, berminyak, kusam, dan sulit disisir [1]. Khususnya dalam profesi tertentu, tubuh yang sehat dan indah termasuk rambut, seringkali dibutuhkan. Rambut rontok pada manusia memiiki beberapa faktor penyebabnya yaitu faktor *genetic* ataupun gaya hidup manusia. Hal ini membuat manusia berusaha untuk membuat rambut mereka tetap bagus dan indah, serta melakukan perawatan untuk menghilangkan rambut rontok tersebut. Dalam hal ini bagaimana membuat alarm atau *trigger* bagi manusia agar dapat mengetahui dan dapat melakukan sedini mungkin perawatan terhadap rambutnya.

Berdasarkan survei yang dirilis oleh *Lembaga Jajak Pendapat (Jakpat)* pada pertengahan 2023, masalah rambut rontok merupakan keluhan paling umum yang dialami oleh masyarakat Indonesia, dengan persentase mencapai 64,7% dari 3.041 responden. Masalah ini

Diterima Redaksi: 05-08-2025 | Selesai Revisi: 15-09-2025 | Diterbitkan Online: 15-10-2025

secara signifikan lebih banyak terjadi pada kelompok usia muda, terutama rentang usia 20-25 tahun. Selain kerontokan, survei tersebut juga memaparkan masalah lainnya secara berurutan, yaitu rambut berketombe yang dialami oleh 44,3% responden, diikuti oleh rambut kering dan kusam (30,8%), rambut berminyak atau lepek (26,1%), serta rambut rusak atau bercabang (18%). Hasil survei ini memberikan gambaran jelas mengenai tingkat kesehatan rambut di masyarakat, sehingga diharapkan dapat mendorong masyarakat untuk lebih bijak dalam memilih metode perawatan yang sesuai untuk menjaga kesehatan rambut mereka [2].

Menurut Dokter Pittara Rambut rontok adalah lepasnya rambut secara berlebihan. Kondisi ini dapat mengakibatkan penipisan rambut atau kebotakan, baik sementara maupun permanen. Rambut rontok juga bisa terjadi sedikit demi sedikit atau banyak secara tiba-tiba [3].

Terdapat beragam metode yang bisa diterapkan untuk klasifikasi rambut rontok, di antaranya adalah Decision Tree, K-Nearest Neighbor (KNN), *Naïve Bayes*, ID3, dan C4.5. Dari kelima metode ini, *Naïve Bayes* merupakan sebuah metode klasifikasi statistik yang sederhana namun efektif. Metode ini dikenal memiliki tingkat akurasi yang baik dan mampu meminimalkan tingkat kesalahan (error rate) dalam proses pengklasifikasian data [4].

Pada penelitian berjudul "Penggunaan *Algoritma Naive Bayes* dan *Particle Swarm Optimization (PSO)* untuk Mendeteksi Stroke" membandingan antara dua metode untuk meningkatkan prediksi stroke otak berdasarkan algoritma Naive Bayes. Pertama, evaluasi kinerja Naive Bayes secara independen tanpa metode optimisasi tambahan. Hasil eksperimen menunjukkan bahwa model Naive Bayes memiliki tingkat akurasi sebesar 86,21%, menunjukkan kemampuan model tersebut dalam memprediksi stroke otak tanpa adanya penyetelan tambahan. Namun, meskipun performa ini menunjukkan potensi, masih ada ruang untuk peningkatan lebih lanjut [5].

Penelitian ini mengevaluasi kinerja algoritma *Naïve Bayes* dalam klasifikasi kebotakan menggunakan *data mining*. Dataset yang digunakan terdiri dari 999 data individu dengan 13 atribut, termasuk atribut kelas yang menentukan kebotakan (kelas 1) atau tidak (kelas 0). Hasil pengujian menunjukkan bahwa akurasi tertinggi adalah 52,33% pada rasio 70:30 dan terendah 49.00% pada rasio 90:10. Pengujian *k-fold cross validation* (K=3,4,5 dan 6) memberikan akurasi tertinggi pada K= 6 dengan akurasi 55.42% [6].

