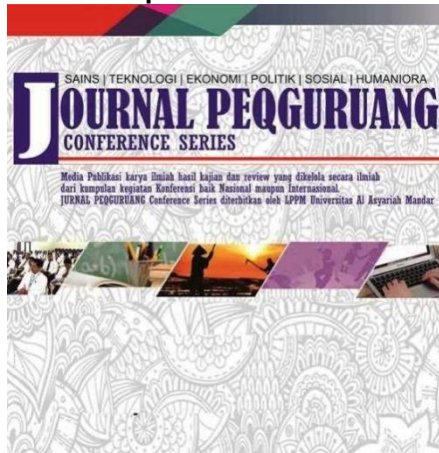


Graphical abstract



ANALISIS SENTIMEN TERHADAP KOMENTAR GAME MOBILE LEGENDSDI PLAY STORE DENGAN MENERAPKAN METODE K-NEAREST NEIGHBORS (KNN)

¹Bustamin, ²Nahya Nur, ³Farid Wajidi.

¹Teknik Informatika STMIK Hasan Sulur Wonomulyo,

²Teknik Informatika Universitas Sulawesi Barat,

³Teknik Informatika Universitas Sulawesi Barat.

*Corresponding author

1bustamin@stmik-hsw.ac.id,

2nahya.nur@unsulbar.ac.id,

3faridwajidi@unsulbar.ac.id

Abstract

Mobile Legends adalah salah satu game mobile terpopuler dengan jutaan pengguna di seluruh dunia, yang menghasilkan banyak ulasan di Google Play Store. Analisis sentimen terhadap komentar pengguna menjadi penting untuk pengembangan dan peningkatan kualitas permainan. Penelitian ini menggunakan metode K-Nearest Neighbors (KNN) untuk mengklasifikasikan komentar-komentar tersebut ke dalam kategori sentimen positif dan negatif. Langkah-langkah penelitian meliputi scraping data, case folding, normalisasi, tokenizing, stemming, dan pelabelan data. Hasil evaluasi model menunjukkan bahwa nilai K terbaik adalah 11, dengan akurasi tertinggi mencapai 65,82%. Proporsi data menunjukkan distribusi sentimen negatif yang lebih tinggi dibandingkan dengan sentimen positif, menandakan adanya kekurangan yang perlu diperhatikan oleh pengembang. Visualisasi data menggunakan word cloud membantu memahami pendapat pengguna secara intuitif. Hasil penelitian ini diharapkan dapat memberikan wawasan yang berguna bagi pengembang dalam meningkatkan kualitas dan kepuasan pengguna terhadap game Mobile Legends.

Keywords: *Mobile Legends, Analisis Sentimen, K-Nearest Neighbors (KNN), Google Play Store.*

Abstrak

Mobile Legends adalah salah satu game mobile terpopuler dengan jutaan pengguna di seluruh dunia, yang menghasilkan banyak ulasan di Google Play Store. Analisis sentimen terhadap komentar pengguna menjadi penting untuk pengembangan dan peningkatan kualitas permainan. Penelitian ini menggunakan metode K-Nearest Neighbors (KNN) untuk mengklasifikasikan komentar-komentar tersebut ke dalam kategori sentimen positif dan negatif. Langkah-langkah penelitian meliputi scraping data, case folding, normalisasi, tokenizing, stemming, dan pelabelan data. Hasil evaluasi model menunjukkan bahwa nilai K terbaik adalah 11, dengan akurasi tertinggi mencapai 65,82%. Diagram rating sampel menunjukkan distribusi sentimen negatif yang lebih tinggi dibandingkan dengan sentimen positif, menandakan adanya kekurangan yang perlu diperhatikan oleh pengembang. Visualisasi data menggunakan word cloud membantu memahami pendapat pengguna secara intuitif. Hasil penelitian ini diharapkan dapat memberikan wawasan yang berguna bagi pengembang dalam meningkatkan kualitas dan kepuasan pengguna terhadap game Mobile Legends.

Kata kunci: *Mobile Legends, Sentiment Analysis, K-Nearest Neighbors (KNN), Google Play Store.*

Article history

DOI: <http://dx.doi.org/10.35329/jp.v7i1>

Received : 2025-03-25 | Received in revised form : 2025-05-02 | Accepted : 2025-05-13

1. PENDAHULUAN

Mobile Legends adalah sebuah game bergenre MOBA (Multiplayer Online Battle Arena) tipe RPG (Role-Playing Game) yang dirancang khusus untuk smartphone. Setiap tim bertujuan untuk menghancurkan basis lawan sambil mempertahankan basis mereka sendiri dan mengendalikan jalur yang diserang. Setiap basis dihubungkan oleh jalur yang dikenal sebagai top, middle, dan bottom[1]. Komentar-komentar ini mencerminkan berbagai perasaan dan pendapat pemain tentang game, mulai dari pujian hingga kritik dan saran untuk perbaikan.

Komentar pengguna di Google Play Store dapat menjadi sumber data yang sangat berharga bagi pengembang game. Analisis sentimen merupakan proses analisis opini seseorang berdasar peristiwa yang sedang terjadi, baik itu positif atau negatif. Biasanya pada analisis sentimen, nilai pada analisis sentimen dibagi menjadi tiga yaitu positif dan negatif. Sehingga dalam proses pembagian data dapat ditemukan sentimen yang positif dan sentimen negatif[2]. Analisis ini tidak hanya membantu dalam memahami tingkat kepuasan pengguna, tetapi juga mengidentifikasi masalah-masalah yang perlu diperbaiki dan area-area yang dapat ditingkatkan.

