

## PENELITIAN ASLI

# PERBANDINGAN ALGORITMA *NAÏVE BAYES* DAN *LONG SHORT-TERM MEMORY* DALAM KLASIFIKASI KUALITAS SUSU

Fajar Rahardika Bahari Putra<sup>1</sup>, Murni<sup>2\*</sup>, Muhammad Surahmanto<sup>3</sup>, Rezki<sup>4</sup>,  
Dimas Adi Suseno<sup>5</sup>

<sup>1,2,3,4,5</sup>Fakultas Teknik, Universitas Muhammadiyah Sorong, Papua Barat Daya, 98416, Indonesia

### Info Artikel

Riwayat Artikel:

Diterima: 21 November 2025

Direvisi : 29 November 2025

Diterima : 08 Desember 2025

Diterbitkan : 20 Desember 2025

**Kata kunci:** Perbandingan Algoritma, Klasifikasi, *Naïve Bayes*, *Long Short-Term Memory*, Kualitas Susu

**Penulis Korespondensi :**

Murni

Email: [murni44@um-sorong.ac.id](mailto:murni44@um-sorong.ac.id)

### Abstrak

**Latarbelakang:** Kualitas susu merupakan aspek penting dalam industri pangan karena secara langsung berkaitan dengan keamanan konsumsi dan kesehatan masyarakat. Pengujian kualitas susu konvensional memiliki keterbatasan dalam efisiensi dan objektivitas, sehingga memerlukan pendekatan berbasis data untuk mendukung proses klasifikasi kualitas susu.

**Tujuan:** Penelitian ini difokuskan pada perbandingan hasil klasifikasi antara algoritma *Naïve Bayes* dan *Long Short-Term Memory (LSTM)* pada data kualitas susu. Dataset dalam penelitian ini berupa data sekunder yang diambil melalui platform *Kaggle*, terdiri dari 844 data dengan tiga kelas kualitas, yaitu *Low*, *Medium*, dan *High*. Semua data digunakan sebagai populasi penelitian, dengan pembagian data pelatihan dan data uji, di mana 212 dataset digunakan sebagai data uji.

**Metode** penelitian yang diterapkan adalah pendekatan eksperimental dengan *simulation-based experiment*, menggunakan dataset yang sama pada kedua algoritma. Penilaian performa model dilakukan dengan memanfaatkan metrik *accuracy*, *Confusion matrix*,

**Hasil:** Berdasarkan hasil pengujian, algoritma *Naïve Bayes* mencapai tingkat *accuracy* sebesar 85%, lebih tinggi dibandingkan *Long Short-Term Memory (LSTM)* yang memperoleh 75%. Analisis per kelas menunjukkan bahwa *Naïve Bayes* lebih efektif dalam mengidentifikasi kelas *Low* dan *Medium*, sementara *LSTM* cenderung menghasilkan nilai *recall* yang lebih baik pada kelas *Medium*, namun menunjukkan penurunan kemampuan klasifikasi pada kelas *Low* dan *High*.

**Kesimpulan:** Berdasarkan hasil tersebut, penelitian ini kemudian menegaskan bahwa pemilihan algoritma klasifikasi perlu disesuaikan dengan karakteristik dataset untuk mendapatkan hasil yang optimal. *precision*, *recall*, dan *F1-score*.



## 1. Pendahuluan

Susu sebagai bahan pangan bergizi tinggi karena mengandung berbagai zat gizi penting, seperti nutrisi esensial, seperti protein, lemak, karbohidrat, mineral, dan vitamin yang dibutuhkan oleh tubuh manusia. Meskipun demikian, susu yang dikonsumsi tanpa memenuhi standar mutu dapat menimbulkan dampak negatif bagi kesehatan, khususnya gangguan pada sistem pencernaan [1]. Dalam industri pangan, kualitas susu memiliki peran penting karena berhubungan langsung dengan keamanan produk dan perlindungan kesehatan konsumen. Oleh karena itu, pengujian mutu susu diperlukan untuk memastikan produk yang beredar telah memenuhi standar kelayakan dan keamanan pangan yang berlaku [2], [3].

Meskipun berbagai metode klasifikasi kualitas susu telah dikembangkan, sebagian besar penelitian masih mengandalkan algoritma konvensional yang memiliki keterbatasan dalam menangani data dengan tingkat kompleksitas dan ketergantungan antar fitur yang tinggi. Kondisi ini berpotensi menurunkan akurasi klasifikasi, yang pada konteks industri pangan dapat berdampak pada kesalahan penilaian mutu dan risiko terhadap keamanan konsumen. Oleh karena itu, diperlukan pendekatan yang mampu menangani karakteristik data yang lebih kompleks tanpa mengorbankan efisiensi komputasi. Dengan kemajuan teknologi, pembelajaran mendalam menawarkan pendekatan yang lebih canggih. Salah satu algoritma tersebut adalah algoritma *Long Short-Term Memory (LSTM)*, yang dikenal karena kemampuannya untuk memproses data sekuensial dan mengingat informasi dalam jangka Panjang [4]. *LSTM* dapat digunakan untuk menganalisis data kualitas susu yang diukur selama periode waktu tertentu, memberikan keuntungan dibandingkan teknik tradisional dalam menangani data yang bergantung pada waktu.