Naïve Bayes adalah metode *klasifikasi* dengan melakukan *preprocessing* pada judul berita, kemudian menghitung *probabilitas* setiap kelasnya. Kelas yang dipakai dalam metode ini adalah kategori berita. Kategori berita meliputi, Olahraga, Hiburan, Kesehatan, Politik, dan Teknologi. Dari 500 data latih yang dijadikan acuan untuk menghitung probabilitas, setelah data uji dimasukkan maka akan dihitung probabilitas setiap kata yang digunakan dan akan menghasilkan suatu kategori, dari 50 data yang diuji sebanyak 43 dokumen yang berhasil sesuai dengan kategori yang tepat yaitu sebesar 86% dan sebanyak 7 dokumen dengan kesalahan kategori sebesar 14% [7].

Metode analisis manual yang telah lama digunakan dalam mendiagnosis penyakit kini tidak lagi memadai. Perkembangan teknologi dan sistem berbasis pengetahuan, khususnya dalam bidang medis, menuntut adanya penerapan sistem komputerisasi yang lebih efektif sebagai sarana analisis. Dengan demikian, pengembangan sistem pengetahuan berbasis komputer modern menjadi kebutuhan mendesak untuk diagnosis penyakit yang lebih efisien [8].

Data Mining adalah proses menemukan informasi yang berguna secara otomatis dalam repositori data yang besar. Teknik *data mining* digunakan untuk menjelajahi kumpulan data yang besar untuk menemukan pola baru dan berguna yang mungkin tidak diketahui. Teknik ini juga memberikan kemampuan untuk memprediksi hasil pengamatan di masa depan, seperti jumlah yang akan dibelanjakan pelanggan di toko online atau toko fisik [9].

Metode ini bekerja berdasarkan *Teorema Bayes* yang menghitung kemungkinan suatu kejadian berdasarkan informasi sebelumnya dengan menggunakan teknik *probabilitas* dan *statistik*. *Naïve Bayes* dapat digunakan untuk memprediksi kategori rambut rontok berdasarkan parameter seperti begadang, tingkat tekanan, konsumsi kopi, durasi kerja otak, tingkat stres, berenang, minyak rambut, ketombe, *libido* (*hormon*). Keunggulan *Naïve Bayes Classifier* ialah dasar namun akurat. Prosedur kategorisasi data dibagi menjadi dua tahap. Langkah pertama adalah berlatih dengan contoh *(Training Example)*. Tahap kedua adalah proses pengkategorian data yang belum diketahui kelasnya [10].

Dengan penerapan metode *Naïve Bayes*, diharapkan dapat diperoleh sistem klasifikasi rambut rontok yang akurat, Oleh karena itu penulis akan melakukan penelitian tentang "*Klasifikasi Rambut rontok menggunakan metode Naïve Bayes*".

2. Metode Penelitian

2.1 Metode Pendekatan Penyelesaian

Data mining atau knowledge discovery in database (KDD) adalah sebuah metode yang dapat digunakan untuk mencari pola informasi baru yang sebelumnya tidak diketahui dari sejumlah besar data atau big data, tetapi valid dan dapat dipahami yang memiliki potensi bermanfaat dalam pengambilan keputusan. Pengolahan big data menjadi bagian awal menemukan tren data yang selanjutnya dapat dijadikan model yang dapat digunakan oleh *controller*

DOI: 10.54259/satesi.v5i2.5612

untuk memantau dan mengendalikan sistem secara adaptif [11].

