

## PERBANDINGAN PERFORMANSI MODEL NAÏVE BAYES DALAM MENGANALISIS SENTIMEN MENGENAI E-SPORT MOBILE LEGENDS

Muhammad Aqil Arkanudin<sup>1\*</sup>, Irving Vitra Paputungan<sup>2</sup>

<sup>1,2</sup>Universitas Islam Indonesia, Indonesia

\*Corresponding author: [muhammad.arkanudin@students.uii.ac.id](mailto:muhammad.arkanudin@students.uii.ac.id)

---

**Abstract:** Sentiment analysis is a natural language processing method used to identify and classify opinions or emotions in a text. This study compares three Naïve Bayes models- Multinomial Naïve Bayes, Gaussian Naïve Bayes, and Bernoulli Naïve Bayes-to analyze sentiment in live chat comments from the E-Sports Mobile Legends tournament on YouTube. The data was collected from the live chat section of the MSC 2023 Grand Final broadcast, followed by preprocessing techniques such as tokenization, stopword removal, stemming, and TF-IDF transformation. The three Naïve Bayes models were applied to the processed dataset, and their performance was evaluated using accuracy, precision, recall, and F1-score metrics. The results show that Bernoulli Naïve Bayes achieved the best performance, with the highest accuracy and F1-score compared to the other models. This indicates that the Bernoulli approach is more effective in handling the characteristics of live chat comments, which are often short and contain informal terms. These findings provide a useful reference for applying sentiment analysis to unstructured text data in the E-Sports industry.

**Keywords:** Sentiment Analysis, Naïve Bayes, Machine Learning, Esports, Mobile Legends

**Abstrak:** Analisis sentimen merupakan salah satu metode dalam pemrosesan bahasa alami yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau emosi dalam suatu teks. Dalam penelitian ini, dilakukan perbandingan tiga model Naïve Bayes- Multinomial Naïve Bayes, Gaussian Naïve Bayes, dan Bernoulli Naïve Bayes-untuk menganalisis sentimen komentar live chat turnamen E-Sports Mobile Legends di YouTube. Data dikumpulkan dari kolom live chat siaran grand final MSC 2023, kemudian dilakukan preprocessing dengan teknik tokenisasi, penghapusan stopwords, stemming, dan transformasi TF-IDF. Ketiga model Naïve Bayes diterapkan pada dataset yang telah diproses, dan kinerjanya dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa Bernoulli Naïve Bayes memberikan performa terbaik dengan akurasi dan F1-score tertinggi dibandingkan dengan model lainnya. Hal ini menunjukkan bahwa pendekatan Bernoulli lebih optimal dalam menangani karakteristik teks komentar live chat yang cenderung pendek dan mengandung banyak istilah informal. Temuan ini dapat menjadi acuan dalam penerapan analisis sentimen untuk data teks tidak terstruktur dalam industri E-Sports.

**Kata kunci:** Analisis Sentimen, Naïve Bayes, Machine Learning, Esports, Mobile Legends

---

Copyright (c) 2025 The Authors. This is an open-access article under the CC BY-SA 4.0 license (<https://creativecommons.org/licenses/by-sa/4.0/>)

---

### PENDAHULUAN

Perkembangan teknologi dan popularitas media sosial telah menciptakan ruang bagi masyarakat untuk mengekspresikan pendapat dan pandangan mereka terhadap berbagai

topik termasuk esport. Salah satu game yang memiliki komunitas besar di dunia esport adalah game Mobile Legends yang menjadi subjek diskusi publik yang intens baik dalam bentuk dukungan maupun kritik (Putra et al., 2023). Dukungan dari para penggemar sering kali terlihat dalam bentuk pujian atas inovasi game, sedangkan kritik biasanya berfokus pada aspek gameplay atau kebijakan pengembang. Oleh karena itu analisis sentimen terhadap game Mobile Legends semakin penting untuk memahami persepsi masyarakat, serta bagaimana game ini mempengaruhi komunitas pemainnya (Sinaga & Jatmoko, 2020).

Analisis sentimen adalah proses penting untuk menggali opini dari teks seperti komentar di media sosial, ulasan, dan postingan forum (Kusuma, Ratnawati, & Setiawan, 2025). Teknik ini biasanya digunakan untuk mengkategorikan sentimen menjadi positif, negatif, atau netral (Sari & Wibowo, 2019). Model Naive Bayes adalah kelompok algoritma probabilistik yang berdasarkan pada Teorema Bayes dengan asumsi independensi antar fitur (Alkhairi, Zer, Batubara, & Nugraha, 2021). Terdapat beberapa model dari Naive Bayes yang digunakan dalam berbagai konteks analisis sentimen antara lain Gaussian Naive Bayes, Multinomial Naive Bayes, dan Bernoulli Naive Bayes. Masing-masing model ini memiliki karakteristik tersendiri dalam mengolah data dan menghasilkan prediksi terutama dalam konteks data teks (Lowd & Domingos, 2005). Meskipun lebih umum digunakan untuk data numerik, Gaussian Naive Bayes terkadang diterapkan pada analisis sentimen untuk memproses fitur yang diturunkan dari teks yang dikonversi menjadi representasi numerik tertentu (Nurul A'ayunnisa, Salim, & Azis, 2022).

Multinomial Naive Bayes sangat cocok untuk data yang berupa frekuensi atau jumlah kejadian, seperti analisis teks di mana dokumen diubah menjadi representasi vektor frekuensi kata. Ini menjadikannya pilihan utama dalam analisis sentimen karena karakteristik data teks yang sering kali direpresentasikan dalam bentuk bag-of-words atau tf-idf (Yuyun, Nurul Hidayah, & Supriadi Sahibu, 2021). Adanya kebutuhan untuk mendapatkan pemahaman yang lebih baik tentang bagaimana persepsi masyarakat terhadap Mobile Legends dapat memengaruhi popularitas dan keberlanjutan game tersebut menjadi motivasi dalam penelitian ini. Pemahaman ini tidak hanya penting bagi pengembang game, tetapi juga bagi penyelenggara turnamen dan pelaku industri esport lainnya untuk mengembangkan strategi yang lebih efektif dalam mengelola komunitas

pemain. Selain itu penelitian ini berusaha mengisi kesenjangan literatur yang ada dengan menyediakan analisis yang komprehensif mengenai model Naive Bayes dalam konteks analisis sentimen mengenai esport. Penelitian ini bertujuan untuk membandingkan performansi ketiga varian utama dari model Naive Bayes dalam menganalisis sentimen terhadap esport Mobile Legends. Dengan mengevaluasi akurasi, presisi, recall, dan metrik evaluasi lainnya, penelitian ini akan membantu menentukan model Naive Bayes mana yang paling efektif dalam memahami persepsi publik terhadap Mobile Legends. Hasil perbandingan ini diharapkan dapat memberikan wawasan mendalam bagi pengembang game, penyelenggara turnamen, serta pelaku industri esport secara keseluruhan dalam merespons sentimen publik dan mengembangkan strategi yang lebih baik.