Sebagai perbandingan, *Naïve Bayes* adalah algoritma probabilistik yang sederhana namun efektif untuk mengklasifikasikan data dengan fitur independen [5], [6]. *Naïve Bayes* banyak digunakan untuk klasifikasi data yang relatif sederhana dan menawarkan keunggulan dalam kecepatan dan kemudahan implementasi, meskipun kemampuannya terbatas pada data dengan ketergantungan fitur yang rendah [7], [8]. Kedua algoritma ini menawarkan pendekatan yang berbeda, dan pilihan yang tepat dapat secara signifikan memengaruhi hasil klasifikasi kualitas susu.

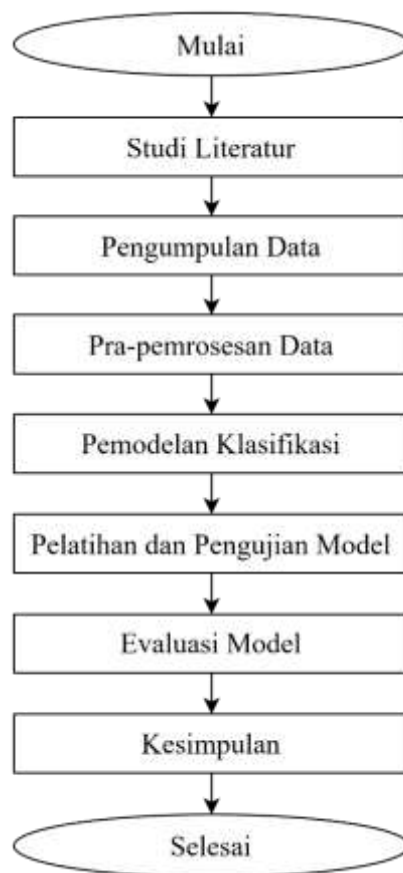
Berbagai penelitian terdahulu yang dilakukan oleh Ni Kadek dkk., menganalisis sentimen pelayanan daring di Fakultas Teknik dan Kejuruan Undiksha menggunakan metode Naïve Bayes dan LSTM, dengan hasil menunjukkan bahwa Naïve Bayes lebih unggul (akurasi 83,69%) dibandingkan LSTM (53,12%), namun penelitian ini terbatas pada konteks pelayanan kampus dan dataset yang relatif kecil [9]. Sementara itu, penelitian Annisa dkk., menerapkan metode Naïve Bayes untuk klasifikasi performa siswa dengan hasil akurasi 83% dan F1-score 89%, tetapi penelitian ini hanya berfokus pada domain pendidikan dan menggunakan satu algoritma saja [10].

Sementara itu, algoritma *Naïve Bayes* banyak digunakan dalam penelitian klasifikasi karena memiliki struktur yang sederhana dan efisiensi komputasi yang baik [11]. Fauzi dkk., melaporkan bahwa *Naïve Bayes* mampu memberikan hasil klasifikasi yang cukup optimal pada data kualitas susu dengan tingkat kompleksitas rendah, khususnya ketika antar fitur memiliki asumsi independensi yang kuat [12]. Akan tetapi, performa algoritma ini cenderung menurun ketika diterapkan pada dataset dengan ukuran yang lebih besar dan pola hubungan antar atribut yang lebih kompleks, sehingga membatasi efektivitasnya pada data yang bersifat dinamis dan heterogen [13].

Berdasarkan hasil kajian terhadap penelitian-penelitian sebelumnya, dapat disimpulkan bahwa sebagian besar studi masih berfokus pada penggunaan algoritma secara terpisah atau diterapkan pada jenis permasalahan yang berbeda, seperti prediksi deret waktu maupun klasifikasi tunggal. Penelitian yang secara khusus membandingkan algoritma klasifikasi sederhana, seperti *Naïve Bayes*, dengan algoritma berbasis deep learning, seperti *Long Short-Term Memory (LSTM)*, dalam konteks klasifikasi kualitas susu menggunakan dataset yang sama masih relatif terbatas. Oleh karena itu, kebaruan penelitian ini terletak pada penyajian analisis perbandingan kinerja antara *Naïve Bayes* dan *LSTM* dalam mengklasifikasikan kualitas susu, guna mengevaluasi kemampuan masing-masing algoritma dalam menangani kompleksitas data serta menentukan metode yang paling efektif untuk permasalahan tersebut.

## 2. Metode

Penelitian ini menerapkan desain eksperimen berbasis simulasi untuk mengevaluasi dan membandingkan kinerja dua algoritma klasifikasi, yaitu *Naïve Bayes* dan *Long Short-Term Memory (LSTM)*, dalam proses pengklasifikasian kualitas susu. Pemilihan pendekatan ini didasarkan pada kebutuhan untuk menguji performa model secara terkontrol melalui skema eksperimen yang sistematis, sebagaimana dirangkum dalam alur penelitian pada Gambar 1.



Gambar 1 Alur Penelitian

### 2.1 Studi literatur

Tahap awal penelitian diawali dengan studi literatur yang bertujuan untuk memperoleh landasan teoritis dan pemahaman konseptual yang relevan dengan topik penelitian [14]. Studi literatur dilakukan dengan mengkaji berbagai sumber acuan, meliputi publikasi ilmiah dalam jurnal, buku referensi, serta artikel-artikel penelitian terdahulu yang berkaitan dengan klasifikasi data, algoritma *Naïve Bayes*, *Long Short-Term Memory (LSTM)*, dan penelitian mengenai klasifikasi kualitas susu.