Sebagai sebuah proses yang terstruktur, KDD memiliki beberapa tahapan sebagai berikut:

a) *Data Selection*

Dalam penelitian ini data yang digunakan adalah 400 data public yang diperoleh dari kaggle. Data terdiri 12 fitur seperti begadang, tingkat tekanan, konsumsi kopi, durasi kerja otak, ujian sekolah, tingkat stres, merk shampo, berenang, mencuci rambut, minyak rambut, ketombe, libido (hormon) di seleksi menjadi 9 fitur yaitu begadang, tingkat tekanan, konsumsi kopi, durasi kerja otak, tingkat stres, berenang, minyak rambut, ketombe, libido (hormon) karena dianggap kurang penting dan memiliki kesamaan dengan fitur lain serta terdiri dari 4 label yaitu few, medium, many, a lot.

Tabel 1. Seleksi Fitur

stay_up_late	swimming
pressure_level	hair_washing
coffee_consumed	hair_grease
brain_working_duration	Dandruff
school_assesssment	shampoo_brand
stress_level	Libido

stay_up_late
pressure_level
coffee_consumed
brain_working_duration
stress_level
swimming
hair_grease
Dandruff
Libido

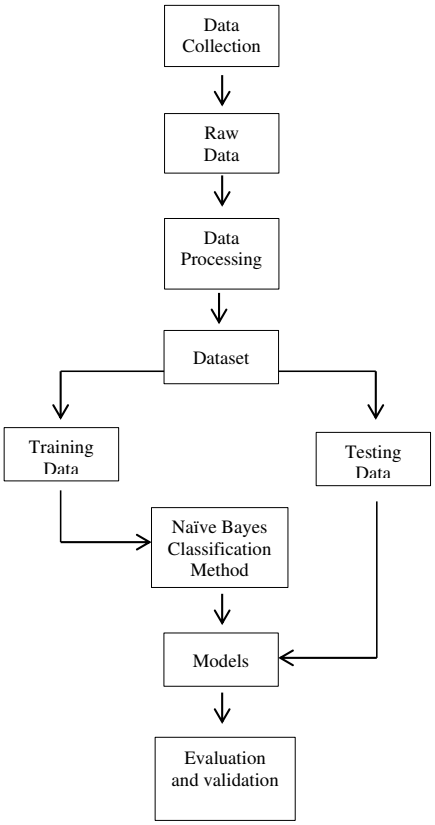
b) *Pre-processing Cleaning*

Sebelum proses *data mining* dilanjutkan, data yang telah diseleksi harus melalui tahap pembersihan (*data cleaning*). Tahap ini krusial untuk menangani berbagai masalah seperti duplikasi data, informasi yang inkonsisten, dan memperbaiki kesalahan pengetikan (*typo*). Dimana atribut-atribut terdiri dari begadang, tingkat tekanan, konsumsi kopi, durasi kerja otak, tingkat stres, berenang, minyak rambut, ketombe, libido (hormon).

c) Transformation

Transformasi data adalah tahap krusial dalam persiapan *data mining* yang berfokus pada pengubahan dan penggabungan data ke dalam format yang optimal untuk analisis. Langkah ini esensial karena data mentah seringkali memiliki format atau struktur yang tidak kompatibel dengan algoritma mining yang akan digunakan [12].

2.2 Tahapan Penelitian



1. Identifikasi Masalah

   Tujuan Penelitian: tujuan dari penelitian ini adalah untuk mengukur efektifitas metode klasifikasi *Naïve Bayes* dalam memprediksi kemungkinan seseorang mengalami rambut rontok

   Rumusan Masalah: apakah metode klasifikasi *Naïve Bayes* efektif digunakan untuk memprediksi rambut rontok ?

2. Studi Literatur
   Mengkaji penelitian terdahulu terkait:

   - Penyebab rambut rontok (genetik, hormonal, stres, pola makan, dll).
   - Penggunaan Naive Bayes dalam klasifikasi data .
   - Menentukan variabel-variabel penting untuk dijadikan fitur.