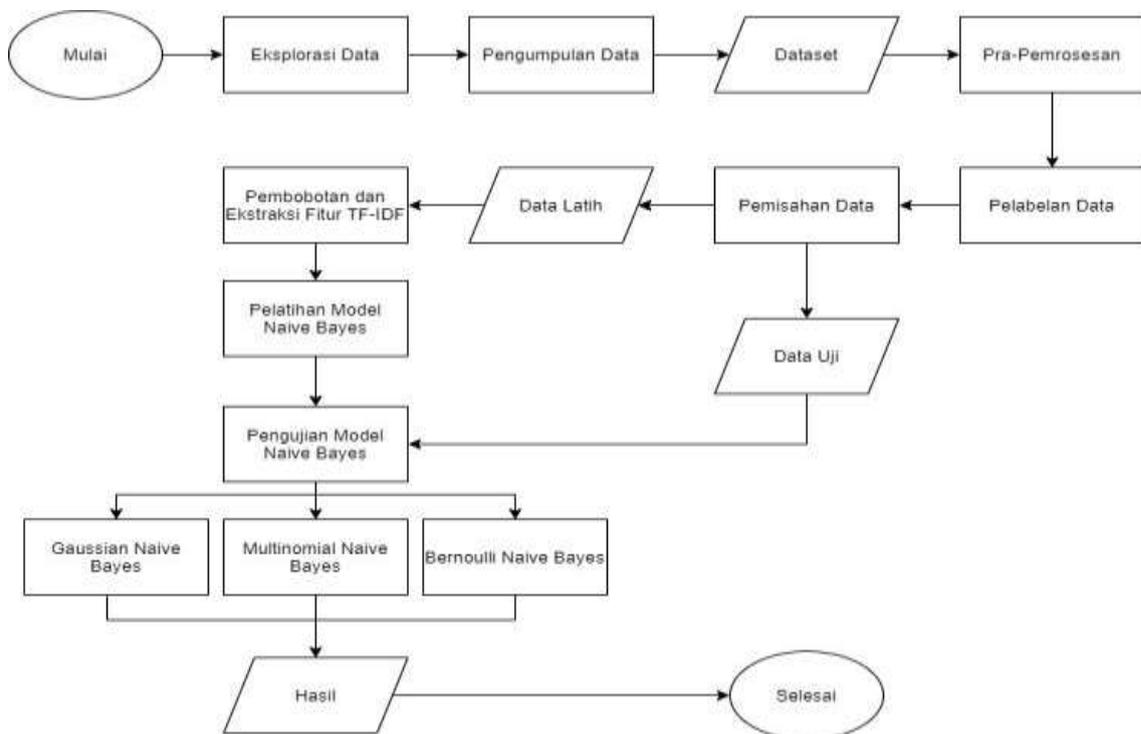
Kontribusi utama dari penelitian ini adalah menyediakan panduan yang jelas mengenai model Naive Bayes mana yang paling efektif untuk analisis sentimen di bidang esport, sehingga dapat digunakan oleh para praktisi dan peneliti di masa mendatang. Dengan semakin berkembangnya esport dan meningkatnya volume opini yang dipublikasikan oleh penggemar, pemain, dan pengamat, penting untuk mengetahui analisis sentimennya, selain itu, masih terdapat adanya kesenjangan dalam literatur yang ada sehingga perlu menyediakan analisis komprehensif mengenai model Naïve Bayes dalam konteks analisis sentimen pada esport. Berdasarkan rumusan masalah tersebut, berikut merupakan pertanyaan dalam penelitian ini a) Bagaimana performa 3 model Naïve Bayes dalam analisis sentiment terhadap esport Mobile Legends? b) Model Naïve Bayes mana yang paling efektif dalam mengklasifikasikan sentimen publik terhadap esport Mobile Legends berdasarkan akurasi, presisi, recall, dan metrik evaluasi lainnya? c) Apa pemahaman dan pengembangan strategi yang perlu dilakukan dalam mengelola komunitas pemain game Mobile Legends dan industri esport?

## **METODE**

Perancangan analisis sentimen merupakan proses-proses yang akan dilakukan sebelum pelaksanaan penelitian (Juniarsih, Ripanti, & Pratama, 2020). Perancangan ini dijadikan acuan selama penelitian agar hasil yang diharapkan dapat tercapai. Pada penelitian ini ada beberapa tahapan atau proses yang dilakukan mulai dari tahap eksplorasi data dari akun Youtube Mobile Legends: Bang Bang. Kemudian dilanjutkan dengan tahap pengumpulan data dengan melakukan scrapping data pada kolom live chat

siaran Grand Final MSC 2023 pada akun Youtube tersebut menggunakan extension berupa CommentStack pada browser Microsoft Edge. Tahap berikutnya adalah pra-pemrosesan data yang meliputi cleaning, case folding, tokenizing, stopword removal, stemming, serta pembersihan data secara manual dan pelabelan data. Setelah pra-pemrosesan data, tahap yang dilakukan adalah melakukan pembobotan menggunakan algoritma Term Frequency – Inverse Document Frequency (TF-IDF). Tahap berikutnya pelatihan model yang merupakan proses melatih model menggunakan data latih yang kemudian akan dievaluasi menggunakan data uji untuk mengukur performanya.

Setelah pelatihan model, tahap yang dilakukan adalah evaluasi model menggunakan Performance Evaluation Measure beserta Cross Validation yang kemudian dapat menampilkan hasil evaluasi dari model tersebut yang merupakan accuracy, precision, dan recall. Tahap terakhir adalah visualisasi data dimana hasil dari analisis sentimen dan evaluasi model disajikan dalam bentuk visual seperti grafik maupun WordCloud. Visualisasi ini membantu dalam menginterpretasikan hasil dan menyampaikan temuan penelitian secara lebih jelas dan informatif. Proses yang akan dilakukan dapat dilihat pada Gambar 1. berikut.



**Gambar 1.** Diagram Alir Analisis Sentimen

### **Eksplorasi Data**

Data yang digunakan untuk penelitian ini merupakan data primer. Data primer merupakan hasil pencarian data yang diambil secara langsung (Qomaruddin & Sa'diyah, 2024). Pada penelitian ini, data diambil secara langsung dari rekaman siaran langsung pertandingan Grand Final MSC (MLBB Southeast Asia Cup) 2023 pada platform Youtube. Pada Langkah pertama adalah memahami struktur data dengan meninjau jumlah total pesan dan tipe data yang ada. Pembersihan data dilakukan dengan menghapus pesan duplikat dan menangani nilai yang hilang untuk memastikan keakuratan analisis. Analisis teks dilakukan dengan tokenisasi untuk memecah pesan menjadi kata-kata atau frasa. Identifikasi pola dan tren melibatkan analisis perubahan sentimen seiring waktu dan korelasi antara katakata tertentu dengan sentimen yang dihasilkan.