### 2.2 Pengumpulan data

Pada tahap ini, dataset kualitas susu dikumpulkan dari platform Kaggle sebagai sumber data sekunder. Dataset tersebut terdiri dari 844 item, dibagi menjadi tiga kelas kualitas, yaitu *High*, *Medium*, dan *Low*. Atribut yang digunakan dalam dataset meliputi nilai *pH*, *temperature*, *taste*, *odor*, *fat*, *turbidity*, *colour*, serta satu atribut kelas yaitu *quality*.

### 2.3 Pra-pemrosesan Data

Dataset kualitas susu awalnya melewati tahap pra-pemrosesan untuk memastikan data tetap konsisten dan siap pakai sebelum diterapkan pada algoritma *Naïve Bayes* serta *Long Short-Term Memory (LSTM)*. Dalam langkah ini, dilakukan penilaian awal mengenai kelengkapan dan keterkaitan fitur, plus penyesuaian jenis data untuk membedakan antara variabel numerik dan

kategorikal sebagai masukan model. Fitur numerik seperti *pH*, *temperature*, dan *colour* dinormalisasi agar rentang nilainya seragam dan menghindari penguasaan oleh fitur tertentu, terutama pada model *LSTM* yang peka terhadap variasi skala data. Seluruh tahap pra-pemrosesan dilakukan tanpa memodifikasi nilai inti atau sebaran kategori dalam dataset, sehingga sifat asli data tetap terjaga. Setelah itu, dataset yang sudah dipersiapkan digunakan dalam proses pemisahan data dan pembentukan model klasifikasi dengan algoritma *Naïve Bayes* dan *LSTM*.

#### 2.4 Pemodelan Klasifikasi

Tahap klasifikasi bertujuan menentukan kategori kualitas susu (*Low*, *Medium*, dan *High*) berdasarkan data uji yang telah melalui prapemrosesan. Data tersebut diproses menggunakan model yang telah dilatih sebelumnya, dengan menerapkan dua pendekatan klasifikasi, yaitu *Naïve Bayes* dan *Long Short-Term Memory (LSTM)*, menggunakan dataset dan skema pengujian yang sama agar hasilnya dapat dibandingkan secara objektif. Dalam algoritma *Naïve Bayes*, pengklasifikasian dilakukan melalui perhitungan probabilitas posterior untuk masing-masing kategori dengan mengasumsikan bahwa setiap fitur bersifat independen satu sama lain, lalu memilih kategori dengan probabilitas paling tinggi sebagai hasil ramalan. Di sisi lain, *LSTM* mengolah data sebagai rangkaian fitur untuk menangkap hubungan antar atribut, dengan menggunakan sistem memory cell untuk mendeteksi pola data yang lebih rumit, sehingga penentuan klasifikasi bergantung pada pemahaman mendalam terhadap data secara keseluruhan.

#### 2.5 Pelatihan dan Pengujian Model

Dataset kualitas susu yang telah melalui tahap pra-pemrosesan dibagi menjadi subset untuk pelatihan dan pengujian dengan rasio tertentu. Subset pelatihan digunakan untuk membangun dan melatih model klasifikasi, sedangkan subset pengujian dimanfaatkan untuk mengevaluasi kemampuan model dalam menggeneralisasi pada data baru yang belum pernah dilatih. Pembagian ini diterapkan secara konsisten pada kedua algoritma agar perbandingan kinerjanya bersifat adil dan objektif. Proses pelatihan dilakukan secara terpisah menggunakan algoritma *Naïve Bayes* dan *Long Short-Term Memory (LSTM)*. *Naïve Bayes* memproses data dalam bentuk tabel dengan asumsi independensi antar atribut, sedangkan *LSTM* merepresentasikan data dalam bentuk sekuensial sesuai dengan karakteristik jaringan saraf berulang. Fase pengujian dilakukan dengan membandingkan hasil prediksi model dengan label kelas aktual untuk mengukur kinerja klasifikasi pada kelas Rendah, Sedang, dan Tinggi, sehingga perbandingan kinerja antar model dapat dibenarkan secara ilmiah.

#### 2.6 Evaluasi Model

Dalam menjalankan evaluasi model untuk mengevaluasi kemampuan *Naïve Bayes* dan *Long Short-Term Memory (LSTM)*, perlu dilakukan klasifikasi kualitas susu ke dalam tiga kelas: *Low*, *Medium*, dan *High*. Evaluasi ini lebih menekankan akurasi dari hasil prediksi dan bagaimana model bisa membedakan antar kelas berdasarkan data uji, dengan *Confusion matrix*

sebagai alat pokok untuk menganalisis performa klasifikasi [15]. *Confusion matrix* menyajikan ringkasan jumlah prediksi yang tepat maupun keliru pada masing-masing kelas, sehingga pola kesalahan klasifikasi dapat dianalisis secara lebih rinci. Selain itu, penilaian kinerja model juga dilengkapi dengan classification report yang mencakup metrik *precision*, *recall*, dan *F1-score*. Nilai *precision* mencerminkan tingkat ketepatan prediksi suatu kelas, *recall* menunjukkan kemampuan model dalam mengidentifikasi seluruh data pada kelas tersebut, sedangkan *F1-score* merepresentasikan keseimbangan antara *precision* dan *recall* dalam mengevaluasi kinerja klasifikasi secara keseluruhan.