3. Pengumpulan Data

   Mengumpulkan data kasus rambut rontok dari dataset public.

   fitur yang dikumpulkan:

DOI: 10.54259/satesi.v5i2.5612

a. Begadang
b. Tingkat Tekanan
c. Konsumsi Kopi
d. Durasi Kerja Otak
e. Tingkat Stres
f. Berenang
g. Minyak Rambut
h. Ketombe
i. Libido (Hormon)

4. Pra-pemrosesan Data

Pembersihan Data: Menghapus atau menambahkan nilai kosong

Transformasi Data: Konversi data kategorikal menjadi numerik (Encoding).

Seleksi Fitur: Menghapus fitur yang tidak relevan.

5. Pembagian Dataset

Membagi data menjadi:

- Training set (70–80%)
- Testing set (20–30%)

6. Penerapan Algoritma Naive Bayes

Pilih jenis Naive Bayes yang sesuai:

- Gaussian Naive Bayes → jika data numerik dan berdistribusi normal.
- Multinomial/Bernoulli → untuk data kategorikal.

Latih model pada data training.

Algoritma *Naïve Bayes* menggunakan teknik percabangan matematis dengan mencari probabilitas terbesar dari sebuah klasifikasi berdasarkan frekuensi dari setiap klasifikasi terhadap data training, yang sering disebut sebagai teori probabilistik. Rumus perhitungan *Naïve Bayes* adalah sebagai berikut:

$$P\ (X \mid Y) = \frac{P(Y \mid X) \times P(X)}{P(Y)}$$

Pada rumus tersebut, P(X) merupakan probabilitas awal dari hipotesis X atau seberapa besar kemungkinan hipotesis tersebut benar tanpa mempertimbangkan data yang ada. Selanjutnya, P(Y|X)P(Y|X)P(Y|X) adalah probabilitas Y berdasarkan kondisi dari hipotesis X, yang merepresentasikan seberapa besar kemungkinan data Y terjadi jika hipotesis X benar. Terakhir, P(Y)P(Y)P(Y) adalah probabilitas Y atau kemungkinan terjadinya data tanpa mempertimbangkan hipotesis. Pendekatan ini memungkinkan algoritma untuk menghitung P(X|Y)P(X|Y)P(X|Y), adalah seberapa besar kemungkinan hipotesis X benar mengingat data Y. Nilai probabilitas ini kemudian digunakan untuk menentukan kelas yang paling mungkin terjadi pada data Y. Naive Bayes mengasumsikan bahwa semua fitur dalam data independen satu sama lain, yang menyederhanakan penghitungan namun tetap memberikan hasil yang efektif di berbagai aplikasi, termasuk prediksi risiko obesitas [13].

7. Evaluasi

Dalam konteks evaluasi klasifikasi, kriteria standar yang lazim digunakan adalah *precision, recall*, dan *accuracy* . Perhitungan untuk setiap metrik tersebut dapat dilakukan dengan mengaplikasikan persamaan berikut pada data yang tersaji di Tabel 2. Lebih lanjut, *F1-measure*, yang merupakan rerata dari *precision* dan *recall*, turut digunakan sebagai tolok ukur performa tambahan [14].

Tabel 2. Tabel Penilaian

	Positif	Negatif
Positif	a	b
Negatif	c	d

Menggunakan metrik evaluasi seperti:

- *Akurasi*

Accuracy adalah metrik yang paling umum, mengukur tingkat kebenaran model secara keseluruhan. Ini adalah rasio antara jumlah prediksi yang benar (baik positif maupun negatif) terhadap total keseluruhan data.

$$Accuracy = (\frac{a+d}{total}) \times 100\%$$

- *Precision*

Dalam klasifikasi, *precision* sering disamakan dengan *positive predictive value*. Metrik ini mengukur seberapa akurat prediksi positif yang dibuat oleh model. Dengan kata lain, dari semua data yang diprediksi sebagai "positif", berapa persen yang sebenarnya memang positif:

$$Precision = (\frac{d}{b+d}) \times 100\%$$

- *Recall*

Recall mengukur kemampuan model untuk menemukan kembali semua data yang relevan atau positif. Dengan kata lain, dari semua data yang seharusnya "positif", berapa persen yang berhasil diidentifikasi oleh model.