Algoritma k-Nearest Neighbor adalah algoritma supervised learning dimana hasil dari instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori k-tetangga terdekat. Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek baru berdasarkan atribut dan sample-sample dari training data. Algoritma k-Nearest Neighbor menggunakan Neighborhood Classification sebagai nilai prediksi dari nilai instance yang baru[3].

Terdapat beberapa penelitian sebelumnya yang telah membahas terkait analisis sentimen pada aplikasi mobile. Terdapat beberapa penelitian sebelumnya yang telah membahas terkait analisis sentimen pada aplikasi mobile. Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor Penelitian ini menganalisis sentimen pengguna Twitter terhadap perpanjangan PPKM menggunakan metode K-Nearest Neighbor (K-NN). Data yang digunakan sebanyak 6408 tweet. Hasil penelitian menunjukkan akurasi sebesar 69,5%, recall 69,5%, dan presisi 68,7%[4].

Selanjutnya Analisis Sentimen Masyarakat Terhadap Penggunaan e-Commerce Menggunakan Algoritma K-Nearest Neighbor, Penelitian ini mengkaji sentimen masyarakat terhadap e-commerce melalui ulasan di Google Play Store. Setelah melakukan preprocessing seperti penghilangan stopwords, tokenisasi, dan stemming, algoritma K-Nearest Neighbor (K-NN) digunakan untuk mengklasifikasikan sentimen menjadi positif atau negatif. Hasil penelitian menunjukkan akurasi sebesar 82%[5].

Analisis Sentimen Terkait Opini Masyarakat Terhadap Perkembangan E-Sport Mobile Di Indonesia Menggunakan K Nearest Neighbor Penelitian ini

bertujuan untuk menganalisis sentimen masyarakat terhadap perkembangan e-sport mobile di Indonesia menggunakan metode K-Nearest Neighbor (K-NN) melalui data tweet. Hasil menunjukkan dominasi sentimen positif. Pada skenario pembagian data 70:30 dan 80:20, skenario 80:20 menghasilkan hasil yang lebih optimal dengan akurasi 84,66%, serta presisi, recall, dan f1-score yang lebih baik dibandingkan skenario 70:30 pada K=5[6].

Analisis Sentimen Ulasan Pengguna Aplikasi Discord Menggunakan Metode K-Nearest Neighbor Penelitian ini menganalisis sentimen ulasan pengguna aplikasi Discord menggunakan metode K-Nearest Neighbor (K-NN). Dari 2000 dataset ulasan yang dikumpulkan, setelah preprocessing dan penghapusan data duplikat, tersisa 1807 dataset. Implementasi K-NN dengan teknik k-5 fold menghasilkan akurasi rata-rata sebesar 82,76%[7].

Penerapan Algoritma KNN Pada Analisis Sentimen Review Aplikasi Peduli Lindungi Penelitian ini menganalisis sentimen ulasan aplikasi Peduli Lindungi menggunakan algoritma K-Nearest Neighbor (K-NN). Dari 200 data review yang terdiri dari sentimen positif dan negatif, penelitian ini menunjukkan akurasi 81,72% dengan AUC 0,856. Pengelompokan nilai AUC 0,856 termasuk dalam kategori Good Classification, membuktikan bahwa metode K-NN efektif dalam menganalisis sentimen review aplikasi Peduli Lindungi[8].

Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP Penelitian ini berfokus pada analisis sentimen pengguna terhadap aplikasi Flip untuk menentukan apakah ulasan pengguna mencerminkan rating positif yang diterima aplikasi. Hasil evaluasi menunjukkan bahwa 77,67% data uji terklasifikasi sebagai ulasan positif dengan precision dan recall masing-masing sebesar 82,67% dan 86,92%. Dalam klasifikasi data ulasan pengguna Flip dengan rasio data latih dan data uji sebesar 80%:20%, algoritma K-Nearest Neighbor mencapai akurasi sebesar 76,68%[9].

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap komentar-komentar yang ditinggalkan oleh pengguna Mobile Legends di Google Play Store menggunakan metode KNN. Melalui analisis ini, diharapkan dapat diperoleh wawasan yang lebih mendalam tentang persepsi dan respons pengguna terhadap game Mobile Legends. Selain itu, hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan dan peningkatan kualitas permainan tersebut, dengan cara mengidentifikasi keluhan utama dalam komentar negatif, memahami faktor-faktor yang menyebabkan kepuasan pengguna dalam komentar positif.

Dengan memahami sentimen pengguna melalui komentar mereka, pengembang Mobile Legends dapat membuat keputusan yang lebih baik dalam pengembangan game, meningkatkan kualitas layanan, dan pada akhirnya meningkatkan kepuasan pengguna.

Analisis sentimen ini merupakan langkah penting untuk memastikan bahwa game terus berkembang dan memenuhi harapan serta kebutuhan pemainnya.

2. METODE PENELITIAN

Penelitian ini merupakan penelitian kuantitatif yang menggunakan metode analisis data untuk mengidentifikasi pola dan hubungan antara komentar-komentar pengguna Mobile Legends dengan sentimen yang terkandung di dalamnya. Melalui pendekatan ini, peneliti dapat mengklasifikasikan komentar-komentar tersebut ke dalam kategori sentimen positif dan negatif dengan menggunakan metode K-Nearest Neighbors (KNN).

1. Pengumpulan Data

Pengumpulan data merupakan tahapan pertama dari penelitian ini. Data yang dikumpulkan diperoleh dari platform Google Play Store. Proses pengumpulan data dilakukan melalui teknik Web scraping untuk mendapatkan komentar pengguna Mobile Legends dari Google Play Store. Web scraping adalah teknik otomatisasi yang digunakan untuk mengekstrak data dari situs web. Sumber data utama dalam penelitian ini adalah komentar pengguna Mobile Legends di Google Play Store, yang mencakup berbagai sentimen, mulai dari positif dan negatif. Untuk mengimplementasikan web scraping, Google Colab dipilih sebagai platform karena kemampuannya dalam menangani komputasi data besar serta dukungannya terhadap berbagai pustaka Python yang diperlukan untuk scraping dan analisis data. Komentar yang ada pada laman tersebut akan diolah melalui beberapa tahapan preprocessing.