## 2.7 Kesimpulan

Tahapan terakhir penelitian difokuskan pada penyusunan kesimpulan yang didasarkan pada hasil evaluasi kinerja model klasifikasi. Kesimpulan itu didapat dari analisis perbandingan antara *Naïve Bayes* dan *Long Short-Term Memory (LSTM)* menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Analisis tersebut bertujuan menilai perbedaan kemampuan kedua algoritma dalam mengklasifikasikan kualitas susu pada setiap kelas yang diuji. Hasil evaluasi selanjutnya dimanfaatkan untuk menentukan algoritma yang menunjukkan kinerja paling sesuai dengan karakteristik dataset yang digunakan, sekaligus mengidentifikasi keunggulan dan keterbatasan masing-masing pendekatan. Kesimpulan yang dihasilkan menjadi dasar dalam menjawab tujuan penelitian serta memberikan rekomendasi terkait pemilihan metode klasifikasi dan arah pengembangan penelitian selanjutnya.

## 3. Hasil

Bagian ini menyajikan hasil penelitian dan menyatakan temuan-temuan yang signifikan, bukannya memberikan data dengan sangat rinci. Hasil penelitian harus jelas dan ringkas. Hasil penelitian dapat dilaporkan dalam bentuk teks dan dilengkapi dengan tabel, grafik (gambar), dan bagan. Harap perkenalkan informasi dalam tabel, grafik (gambar), dan bagan.

### 3.1 Persiapan Dataset

Persiapan dataset dilakukan dengan memastikan seluruh data kualitas susu berada dalam kondisi lengkap dan layak digunakan sebagai masukan bagi algoritma *Naïve Bayes* maupun *Long Short-Term Memory (LSTM)*. Dataset terdiri atas 844 sampel tanpa nilai kosong (missing value), sehingga seluruh data dapat dimanfaatkan secara penuh pada tahap pelatihan maupun pengujian model. Atribut numerik, yaitu pH, Temperature, dan Colour, terlebih dahulu melalui proses normalisasi untuk menyeragamkan skala data, khususnya guna mendukung kestabilan proses pembelajaran pada model *LSTM*. Sementara itu, atribut biner seperti Taste, Odor, Fat, dan Turbidity dipertahankan dalam bentuk aslinya karena telah merepresentasikan nilai kategorikal secara langsung. Kolom Quality berfungsi sebagai label target dengan tiga kategori, yaitu *High*, *Medium*, dan *Low*, yang menjadi acuan utama dalam proses klasifikasi. Pada penerapan model *Naïve Bayes*, data diproses dalam format tabular dengan asumsi bahwa setiap atribut bersifat independen dalam menentukan kelas kualitas susu. Sebaliknya, pada model *LSTM*, data direpresentasikan dalam bentuk urutan fitur yang disesuaikan dengan

karakteristik arsitektur jaringan, tanpa melakukan perubahan terhadap jumlah maupun nilai data yang digunakan. Selanjutnya, dataset dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan pendekatan stratified sampling untuk menjaga keseimbangan proporsi kelas. Dengan demikian, hasil klasifikasi yang dihasilkan oleh kedua model dapat dibandingkan secara objektif, konsisten, dan dapat dipertanggungjawabkan secara ilmiah.

Tabel 1 Dataset Kualitas Susu

No	pH	Temperature	Taste	Odor	Fat	Turbidity	Colour	Quality
1	6.6	35	1	1	0	0	254	High
2	6.8	34	1	0	1	1	253	Medium
3	6.5	36	0	1	0	1	255	Low
4	6.7	33	1	1	1	0	252	High
5	6.4	37	0	0	0	1	256	Medium

### 3.2 Pra-Pemrosesan Data

Sebelum digunakan dalam proses pemodelan, dataset kualitas susu dipersiapkan melalui tahap preprocessing atau prapemrosesan guna menjamin konsistensi data dan keselarasan dengan kebutuhan algoritma *Naïve Bayes* serta *Long Short-Term Memory (LSTM)*. Hasil pemeriksaan awal menunjukkan tidak adanya missing value, sehingga seluruh data dapat digunakan dalam proses pelatihan dan pengujian. Atribut numerik seperti *pH*, *temperature*, dan *colour* dinormalisasi untuk menyamakan rentang nilai, sedangkan atribut biner seperti *taste*, *odor*, *fat*, dan *turbidity* dipertahankan dalam bentuk aslinya karena telah berada pada skala yang sesuai. Pada pemodelan *Naïve Bayes*, data digunakan dalam bentuk tabular sesuai struktur atribut awal. Sebaliknya, pada pemodelan *LSTM*, dilakukan penyesuaian struktur data dengan menyusun atribut ke dalam format sekuens fitur menggunakan penamaan generik. Penyesuaian ini bersifat struktural tanpa mengubah nilai maupun jumlah data, sehingga perbandingan hasil klasifikasi antar algoritma tetap objektif dan dapat dipertanggungjawabkan secara ilmiah.