$$Recall = (\frac{d}{c+d}) \times 100\%$$

- F1-score

F1-Measure adalah rata-rata harmonik dari *precision* dan *recall*. Metrik ini memberikan skor tunggal (dari 0 hingga 1) yang menyeimbangkan kedua metrik

DOI: 10.54259/satesi.v5i2.5612

tersebut, di mana nilai 1 adalah yang terbaik. Ini sangat berguna ketika kita ingin memastikan model memiliki *precision* dan *recall* yang baik secara bersamaan [15].

$$F1 = \frac{2xPrecision\ x\ Recall}{Precision+Recall}$$

Evaluasi hasil klasifikasi: misalnya apakah model bisa membedakan antara rambut rontok ringan, sedang, dan parah.

8. Kesimpulan dan Saran

Menyimpulkan efektivitas Naive Bayes dalam klasifikasi rambut rontok.

Memberikan rekomendasi untuk peningkatan model, misalnya:

- Penambahan data
- Eksperimen dengan algoritma lain (SVM, Decision Tree)

3. Hasil dan Pembahasan

Pada penelitian ini menggunakan algoritma *Naive Bayes* untuk mengklasifikasikan tingkat kerontokan rambut yang terdiri dari 9 fitur seperti begadang, tingkat tekanan, konsumsi kopi, durasi kerja otak, tingkat stres, berenang, minyak rambut, ketombe, libido (hormon) dan 4 label yang terdiri dari few, medium, many, a lot. Dataset yang digunakan terdiri dari 400 data sampel yang telah melalui proses prapemrosesan seperti pembersihan data, pembobotan fitur dan diimplementasikan pada *tools Google Collabs* sehingga memperoleh tingkat akurasi 76% dengan rasio 70% data training dan 30% data testing.

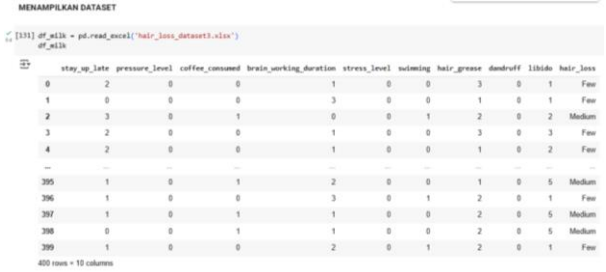
Implementasi dengan *Tools Google Collabs:*

IMPORT LIBRARY

```
[130] import pandas as pd
      from sklearn.naive_bayes import GaussianNB
      from sklearn import metrics
      from sklearn.metrics import classification_report
      from sklearn.metrics import ConfusionMatrixDisplay
      from sklearn.model_selection import train_test_split
```

Gambar 1. Import Library

MENAMPILKAN DATASET



Gambar 2. Menampilkan Dataset

```
df_milk.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 10 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   stay_up_late            400 non-null    int64
 1   pressure_level          400 non-null    int64
 2   coffee_consumed         400 non-null    int64
 3   brain_working_duration  400 non-null    int64
 4   stress_level            400 non-null    int64
 5   swimming                400 non-null    int64
 6   hair_grease             400 non-null    int64
 7   dandruff                400 non-null    int64
 8   libido                  400 non-null    int64
 9   hair_loss               400 non-null    object
dtypes: int64(9), object(1)
memory usage: 31.4+ KB
```

Gambar 3. Manampilkan info dataset

MENAMPILKAN PROBABILITAS DATA CLASS/LABEL

```
df_milk['hair_loss'].value_counts()
```

hair_loss	count
Few	169
Medium	167
Many	42
A lot	22

dtype: int64

Gambar 4. Mencari Probabilitas Data Class/Label

Featur Scaling dilakukan jika jarak antara value-value pada setiap fitur terlihat berjauhan, maka dilakukan penskalaan agar terlihat tidak jauh jarak antar nilainya.

