

PERBANDINGAN KINERJA NAÏVE BAYES, SUPPORT VECTOR MACHINE, DAN K-NEAREST NEIGHBOR DALAM ANALISIS SENTIMEN MOBILE LEGENDS

Hikmawan Alvin Zikirlah, Iltavera Paula, Muhammad Fazilla, Riski Annisa,
Lady Agustin Fitriana[✉]

Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Pontianak, Indonesia

Email: lady.lag@bsi.ac.id

ABSTRACT

The rapid advancement of information and communication technology has significantly increased the popularity of online games in Indonesia, one of which is Mobile Legends: Bang Bang (MLBB) with millions of active users. The abundance of user reviews on digital platforms provides valuable data for analysis using text mining and natural language processing (NLP) approaches. Sentiment analysis is applied to classify user opinions into positive, negative, and neutral categories, offering insights into player satisfaction and perceptions of game quality. This study compares the performance of three classification algorithms Naïve Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) in analyzing sentiment from Mobile Legends user reviews on the Google Play Store. A total of 5,000 reviews were collected using the web scraping technique and processed through the Knowledge Discovery in Databases (KDD) framework, which includes cleaning, case folding, tokenization, normalization, and stopword removal. Sentiment labeling was performed using a lexicon-based approach with the InSet sentiment lexicon. The dataset was divided into training and testing sets with an 80:20 ratio and evaluated using accuracy, precision, recall, and f1-score metrics. The results show that the SVM algorithm achieved the highest accuracy of 88.1%, followed by KNN at 65.1% and NB at 62.6%. Thus, SVM is recommended as the most effective model for sentiment analysis of Mobile Legends user reviews.

Keywords: Sentiment Analysis, Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Mobile Legends.

ABSTRAK

Perkembangan teknologi informasi dan komunikasi mendorong peningkatan popularitas permainan daring di Indonesia, salah satunya Mobile Legends: Bang Bang (MLBB) dengan jutaan pengguna aktif. Banyaknya ulasan pemain di platform digital menjadi sumber data penting untuk dianalisis menggunakan pendekatan text mining dan natural language processing (NLP). Analisis sentimen berfungsi untuk mengklasifikasikan opini pengguna menjadi kategori positif, negatif, dan netral, sehingga dapat menggambarkan tingkat kepuasan serta persepsi terhadap kualitas permainan. Penelitian ini membandingkan performa tiga algoritma klasifikasi, yaitu Support Vector Machine (SVM), Naïve Bayes Classifier (NBC) dan K-Nearest Neighbor (KNN) dalam menganalisis sentimen ulasan pengguna Mobile Legends di Google Play Store. Sebanyak 5.000 ulasan dikumpulkan melalui teknik web scraping dan diproses dengan tahapan Knowledge Discovery in Databases (KDD) yang mencakup cleaning, case folding, tokenization, normalization, dan stopword removal. Pelabelan sentimen dilakukan dengan metode lexicon-based menggunakan kamus InSet. Dataset dibagi menjadi data pelatihan dan pengujian dengan rasio 80:20, kemudian dievaluasi menggunakan metrik accuracy, precision, recall, dan f1-score. Hasil penelitian menunjukkan bahwa algoritma SVM memberikan akurasi tertinggi sebesar 88,1%, diikuti KNN sebesar 65,1% dan NB sebesar 62,6%. Dengan demikian, SVM direkomendasikan sebagai model paling efektif untuk analisis sentimen ulasan pengguna Mobile Legends.

Kata Kunci: Analisis Sentimen, Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Mobile Legends.

PENDAHULUAN

Perkembangan pesat teknologi informasi dan komunikasi telah memengaruhi kemajuan berbagai sektor, termasuk industri digital mendorong meningkatnya popularitas permainan daring (*mobile game*) di Indonesia, salah satunya Mobile Legends: Bang Bang (MLBB) yang memiliki jutaan pengguna

aktif di berbagai platform digital. Banyaknya ulasan dan komentar dari pengguna mengenai pengalaman bermain game ini, baik positif maupun negatif, menjadi sumber data yang berharga untuk dianalisis guna mengetahui persepsi publik terhadap kualitas.

Analisis sentimen sebagai *text mining* dan *natural language processing (NLP)* bertujuan untuk

mengenali dan mengelompokkan opini maupun emosi yang terkandung dalam teks ke dalam kategori tertentu, yaitu positif, negatif, atau netral (Agustin Fitriana et al., 2025). Dalam konteks MLBB, analisis sentimen dapat membantu pengembang memahami tingkat kepuasan pemain, respons terhadap pembaruan fitur, hingga permasalahan teknis yang sering dikeluhkan (Sinaga & Jatmoko, 2020).

