



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 17028- 17037

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Algoritma Naïve Bayes untuk Analisis Sentimen Pengguna Game Honor of King Berdasarkan Ulasan di Google Play Store

Mochammad Dayke Xavier Lumankun¹, Ghofar Taufik²

¹²Informatika, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

¹mochammaddayke@gmail.com , ²ghofar.gft@bsi.ac.id

Abstrak

Game Honor of Kings merupakan salah satu permainan mobile yang banyak memperoleh ulasan dari pengguna di Google Play Store. Banyaknya ulasan tersebut mengandung berbagai opini yang dapat dimanfaatkan untuk mengetahui tingkat kepuasan pengguna, namun analisis secara manual memerlukan waktu yang lama dan kurang efisien. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna game Honor of Kings berbahasa Indonesia secara otomatis menggunakan metode klasifikasi berbasis machine learning. Data penelitian diperoleh melalui teknik web scraping dari Google Play Store. Tahapan penelitian meliputi text preprocessing yang terdiri atas cleaning, case folding, tokenizing, stopword removal, dan stemming untuk menghasilkan data yang bersih dan terstruktur. Selanjutnya dilakukan ekstraksi fitur menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) sehingga data teks dapat direpresentasikan dalam bentuk numerik. Proses klasifikasi dilakukan menggunakan algoritma Multinomial Naïve Bayes dengan pembagian data latih dan data uji sebesar 80:20. Evaluasi model dilakukan menggunakan confusion matrix untuk mengukur kinerja klasifikasi. Hasil penelitian menunjukkan bahwa model Multinomial Naïve Bayes mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi sebesar 79%, sehingga cukup efektif dalam membedakan sentimen positif dan negatif. Selain itu, hasil analisis menunjukkan bahwa mayoritas ulasan pengguna memiliki sentimen positif terhadap game Honor of Kings. Temuan penelitian ini diharapkan dapat menjadi masukan bagi pengembang dalam meningkatkan kualitas layanan dan pengalaman pengguna.

Kata Kunci: Analisis Sentimen, Naïve Bayes, Web Scraping, Honor of Kings.

1. Latar Belakang

Perkembangan teknologi digital telah mendorong pertumbuhan pesat industri game, khususnya berbasis perangkat bergerak (*mobile gaming*). Kemudahan akses melalui smartphone menjadikan *mobile gaming* sebagai salah satu sektor hiburan digital dengan jumlah pengguna terbesar di dunia. Salah satu gim yang mendapatkan perhatian luas adalah *Honor of Kings*, yaitu permainan bergenre *Multiplayer Online Battle Arena* (MOBA) yang dikembangkan oleh TiMi Studio Group dan dirilis secara global oleh Level Infinite. Game ini berhasil menarik minat jutaan pemain karena kualitas grafis yang baik, mekanik permainan yang kompetitif, serta variasi karakter atau *hero* yang beragam [1].

Sebagai platform distribusi aplikasi digital, Google Play Store menyediakan fitur ulasan pengguna yang memungkinkan pemain memberikan penilaian dan pengalaman mereka setelah menggunakan suatu aplikasi. Ulasan tersebut mengandung berbagai opini yang bersifat positif maupun negatif mengenai berbagai aspek permainan, seperti kualitas grafis, stabilitas sistem, performa server, keseimbangan karakter, hingga fitur-fitur yang tersedia [2]. Informasi yang terkandung dalam ulasan pengguna dapat menjadi sumber data yang berharga bagi pengembang untuk mengevaluasi kualitas aplikasi serta meningkatkan pengalaman pengguna.

Namun, jumlah ulasan pengguna yang sangat besar dan terus bertambah setiap waktu menyebabkan proses analisis secara manual menjadi tidak efektif dan memerlukan waktu yang lama. Oleh karena itu, diperlukan suatu pendekatan otomatis yang mampu mengolah dan menganalisis opini pengguna secara sistematis. Salah satu pendekatan yang dapat digunakan adalah analisis sentimen, yaitu teknik dalam *text mining* yang bertujuan untuk mengidentifikasi dan mengelompokkan opini pengguna ke dalam kategori sentimen tertentu, seperti positif, negatif, atau netral.