### **Pengumpulan Data**

Metode yang digunakan untuk pengumpulan data dari penelitian ini adalah penggunaan CommentStack untuk perekaman live chat Youtube. Pada penelitian ini platform CommentStack digunakan sebagai alat untuk mengumpulkan data dari rekaman live chat Youtube. CommentStack memungkinkan peneliti untuk mengakses dan merekam percakapan yang terjadi selama siaran live di Youtube yang kemudian dapat langsung diekspor tanpa perlu menyalin secara manual. Pengumpulan data yang digunakan dalam penelitian ini dilakukan pada tanggal 27 Agustus 2023.

### **Pra-pemrosesan Data**

Tahap pra-pemrosesan merupakan langkah yang dilakukan setelah proses pengumpulan data. Pre-processing adalah tahap persiapan data dalam analisis atau pemrosesan data sebelum diterapkan pada algoritma atau model tertentu. Kualitas data memainkan peran penting dalam menentukan kualitas keluaran yang dihasilkan ketika menggunakan algoritma apa pun (Roy, Sharma, Nath, Bhattacharyya, & Kalita, 2018). Tujuan dari pre-processing sendiri adalah untuk mengubah data mentah menjadi bentuk yang lebih sesuai untuk analisis dan model yang akan diterapkan.

### **Pembobotan TF-IDF**

Setelah melewati tahap pre-processing, data yang akan digunakan harus berbentuk angka agar data yang masih berupa kata-kata dapat diolah dan dihitung. Data yang masih belum berupa angka selanjutnya akan diubah menjadi vector kemudian diberi nilai dan

bobot untuk setiap katanya yang kemudian akan diolah menggunakan algoritma prediksi. Pada tahap pembobotan kata, algoritma yang digunakan adalah Term Frequency-Inverse Document Frequency (TF-IDF).

### **Pelatihan Model**

Pada proses pelatihan model ini, model dilatih menggunakan data latih yang disiapkan dan akan dievaluasi menggunakan data uji untuk mengukur performanya. Algoritma yang digunakan adalah Naïve Bayes Classifier dengan menggunakan model Gaussian Naïve Bayes, Multinomial Naïve Bayes, serta Bernoulli Naïve Bayes.

### **Evaluasi Model**

Untuk membandingkan Gaussian, Multinomial, dan Bernoulli Naive Bayes didasarkan pada perbedaan karakteristik dari masing-masing algoritma yang cocok untuk tipe data yang berbeda. Perbandingan ini membantu memahami bagaimana berbagai jenis Naive Bayes bekerja pada dataset tertentu dan memilih model terbaik berdasarkan kinerja dan karakteristik data.

### **Visualisasi Data**

Menurut penemu aplikasi Tableau, menjelaskan visualisasi data atau data visualization adalah tampilan berupa grafis atau visual dari informasi dan data. Dengan kata lain, data visualization mengubah kumpulan data menjadi hal lebih sederhana untuk ditampilkan. Teknik visualisasi data yang efektif akan memberikan dampak signifikan pada keberhasilan komunikasi dengan target audiens. Tidak hanya mempermudah pengenalan pola pada data, visualisasi data juga dapat memperkuat pesan yang ingin disampaikan (Johannes Kurniawan, 2023).

## **HASIL DAN PEMBAHASAN**

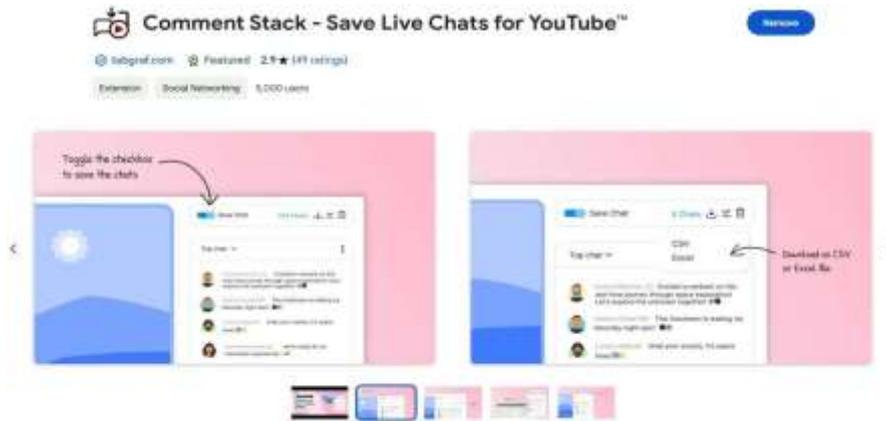
### **Eksplorasi data**

Data yang digunakan dalam penelitian ini merupakan data yang diambil secara langsung dari rekaman siaran langsung pertandingan Grand Final MSC 2023 yang berjudul “Bahasa Indonesia | Grand Final MSC | MSC 2023” pada kanal Youtube Mobile Legends: Bang Bang, yang mempertemukan antara tim Onic Indonesia dengan tim Blacklist International pada tanggal 18 Juni 2023.

### **Pengumpulan Data**

Data pesan yang dikumpulkan untuk dapat dipergunakan dalam penelitian ini diambil dengan cara melakukan perekaman live chat pada kolom live chat YouTube

dengan memanfaatkan bantuan dari extension CommentStack. Output dari proses ini ialah sebuah file .csv yang kemudian dapat dilakukan pembersihan dan pelabelan data.



**Gambar 2.** Menambahkan Extension CommentStack di browser

Langkah pertama yang dilakukan yaitu mengunduh dan menambahkan extension CommentStack di browser. Setelah CommentStack terpasang pada browser, buka rekaman siaran langsung pada platform YouTube. Dalam penelitian ini, data diambil dari rekaman siaran langsung pertandingan Grand Final MSC 2023 pada kanal Youtube Mobile Legends: Bang Bang. Langkah selanjutnya adalah merekam live chat menggunakan CommentStack, kemudian menyimpan live chat tersebut dan mengekspornya ke dalam bentuk file .csv.