Penelitian sebelumnya membandingkan algoritma *Naïve Bayes*, *SVM* untuk analisis sentimen 52.651 ulasan *Mobile Legends* di Google Play. Dengan praproses teks dan representasi *TF-IDF*, hasil menunjukkan *SVM* dan *Logistic Regression* mencapai akurasi 90–91%, dengan *SVM* paling stabil (Alengka et al., 2025).

Dalam perkembangan pesat industri *e-sports* di Indonesia, *Mobile Legends: Bang Bang* menjadi salah satu permainan daring dengan komunitas pemain terbesar dan tingkat interaksi tertinggi. Ulasan yang diunggah oleh pengguna melalui platform digital mencerminkan persepsi publik terhadap kualitas permainan (Ardianto et al., 2020).

Penelitian ini diharapkan dapat memberikan kontribusi akademis dan praktis dalam menentukan algoritma terbaik untuk klasifikasi sentimen di game. Pengujian dilakukan dengan Penelitian ini melakukan perbandingan terhadap kinerja tiga algoritma *machine learning*, yaitu *K-Nearest Neighbor* (KNN), *Naïve Bayes Classifier* (NBC), dan *Support Vector Machine* (SVM). Berdasarkan hasil pengujian menggunakan metode pembagian data, algoritma KNN memperoleh performa paling unggul dengan nilai akurasi sebesar 87,8%, presisi 87,8%, serta recall mencapai 84% (Putri et al., 2023).

Penelitian selanjutnya membandingkan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) untuk analisis sentimen 6.000 ulasan *Mobile Legends*. Setelah dilakukan pemrosesan teks dan tahap pelabelan sentimen, diperoleh hasil 95,5%, sedangkan *Naïve Bayes* 76,08%. Berdasarkan *confusion matrix*, SVM terbukti lebih efektif dalam mengklasifikasikan aplikasi tersebut (Yani et al., 2025).

Sering dengan meningkatnya penggunaan aplikasi digital dan platform interaktif, analisis opini pengguna menjadi aspek krusial dalam memahami persepsi serta tingkat kepuasan terhadap suatu produk. Penerapan metode *machine learning* dan *Natural Language Processing* (NLP) memungkinkan proses analisis sentimen dilakukan secara otomatis, sistematis, dan efisien. Oleh sebab itu, penelitian ini berfokus pada pengembangan model yang mampu mengklasifikasikan data secara lebih akurat dan efisien analisis sentimen yang mampu mengklasifikasikan

ulasan pengguna secara akurat guna mendukung proses pengambilan keputusan berbasis data. (Azanda et al., 2025).

Berdasarkan penjelasan penelitian terkait diatas, ketiga algoritma, yaitu *Naïve Bayes* (NB), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (K-NN). Pemilihan ketiga algoritma tersebut didasarkan pada performa yang baik dalam Pengolahan data teks dan analisis sentimen digunakan untuk mengevaluasi persepsi publik terhadap dinamika perkembangan dunia *e-sport* di Indonesia (Alamsyah & Pratiwi, 2024).

TINJAUAN PUSTAKA

Naïve bayes

Algoritma *Naïve Bayes* termasuk metode pembelajaran mesin yang sederhana namun memiliki tingkat efisiensi tinggi. Pendekatan ini beroperasi sebuah data masuk ke dalam kelas tertentu berdasarkan fitur-fiturnya, dan dikenal karena kemampuannya membangun model dengan cepat serta memberikan hasil prediksi yang efektif (Ridwan, 2020).

Pendekatan ini merupakan metode pembelajaran yang didasari oleh prinsip teorema Bayes. Algoritma *Naïve Bayes Classifier* dikenal luas karena sifatnya yang sederhana namun efektif, mampu memberikan hasil klasifikasi yang sebanding dengan metode lain seperti *decision tree* dan *neural network*, serta memiliki keunggulan dalam kecepatan pemrosesan data berjumlah besar (Asnawi et al., 2021).

Jadi dapat di sumpulkan Naive bayes adalah Algoritma *Naïve Bayes* adalah metode teorema Bayes yang sederhana, cepat, dan efektif dalam mengklasifikasi data berdimensi tinggi, dengan hasil sebanding metode lain seperti *decision tree* dan *neural network*.