2. Case folding

Case folding adalah proses mengubah semua huruf dalam teks menjadi huruf kecil. Proses ini sering digunakan dalam pemrosesan teks, terutama dalam pencarian teks dan analisis teks, untuk memastikan bahwa perhitungan teks dilakukan tanpa memperhatikan perbedaan antara huruf besar dan kecil. Misalnya, kata "Game", "game", dan "GAME" akan dianggap sama setelah melalui proses pelipatan kasus.

3. Normalisasi

Normalisasi dalam konteks preprocessing data adalah langkah-langkah untuk mengubah data mentah menjadi format yang lebih bersih, seragam, dan siap untuk analisis atau penggunaan dalam algoritma pembelajaran mesin (machine learning). Proses normalisasi ini mencakup berbagai teknik dan metode tergantung pada jenis data yang sedang diproses, baik itu teks, numerik, atau lainnya.

4. Tokenizing

Tokenizing adalah proses memecah teks menjadi unit-unit diskrit yang disebut token. Token bisa berupa kata, frasa, atau karakter terpisah, tergantung pada aturan yang digunakan dalam proses tokenisasi. Ini adalah langkah awal dalam analisis teks dan pemrosesan bahasa alami, memungkinkan komputer untuk memahami struktur dan makna dari teks yang diberikan.

5. Remove Stopwords

Remove stopwords adalah proses menghapus kata-kata umum yang sering muncul dalam teks tetapi tidak memberikan banyak informasi penting dalam konteks analisis teks atau pemrosesan bahasa alami (Natural Language Processing/NLP). Kata-kata ini dikenal sebagai "stopwords." Contoh stopwords dalam bahasa Inggris termasuk kata-kata seperti "the," "is," "in," "and," "of," "to" dan sebagainya. Dalam bahasa Indonesia, contoh stopwords termasuk "dan", "yang", "di", "ke", "dari", "adalah", dan sebagainya.

6. Stemming

Stemming adalah proses normalisasi teks dalam pemrosesan bahasa alami yang menghapus infleksi kata untuk menghasilkan bentuk dasar kata, disebut sebagai "stems". Hal ini membantu mengurangi variasi kata yang sama sehingga mempermudah analisis dan pencarian.

7. Pelabelan Data

Pelabelan data adalah proses menetapkan label atau kategori tertentu kepada setiap contoh atau entitas dalam kumpulan data. Tujuannya adalah untuk memberikan identifikasi atau penjelasan tentang setiap data sehingga dapat digunakan untuk tujuan analisis, pemodelan, atau pembelajaran mesin. Misalnya, dalam klasifikasi email sebagai "spam" atau "non-spam", setiap email akan diberi label sesuai dengan kategorinya.

3. HASIL DAN PEMBAHASAN

A. Scraping Data

Tahapan dimulai dengan melakukan pengumpulan data. Pengumpulan data menggunakan Google Colab dengan algoritma Web scraping yang menghasilkan data sebanyak 995 komentar. Gambar 1 merupakan sampel data yang diperoleh melalui proses scraping.

	username	score	at	content
0	Pasha Fadillah Ardana	1	5/14/2024 13:46	Hero rusak mm ngecounter lesley pakai item apa...
1	Boxx Kah	5	5/14/2024 13:45	Monton babi kenapa di kasi tim tolot trus arjink
2	Akun Utama	5	5/14/2024 13:45	Geam ini sangat seru dan asik dan permainan in...
3	Muhammad Nizham permana	1	5/14/2024 13:45	Gw udah lama main mobeli legends dan gua selal...
4	M. Hafiz	5	5/14/2024 13:45	Lol, gg montoon
...	...	...	...	...
990	Kis wanto	4	5/14/2024 6:16	selalu dapat team yg cupu sedangkan musuh nya ...
991	Sri Rahayu	5	5/14/2024 6:16	gilakkk sanzxxxxx
992	Loreza Maulana AMV	1	5/14/2024 6:14	game sangat ajiing" match tidak sesuai dengan t...
993	Riko Tayco	5	5/14/2024 6:13	Kren
994	Rafli Hamdi	4	5/14/2024 6:13	Bagus Tapi Ada Sebagian Yg Kurang contohnya s...

995 rows x 5 columns

Gambar 1. Sampel Data

Gambar 1 menampilkan username, score, at, dan content, isi dari data di atas adalah ulasan pengguna pada aplikasi Mobile Legends, yang menampilkan gambar kolom pertama usedrname berisi nama pengguna atau username dari orang yang memberikan ulasan. Kolom score menunjukkan score atau rating yang di berikan oleh pengguna, kemungkinan besar dalam skala 1 sampai 5. Kolom selanjutnya at mencatat waktu kapan ulasan tersebut di buat, yang berisi tanggal dan jam. Kolom terakhir berisi content atau isi dari ulasan yang ditulis oleh pengguna.

B. Case folding

Pada tahapan ini, teks dipecah menjadi unit yang lebih kecil, seperti karakter, kata, atau frasa tergantung pada kebutuhan. Kemudian setiap huruf dalam teks dikonversi menjadi huruf kecil untuk memastikan konsistensi dalam analisis teks dan menghindari masalah perbedaan kapitalisasi yang dapat mengakibatkan variasi kata yang sebenarnya sama dianggap sebagai entitas yang berbeda. Misalnya, dalam bahasa Inggris, semua huruf dapat dikonversi menjadi huruf kecil. Tabel 1 merupakan hasil tahapan case folding.