### Implementasi Pra-pemrosesan Data

Data rekaman live chat yang telah didapatkan belum dapat digunakan karena masih merupakan data mentah yang belum diolah. Agar dapat digunakan, data rekaman live chat tersebut harus melalui pra-pemrosesan terlebih dahulu. Pra-pemrosesan pada penelitian ini dilakukan menggunakan platform Google Colab untuk menjalankan kode Python. Langkah yang pertama dilakukan adalah dengan menginstall package serta mengimport library yang akan digunakan.

```
!pip install Sastrawi
!pip install nltk
!pip install neattext
```

**Gambar 3.** Install Package

Kode diatas merupakan perintah yang digunakan dalam Google Colab untuk menginstall paket-paket Python menggunakan 'pip', yang merupakan package manager untuk Python. Perintah-perintah diatas dijalankan untuk menginstall paket yang

diperlukan untuk analisis teks dan pemrosesan bahasa alami atau biasa dikenal dengan NLP (Neural Language Processing). Paket pertama yang diinstal adalah Sastrawi yang merupakan sebuah pustaka Python untuk stemming dan stopword removal dalam bahasa Indonesia. Paket kedua adalah NLTK (Natural Language Toolkit) yang juga merupakan pustaka Python yang menyediakan alat-alat untuk bekerja dengan teks dalam pengolahan bahasa alami.

```
import pandas as pd
import string
import nltk
nltk.download('punkt')
from nltk.tokenize import word_tokenize
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory
from google.colab import drive
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
```

**Gambar 4.** Import Library

Kode diatas merupakan rangkaian perintah untuk mengimpor library. Kode pada baris yang pertama berfungsi untuk mengimpor library 'pandas', yang digunakan untuk manipulasi data dan analisis dalam bentuk DataFrame. 'pandas' sangat berguna untuk menangani data dalam format tabel seperti CSV. Kode pada baris kedua berfungsi untuk mengimpor module string, yang menyediakan konstanta dan fungsi untuk operasi string, seperti daftar karakter alfabet dan simbol. Kode pada baris yang ketiga berfungsi untuk mengimpor library nltk (Natural Language Toolkit), yang digunakan untuk pemrosesan bahasa alami (NLP). nltk menyediakan alat untuk tokenisasi, stemming. Kode pada baris yang keempat berfungsi untuk mengunduh paket 'punkt' dari nltk, yang berisi tokenizer. Tokenizer ini digunakan untuk memecah teks menjadi kalimat atau kata. Kode pada baris kelima berfungsi untuk mengimpor fungsi word\_tokenize dari nltk.tokenize. Fungsi ini digunakan untuk memecah teks menjadi token (kata). Kode pada baris keenam berfungsi untuk mengimpor StopWordRemoverFactory dari library Sastrawi. Sastrawi adalah library untuk pemrosesan bahasa Indonesia yang menyediakan alat untuk menghilangkan kata-kata umum (stopwords) dan melakukan stemming. StopWordRemoverFactory digunakan untuk membuat objek yang dapat menghapus stopwords dari teks bahasa Indonesia. Kode pada baris ketujuh berfungsi untuk mengimpor module drive dari google.colab. Ini memungkinkan akses ke Google Drive dari lingkungan Google Colab. Anda dapat menggunakan ini untuk membaca dan menyimpan file di Google Drive, memudahkan integrasi dengan file yang disimpan di cloud. Kode pada baris terakhir

berfungsi untuk mengimpor StemmerFactory dari Sastrawi. StemmerFactory digunakan untuk melakukan stemming, yaitu proses mengubah kata menjadi bentuk dasarnya dalam bahasa Indonesia. Stemming membantu mengurangi variasi kata sehingga analisis teks menjadi lebih konsisten.

```
drive.mount('/content/gdrive')
df=pd.read_csv('gdrive/My Drive/DATASET/DATASET.csv')
```

**Gambar 5.** Impor StemmerFactory dari Sastrawi

### **Implementasi Pembagian Data dan Pembobotan TF-IDF**

Setelah melewati tahap pra-pemrosesan data, data yang digunakan harus melalui tahap pembobotan agar data yang masih berupa kata-kata dapat diolah dan dihitung. Data yang belum berupa angka diubah menjadi vector dan kemudian diberi nilai dan bobot untuk setiap kata atau term-nya. Algoritma yang digunakan dalam pembobotan adalah Term Frequency-Inverse Document Frequency (TF-IDF). Sebelum dilakukan pembobotan, data yang digunakan akan melalui proses splitting data atau pembagian data untuk memisahkan dataset menjadi dua bagian yaitu data latih dan data uji. Data latih merupakan data yang akan dilatihkan kepada model dari Naïve Bayes, sedangkan data uji merupakan data yang akan digunakan untuk menguji model Naïve Bayes yang telah dilatih menggunakan data latih. Rasio pembagian data yang digunakan dalam penelitian ini adalah 80:20. 80% dari dataset digunakan sebagai data latih dan 20% sisanya digunakan menjadi data uji.

Langkah pertama yang dilakukan adalah mengimpor pustaka Python yang digunakan dalam pengolahan data, ekstraksi fitur, pembuatan model, pelatihan model, evaluasi model, dan visualisasi data. Kode yang digunakan adalah sebagai berikut:

```
import numpy as np
import pandas as
pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import GaussianNB, MultinomialNB,
BernoulliNB from sklearn.metrics import accuracy_score,
classification_report
from sklearn.preprocessing import
LabelEncoder from sklearn.pipeline
import make_pipeline from sklearn import
```

**Gambar 6.** Impor Pustaka Python

Kode di atas digunakan untuk mengimpor pustaka-pustaka yang digunakan seperti TfidfVectorizer, Gaussian Naïve Bayes, Multinomial Naïve Bayes, Bernoulli Naïve Bayes, accuracy\_score dan classification\_report yang digunakan untuk mengukur kinerja model klasifikasi, metrics yang menyediakan berbagai fungsi untuk mengevaluasi model,

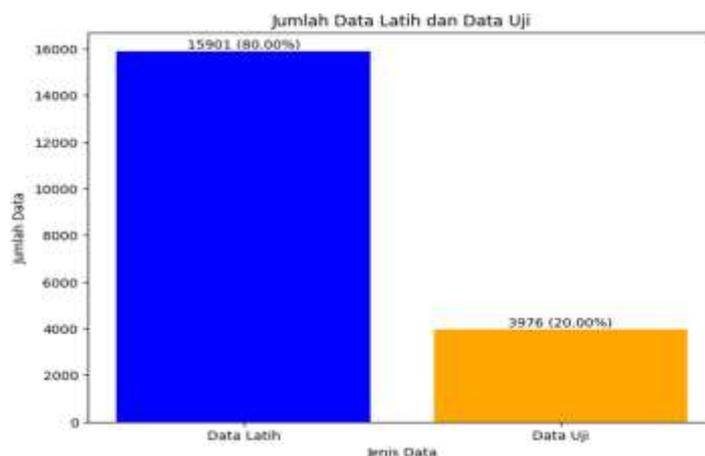
cross\_val\_score yang digunakan untuk mengevaluasi model menggunakan teknik crossvalidation, dan yang terakhir adalah mengimpor seaborn yang merupakan pustaka visualisasi data yang digunakan untuk pembuatan grafik statistik yang informatif. Kode di atas digunakan untuk menyediakan alat dan metode yang digunakan untuk mengolah data teks dan mengubahnya menjadi fitur berbobot, membuat dan melatih tiga model Naïve Bayes untuk klasifikasi teks, mengukur dan melaporkan kinerja model menggunakan metrik evaluasi, mengubah label kategori menjadi label numerik, membuat pipeline untuk alur kerja pemodelan yang lebih efisien, melakukan cross-validation untuk mengevaluasi stabilitas dan kinerja model, serta alat untuk membuat visualisasi data untuk analisis yang lebih mendalam dan informatif.

```

from sklearn.model_selection import train_test_split
# Bagi data menjadi data pelatihan dan pengujian
X_train, X_test, y_train, y_test = train_test_split(data['Message'],
data['Label'], test_size=0.2, random_state=42)
    
```