Support Vector Macine (SVM)

Support Vector Machine (SVM) adalah metode pembelajaran mesin yang memanfaatkan fungsi linear pada ruang fitur berdimensi tinggi dan dilatih melalui pendekatan berbasis teori optimasi. Esensi utama dari SVM terletak pada pencarian *hyperplane* optimal yang mampu memisahkan data ke dalam dua kelas dengan margin maksimum atau pemisah antar kelas tersebut. Algoritma ini telah diterapkan secara luas dalam berbagai studi pada bidang terkait data mining dan text mining karena kemampuannya menunjukkan performa yang unggul dalam klasifikasi data kompleks dan dimensi tinggi. (M Riski Qisthiano, 2022)

K-Nearest Neighbor

K-Nearest Neighbor (K-NN) mengklasifikasi data baru dengan mengacu pada data pelatihan

yang telah berlabel. Keakuratan K-NN sangat bergantung pada karakteristik data dan pemilihan fitur yang tepat. Oleh karena itu, banyak Penelitian yang berfokus pada penerapan teknik seleksi fitur dan pemberian bobot guna meningkatkan kinerja algoritma ini dalam proses klasifikasi. Pendekatan ini bertujuan mengoptimalkan hasil klasifikasi dengan mengadaptasi parameter dan fitur agar algoritma K-NN dapat bekerja lebih efisien dan akurat (M Riski Qisthiano, 2022)

Analisis Sentimen

Analisis sentimen atau yang dikenal juga dengan *opinion mining*, merupakan bagian dari bidang *natural language processing*, komputasi linguistik, serta penambangan teks yang berfokus pada pengidentifikasi dan pemahaman opini, sikap, maupun emosi individu baik penulis maupun pembicara terhadap suatu topik, produk, layanan, organisasi, atau aktivitas tertentu tertentu. Pendekatan ini digunakan untuk memahami dan mengklasifikasikan ekspresi emosional dalam teks sehingga dapat menilai apakah sentimen yang terkandung bersifat positif, negatif, atau netral.(M Riski Qisthiano, 2022)

PENELITIAN TERDAHULU

Penelitian oleh bustamin,Nahya Nur dan Farid Wajidi mebandingkan ulasan komentar Mobile Legends menggunakan KNN mengklasifikasikan ulasan menjadi positif dan negatif dengan akurasi 65,82%. Data menunjukkan sentimen negatif lebih dominan, memberikan insight penting bagi pengembang untuk peningkatan kualitas game (Bustamin et al., 2025)

Penelitian oleh Son Gohan Alengka1 Jordy Lasmana Putra, Tyas Setiyorini Mobile Legends menghasilkan jutaan ulasan, sehingga analisis manual tidak efektif. Penelitian membandingkan Naïve Bayes, SVM, dan Logistic Regression pada 52.651 ulasan, dengan SVM dan Logistic Regression mencapai akurasi sekitar 90-91%. Model SVM dipilih karena stabilitasnya dan dapat digunakan untuk memantau ulasan secara real-time, membantu pengembang dalam meningkatkan kualitas game. Hasil ini khusus berlaku pada ulasan Google Play Mobile Legends dan perlu validasi lintas aplikasi(Alengka et al., 2025)

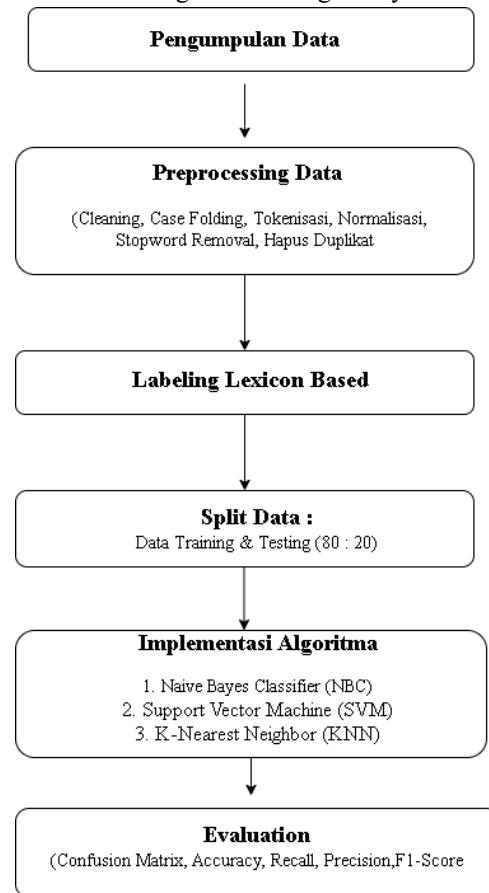
Penelitian oleh Nikolas Orlando, Benedikto Nugroho Prasetya, Latius Hermawan K-Means mengelompokkan keluhan pengguna Mobile Legends menjadi kategori utama untuk membantu pengembang meningkatkan pengalaman bermain.K-Means mengelompokkan keluhan pengguna Mobile Legends berdasarkan pola untuk membantu pengembang

meningkatkan pengalaman dan kualitas permainan.(Orlando et al., 2025)