#### a. Pra-Pemrosesan untuk *Naïve Bayes*

Pada penerapan model *Naïve Bayes*, dataset kualitas susu dipertahankan dalam format tabular sesuai struktur awal. Atribut numerik seperti *pH*, suhu, dan warna diperlakukan sebagai nilai kontinu, sementara atribut rasa, bau, lemak, dan kekeruhan direpresentasikan dalam bentuk data biner bernilai 0 dan 1. Hasil pemeriksaan menunjukkan tidak terdapat data kosong, sehingga tahap imputasi tidak diperlukan. Label kualitas susu yang mencakup *Low*, *Medium*, dan *High* selanjutnya dikodekan dalam bentuk numerik. Dataset yang telah diproses kemudian dibagi dengan proporsi 80% sebagai data latih dan 20% sebagai data uji, sehingga vektor fitur numerik yang dihasilkan siap digunakan dalam proses pelatihan dan pengujian model *Naïve Bayes*.

```

from sklearn.naive_bayes import GaussianNB

nb = GaussianNB()
nb.fit(X_nb_train, y_nb_train)

```

▼ GaussianNB ⓘ ?

GaussianNB()

Gambar 2 Kode Pra-Pemrosesan untuk Naïve Bayes

b. Pra-Pemrosesan untuk *LSTM*

Dalam model *LSTM*, data kualitas susu disusun dalam bentuk vektor fitur yang direpresentasikan sebagai urutan fitur. Atribut numerik seperti *pH*, suhu, dan warna dinormalisasi ke dalam rentang 0 hingga 1 agar sesuai dengan kebutuhan jaringan saraf. Sementara itu, atribut biner seperti rasa, aroma, lemak, dan kekeruhan digunakan dalam bentuk aslinya tanpa perubahan nilai. Label kelas *Low*, *Medium*, dan *High* dikonversi menjadi bentuk numerik dan *one-hot encoding* agar bisa diproses oleh *LSTM*. Setelah itu, dataset dibagi menjadi 80% untuk data latih dan 20% untuk data uji. Dengan langkah-langkah ini, data sudah siap digunakan sebagai input model *LSTM* dalam proses pelatihan dan pengujian.

```

from sklearn.preprocessing import MinMaxScaler, LabelEncoder
from sklearn.model_selection import train_test_split

# ===== LOAD DATASET LSTM =====
df_lstm = pd.read_csv("kualitas_susu_lstm.csv")

# ===== LABEL ENCODING (MAJIE) =====
le_lstm = LabelEncoder()
df_lstm["Quality"] = le_lstm.fit_transform(df_lstm["Quality"])

# ===== SPLIT FITUR & LABEL =====
X_lstm = df_lstm.drop("Quality", axis=1).values
y_lstm = df_lstm["Quality"].values

# ===== SCALING =====
scaler = MinMaxScaler()
X_lstm = scaler.fit_transform(X_lstm)

# ===== RESHAPE KE LSTM =====
X_lstm = X_lstm.reshape(X_lstm.shape[0], X_lstm.shape[1], 1)

# ===== TRAIN TEST SPLIT =====
X_train_lstm, X_test_lstm, y_train_lstm, y_test_lstm = train_test_split(
    X_lstm,
    y_lstm,
    test_size=0.2,
    random_state=42,
    stratify=y_lstm
)

```

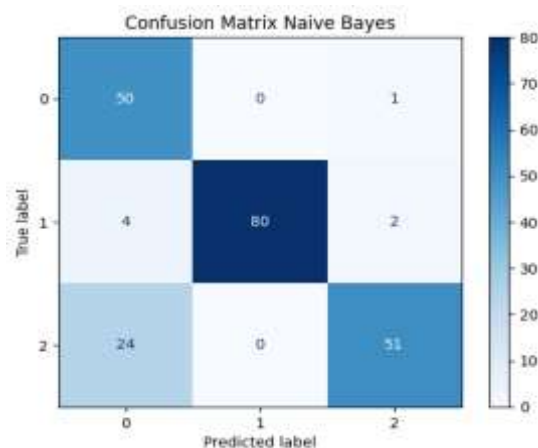
Gambar 3 Kode Pra-Pemrosesan untuk LSTM



### 3.3 Hasil Klasifikasi

#### a. Hasil Klasifikasi Algoritma Naive Bayes

Penilaian kinerja model *Naïve Bayes* dilaksanakan dengan menggunakan *Confusion matrix* dan *classification report* sebagai metode evaluasi. Gambar 4 menyajikan *Confusion matrix* yakni *output* dari proses klasifikasi menggunakan algoritma *Naïve Bayes*, yang menggambarkan sebaran prediksi yang tepat maupun keliru pada setiap kelas kualitas susu.