**Gambar 7.** Membagi Dataset Menjadi Data Latih dan Data Uji

Kode di atas digunakan untuk membagi dataset menjadi data latih dan data uji dengan rasio 80:20. Dengan rasio tersebut proporsi untuk data uji adalah 20% dari dataset dan proporsi untuk data latih adalah 80% dari dataset. Kolom pesan ‘Message’ digunakan sebagai fitur untuk pelatihan model. Kolom label ‘Label’ digunakan sebagai target dalam pelatihan model. Setelah dilakukan pembagian data dari dataset yang berisikan 19877 pesan, terdapat data latih dengan jumlah 15901 data/pesan dan data uji yang berjumlah 3976 data/pesan.



**Gambar 8.** Jumlah Data Latih dan Data Uji

Setelah pembagian data, proses berikutnya adalah pembobotan menggunakan algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Pembobotan TF-IDF

dilakukan dengan menjalankan kode berikut pada Google Colab:

```
# Mengganti NaN dengan string kosong pada data latih dan uji
X_train = X_train.fillna('')
X_test = X_test.fillna('')

# Membuat TfidfVectorizer
vectorizer = TfidfVectorizer()

# Melakukan vektorisasi pada data latih dan uji
X_train_vectorized = vectorizer.fit_transform(X_train)
X_test_vectorized = vectorizer.transform(X_test)
```

**Gambar 9.** Pembobotan Menggunakan Algoritma Term Frequency – Inverse Document Frequency (TF-IDF)

Berikut merupakan beberapa contoh hasil dari pembobotan TF-IDF yang telah dilakukan:

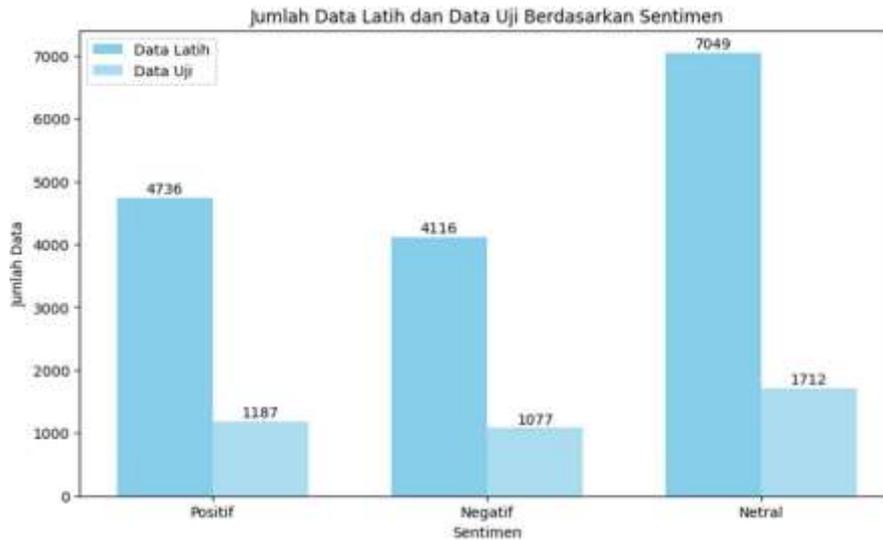
**Tabel 1.** Hasil Pembobotan TF-IDF

meta pause	
meta	0.720798
pause	0.693145
meta pause kalah	
meta	0.589048
pause	0.566449
kalah	0.576331
kelas gg onic	
kelas	0.751738
gg	0.550725
onic	0.362757
kelas onic	
kelas	0.900623
onic	0.434602

### Pelatihan Model Naïve Bayes

Pada proses pelatihan model ini, model dilatih menggunakan data latih yang telah disiapkan dan akan dievaluasi menggunakan data uji untuk mengukur performanya. Algoritma yang digunakan adalah Naïve Bayes Classifier dengan menggunakan model

Gaussian Naïve Bayes, Multinomial Naïve Bayes, dan Bernoulli Naïve Bayes. Data yang digunakan dalam pelatihan model merupakan data latih yang berjumlah 15.901 data dengan rincian untuk masing-masing kelas positif, negatif dan netral adalah berjumlah 4.736 data, 4.116 data, dan 7.049 data.