FEATUR SCALING (JIKA DIPERLUKAN)

```
[134] from sklearn.preprocessing import StandardScaler

      sc = StandardScaler()
      x_train = sc.fit_transform(x_train)
      x_test = sc.fit_transform(x_test)

[136] print(x_train)

[[ 0.79322978 -0.56635211 -0.02157282 ... -1.21898252 -0.53073452
  -1.01980776]
 [ 0.79322978 -0.56635211 -0.57069905 ...  0.42195549 -0.53073452
  -1.01980776]
 [ 0.13220496 -0.56635211 -0.57069905 ... -0.39851352 -0.53073452
  -1.01980776]
 ...
 [ 0.79322978 -0.56635211 -0.57069905 ...  0.42195549 -0.53073452
  -0.46962246]
 [-1.18984467 -0.56635211 -0.57069905 ... -0.39851352 -0.53073452
  -0.46962246]
 [ 4.09835386  2.93170506  4.92056329 ...  2.0628935   2.88548769
  -1.56999307]]
```

Gambar 5. Feature Scaling

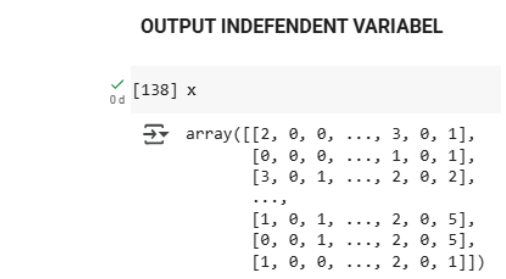
Memisahkan variable Independent yang terdiri dari 9 fitur dan variable dependent yang terdiri dari 4 Label untuk diproses.

MEMISAHKAN VARIABEL INDEPENDENT DAN DEPENDENT (LABEL)

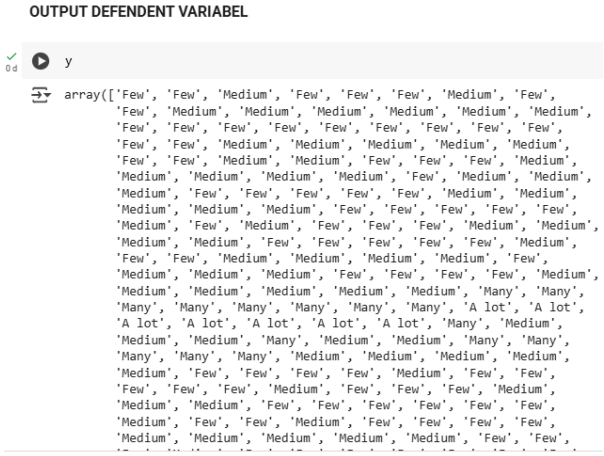
```
[137] x = df_milk.iloc[:,:9].values
      y = df_milk.iloc[:,9].values
```

Gambar 6. Memisahkan variable Independent dan dependent

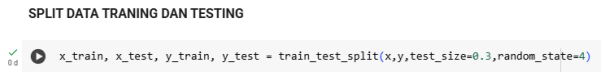
DOI: 10.54259/satesi.v5i2.5612



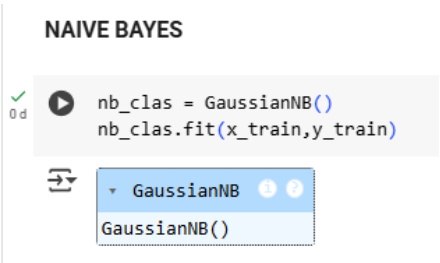
Gambar 7. Output Variable Independent



Gambar 8. Output Variabel Dependent



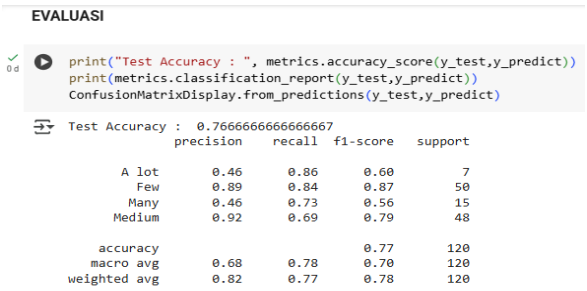
Gambar 9. membagi data traning dan testing