Dalam proses analisis sentimen, berbagai metode klasifikasi dapat digunakan, salah satunya adalah algoritma Naïve Bayes. Metode ini merupakan algoritma klasifikasi berbasis probabilitas yang didasarkan pada Teorema

Bayes dan dikenal memiliki performa yang baik dalam pengolahan data teks[3]. Selain itu, Naïve Bayes memiliki keunggulan dalam hal kesederhanaan, efisiensi komputasi, serta kemampuan dalam menangani dataset berukuran besar. Meskipun demikian, evaluasi terhadap tingkat akurasi model tetap diperlukan untuk memastikan bahwa proses klasifikasi sentimen dapat dilakukan secara tepat dan andal.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pengguna game *Honor of Kings* pada platform Google Play Store menggunakan algoritma Naïve Bayes. Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap kualitas permainan serta menjadi bahan evaluasi bagi pengembang dalam meningkatkan kualitas layanan dan pengalaman bermain bagi pengguna.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dalam proses analisis data. Meskipun data awal berupa teks yang berasal dari ulasan pengguna, data tersebut terlebih dahulu melalui tahapan *text preprocessing* untuk membersihkan dan menormalisasi data agar dapat diolah secara komputasional. Tahapan *preprocessing* meliputi proses *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Selanjutnya, teks diubah menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sehingga menghasilkan fitur yang dapat digunakan dalam proses pembelajaran mesin. Representasi fitur tersebut kemudian digunakan oleh algoritma klasifikasi untuk melakukan analisis sentimen terhadap ulasan pengguna. Hasil dari proses klasifikasi selanjutnya dievaluasi menggunakan beberapa metrik kuantitatif, yaitu akurasi, *confusion matrix*, serta *classification report* yang mencakup nilai *precision*, *recall*, dan *f1-score* untuk mengukur kinerja model yang dihasilkan.

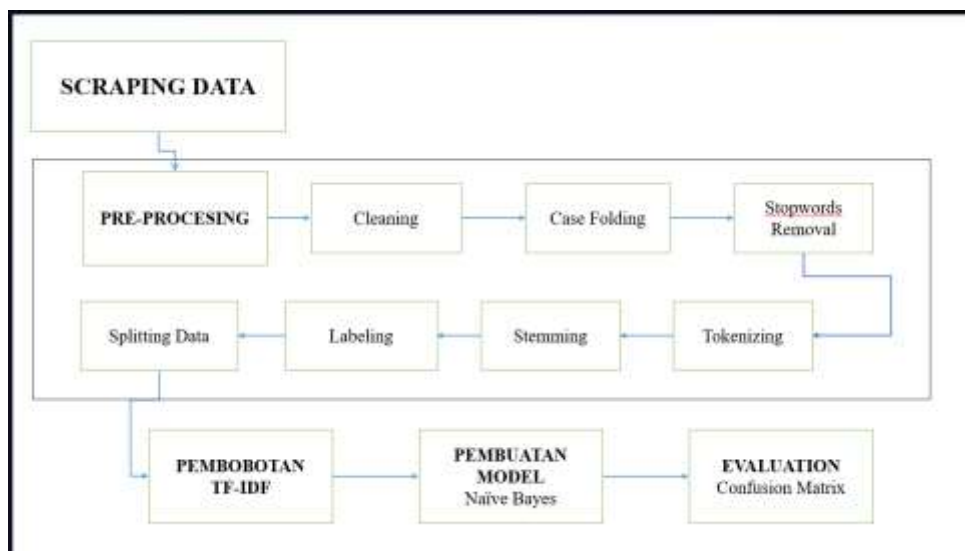
2.1 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah *web scraping*, yaitu metode otomatis untuk mengambil data dari halaman web dengan bantuan skrip pemrograman [4]. Dalam penelitian ini, *web scraping* diterapkan pada halaman aplikasi *Honor of Kings* di Google Play Store untuk memperoleh data ulasan pengguna. Data yang diambil meliputi tanggal ulasan, teks ulasan, serta rating yang diberikan oleh pengguna terhadap aplikasi tersebut [4].

Proses pengambilan data dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka *google-play-scraper* untuk mengekstraksi konten dari halaman Google Play Store. Melalui pustaka tersebut, data ulasan pengguna dapat dikumpulkan secara otomatis dan terstruktur. Setelah proses *scraping* selesai, data yang diperoleh kemudian disimpan dalam bentuk file CSV untuk memudahkan proses pengolahan data pada tahapan selanjutnya, yaitu *preprocessing* dan analisis sentimen menggunakan algoritma Naïve Bayes [5].

2.2 Teknik Analisis Data

Alur pada penelitian ini ditunjukkan pada Gambar 1 berikut :



Gambar 1. Alur Penelitian

Teknik analisis data dalam penelitian ini dilakukan melalui beberapa tahapan sistematis yang dimulai dari proses pre-processing data hingga evaluasi performa model klasifikasi.

Tahap *preprocessing* merupakan proses penting dalam pengolahan data teks yang bertujuan untuk mengubah data teks mentah menjadi bentuk yang lebih terstruktur dan siap digunakan dalam proses analisis [6]. Pada tahap ini dilakukan beberapa proses pengolahan data untuk membersihkan serta menormalisasi teks agar dapat diolah oleh algoritma pembelajaran mesin [7].