**Gambar 10.** Jumlah Data Latih dan Data Uji Berdasarkan Sentimen

### Evaluasi Model Naïve Bayes

Penghitungan dilakukan pada model yang telah dilatih pada tahap pelatihan model untuk dapat dilakukan penilaian terhadap model-model tersebut. Evaluasi model dilakukan untuk memastikan apakah model yang telah dilatih dapat membuat prediksi dengan benar. Evaluasi model yang dilakukan adalah menggunakan Performance Evaluation Measure dengan confusion matrix dan juga cross validation. Terdapat beberapa aspek dalam penilaian, seperti accuracy, precision, recall, serta f1-score. Accuracy sendiri adalah ketepatan model dalam melakukan prediksi terhadap data. Precision adalah tingkat kepastian dalam melakukan prediksi atau klasifikasi data. Recall digunakan untuk mengukur seberapa baik model dapat memprediksi atau melakukan klasifikasi data. Sedangkan f1-score merupakan rata-rata harmonik dari precision dan recall. Data yang digunakan dalam evaluasi model merupakan data uji yang berjumlah 3.986 data dengan rincian untuk masing-masing kelas positif, negatif dan netral adalah berjumlah 1.187 data, 1.077 data, dan 1.712 data.

```
from sklearn.metrics import confusion_matrix

# Evaluasi model
y_pred_gnb = gnb.predict(X_test_vectorized.toarray())
y_pred_mnb = mnb.predict(X_test_vectorized)
y_pred_bnb = bnb.predict(X_test_vectorized)

# Lakukan cross-validation dengan 5 folds
scores_gnb = cross_val_score(gnb, X_train_vectorized.toarray(),
y_train, cv=5)
scores_mnb = cross_val_score(mnb, X_train_vectorized, y_train, cv=5)
scores_bnb = cross_val_score(bnb, X_train_vectorized, y_train, cv=5)

# Menampilkan confusion matrix untuk setiap model
print("Gaussian Naive Bayes Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_gnb))

# Cetak akurasi rata-rata dari cross-validation
print("Mean cross-validation accuracy for Gaussian NB:",
scores_gnb.mean())

print("\nMultinomial Naive Bayes Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_mnb))

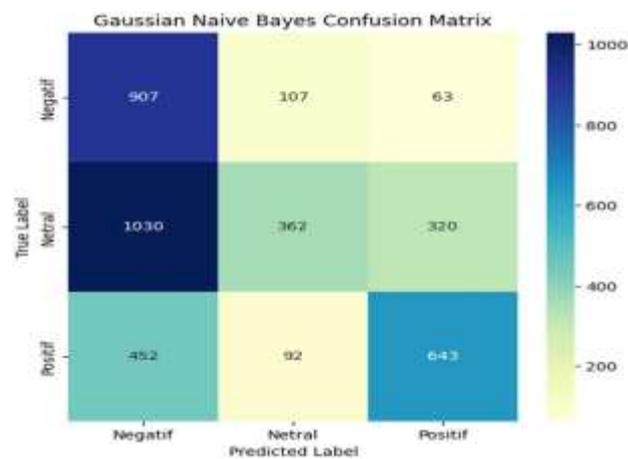
# Cetak akurasi rata-rata dari cross-validation
print("Mean cross-validation accuracy for Multinomial NB:",
scores_mnb.mean())

print("\nBernoulli Naive Bayes Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_bnb))

# Cetak akurasi rata-rata dari cross-validation
print("Mean cross-validation accuracy for Bernoulli NB:",
scores_bnb.mean())
```

**Gambar 11.** Evaluation Measure dengan Confusion Matrix dan Cross Validation

Confusion matrix didapatkan pada saat melakukan evaluasi model. Confusion matrix memberikan gambaran tentang bagaimana model melakukan prediksi dengan mengklasifikasikan hasil menjadi empat kategori yaitu True Positif (TP), False Positif (FP), True Negatif (TN), dan False Negatif (FN). Dari confusion matrix tersebut, nilai akurasi, presisi, recall, serta f1-score dapat dihitung. Akurasi mengukur proporsi prediksi benar dari total prediksi, memberikan gambaran umum mengenai kinerja model. Presisi mengukur keakuratan prediksi positif atau benar. Recall, atau sensitivitas, mengukur kemampuan model dalam mendeteksi semua sampel positif yang sebenarnya. Sedangkan F1-Score merupakan rata-rata harmonis dari presisi dan recall. Dengan menggunakan confusion matrix, performa model dapat dievaluasi dan dipahami secara komprehensif.

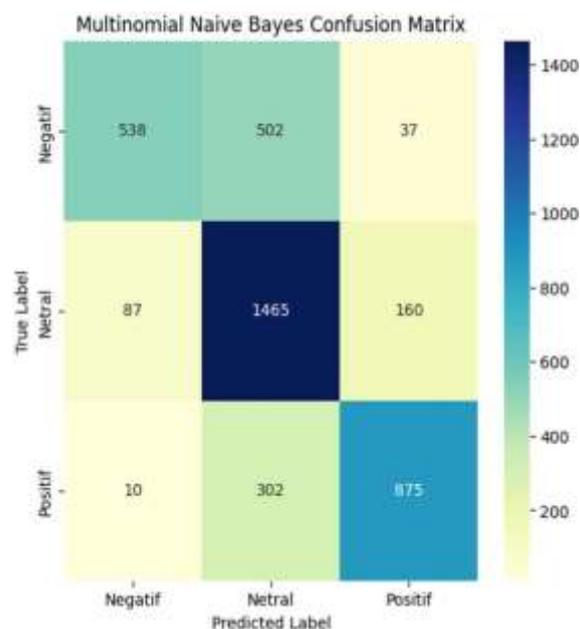


**Gambar 12.** Gaussian Naive Bayes Confusion Matrix

Dari nilai confusion matrix model Gaussian Naïve Bayes diatas, hasil penghitungan untuk accuracy, precision, recall, serta f1 score adalah sebagai berikut:

**Tabel 2.** Hasil Penghitungan untuk Accuracy, Precision, Recall, Serta F1 Score Gaussian Naive Bayes Confusion Matrix

Gaussian Naïve Bayes	
Accuracy	0.48
Precision	0.55
Recall	0.53
F1-Score	0.47

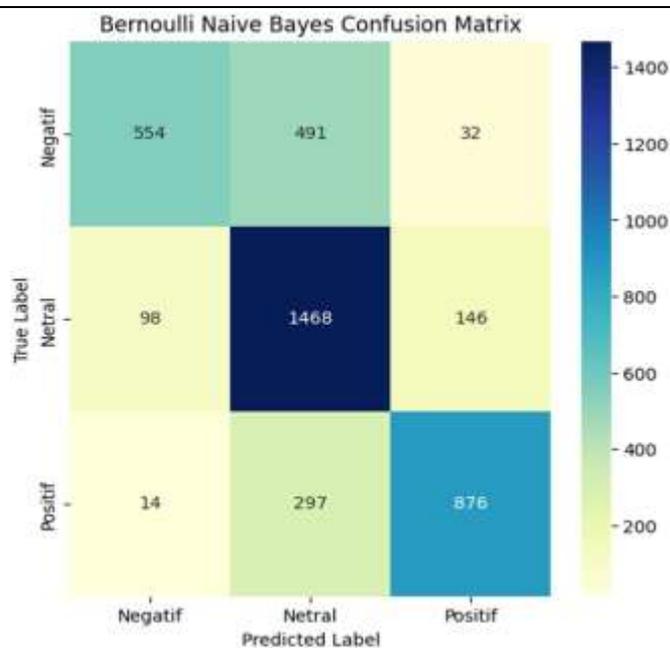


**Gambar 13.** Multinomial Naïve Bayes Confusion Matrix

Dari nilai confusion matrix model Multinomial Naïve Bayes diatas, hasil penghitungan untuk accuracy, precision, recall, serta f1 score adalah sebagai berikut:

**Tabel 3.** Hasil Penghitungan untuk Accuracy, Precision, Recall, Serta F1 Score Multinomial Naïve Bayes Confusion Matrix

<b>Multinomial Naïve Bayes</b>	
Accuracy	0.72
Precision	0.77
Recall	0.70
F1-Score	0.71



**Gambar 14.** Bernoulli Naïve Bayes Confusion Matrix

Dari nilai confusion matrix model Bernoulli Naïve Bayes diatas, hasil penghitungan untuk accuracy, precision, recall, serta f1 score adalah sebagai berikut:

**Tabel 4.** Hasil Penghitungan untuk Accuracy, Precision, Recall, Serta F1 Score Bernoulli Naïve Bayes Confusion Matrix

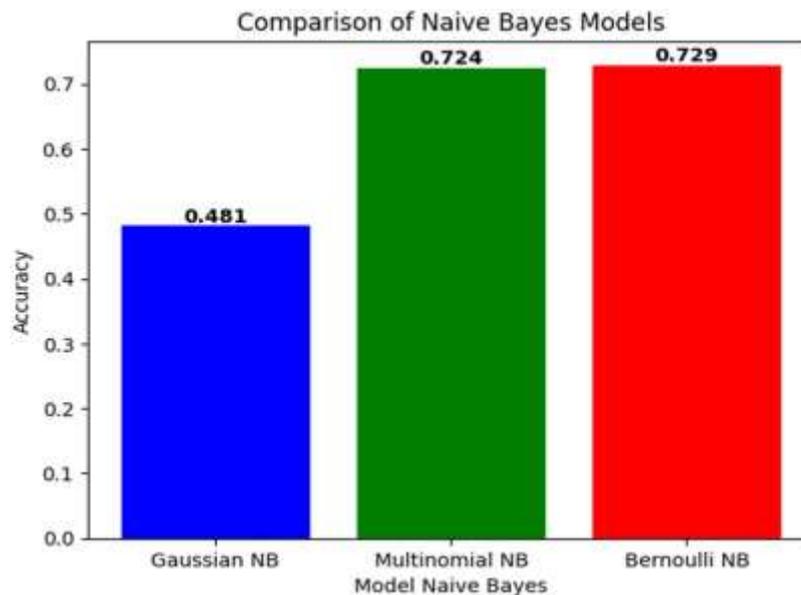
<b>Bernoulli Naïve Bayes</b>	
Accuracy	0.73
Precision	0.77
Recall	0.70
F1-Score	0.72

Dalam penelitian ini, tiga model Naïve Bayes, yaitu Gaussian Naïve Bayes, Multinomial Naïve Bayes, dan Bernoulli Naïve Bayes, digunakan untuk melakukan analisis sentimen. Hasil evaluasi menunjukkan bahwa kinerja setiap model bervariasi berdasarkan confusion matrixnya yang menghasilkan metrics akurasi, presisi, recall, dan f1-score. Model Gaussian Naïve Bayes menunjukkan kinerja dengan hasil yang paling rendah dengan akurasi sebesar 0.48. Presisi model Gaussian Naïve Bayes ini adalah 0.55, yang menunjukkan bahwa dari semua prediksi positif atau benar yang dihasilkan, hanya 55% di antaranya yang benar-benar positif. Recall dari model Gaussian Naïve Bayes ini adalah 0.53, yang berarti model ini berhasil mendeteksi 53% dari semua contoh positif atau benar yang ada. F1-score yang merupakan rata-rata harmonis dari presisi dan recall, pada model Gaussian Naïve Bayes ini berada pada nilai 0.47, mencerminkan keseimbangan antara presisi dan recall yang cukup rendah.

Model Bernoulli Naïve Bayes memiliki kinerja yang sedikit lebih baik dibandingkan model Multinomial Naïve Bayes, dengan akurasi sebesar 0.73. Presisi dari model Bernoulli Naïve Bayes ini adalah 0.77, menunjukkan bahwa 77% dari prediksi positif atau benar adalah prediksi yang benar. Recall dari model Bernoulli Naïve Bayes ini adalah 0.70, yang berarti model ini dapat mendeteksi 70% dari semua contoh positif atau benar yang ada. F1-score pada model Bernoulli Naïve Bayes ini memiliki nilai sebesar 0.71, menunjukkan performa yang seimbang dan sedikit lebih baik dibandingkan dengan Multinomial Naïve Bayes. Selain itu hasil cross-validation menunjukkan mean cross-validation accuracy untuk setiap model. Nilai mean cross validation accuracy untuk model Gaussian Naïve Bayes adalah sebesar 0.484, untuk model Multinomial Naïve Bayes adalah 0.723, dan untuk Bernoulli Naïve Bayes adalah 0.726. Hasil ini menunjukkan konsistensi kinerja model pada dataset yang berbeda-beda selama proses cross-validation, dengan Bernoulli dan Multinomial Naïve Bayes menunjukkan performa yang lebih baik dan stabil dibandingkan dengan model Gaussian Naïve Bayes.

Performa tiap model dalam mengklasifikasikan sentimen tertentu seperti sentimen negatif, netral, maupun sentimen positif dapat dinilai dari nilai f1-score nya. Model Gaussian Naïve Bayes, pada kelas negatif mendapatkan nilai f1-score sebesar 0.52, sedangkan pada kelas netral mendapatkan nilai sebesar 0.32, dan pada kelas positif mendapatkan nilai sebesar 0.58. Model Multinomial Naïve Bayes, pada kelas negatif mendapatkan nilai f1-score sebesar 0.63, sedangkan pada kelas netral mendapatkan nilai

sebesar 0.74, dan pada kelas positif mendapatkan nilai sebesar 0.77. Model Bernoulli Naïve Bayes, pada kelas negatif mendapatkan nilai f1-score sebesar 0.63, sedangkan pada kelas netral mendapatkan nilai sebesar 0.74, dan pada kelas positif mendapatkan nilai sebesar 0.78. Hasil tersebut dapat menunjukkan bahwa model-model Naïve Bayes cenderung lebih efektif dalam mengklasifikasikan sentimen positif dan negatif dibandingkan dengan sentimen netral. Secara keseluruhan, model Bernoulli dan Multinomial Naïve Bayes menunjukkan performa yang superior dalam analisis sentiment ini, dengan kinerja yang lebih baik dalam hal akurasi, presisi, recall, dan f1-score dibandingkan dengan model Gaussian Naïve Bayes. Hasil ini mengindikasikan bahwa Bernoulli dan Multinomial Naïve Bayes lebih efektif dalam menangkap pola dan sentimen kata dalam data teks yang digunakan dalam penelitian ini.



**Gambar 15.** Comparison of Naïve Bayes Models

### Visualisasi Data

Visualisasi data dilakukan terhadap keseluruhan data. Pada wordcloud positif hanya terdapat kata yang paling sering muncul dalam kelas berlabel positif, pada wordcloud negatif hanya terdapat kata yang paling sering muncul dalam kelas berlabel negatif, sedangkan pada wordcloud netral hanya terdapat kata yang paling sering muncul dalam kelas berlabel netral. Dalam penelitian ini, wordcloud digunakan untuk memberikan gambaran visual tentang katakata yang paling sering muncul dalam data yang dianalisis. Ukuran setiap kata dalam wordcloud berbanding lurus dengan frekuensi kemunculannya, sehingga kata-kata yang lebih sering muncul akan ditampilkan lebih besar dan lebih

menonjol. Teknik ini tidak hanya memudahkan dalam mengidentifikasi kata-kata kunci atau tema utama dalam data, tetapi juga membantu dalam memahami distribusi frekuensi kata secara keseluruhan. Wordcloud dapat menjadi alat yang sangat berguna dalam eksplorasi data teks, karena mampu menyajikan informasi secara ringkas dan langsung.