Ketiga penelitian sebelumnya membahas analisis sentimen Mobile Legends menggunakan KNN, Naïve Bayes, SVM, serta K-Means untuk mengelompokkan ulasan dan keluhan pengguna. secara langsung membandingkan kinerja algoritma Naïve Bayes, Support Vector Machine, dan K-Nearest Neighbor dalam konteks analisis sentimen Mobile Legends. Oleh karena itu, penulis tertarik mengangkat topik ini untuk mengisi kekosongan penelitian sekaligus memberikan kontribusi baru dalam pemilihan algoritma terbaik untuk analisis sentimen game Mobile Legends.

METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja *Knowledge Discovery in Databases* (KDD) sebagai pendekatan sistematis untuk mengekstraksi informasi dan pengetahuan dari kumpulan data yang dianalisis dari data Mobile Legends di Google Play Store.



Gambar 1. Tahapan penelitian

Tahapan ini diawali dengan proses pengumpulan data diperoleh dari Google Play Store melalui metode web scraping menggunakan kata kunci "Mobile Legends". Proses scraping ini dilakukan untuk mengumpulkan sebanyak 5.000 baris data ulasan

pengguna terhadap aplikasi tersebut. Data yang diambil berupa teks ulasan, rating, serta informasi tambahan seperti tanggal dan nama pengguna yang kemudian dianonimkan untuk menjaga privasi. Tujuan dari tahapan ini adalah memperoleh sumber data yang representatif dan relevan sebagai dasar dalam melakukan analisis sentimen terhadap aplikasi Mobile Legends.

Pada tahap preprocessing, dilakukan beberapa tahapan sistematis untuk mempersiapkan data sebelum digunakan dalam proses analisis mempersiapkan data teks. Proses *preprocessing* diawali dengan penghapusan elemen-elemen nonteks seperti simbol, angka, dan emotikon. Tahap berikutnya adalah *case folding*, yaitu mengonversi seluruh huruf menjadi format kecil (*lowercase*) agar konsisten dalam analisis. Setelah itu dilakukan *tokenisasi*, yang berfungsi memecah kalimat menjadi satuan kata atau *token* untuk mempermudah pemrosesan. Tahap *normalisasi* dilakukan guna menyeragamkan kata tidak baku maupun singkatan ke bentuk ejaan yang sesuai. Kemudian dilanjutkan dengan *stopword removal* untuk menghapus kata-kata umum yang tidak memberikan makna penting terhadap konteks, seperti “dan”, “yang”, atau “di”. Terakhir, penghapusan data duplikat diterapkan agar tidak terjadi pengulangan ulasan yang dapat memengaruhi validitas hasil analisis secara bias.

Pada tahap pelabelan, penelitian ini menggunakan metode *lexicon-based* untuk mengidentifikasi dan menentukan polaritas sentimen pada setiap ulasan pengguna. Metode tersebut berfungsi dengan mencocokkan kata-kata dalam teks ulasan terhadap daftar kosakata (*lexicon*) yang telah memiliki nilai sentimen positif, negatif, maupun netral. Setiap kata kemudian diberikan skor sesuai dengan kategori sentimennya untuk menghasilkan penilaian akhir, kemudian skor total dari seluruh kata dalam satu ulasan digunakan untuk menentukan label akhir. Jika nilai skor cenderung positif maka ulasan dikategorikan sebagai positif, apabila cenderung negatif maka dikategorikan sebagai negatif, dan jika seimbang maka dikategorikan sebagai netral. Pendekatan ini dipilih karena efektif dalam menganalisis teks berbahasa Indonesia dan tidak memerlukan proses pelatihan model seperti metode berbasis machine learning.