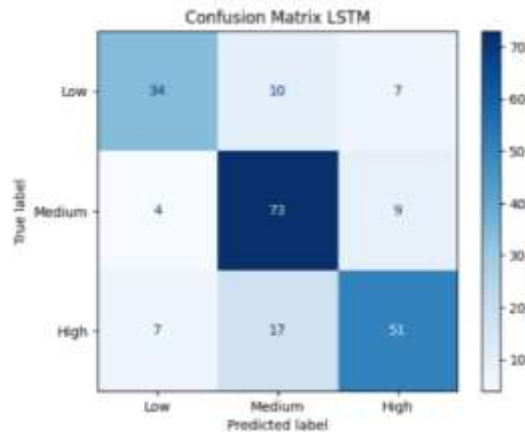


Gambar 4 *Confusion matrix Naïve Bayes*

*Confusion matrix* pada model *Naïve Bayes* menunjukkan bagaimana algoritma ini melakukan klasifikasi terhadap kualitas susu dalam tiga kategori, yaitu *Low* (0), *Medium* (1), dan *High* (2). Dari hasil pengujian, terlihat bahwa model mampu mengenali kelas *Medium* dengan sangat baik, ditunjukkan oleh 80 prediksi benar dari total 86 sampel, meskipun masih terdapat 4 sampel yang salah diklasifikasikan sebagai *Low* dan 2 sampel sebagai *High*. Untuk kelas *Low*, sebanyak 50 sampel berhasil diklasifikasikan dengan benar, namun terdapat 1 sampel yang salah masuk ke kelas *High*. Sementara itu, pada kelas *High*, model berhasil mengklasifikasikan 51 sampel dengan benar, tetapi masih terdapat kesalahan cukup besar, yaitu 24 sampel yang salah diklasifikasikan sebagai *Low*. Temuan ini mengindikasikan bahwa *Naïve Bayes* memiliki performa yang kuat dalam membedakan kelas *Medium*, namun kurang optimal dalam memisahkan kelas *Low* dan *High*, kemungkinan karena adanya kemiripan pola atribut di antara kedua kelas tersebut. Secara keseluruhan, hasil ini memberikan gambaran bahwa meskipun *Naïve Bayes* efektif dalam menangani sebagian besar data, terdapat keterbatasan dalam akurasi klasifikasi pada kelas yang memiliki karakteristik mirip.

#### b. Hasil Klasifikasi Algoritma LSTM

Evaluasi kinerja model *Long Short-Term Memory (LSTM)* dilakukan menggunakan pendekatan yang sama, yaitu *Confusion matrix* dan *classification report*. Gambar 5 menunjukkan *Confusion matrix* hasil klasifikasi kualitas susu menggunakan algoritma *LSTM*.



Gambar 5 *Confusion matrix LSTM*

*Confusion matrix* pada model *Long Short-Term Memory (LSTM)* menunjukkan performa klasifikasi kualitas susu dalam tiga kategori: *Low*, *Medium*, dan *High*. Model berhasil mengklasifikasikan 34 sampel kelas *Low* dengan benar, namun masih terdapat 10 sampel yang salah diklasifikasikan sebagai *Medium* dan 7 sebagai *High*. Untuk kelas *Medium*, sebanyak 73 sampel diklasifikasikan dengan tepat, sementara 4 sampel salah sebagai *Low* dan 9 sebagai *High*. Pada kelas *High*, model mengidentifikasi 51 sampel dengan benar, namun masih terjadi kesalahan klasifikasi sebanyak 7 sampel ke kelas *Low* dan 17 ke kelas *Medium*. Secara keseluruhan, *LSTM* menunjukkan performa yang cukup baik dan lebih seimbang dibandingkan *Naïve Bayes*, terutama dalam mengenali pola sekuensial antar fitur, meskipun masih terdapat tantangan dalam membedakan kelas yang memiliki karakteristik mirip seperti *Medium* dan *High*.

### 3.4 Evaluasi

Kemudian dilakukan evaluasi menggunakan metrik precision, recall, f1-score, dan accuracy yang dapat dilihat pada gambar 6.

--- Naive Bayes ---				
	precision	recall	f1-score	support
0	0.64	0.98	0.78	51
1	1.00	0.93	0.96	86
2	0.94	0.68	0.79	75
accuracy			0.85	212
macro avg	0.86	0.86	0.84	212
weighted avg	0.89	0.85	0.86	212
--- LSTM ---				
	precision	recall	f1-score	support
0	0.76	0.67	0.71	51
1	0.73	0.85	0.78	86
2	0.76	0.68	0.72	75
accuracy			0.75	212
macro avg	0.75	0.73	0.74	212
weighted avg	0.75	0.75	0.74	212

Gambar 6 Hasil Akurasi Model Naive Bayes dan LSTM

Gambar 6 diatas menunjukkan hasil evaluasi perbandingan metrik performa antara model *Naïve Bayes* dan *Long Short-Term Memory (LSTM)* dalam klasifikasi kualitas susu. Pada model *Naïve Bayes*, akurasi keseluruhan

mencapai 85%, dengan nilai precision tertinggi pada kelas *Medium* (1.00) dan f1-score sebesar 0.96, menunjukkan kemampuan yang sangat baik dalam mengenali kelas tersebut. Namun, recall pada kelas *High* hanya 0.68, menandakan adanya kesulitan dalam mendeteksi seluruh sampel yang benar-benar termasuk kategori *High*. Sebaliknya, model *LSTM* menghasilkan akurasi keseluruhan sebesar 75%, dengan nilai precision, recall, dan f1-score yang relatif seimbang di semua kelas (sekitar 0.70–0.76). Hal ini menunjukkan bahwa meskipun *LSTM* memiliki akurasi lebih rendah dibandingkan *Naïve Bayes*, distribusi performanya lebih merata antar kelas. Dengan demikian, gambar evaluasi ini menegaskan bahwa *Naïve Bayes* unggul dalam akurasi khususnya pada kelas dominan, sedangkan *LSTM* lebih konsisten dalam menjaga keseimbangan prediksi antar kelas, meskipun dengan tingkat akurasi yang lebih rendah.

#### 4. Pembahasan

Penelitian ini berhasil membandingkan dua algoritma klasifikasi, yaitu *Naïve Bayes* dan *Long Short-Term Memory (LSTM)*, dalam menentukan kualitas susu berdasarkan atribut numerik dan biner yang tersedia. Hasil penelitian ini secara langsung berkaitan dengan tujuan awal yang diuraikan pada bagian latar belakang, yaitu mencari metode klasifikasi yang paling sesuai untuk dataset kualitas susu. Dengan capaian akurasi sebesar 85% pada *Naïve Bayes* dan 75% pada *LSTM*, penelitian ini menunjukkan bahwa model sederhana berbasis probabilistik justru lebih efektif dibandingkan model berbasis jaringan saraf yang kompleks, terutama pada data tabular tanpa pola temporal yang dominan.