Gambar 10. membuat model / rumus *naïve bayes*

Gambar 11. membuat variabel testing (random)

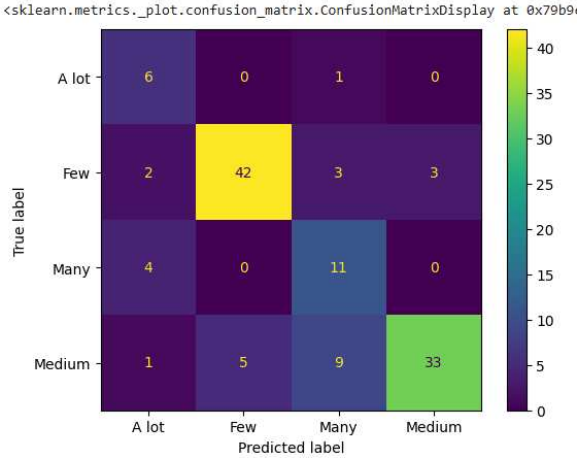
Dari hasil evaluasi dibawah ini mendapatkan tingkat akurasi sebesar 76% dengan data testing 30% yaitu 120 data, terbagi menjadi 7 data a lot, 50 data few, 15 data many dan 48 data medium.



Gambar 12. Evaluasi

Hasil pada diagram metric dibawah ini menunjukan bahwa:

a. dari 7 data a lot memperoleh prediksi 6 data sesuai dan 1 data masuk pada label many.
b. pada label few terdapat 50 data yang sesuai sebanyak 42 data dan 2 data masuk label a lot, 3 data masuk label many dan 3 data masuk label medium.
c. pada label many terdapat 15 data yang sesuai sebanyak 11 data dan 4 data masuk label a lot
d. pada label medium terdapat 48 data, yang sesuai sebanyak 33 data, 1 data masuk label a lot, 5 data masuk label few dan 9 data masuk label many.



Gambar 13. Diagram Metrik

Gambar 14. Input contoh data baru

DOI: 10.54259/satesi.v5i2.5612

4. Kesimpulan

Penelitian ini mengevaluasi kinerja algoritma *Naïve Bayes* dalam klasifikasi rambut rontok menggunakan *data mining*. Dataset yang digunakan terdiri dari 400 data individu dengan 10 atribut, termasuk atribut kelas yang menentukan tingkat rambut rontok (*few, medium, many, a lot*). Hasil pengujian menunjukkan bahwa akurasi tertinggi adalah 76,67% pada rasio 70:30 dan terendah 55.00% pada rasio 90:10.

Berdasarkan hasil penelitian yang telah dilakukan, ada beberapa saran untuk peneliti yang ingin mengembangkan penelitian ini yaitu:

1. Penelitian ini hanya menggunakan satu algoritma klasifikasi yang dioptimasi menggunakan metode feature selection atau optimasi, sehingga untuk peneliti selanjutnyaa dapat dikembangkan dengan menggunakan beberapa algoritma klasifikasi untuk mendapatkan algoritma yang paling akurat.
2. Diharapkan pada penelitian berikutnya selain menggunakan *Google Collabs* juga menggunakan *tools* lainya seperti *Rapidminer* dan *MATLAB*.

SUMBER RUJUKAN
Referensi

[1] Zaradiya Audrey Tritania, "Analisis penggunaan Jilbab dan Perawatan Rambut terhadap Kesehatan Kulit Kepala Dan Rambut Pada Mahasiwi Berjilbab," vol. 12, no. 2, pp. 88–94, 2023, doi: https://doi.org/10.26740/jtr.v12n2.53889.