Pada tahap *split data*, dataset yang telah diberi label dibagi menjadi dua subset, yaitu data pelatihan dan data pengujian dengan perbandingan 80:20. Pembagian ini bertujuan agar sebagian besar data (80%) digunakan untuk melatih model dalam mengenali pola dan karakteristik sentimen, sedangkan 20% sisanya dimanfaatkan untuk menguji kinerja model terhadap data yang belum pernah dipelajari

sebelumnya. kinerja model yang telah dibangun. Dengan proporsi tersebut, model dapat mempelajari berbagai variasi data secara optimal sekaligus dievaluasi kemampuannya dalam memprediksi data baru yang belum pernah dipelajari sebelumnya. Proses pembagian dilakukan secara acak guna menjaga keseimbangan distribusi antara kelas positif, negatif, dan netral, sehingga hasil evaluasi model menjadi lebih akurat dan representatif.

Pada tahap implementasi algoritma, penelitian ini menggunakan tiga metode klasifikasi, yaitu Support Vector Machine (SVM), Naïve Bayes Classifier (NBC), dan K-Nearest Neighbor (KNN) untuk menganalisis sentimen ulasan pengguna aplikasi Mobile Legends. Algoritma Naïve Bayes Classifier bekerja dengan menghitung probabilitas kemunculan kata dalam setiap kelas sentimen berdasarkan prinsip teorema Bayes, di mana setiap fitur dianggap independen satu sama lain. Sementara itu, Support Vector Machine (SVM) berfungsi dengan mencari hyperplane terbaik yang mampu memisahkan data ke dalam kelas sentimen positif, negatif, dan netral dengan margin maksimal, sehingga hasil klasifikasi menjadi lebih akurat. Adapun K-Nearest Neighbor (KNN) mengklasifikasikan data berdasarkan jarak terdekat antara data uji dan sejumlah tetangga terdekatnya (nilai k), di mana label sentimen ditentukan berdasarkan mayoritas kelas tetangga tersebut. Ketiga algoritma ini diimplementasikan untuk dibandingkan kinerjanya dalam mengidentifikasi sentimen pengguna, sehingga dapat diketahui algoritma mana yang memberikan hasil paling optimal terhadap dataset yang digunakan.

Pada tahap evaluasi, kinerja ketiga algoritma, yaitu Naïve Bayes Classifier (NBC), Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN) diukur menggunakan metrik evaluasi berbasis confusion matrix yang mencakup nilai accuracy, recall, precision, dan F1-score. Confusion matrix digunakan untuk menunjukkan perbandingan antara hasil prediksi model dengan label sebenarnya pada setiap kategori sentimen. Nilai accuracy merepresentasikan tingkat ketepatan model dalam mengklasifikasikan seluruh data secara benar, sedangkan precision menggambarkan proporsi prediksi positif yang benar-benar relevan. Sementara itu, recall menunjukkan kemampuan model dalam mendeteksi seluruh data positif yang sesungguhnya. Adapun F1-score merupakan rata-rata harmonis antara precision dan recall yang memberikan gambaran menyeluruh mengenai performa model. Melalui evaluasi ini, dapat diketahui algoritma mana yang menunjukkan kinerja paling optimal dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Mobile Legends.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pada tahapan pengumpulan data ini menggunakan pustaka google play scraper untuk mengambil ulasan dari Google Play Store secara otomatis dengan kata kunci atau app id= *com.mobilelegends* seperti yang ditampilkan pada gambar 2. Melalui fungsi `get_reviews()`, program mengumpulkan hingga 5000 ulasan terbaru, lalu menampilkan jumlah serta contoh ulasan yang ditampilkan gambar 3.

```
from google_play_scraper import reviews, Sort
app_id = 'com.mobile.legends'

def get_reviews(app_id, lang='id', count=5000):
    try:
        result, continuation_token = reviews(
            app_id,
            lang=lang,
            country='id',
            sort=sort,
            count=count,
            filter_score_with=filter_score_with,
            filter_device_with=filter_device_with,
            continuation_token=continuation_token
    )
```

Gambar 2. Pengumpulan Data

Review Text
Ngasi tim yg seimbang lah kocak.
baguss
permain. yang. bagus

Gambar 3. Hasil Scrapping Data

Preprocessing Data

Pada tahapan data ini yang telah di scraping akan melalui preprocessing data, tahapan pertama melalui proses cleaning seperti ditunjukkan tabel 1.