Tabel 1. Hasil proses case folding

Sebelum Case Folding	Sebelum Case Folding
Gw udah lama main mobeli legends dan gua selalu stack di rank epic 3 dan legend 5 dan 3 Setiap satu bintang lagi naik gw selau kalah sampai gw mau pindah game tapi di pertandingan terakhir dikasih menang, tolong kasih musuh yang gak pro amat, ngasih musuh yang sebanding lah jangan ngasih tim keturunan immortal tolong kurangi dark sistem	gw udah lama main mobeli legends dan gua selalu stack di rank epic 3 dan legend 5 dan 3 setiap satu bintang lagi naik gw selau kalah sampai gw mau pindah game tapi di pertandingan terakhir dikasih menang, tolong kasih musuh yang gak pro amat, ngasih musuh yang sebanding lah jangan ngasih tim keturunan immortal tolong kurangi dark sistem
PLEASE MOONTON KENAPA SAYA GAK	please moonton kenapa saya gak bisa login???

BISA LOGIN???	kenapa bisa gameplay tidak di temukan di server??? astaga!!!!
KENAPA BISA GAMEPLAY TIDAK DI TEMUKAN DI SERVER???	
ASTAGA!!!!	
Gamenya bagus banyak banget Hero bagus	gamenya bagus banyak banget hero bagus
Saya minta tolong MLBB KALO KASIH musuh yang setara	saya minta tolong mlbb kalo kasih musuh yang setara

Tabel 1 menunjukkan hasil proses normalisasi dari teks percakapan dalam permainan Mobile Legends. Di kolom "Sebelum Normalisasi", terdapat ungkapan-ungkapan pengguna yang menggunakan bahasa sehari-hari dan singkatan yang sering ditemukan dalam komunikasi informal. Di kolom "Sesudah Normalisasi", teks tersebut telah diubah menjadi bentuk yang lebih formal dan baku. Contohnya, "gw udah lama main mobeli legends" menjadi "saya sudah lama main Mobile Legends" dan "tolong kurangi dark sistem" menjadi "tolong kurangi dark sistem". Normalisasi ini dilakukan untuk meningkatkan keterbacaan dan pemahaman teks dengan menghilangkan singkatan, menggunakan bahasa yang lebih formal, dan memperbaiki tata bahasa.

C. Normalisasi

Tahapan ini melibatkan penghapusan karakter yang tidak diinginkan atau tidak relevan dari teks, seperti tanda baca yang tidak penting, karakter spasi yang berlebihan, atau karakter khusus yang tidak relevan.

Tabel 2. Hasil proses normalisasi

Sebelum Normalisasi	Sesudah normalisasi
gw udah lama main mobeli legends dan gua selalu stack di rank epic 3 dan legend 5 dan 3 setiap satu bintang lagi naik gw selau kalah sampai gw mau pindah game tapi di pertandingan terakhir dikasih menang, tolong kasih musuh yang gak pro amat, ngasih musuh yang sebanding lah jangan ngasih tim keturunan immortal tolong kurangi dark sistem	saya sudah lama main mobeli legends dan saya selalu stack di rank epic dan Mobile Legends dan setiap satu bintang lagi naik saya selau kalah sampai saya mau pindah game tapi di pertandingan terakhir dikasih menang tolong kasih musuh yang tidak pro amat ngasih musuh yang sebanding lah jangan ngasih tim keturunan immortal tolong kurangi dark sistem
please moonton kenapa saya gak bisa login??? kenapa bisa gameplay	tolong moonton kenapa saya tidak bisa login kenapa bisa gameplay

tidak di temukan di server??? astaga!!!!	tidak di temukan di server astaga
gamenya bagus banyak banget hero bagus	gamenya bagus banyak banget hero bagus
saya minta tolong mlbb kalo kasih musuh yang setara	saya minta tolong mobilelegend bang bang kalo kasih musuh yang setara

Tabel 2 menunjukkan hasil proses case folding pada teks percakapan dalam permainan Mobile Legends. Proses case folding adalah langkah untuk mengubah semua huruf dalam teks menjadi huruf kecil (lowercase) guna menyederhanakan pengolahan teks. Pada kolom "Sebelum Case Folding", teks masih mengandung huruf kapital yang tidak konsisten, seperti "Gw" dan "PLEASE MOONTON KENAPA SAYA GAK BISA LOGIN???". Setelah proses case folding di kolom "Setelah Case Folding", semua huruf diubah menjadi huruf kecil sehingga lebih seragam, seperti "gw udah lama main mobeli legends" dan "please moonton kenapa saya gak bisa login???". Proses ini penting dalam pemrosesan teks untuk memastikan bahwa perbedaan kapitalisasi tidak mempengaruhi analisis lebih lanjut.

D. Stopwords Removal

Untuk melakukan proses stopwords removal diperlukan daftar kata-kata yang akan dihapus dari teks. Daftar ini dapat berupa daftar bawaan dari library atau dapat dibuat sendiri berdasarkan bahasa dan konteks yang relevan. Tahapan ini penting dalam analisis teks karena dapat meningkatkan kualitas hasil analisis dan mempercepat proses komputasi dengan mengurangi jumlah kata yang harus diproses. Namun, perlu diingat bahwa dalam beberapa konteks atau jenis analisis, stopwords mungkin tetap relevan dan tidak perlu dihapus sepenuhnya. Tabel 3 Menunjukkan hasil penghapusan stopword.