Secara ilmiah, temuan ini dapat dijelaskan melalui karakteristik masing-masing algoritma. *Naïve Bayes* bekerja dengan asumsi independensi antar atribut, sehingga mampu memanfaatkan distribusi probabilitas sederhana untuk menghasilkan klasifikasi yang stabil. Hal ini terbukti efektif karena atribut dalam dataset kualitas susu relatif berdiri sendiri dan tidak membentuk urutan sekuensial. Sebaliknya, *LSTM* dirancang untuk menangani data berurutan dengan pola temporal yang kuat, sehingga penerapannya pada dataset tabular menyebabkan performa yang kurang optimal. Interpretasi ini menegaskan bahwa kompleksitas arsitektur model tidak selalu menjamin peningkatan kinerja, melainkan harus disesuaikan dengan karakteristik data yang digunakan.

Jika dibandingkan dengan penelitian lain, hasil ini konsisten dengan temuan beberapa studi yang melaporkan bahwa algoritma sederhana seperti *Naïve Bayes* atau Decision Tree sering kali lebih unggul pada dataset tabular dengan jumlah fitur terbatas. Namun, terdapat juga penelitian yang menunjukkan keunggulan *LSTM* pada data dengan dimensi temporal, seperti analisis sensor waktu nyata atau data deret waktu. Perbedaan ini menegaskan bahwa pemilihan algoritma harus mempertimbangkan sifat data, bukan hanya kompleksitas model.

Dengan demikian, penelitian ini memberikan kontribusi penting dalam

memperjelas bahwa untuk klasifikasi kualitas susu dengan dataset tabular, *Naïve Bayes* merupakan pilihan yang lebih sesuai dibandingkan *LSTM*. Namun, *LSTM* tetap memiliki potensi besar apabila diterapkan pada dataset dengan pola temporal yang jelas atau kompleksitas yang lebih tinggi.

## 5. Kesimpulan

Berdasarkan hasil penelitian, diperoleh bahwa algoritma *Naïve Bayes* menunjukkan performa klasifikasi yang lebih baik dengan capaian akurasi sebesar 85%, sedangkan model *Long Short-Term Memory (LSTM)* menghasilkan akurasi sebesar 75%. Jika ditinjau lebih lanjut pada tingkat kelas, *Naïve Bayes* mampu mengidentifikasi kelas *Low* dan *Medium* secara lebih optimal, yang tercermin dari tingginya nilai recall pada kelas *Low* serta nilai precision yang sangat baik pada kelas *Medium*. Di sisi lain, *LSTM* memperlihatkan performa yang relatif konsisten pada kelas *Medium*, namun masih menghadapi kendala dalam mengenali kelas *Low* dan *High* secara akurat. Temuan ini menegaskan bahwa peningkatan kompleksitas arsitektur model tidak selalu berbanding lurus dengan peningkatan kinerja klasifikasi. Meskipun *LSTM* dirancang untuk menangani data yang bersifat kompleks dan berurutan, karakteristik dataset kualitas susu yang berbentuk tabular serta tidak memiliki pola temporal yang dominan menyebabkan keunggulan model tersebut tidak termanfaatkan secara maksimal. Sebaliknya, pendekatan *Naïve Bayes* yang lebih sederhana justru mampu memberikan hasil klasifikasi yang lebih stabil, efisien, dan sesuai dengan karakteristik data yang digunakan. Berdasarkan keseluruhan analisis, dapat disimpulkan bahwa *Naïve Bayes* merupakan metode yang lebih tepat untuk klasifikasi kualitas susu pada dataset dengan karakteristik seperti dalam penelitian ini. Adapun *LSTM* tetap memiliki potensi untuk menghasilkan performa yang lebih optimal apabila diterapkan pada data dengan tingkat kompleksitas lebih tinggi atau pola temporal yang jelas. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset dengan skala dan kompleksitas yang lebih besar, serta mengkaji penerapan strategi optimasi model guna meningkatkan akurasi dan konsistensi hasil klasifikasi.

## 6. Ucapan Terimakasih

Penulis menyampaikan penghargaan dan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah berkontribusi dalam mendukung terlaksananya penelitian ini. Ucapan terima kasih khusus ditujukan kepada penyedia dataset, institusi terkait, serta rekan-rekan sejawat yang senantiasa memberikan masukan berharga dan dukungan penuh selama proses penelitian hingga penyusunan artikel ini selesai. Dukungan, saran, dan kerja sama yang diberikan telah menjadi bagian penting dalam keberhasilan penelitian ini.