Tabel 1. Cleaning Data

Review Text	Cleaning
Ngasi tim yang seimbang lah Kocak.	Ngasi tim yang seimbang lah Kocak
Pemain. yang. bagus	Pemain. yang. bagus
Seru banget.	Seru banget

Bagian ini menampilkan hasil dari proses cleaning atau pembersihan teks terhadap data ulasan pengguna. Tahapan ini untuk menghilangkan Komponen yang tidak diperlukan, seperti tanda baca,

penggunaan huruf kapital, spasi ganda, serta karakter yang tidak memiliki makna kontekstual, dihapus dari teks dalam konteks analisis. Proses ini bertujuan agar teks ulasan menjadi lebih terstruktur, konsisten, dan mudah diproses pada tahap analisis berikutnya. Hasil dari proses pembersihan tersebut.

cleaning	case_folding
Ngasi tim yg seimbang lah kocak	ngasi tim yg seimbang lah kocak
baguss	baguss
permain yang bagus	permain yang bagus

Gambar 4. Case Folding

Bagian ini menampilkan hasil dari proses case folding. Tahap *case folding* berfungsi untuk menyeragamkan format teks dengan mengonversi seluruh karakter menjadi huruf kecil agar lebih konsisten dan mudah diolah seperti gambar 4. Proses ini memastikan kata dengan bentuk huruf berbeda memiliki makna yang sama, sehingga meningkatkan konsistensi data teks dan akurasi pada tahap analisis sentimen berikutnya.

normalisasi	tokenize
mengasih tim yang seimbang lah kocak	[mengasih, tim, yang, seimbang, lah, kocak]
bagus	[bagus]
permain yang bagus	[permain, yang, bagus]

Gambar 5. Tokenisasi

Pada tahap *tokenisasi* kata tidak baku atau singkatan disesuaikan menjadi bentuk yang mengikuti kaidah bahasa yang benar, misalnya kata “ngasi” diubah menjadi “mengasih” dan “yg” menjadi “yang” seperti gambar 5.

cleaning	case_folding
Ngasi tim yg seimbang lah kocak	ngasi tim yg seimbang lah kocak
baguss	baguss
permain yang bagus	permain yang bagus

Gambar 6. Normalisasi

Proses normalisasi pada Gambar 6 menunjukkan perubahan Kata tidak baku disesuaikan ke dalam bentuk yang mengikuti aturan ejaan Bahasa Indonesia yang tepat, seperti “ngasi” menjadi “mengasih”, “yg” menjadi “yang”, dan “udah” menjadi “sudah”.

tokenize	stopword removal
[mengasih, tim, yang, seimbang, lah, kocak]	[mengasih, tim, seimbang, kocak]
[bagus]	[bagus]
[permain, yang, bagus]	[permain, bagus]

Gambar 7. Stopword Removal

Bagian ini menampilkan hasil dari proses stopword removal pada data ulasan pemain Mobile Legends. Pada tahap ini, kata-kata umum yang memiliki frekuensi tinggi namun tidak memiliki kontribusi makna yang berarti terhadap konteks kalimat, seperti “yang”, “lah”, dan “dan”, dihapus dari teks seperti gambar 7.

Labelling Lexicon Based

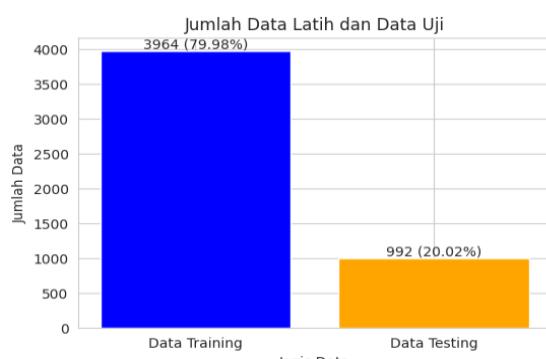
Metode lexicon-based menganalisis sentimen dengan membandingkan jumlah kata positif dan negatif dalam ulasan. Jika kata positif lebih banyak, ulasan dikategorikan positif, dan sebaliknya. Pendekatan ini sederhana namun efektif, terutama menggunakan kamus sentimen InSet untuk teks berbahasa Indonesia hasil proses pelebelan dapat dilihat Gambar 8.

stemming_data	Rating	Score	Sentiment
asih tim imbang kocak	1	-2	Negatif
bagus	5	0	Netral
main bagus	5	0	Netral
moonton pelit kasih lawan jago jago tim dark s...	1	0	Netral

Gambar 8. Labelling Lexicon Based

Split Data

Proporsi data dibagi sebesar 80:20, di mana 80% digunakan untuk pelatihan dan 20% untuk pengujian, sehingga model memiliki data yang memadai untuk mempelajari pola sentimen sekaligus tetap memiliki data terpisah untuk menguji kinerjanya seperti gambar 9.