Tabel 3. Hasil proses Remove stopwords

Sebelum Stopwords Removal	Sesudah Stopwords Removal
saya sudah lama main mobeli legends dan saya selalu stack di rank epic dan Mobile Legends dan setiap satu bintang lagi naik saya selau kalah sampai saya mau pindah game tapi di pertandingan terakhir dikasih menang tolong kasih musuh yang tidak pro amat ngasih musuh yang sebanding lah jangan ngasih tim keturunan	saya sudah lama main mobeli Mobile Legends dan saya selalu stack di rank epic dan Mobile Legends dan setiap satu bintang lagi naik saya selalu kalah sampai saya mau pindah game tapi di pertandingan terakhir dikasih menang tolong kasih musuh yang tidak pro amat kasih musuh yang sebanding lah jangan kasih tim keturunan immortal

immortal tolong kurangi dark sistem	tolong kurangi dark sistem
tolong moonton kenapa saya tidak bisa login kenapa bisa gameplay tidak di temukan di server astaga	tolong moonton kenapa saya tidak bisa login kenapa bisa gameplay tidak di temukan di server astaga
gamenya bagus banyak banget hero bagus	gamenya bagus banyak banget hero bagus
saya minta tolong mobilelegend bang bang kalo kasih musuh yang setara	saya minta tolong Mobile Legends bang bang kalo kasih musuh yang setara

Tabel 3 menunjukkan hasil proses penghapusan stopword pada teks keluhan dan permintaan pengguna terhadap game Mobile Legends. Proses ini menghasilkan teks yang lebih ringkas dan fokus, hanya memuat kata-kata yang relevan dan penting untuk analisis. Kata-kata yang tidak memiliki makna signifikan, seperti "saya," "dan," "di," dan "yang," telah dihapus. Dimana kalimat yang awalnya "saya sudah lama main Mobile Legends dan saya selalu stack di rank epic" menjadi "sudah lama main Mobile Legends selalu stack rank epic." Penghapusan stopword ini menyederhanakan teks, memudahkan pengelompokan keluhan atau saran pengguna secara lebih efektif dan efisien, tanpa gangguan dari kata-kata yang tidak relevan.

E. Tokenizing

Tokenizing dimulai dengan memisahkan teks berdasarkan spasi. Dengan memecah teks menjadi token, mesin dapat lebih mudah melakukan analisis dan ekstraksi informasi dari teks. Tabel 4 merupakan hasil proses tokenizing.

Tabel 4. Hasil proses tokenizing

Sebelum Tokenizing	Sesudah Tokenizing
saya sudah lama main mobeli Mobile Legends dan saya selalu stack di rank epic dan Mobile Legends dan setiap satu bintang lagi naik saya selalu kalah sampai saya mau pindah game tapi di pertandingan terakhir dikasih menang tolong kasih musuh yang tidak pro amat kasih musuh yang sebanding lah jangan kasih tim keturunan immortal tolong kurangi dark sistem	['saya', 'sudah', 'lama', 'main', 'mobeli', 'mobile', 'legends', 'dan', 'saya', 'selalu', 'stack', 'di', 'rank', 'epic', 'dan', 'mobile', 'legends', 'dan', 'setiap', 'satu', 'bintang', 'lagi', 'naik', 'saya', 'selalu', 'kalah', 'sampai', 'saya', 'mau', 'pindah', 'game', 'tapi', 'di', 'pertandingan', 'terakhir', 'dikasih', 'menang', 'tolong', 'kasih', 'musuh', 'yang', 'tidak', 'pro', 'amat', 'kasih', 'musuh', 'yang', 'sebanding', 'lah', 'jangan', 'kasih', 'tim', 'keturunan', 'immortal', 'tolong', 'kurangi', 'dark', 'sistem']

	'jangan', 'kasih', 'tim', 'keturunan', 'immortal', 'tolong', 'kurangi', 'dark', 'sistem']
tolong moonton kenapa saya tidak bisa login kenapa bisa gameplay tidak di temukan di server astaga	['tolong', 'moonton', 'kenapa', 'saya', 'tidak', 'bisa', 'login', 'kenapa', 'bisa', 'gameplay', 'tidak', 'di', 'temukan', 'di', 'server', 'astaga']
gamenya bagus banyak banget hero bagus	['gamenya', 'bagus', 'banyak', 'banget', 'hero', 'bagus']
saya minta tolong Mobile Legends bang kalo kasih musuh yang setara	['saya', 'minta', 'tolong', 'mobile', 'legends', 'bang', 'bang', 'kalo', 'kasih', 'musuh', 'yang', 'setara']

Tabel 4 menunjukkan Proses tokenisasi ini memecah teks menjadi unit-unit kata yang lebih kecil, atau token, yang memudahkan analisis lebih lanjut. Dimana kalimat "saya sudah lama main Mobile Legends dan saya selalu stack di rank epic" dipecah menjadi token-token individu: ['saya', 'sudah', 'lama', 'main', 'mobile', 'legends', 'dan', 'saya', 'selalu', 'stack', 'di', 'rank', 'epic']. Hasil ini menunjukkan semua kata dalam teks asli dalam bentuk terpisah, termasuk kata-kata umum dan stopword. Tokenisasi ini membantu dalam pemrosesan teks lebih lanjut, seperti analisis frekuensi kata, identifikasi pola, dan pengelompokan kata-kata kunci yang sering muncul dalam keluhan atau permintaan pengguna.