[2] Nada Naurah, "Survei: Sebagian Besar Orang Indonesia Alami Permasalahan Rambut Rontok," goodstats.id. [Online]. Available: https://goodstats.id/article/survei-sebagian-besar-orang-indonesia-alami-rambut-rontok-ojtfz

[3] dr. Pittara, "Pengertian Rambut Rontok," alodokter.com. [Online]. Available: https://www.alodokter.com/rambut-rontok

[4] O. Saad, A. Darwish, and R. Faraj, "A survey of machine learning techniques for Spam filtering," *J. Comput. Sci.*, vol. 12, no. 2, pp. 66–73, 2012.

[5] M. S. Hasibuan, D. Fransisca, J. Magister, T. Informatika, and F. I. Komputer, "Penggunaan Algoritma Naive Bayes dan Particle Swarm Optimization (PSO) untuk Mendeteksi Stroke," pp. 109–118, 2022, doi: https://doi.org/10.31284/j.integer.0.v9i1.5738.

[6] Darussalam, "Implementasi Algoritma Naive Bayes untuk Klasifikasi Kesehatan Rambut," Universitas Sulawesi Barat, 2024.

[7] H. Hartono, A. Hajjah, and Y. N. Marlim, "Penerapan Metode *Naïve Bayes* Classifier Untuk Klasifikasi Judul Berita Application of the *Naïve Bayes* Classifier Method for News Title *Classification*," vol. 12, no. 1, pp. 37–46, 2023, doi: https://doi.org/10.21107/simantec.v12i1.19398.

[8] R. Anggriawan and H. W. Nugroho, "Komparasi Algoritma C4.5 dan Naive Bayes Dalam Prediksi Penderita Penyakit Gagal Jantung," *J. SIMADA (Sistem Inf. dan Manaj. Basis Data)*, vol. 6, no. 1, pp. 82–91, 2023, doi: 10.30873/simada.v6i1.3425.

[9] N. Shah and K. Shah, "Introduction to *Data Mining*," *Pract. Data Min. Tech. Appl.*, pp. 1–6, 2023, doi: 10.1201/9781003390220-1.

[10] A. Pebdika *et al.*, "Klasifikasi Menggunakan Metode Naive Bayes untuk Menentukan Calon Penerima PIP," vol. 7, no. 1, pp. 452–458, 2023, doi: https://doi.org/10.36040/jati.v7i1.6303.

[11] Y. Gamaliel, T. A. Nugroho, and L. C. Aguskin, "Knowledge Discovery in Database dengan Multivariate Linear Regression pada Sistem Pertanian Hidroponik Berbasis Internet of Things," *J. Telemat.*, vol. 17, no. 2, pp. 94–100, 2023, doi: 10.61769/telematika.v17i2.542.

[12] S. Melda, "Penerapan *Data Mining* dalam Perancangan Sistem Pendukung Keputusan Seleksi Penerimaan Beasiswa Menggunakan Naive Bayes Classifier (Studi Kasus: IIB Darmajaya)," *J. Tek.*, pp. 165–174, 2020, doi: https://doi.org/10.5281/zenodo.13369413.

[13] M. Andani, J. Triloka, S. Y. Irianto, and H. W. Nugroho, "Performance Comparison of K-Nearest Neighbor , Naive Bayes , and Random Forest Algorithms in Obesity Prediction," vol. 9, no. 1, pp. 502–510, 2025, doi: 10.33395.

[14] A. Shiri, "Introduction to Modern Information Retrieval (2nd edition)," *Libr. Rev.*, vol. 53, no. 9, pp. 462–463, 2004, doi: 10.1108/00242530410565256.

[15] M. M. Muttaqin, Wahyu Wijaya Widiyanto, A. W. Green Ferry Mandias, Stenly Richard Pungus, S. A. H. Wiranti Kusuma Hapsari, E. F. B. Aslam Fatkhudin, Pasnur, and N. S. Mochammad Anshori, Suryani, *Pengenalan Data Mining*, no. July. Yayasan Kita Menulis, 2023.

DOI: 10.54259/satesi.v5i2.5612