Gambar 9. Data Training & Data Testing

Implementasi Algoritma

Algoritma Naïve Bayes (NB) mencapai 62,6% dengan rendah pada gambar 10. Berdasarkan prinsip probabilitas dan asumsi independensi fitur, NB kurang optimal pada data teks yang kompleks. Meskipun demikian, algoritma ini efisien secara komputasi dan sesuai digunakan sebagai model baseline pada dataset

Naive Bayes Confusion Matrix		
Actual		
	Predicted	
	Negatif	Netral
Negatif	237	17
Netral	87	53
Positif	44	13
		331

Gambar 10. Naïve Bayes

Algoritma Support Vector Machine (SVM) mencapai akurasi 88,1% dengan performa terbaik dalam klasifikasi sentimen. SVM menunjukkan keseimbangan tingkat precision, recall, dan f1-score pada setiap kategori kelas serta mampu memisahkan data berdimensi tinggi secara efektif. Hasil evaluasi membuktikan bahwa SVM akurat, stabil, dan merupakan model paling andal untuk analisis sentimen pada dataset ini seperti gambar 11.

SVM Confusion Matrix		
Actual		
	Predicted	
	Negatif	Netral
Negatif	267	34
Netral	22	251
Positif	1	31
		356

Gambar 11. Support Vector Machine

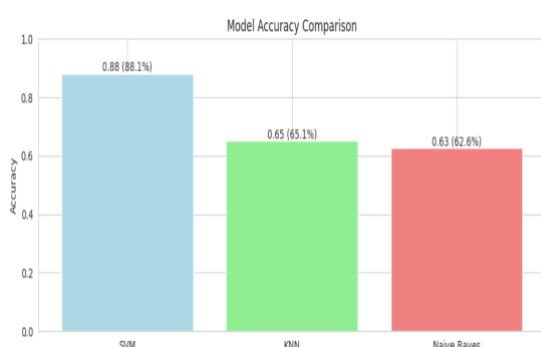
Algoritma K-Nearest Neighbors (KNN) memperoleh akurasi 65,1%, dengan performa baik pada kelas positif namun rendah pada kelas negatif dan netral. KNN mengklasifikasikan data berdasarkan kedekatan dengan data pelatihan, sehingga sensitif terhadap noise dan ketidakseimbangan data. Meski demikian, algoritma ini tetap efisien dan kompetitif pada dataset kecil dengan preprocessing optimal seperti gambar 12.

KNN Confusion Matrix		
		Predicted
Actual	Negatif	149
	Netral	219
Positif	12	98
	Negatif	36
	Netral	45
	Positif	278

Gambar 12. K-Nearest Neighbors

Evaluation

Hasil evaluasi pada Support Vector Machine (SVM) tingkat akurasi paling tinggi, yaitu 88,1%, menandakan kemampuan yang unggul dalam mengklasifikasikan data sentimen dengan konsistensi tinggi; sedangkan K-Nearest Neighbors (KNN) memperoleh akurasi 65,1%, cukup baik namun cenderung kurang stabil pada data yang kompleks, dan Naive Bayes (NB) mencatat akurasi terendah, 62,6%, yang meskipun cepat dan efisien, masih kalah akurat dibandingkan dua model lain, sehingga SVM direkomendasikan sebagai pilihan utama untuk analisis sentimen berdasarkan hasil perbandingan seperti pada gambar 13.



Gambar 13. Model Accuracy Comparison

Tabel 2. Perbandingan model Analisis Sentimen

Metrik	SVM	KNN	NB
Accuracy	0.881	0.651	0.626
Macro Precision	0.879	0.679	0.631
Macro Recall	0.877	0.646	0.603
Macro F1	0.878	0.645	0.565
Weighted F1	0.882	0.654	0.579

KESIMPULAN

Berdasarkan hasil analisis terhadap 5.000 ulasan pengguna *Mobile Legends* yang diambil dari Google Play Store, diketahui bahwa algoritma *Support Vector Machine* (SVM) memberikan kinerja paling unggul dibandingkan dengan *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN) dalam pengklasifikasian sentimen. Algoritma SVM berhasil mencapai tingkat akurasi tertinggi sebesar 88,1%, sedangkan *K-Nearest Neighbor* memperoleh akurasi 65,1%, dan *Naïve Bayes* hanya mencapai 62,6%.