## 7. Referensi

- [1] A. H. Nevia Wiranti, Veronika Wanniatie, “Kualitas Susu Sapi Segar Pada Pemerahan Pagi Dan Sore Quality of Fresh Cow’s Milk at Morning and Afternoon Milking,” *J. Ris. dan Inov. Peternak.*, vol. 6, no. 2, pp. 123–128, 2022, doi: 10.23960/jrip.2022.6.2.123-128.
- [2] N. S. Anindita and D. S. Soyi, “Studi kasus Pengawasan Kualitas Pangan Hewani

- melalui Pengujian Kualitas Susu Sapi yang Beredar di Kota Yogyakarta,” *J. Peternak. Indones.*, vol. 19, no. 2, pp. 96–105, 2017, doi: <https://doi.org/10.25077/jpi.19.2.93-102.2017>.
- [3] P. Adi, R. Mulyani, and L. N. Khabibah, “Kajian keamanan pangan pada industri pengolahan susu di jawa tengah dengan menggunakan metode good manufacturing practices (gmp) food safety evaluation in the dairy processing industry in central java using good manufacturing practices (gmp) methods,” *J. Teknol. Ind. Pertan.*, vol. 33, no. 3, pp. 305–316, 2023, doi: <https://doi.org/10.24961/j.tek.ind.pert.2023.33.3.305>.
  - [4] A. Azrul, A. I. Purnamasari, and I. Ali, “Analisis Sentimen Pengguna Twitter Terhadap Perkembangan Artificial Intelligence dengan Penerapan Algoritma Long Short-Term Memory (LSTM),” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 413–421, 2024, doi: <https://doi.org/10.36040/jati.v8i1.8416>.
  - [5] Andika, Syarli, and C. R. Sari, “Data Mining Klasifikasi Kelulusan Mahasiswa Menggunakan Metode Naïve Bayes,” *J. PEQGURUANG*, vol. 4, no. 1, pp. 423–428, 2022, doi: <http://dx.doi.org/10.35329/jp.v4i1.2358>.
  - [6] M. Z. Haq, C. S. Octiva, Ayuliana, U. W. Nuryanto, and D. Suryadi, “Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial,” *J. Minfo Polgan*, vol. 13, no. 1, pp. 1079–1084, 2024, doi: <https://doi.org/10.33395/jmp.v13i1.13937>.
  - [7] Sarman and D. S. Bahri, “Analisis Perbandingan Algoritma Naïve Bayes dengan SVM (Support Vektor Machine) Dalam Mendiagnosis Penyakit Asma,” *J. Janitra Inform. dan Sist. Inf.*, vol. 5, no. 2, pp. 189–198, 2025, doi: [10.59395/m9qa0357](https://doi.org/10.59395/m9qa0357).
  - [8] P. Ramadani, R. Fadillah, Q. Adawiyah, Suerni, and B. R. Al Ghazali, “Perbandingan Algoritma Naïve Bayes , C4 . 5 , dan K-Nearest Neighbor untuk Klasifikasi Kelayakan Program Keluarga Harapan Jurnal Media Informatika [ JUMIN ],” *J. MEDIA Inform. [JUMIN]*, vol. 6, no. 1, pp. 775–782, 2024, doi: <https://doi.org/10.55338/jumin.v6i1.6083>.
  - [9] N. K. T. A. Saputri, I. G. A. Gunadi, and I. M. G. Sunarya, “Perbandingan Metode Analisis Sentimen Pelayanan Daring di Fakultas Teknik dan Kejuruan Universitas Pendidikan Ganesha Menggunakan Algoritma Naïve Bayes dan LSTM,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 1120–1129, 2024, doi: <https://doi.org/10.57152/malcom.v4i3.1336>.
  - [10] A. Saninah, W. Prihartono, and C. L. Rohmat, “Analisis Sentimen Ulasan Pengguna Terhadap Aplikasi Pembelajaran Berbahasa Duolingo Menggunakan Algoritma Naïve Baiyes Classifier,” *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 1, pp. 619–628, 2025, doi: <https://doi.org/10.23960/jitet.v13i1.5691>.
  - [11] G. Dwilestari and T. A. Afifah, “Perbandingan Kinerja Algoritma Naive Bayes Dan Decision Tree Dalam Klasifikasi Kanker Paru-Paru,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 801–807, 2025, doi: <https://doi.org/10.36040/jati.v9i1.12463>.
  - [12] F. W. Ali, I. W. Sumarjaya, and E. N. Kencana, “Seleksi Fitur Information Gain Untuk Klasifikasi Kualitas Susu Sapi Menggunakan Metode K- Nearest Neighbor Dan Naïve Bayes,” *Innov. J. Soc. Sci. Res. Vol.*, vol. 5, no. 1, pp. 422–435, 2025, doi: <https://doi.org/10.31004/innovative.v5i1.17477>.
  - [13] A. P. Wijaya, “Perbandingan Algoritma Klasifikasi Random Foresst dengan Naïve Bayes Classifier pada Studi Penyakit Berdasarkan Pola Nutrisi,” *Remik*

- Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 9, no. 1, pp. 429–438, 2025, doi: <http://doi.org/10.33395/remik.v9i1.14652>.
- [14] M. Waruwu, S. N. Pu`at, P. R. Utami, E. Yanti, and M. Rusydiana, “Metode Penelitian Kuantitatif: Konsep, Jenis, Tahapan dan Kelebihan,” *J. Ilm. Profesi Pendidik.*, vol. 10, no. 1, pp. 917–932, 2025, doi: [10.29303/jipp.v10i1.3057](https://doi.org/10.29303/jipp.v10i1.3057).
- [15] S. Arti and E. Suherlan, “Evaluasi Kinerja Machine Learning dalam Memprediksi Kemampuan Adaptasi Mahasiswa pada Lingkungan Pembelajaran Daring,” *J. Pustaka AI*, vol. 5, no. 1, pp. 50–57, 2025, doi: <https://doi.org/10.55382/jurnal.pustakaai.v5i1.901>.