F. Stemming

Langkah selanjutnya adalah stemming di mulai dengan kata-kata disederhanakan menjadi bentuk dasar mereka. Proses ini bisa dilakukan dengan menggunakan aturan-aturan linguistik atau dengan pendekatan berbasis aturan yang lebih umum

Tabel 5. Hasil proses Stemming

Sebelum Stemming	Sesudah Stemming
['saya', 'sudah', 'lama', 'main', 'mobeli', 'mobile', 'legends', 'dan', 'saya', 'selalu', 'stack', 'di', 'rank', 'epic', 'dan', 'mobile', 'legends', 'dan', 'setiap', 'satu', 'bintang', 'lagi', 'naik', 'saya', 'selalu', 'kalah', 'sampai', 'saya', 'mau', 'pindah', 'game', 'tapi', 'di', 'pertandingan',	saya sudah lama main mobe mobil legend dan saya lalu stack di rank epic dan mobil legend dan tiap satu bintang lagi naik saya lalu kalah sampai saya mau pindah game tapi di tanding terakhir dikasih menang tolong kasih musuh yang tidak pro amat kasih musuh yang seband lah

'terakhir', 'dikasih', 'menang', 'tolong', 'kasih', 'musuh', 'yang', 'tidak', 'pro', 'amat', 'kasih', 'musuh', 'yang', 'sebanding', 'lah', 'jangan', 'kasih', 'tim', 'keturunan', 'immortal', 'tolong', 'kurangi', 'dark', 'sistem']	jangan kasih tim turunan immort tolong kurangi dark sistem
['tolong', 'moonton', 'kenapa', 'saya', 'tidak', 'bisa', 'login', 'kenapa', 'bisa', 'gameplay', 'tidak', 'di', 'temukan', 'di', 'server', 'astaga']	tolong moonton napa saya tidak bisa login napa bisa gameplay tidak di temu di server astaga
['gamenya', 'bagus', 'banyak', 'banget', 'hero', 'bagus']	game bagus banyak banget hero bagus
['saya', 'minta', 'tolong', 'mobile', 'legends', 'bang', 'bang', 'kalo', 'kasih', 'musuh', 'yang', 'setara']	saya minta tolong mobil legend bang bang kalo kasih musuh yang setara.

Tabel 5 menunjukkan proses mengubah kata-kata ke bentuk dasarnya untuk menyederhanakan dan mengkonsolidasikan berbagai bentuk kata. Dimana kalimat "saya sudah lama main Mobile Legends dan saya selalu stack di rank epic" setelah stemming menjadi "saya sudah lama main mobe mobil legend dan saya lalu stack di rank epic dan mobil legend." Proses ini mengurangi variasi kata sehingga analisis teks menjadi lebih fokus dan efisien. Dengan mengubah kata-kata ke bentuk dasar, stemming membantu mengidentifikasi tema atau keluhan utama tanpa terpengaruh oleh berbagai bentuk kata yang berbeda. Hasil ini berguna untuk analisis teks yang lebih mendalam, seperti pemetaan pola keluhan atau permintaan yang sering muncul di antara pengguna.

G. Pelabelan Data

Pelabelan ini dilakukan dengan menambahkan atribut label terlebih dahulu kemudian melakukan pelabelan manual menggunakan angka 0 dan 1 dimana 0 adalah negatif dan 1 sama dengan positif. Selanjutnya proses pelabelan akan mengubah nilai 0 menjadi negatif dan nilai 1 menjadi positif. Tujuan dari proses ini adalah untuk meningkatkan interpretabilitas data serta memungkinkan analisis yang lebih mudah dan lebih intuitif, terutama dalam konteks evaluasi sentimen. Proses pelabelan ini merupakan langkah penting dalam *preprocessing* data untuk berbagai aplikasi analitik dan *machine learning*. Hasil proses pelabelan data dapat dilihat pada tabel 6.

Tabel 6. Hasil proses Pelabelan

Content	Label Numerik	Label Teks
---------	---------------	------------

saya sudah lama main mobe mobil <i>legend</i> dan saya lalu stack di <i>rank epic</i> dan mobil <i>legend</i> dan tiap satu bintang lagi naik saya lalu kalah sampai saya mau pindah game tapi di tanding terakhir dikasih menang tolong kasih musuh yang tidak pro amat kasih musuh yang seband lah jangan kasih tim turunan immort tolong kurangi dark sistem	0	Negatif
tolong moonton napa saya tidak bisa login napa bisa gameplay tidak di temu di server astaga	0	Negatif
game bagus banyak banget hero bagus	1	Positif
saya minta tolong mobil <i>legend</i> bang bang kalo kasih musuh yang setara	1	Positif

Tabel 6 menunjukkan hasil proses pelabelan sentimen pada komentar pengguna mengenai permainan "*Mobile Legends*". Setiap komentar awalnya diberi label numerik "0" untuk sentimen negatif dan "1" untuk sentimen positif. Label numerik tersebut kemudian dikonversi menjadi label teks, dengan "0" menjadi "Negatif" dan "1" menjadi "Positif". Konversi ini dilakukan untuk memudahkan interpretasi hasil analisis sentimen, sehingga lebih intuitif dan mudah dipahami. komentar yang berisi keluhan tentang masalah teknis atau pengalaman bermain yang kurang menyenangkan diberi label "Negatif", sementara komentar yang memuji aspek positif permainan diberi label "Positif".

H. K Nearest Neighbor

Metode *K-Nearest Neighbor (KNN)* adalah algoritma dalam *machine learning* yang digunakan untuk klasifikasi dan regresi. *K-NN* beroperasi dengan mencari sejumlah tetangga terdekat (K) dari sebuah titik data yang belum diklasifikasikan atau diprediksi, kemudian mengambil mayoritas kelas atau nilai target terdekat. Metode ini sederhana dan mudah diimplementasikan, namun sangat sensitif terhadap pemilihan parameter K [10]. Menurut A. D. Adhi Putra, Algoritma *k-nearest neighbors* memiliki tahapan:

1. Menentukan parameter k (jumlah tetangga paling dekat).
2. Tentukan bobot untuk setiap term dengan menggunakan Term Weighting *TF-IDF*.
3. Hitung kemiripan antar dokumen dengan menggunakan cosine similarity.
4. Urutkan hasil perhitungan cosine similarity dari besar ke kecil.