Gambar 16. Wordcloud Sentimen Positif

Kata-kata yang paling sering masuk ke dalam kelas dengan sentimen berlabel positif pada kolom live chat siaran Grand Final MSC 2023 pada kanal Youtube Mobile Legends: Bang Bang adalah kata “gg”, “onic”, “salam”, “semangat”, “go onic”, “juara”, dan lain-lain sesuai dengan wordcloud di atas. Ukuran setiap kata dalam wordcloud berbanding lurus dengan frekuensi kemunculannya, sehingga kata-kata yang lebih sering muncul akan ditampilkan lebih besar dan lebih menonjol. Hal itu berarti bahwa penonton sering menggunakan kata-kata tersebut sebagai topik yang sering dibicarakan pada kolom live chat siaran Grand Final MSC 2023 pada kanal Youtube Mobile Legends: Bang Bang.

Berikut merupakan visualisasi data dalam kelas berlabel negatif menggunakan wordcloud:



Gambar 17. Wordcloud Sentimen Negatif





## **SIMPULAN**

Berdasarkan penelitian terhadap kolom live chat rekaman siaran langsung Grand Final MSC 2023, didapatkan kesimpulan yaitu Gaussian, Multinomial, dan Bernoulli Naïve Bayes dapat mengklasifikasikan komentar ke dalam tiga kelas label sentimen dengan hasil performansi yang bervariasi. Berdasarkan hasil evaluasi, model Bernoulli dan Multinomial Naïve Bayes terbukti memberikan performa yang lebih baik dibandingkan Gaussian Naïve Bayes dengan akurasi masing-masing sebesar 0,73 dan 0,72 sementara Gaussian hanya mendapatkan akurasi sebesar 0,48. Bernoulli dan Multinomial Naïve Bayes juga menunjukkan keseimbangan antara presisi dan recall yang baik, dengan f1-score sebesar 0,71. Crossvalidation juga mengonfirmasi konsistensi performa Multinomial dan Bernoulli, dengan mean cross-validation accuracy masing-masing sebesar 0,723 dan 0,726, sedangkan Gaussian hanya mencapai 0,484. Secara keseluruhan, model Bernoulli Naïve Bayes memperlihatkan kinerja terbaik dengan performa sedikit lebih tinggi daripada Multinomial Naïve Bayes. Hasil ini menunjukkan bahwa Bernoulli lebih sesuai dalam menangani pola sentimen pada data teks pendek seperti komentar live chat, sehingga dapat menjadi opsi utama dalam analisis sentimen untuk jenis data serupa. Berdasarkan evaluasi sentimen penonton, strategi yang perlu dikembangkan dalam mengelola komunitas esport dan Mobile Legends adalah mempertahankan hal-hal positif seperti performa tim serta sportivitas. Peningkatan diperlukan dalam menjaga kualitas penyiaran untuk menghindari masalah-masalah teknis. Edukasi juga perlu dilakukan untuk menjaga kondusifitas ekosistem industri esports dan Mobile Legends ini.

## **UCAPAN TERIMA KASIH**

Terima kasih kepada pihak Universitas Islam Indonesia atas dukungan dalam penelitian ini. Ucapan terima kasih juga disampaikan kepada dosen pembimbing penulis atas bimbingan yang telah diberikan dalam penyusunan penelitian ini. Penulis juga sangat menghargai bantuan dari rekan-rekan yang turut memberikan dukungan serta kontribusi dalam proses penelitian ini.

## **DAFTAR PUSTAKA**

Alkhairi, F. I. R. H., Zer, E. R., Batubara, E. R., & Nugraha, F. (2021). Pengenalan Pola Kemampuan Pelanggan Dalam Membayar Air PDAM Menggunakan Algoritma Naive Bayes. *JURNAL TIMES (Technology Informatic & Computer System, X (2), 29–38*.

- Johannes Kurniawan. (2023). Analisis Dan Visualisasi Data. In *Angewandte Chemie International Edition*, 6(11), 951–952. (Vol. 3).
- Juniarsih, S., Ripanti, E. F., & Pratama, E. E. (2020). Implementasi Naive Bayes Classifier pada Opinion Mining Berdasarkan Tweets Masyarakat Terkait Kinerja Presiden dalam Aspek Ekonomi. *J. Sist. Dan Teknol. Inf*, 8(3), 239.
- Kusuma, S. A., Ratnawati, D. E., & Setiawan, N. Y. (2025). Analisis Sentimen Pengguna Sosial Media Twitter/X Terhadap Acara Clash of Champions Menggunakan Metode Multinomial Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 9(3).
- Lowd, D., & Domingos, P. (2005). Naive Bayes models for probability estimation. *ICML 2005 - Proceedings of the 22nd International Conference on Machine Learning*, 529–536. <https://doi.org/10.1145/1102351.1102418>
- Nurul A'ayunnisa, Salim, Y., & Azis, H. (2022). Analisis Performa Metode Gaussian Naïve Bayes untuk Klasifikasi Citra Tulisan Tangan Karakter Arab. *Indonesian Journal of Data and Science*, 3(3), 115–121. <https://doi.org/10.56705/ijodas.v3i3.54>
- Putra, E. P., Safitri, A., Kurniawan, M. A., Vadilah, M. I., Prasetyo, M. A., Wijaya, M. R., ... Fansyuri, M. (2023). Merubah Perspektif Masyarakat Terhadap Game Online. *Abdi Jurnal Publikasi*, 1(6), 485–488.
- Qomaruddin, Q., & Sa'diyah, H. (2024). Kajian Teoritis tentang Teknik Analisis Data dalam Penelitian Kualitatif: Perspektif Spradley, Miles dan Huberman. *Journal of Management, Accounting, and Administration*, 1(2), 77–84. <https://doi.org/10.52620/jomaa.v1i2.93>
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Sinaga, D., & Jatmoko, C. (2020). Analisis Sentimen Untuk Mengetahui Kesan Player Game Mobile Legends Menggunakan Naïve Bayes Classifier. *Semnas LPPM UMP*, V, 540–547.
- Yuyun, Nurul Hidayah, & Supriadi Sahibu. (2021). Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4), 820–826. <https://doi.org/10.29207/resti.v5i4.3146>