SVM terbukti unggul dalam memisahkan data akurasi yang stabil pada ketiga kelas sentimen (positif, negatif, dan netral). Naïve Bayes efisien komputasi, tetapi kurang akurat pada data dengan kompleksitas tinggi. Sedangkan KNN cukup mudah diimplementasikan namun sensitif terhadap distribusi data dan pemilihan neighbor parameter (k).

Oleh karena itu, SVM direkomendasikan sebagai algoritma paling efektif untuk analisis sentimen ulasan *Mobile Legends*, terutama ketika berhadapan dengan dataset besar dan beragam.

Penelitian selanjutnya disarankan mengembangkan model dengan algoritma lain atau teknik *feature selection* untuk meningkatkan akurasi, serta memperluas sumber data agar hasil analisis sentimen lebih representatif dan dapat diterapkan secara real-time.

DAFTAR PUSTAKA

- Agustin Fitriana, Lady, Fahmi Julianto, M., Dahlia, R., Rifqi Firdaus, M., & Fazriansyah, A. (2025). Analisis Ulasan Konsumen sebagai Data Non-Kuangan dalam Sistem Informasi Akuntansi. *Profitabilitas*, 5(1), 64–74.
<https://doi.org/10.31294/profitabilitas.v5i1.8269>
- Alamsyah, M. K., & Pratiwi, N. (2024). *Analisis sentimen terkait opini masyarakat terhadap perkembangan e-sport mobile di indonesia menggunakan k nearest neighbor*. 9(1), 349–359.
- Alengka, S. G., Putra, J. L., & Setiyorini, T. (2025). *Analisis Sentimen Ulasan Mobile Legend Menggunakan Algoritma Naïve Bayes , SVM ,*

- Logistic Regression.* 5(3), 439–449.
<https://doi.org/10.35957/algoritme.v5i3.12915>
- Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. (2020). *Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)*. 2.
- Asnawi, M. H., Firmansyah, I., Novian, R., & Pontooh, R. S. (2021). Comparison of Naïve Bayes, K-NN, and SVM Algorithms in Social Media Sentiment Classification. *Seminar Nasional Statistika X*, 10(1).
- Azanda, O. K. A., Fadzilah Prayoganing, G., & Fredi, W. (2025). Analisis Sentimen Ulasan Pengguna Aplikasi Mobile Legends Pada Google Playstore Menggunakan Naïve Bayes. *Teknologi Informasi &*, 4(1), 213–227.
- Bustamin, B., Nur, N., & Wajidi, F. (2025). Analisis Sentimen Terhadap Komentar Game Mobile Legendsdi Play Store Dengan Menerapkan Metode K-Nearest Neighbors (Knn). *Journal Peqguruang: Conference Series*, 7(1), 75. <https://doi.org/10.35329/jp.v7i1.6059>
- M Riski Qisthiano. (2022). Klasifikasi Terhadap Prediksi Kelulusan Mahasiswa Dengan Menggunakan Metode Support Vector Machine (Svm). *Seminar Nasional Teknologi Dan Multidisiplin Ilmu (SEMNASTEKMU)*, 2(2), 203–207. <https://doi.org/10.51903/semnastekmu.v2i1.170>
- Orlando, N., Prasetya, B. N., & Hermawan, L. (2025). Analisis Sentimen Keluhan Pengguna pada Game Mobile Legends Menggunakan Metode K-Means Clustering. *MDP Student Conference*, 4(1), 292–298. <https://doi.org/10.35957/mdp-sc.v4i1.11177>
- Putri, A., Hardiana, C. S., Novfuja, E., Siregar, F. T. P., Rahmaddeni, R., Fatma, Y., & Wahyuni, R. (2023). Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 20–26. <https://doi.org/10.57152/malcom.v3i1.610>
- Ridwan, A. (2020). Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 4(1), 15–21. <https://doi.org/10.47970/siskom-kb.v4i1.169>
- Sinaga, D., & Jatmoko, C. (2020). Analisis Sentimen Untuk Mengetahui Kesan Player Game Mobile Legends Menggunakan Naïve Bayes Classifier. *Semnas LPPM UMP*, V, 540–547.
- Yani, F., Imelda, K., Kurniadi, N. T., Informatika, T., Teknik, F., & Bangsa, U. P. (2025). *Analisis Sentiment Pengguna Aplikasi Mobile Legend Di Playstore Dengan Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine (SVM)*. 6(2), 810–819. <https://doi.org/10.46576/djtechno>