5. Ambil sebanyak K yang paling tinggi kemiripannya dengan dokumen yang diklasifikasikan, tentukan kelasnya [10].

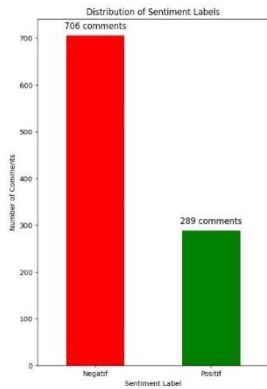
Analisis sentimen pada aplikasi *Mobile Legends* menggunakan metode *K-Nearest Neighbors (KNN)* melibatkan sebuah proses teks ulasan pengguna aplikasi tersebut untuk mengidentifikasi sentimen positif dan negatif terhadap aplikasi *Mobile Legends*. Proses awal di mulai dengan pemrosesan teks, termasuk *case folding*, normalisasi, *tokenizing*, *stemming* dan pelabelan data. Setiap ulasan kemudian direpresentasikan sebagai vektor fitur dalam ruang dimensi tinggi. Metode *KNN* yang mengklasifikasikan data berdasarkan pendekatan dengan data yang sudah diberi label. *KNN* memberikan klasifikasi sentimen yang efektif, membantu pengembang memahami umpan balik pengguna dan meningkatkan pengalaman bermain.

Tabel 7. Penentuan nilai K terbaik *KNN*

Nilai K	Akurasi
K = 8	64.82%
K = 10	65.32%
K = 11	65.82%
K = 17	64.82%

Dari hasil performa model *KNN* dengan variasi nilai K (K=8, 10, 11, dan 17), disimpulkan bahwa nilai K=11 merupakan nilai K yang paling optimal karena menghasilkan nilai akurasi yang tinggi dari nilai K lainnya. Nilai K yang rendah maupun tinggi menunjukkan penurunan satu atau lebih metrik evaluasi, menandakan bahwa nilai K=11 adalah pilihan terbaik untuk model *KNN* dalam konteks ini, menghasilkan prediksi yang lebih andal dan konsisten. Rasio data yang digunakan untuk pelatihan dan pengujian adalah 80 banding 20 atau dengan rasio 0.2. Artinya, 80% data digunakan untuk melatih model, sedangkan 20% sisanya digunakan untuk menguji performa model. Rasio ini memungkinkan model untuk belajar dari jumlah data yang cukup besar dan tetap diuji secara efektif dengan subset data yang memadai, sehingga hasil evaluasi lebih *representatif* dan dapat diandalkan. Dengan demikian, pemilihan K=11 pada rasio 80:20 ini memberikan keseimbangan optimal antara berbagai metrik evaluasi, memastikan model *KNN* yang digunakan mampu memberikan prediksi yang lebih baik dan konsisten. Model *KNN* yang dibangun menunjukkan performa yang kurang baik pada kelas negatif. Meskipun model ini optimal secara umum, terdapat kelemahan dalam mengklasifikasikan data dengan sentimen negatif. Hal ini dapat mengakibatkan ketidakseimbangan dalam hasil prediksi dan menandakan perlunya pendekatan atau penyesuaian lebih lanjut untuk meningkatkan akurasi prediksi pada kelas negatif. Dengan demikian, meskipun K=11 dipilih sebagai pilihan terbaik dalam parameter yang diberikan, namun perlu diperhatikan pada kemampuan model dalam menangani kelas negatif untuk memastikan kinerja yang lebih seimbang dan dapat diandalkan di semua kategori.

Dari 995 data ulasan yang telah diambil dari *platformGoogle Play Store* menggunakan metode *Web scraping* dihasilkan jumlah data positif dan negatif sebagai berikut:

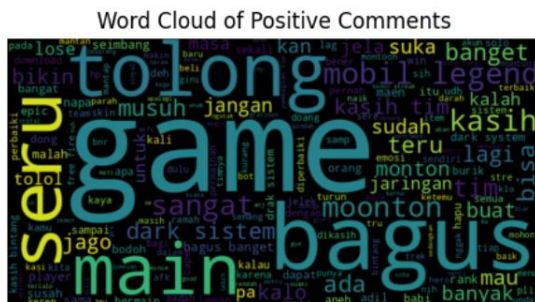


Gambar 2. *Proporsi Data*

Gambar 2 menunjukkan hasil bahwa tanggapan negatif memiliki data yang lebih banyak dibandingkan tanggapan positif. Hal tersebut menjadi indikator bahwa terdapat beberapa masalah atau kekurangan yang dirasakan oleh pengguna aplikasi tersebut sehingga perlu untuk diperhatikan oleh pengembang aplikasi *Mobile Legends*. Pemahaman terkait distribusi sentimen ini sangat penting untuk pengembang aplikasi agar dapat dilakukan perbaikan dan peningkatan aplikasi, agar dapat melakukan perbaikan dan peningkatan kualitas aplikasi berdasarkan umpan balik dari penggunaanya.

J. Word Cloud

Visualisasi *wordcloud* ini membantu dalam mengidentifikasi kata-kata kunci yang paling signifikan dan memperoleh wawasan yang berguna dalam analisis sentimen pada dataset tersebut[5]. *Word cloud* pada analisis sentimen aplikasi *Mobile Legends* digunakan untuk memahami pendapat pengguna pada aplikasi tersebut. Visualisasi data teks ini akan menampilkan kata yang paling sering muncul dalam ulasan aplikasi tersebut. Semakin besar ukuran *font* tersebut, maka semakin besar pula frekuensi kemunculan kata tersebut. Gambar 3 dan 4 adalah *word cloud* pada analisis sentiment aplikasi *Mobile Legends*.



Gambar 3. *Word cloud* Analisis Sentimen positif

Gambar 3 menunjukkan kata-kata yang paling menonjol dalam komentar positif pengguna mengenai permainan "*Mobile Legends*". Kata-kata seperti "game," "tolong," "main," dan "bagus" menonjol, menunjukkan bahwa banyak pengguna memiliki permintaan khusus serta menganggap permainan tersebut menarik dan berkualitas. Kata-kata seperti "bagus," "seru," dan "sangat" memberikan gambaran bahwa pengguna mengalami bermain yang diberikan dianggap positif oleh pengguna. Selain itu, kata-kata seperti "tolong," "mobil," "*legend*," dan "moonton" juga sering muncul, menunjukkan elemen atau fitur spesifik dari permainan yang mendapatkan perhatian lebih. Kata-kata ini mengindikasikan aspek-aspek tertentu dari permainan yang dihargai oleh pengguna, termasuk fitur-fitur dalam permainan, perusahaan pengembang (Moonton), dan pengalaman bermain yang menyenangkan. Secara keseluruhan, *word cloud* ini mengilustrasikan umpan balik yang positif dan kepuasan pengguna terhadap permainan tersebut.



Gambar 4. *Word cloud* Analisis Sentimen negatif

Gambar 4 menunjukkan *word cloud* dari komentar negatif terkait permainan "*Mobile Legends*". Kata-kata yang paling menonjol seperti "game," "main," "moonton," dan "bagus" mencerminkan isu yang sering diangkat oleh pengguna. Meskipun "bagus" muncul dalam *word cloud*, konteksnya kemungkinan besar negatif atau kritis, misalnya dalam perbandingan atau ekspektasi yang tidak terpenuhi. Kata-kata seperti "tolong," "tim," "kalah," "sistem," dan "bikin" mengindikasikan keluhan spesifik terkait kinerja tim, kekalahan dalam permainan, atau kritik terhadap pengembang (Moonton). Selain itu, kata-kata seperti "sistem," "dark," "rank," dan "jaringan" menunjukkan masalah teknis dan ketidakpuasan dalam pengalaman bermain. Gambar ini secara keseluruhan menggambarkan berbagai aspek dan fitur permainan yang menjadi sumber ketidakpuasan bagi pengguna, memberikan wawasan tentang area yang perlu diperbaiki.

4. SIMPULAN

Hasil dari penelitian tentang analisis sentimen terhadap aplikasi Mobile Legends menggunakan metode KNN, dapat diperoleh melalui serangkaian tahapan, mulai dari pengumpulan data, pemrosesan teks seperti Case folding, normalisasi, tokenizing, dan stemming, pelabelan data, serta penerapan metode K-Nearest Neighbors (KNN) untuk klasifikasi sentimen. Nilai K terbaik yang dipilih adalah 11 dengan hasil akurasi 65,82%. Selain itu, hasil menunjukkan bahwa distribusi sentiment negatif yang lebih tinggi, menandakan adanya masalah atau kekurangan yang perlu diperhatikan oleh pengembang. Dengan demikian, analisis ini memberikan wawasan berharga bagi pengembang untuk meningkatkan kualitas dan pengalaman pengguna aplikasi Mobile Legends.

DAFTAR PUSTAKA

- LASAMAHU FEBRIAN POLIANDRES, "ANALISIS SENTIMEN PENGGUNA GAME ONLINE MOBILE LEGENDS DENGAN NAÏVE BAYES," 2021.
- M. K. Alamsyah and N. Pratiwi, "ANALISIS SENTIMEN TERKAIT OPINI MASYARAKAT TERHADAP PERKEMBANGAN E-SPORT MOBILE DI INDONESIA MENGGUNAKAN K NEAREST NEIGHBOR," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 349–359, Feb. 2024, doi: 10.29100/jipi.v9i1.4927.
- Ramadhan Restu, Afdal M, Permana Inggih, and Jazman Muhammad, "Analisis Sentimen pada Ulasan Aplikasi Maxim di Google Play Store dengan K-Nearest Neighbor," 2023.
- Asro'i Arief and Februariyanti Herny, "Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor," *JURNAL KHATULISTIWA INFORMATIKA*, vol. 10, no. 1, pp. 17–24, 2022.
- I. Habib Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 8, no. 3, pp. 302–307, 2023.
- M. K. Alamsyah and N. Pratiwi, "ANALISIS SENTIMEN TERKAIT OPINI MASYARAKAT TERHADAP PERKEMBANGAN E-SPORT MOBILE DI INDONESIA MENGGUNAKAN K NEAREST NEIGHBOR," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 349–359, Feb. 2024, doi: 10.29100/jipi.v9i1.4927.
- R. Q. Rohmansa, N. Pratiwi, and M. J. Palepa, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI DISCORD MENGGUNAKAN METODE K-NEAREST NEIGHBOR," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 368–378, Feb. 2024, doi: 10.29100/jipi.v9i1.4943.
- P. Astuti and N. Nuris, "Penerapan Algoritma KNN Pada Analisis Sentimen Review Aplikasi Peduli Lindungi," 2022. [Online]. Available: <https://www.pedulilindungi.id/>
- S. Rahayu, Y. MZ, J. E. Bororing, and R. Hadiyat, "Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP," *Edumatic: Jurnal Pendidikan Informatika*, vol. 6, no. 1, pp. 98–106, Jun. 2022, doi: 10.29408/edumatic.v6i1.5433.
- I. Alpian Novansyah and T. Suprpti, "ANALISIS SENTIMEN PENGGUNA PADA APLIKASI LINGOKIDS MENGGUNAKAN METODE K-NEAREST NEIGHBOUR," 2023.