Tahap *cleaning* merupakan salah satu langkah penting dalam proses *text preprocessing* pada analisis sentimen. Data ulasan yang diperoleh dari Google Play Store umumnya masih mengandung berbagai bentuk *noise*, seperti karakter khusus, tanda baca berlebih, tautan (*URL*), angka, maupun simbol lain yang tidak memiliki kontribusi terhadap proses analisis [8]. Oleh karena itu, proses *cleaning* dilakukan untuk menghapus elemen-elemen tersebut sehingga teks menjadi lebih bersih dan relevan. Tahap ini bertujuan untuk meningkatkan kualitas data serta membantu menghasilkan representasi teks yang lebih akurat dalam proses analisis sentimen pada tahap selanjutnya.

Tahap *case folding* merupakan proses normalisasi teks yang bertujuan untuk menyeragamkan bentuk huruf dalam data [8]. Pada tahap ini, seluruh huruf pada teks diubah menjadi huruf kecil (*lowercase*). Proses ini dilakukan dengan mengonversi seluruh karakter huruf A–Z menjadi a–z sehingga tidak terdapat perbedaan antara huruf besar dan huruf kecil dalam proses analisis [9]. Dengan demikian, kata yang memiliki makna sama tetapi ditulis dengan variasi huruf besar dan kecil dapat diperlakukan sebagai satu bentuk kata yang sama dalam proses pengolahan data teks.

Tahap *stopword removal* merupakan proses penghapusan kata-kata umum yang tidak memiliki makna penting atau kontribusi signifikan dalam suatu kalimat [9][10]. Kata-kata tersebut biasanya sering muncul dalam teks, seperti kata penghubung atau kata depan, namun tidak memberikan informasi yang relevan terhadap proses analisis sentimen. Oleh karena itu, proses *stopword removal* dilakukan untuk mengurangi jumlah kata yang tidak diperlukan sehingga dapat meningkatkan efektivitas dan kualitas proses analisis data teks pada tahap selanjutnya.

Tahap *tokenizing* merupakan proses pemisahan teks menjadi unit-unit yang lebih kecil yang disebut sebagai *token* [11]. Pada tahap ini, teks dipecah menjadi kata-kata individu sehingga memudahkan proses analisis lebih lanjut. Dalam proses *tokenizing*, berbagai elemen seperti kata, angka, simbol, tanda baca, maupun entitas lain yang terdapat dalam teks dapat diperlakukan sebagai *token*. Proses ini bertujuan untuk mengubah teks menjadi struktur data yang lebih terorganisasi sehingga dapat diproses secara efektif dalam tahap analisis sentimen.

Tahap *stemming* merupakan proses mengubah kata menjadi bentuk dasar atau akar kata dengan menghilangkan imbuhan, baik awalan, sisipan, maupun akhiran [12]. Proses ini bertujuan untuk menyederhanakan variasi kata yang memiliki makna serupa sehingga dapat direpresentasikan dalam satu bentuk dasar. Dengan demikian, tahap *stemming* membantu mengurangi kompleksitas kata dalam data teks serta memudahkan proses analisis pada tahap selanjutnya.

Tahap *labeling* atau pelabelan merupakan proses penting dalam analisis sentimen yang bertujuan untuk memberikan kategori tertentu pada data ulasan berdasarkan sentimen yang terkandung di dalamnya [13]. Proses pelabelan dilakukan untuk menentukan nilai sentimen dari setiap ulasan dengan memanfaatkan nilai skor (*rating*) yang diberikan oleh pengguna pada aplikasi di Google Play Store. Dalam penelitian ini, ulasan dengan nilai skor 1–2 diklasifikasikan sebagai sentimen negatif, sedangkan ulasan dengan nilai skor 4–5 diklasifikasikan sebagai sentimen positif. Proses pelabelan ini bertujuan untuk menghasilkan dataset yang telah memiliki kategori sentimen sehingga dapat digunakan dalam proses pelatihan dan pengujian model klasifikasi pada tahap analisis selanjutnya.

Tahap *data splitting* merupakan proses pembagian dataset menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*) [14]. Data latih digunakan untuk melatih model klasifikasi agar dapat mempelajari pola dari data yang tersedia, sedangkan data uji digunakan untuk mengevaluasi kinerja model yang telah dibangun. Melalui proses pembagian data ini, model dapat diuji kemampuannya dalam mengklasifikasikan data baru yang belum pernah digunakan pada tahap pelatihan sehingga hasil evaluasi yang diperoleh menjadi lebih objektif.

Data latih digunakan untuk melatih model klasifikasi Naïve Bayes agar mampu mempelajari pola dari dataset yang tersedia [12], sedangkan data uji digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya. Pada penelitian ini, dataset dibagi dengan proporsi 80% sebagai data latih dan 20% sebagai data uji (*test size* = 0,20). Pembagian data ini bertujuan untuk memastikan bahwa proses pelatihan dan evaluasi model dapat dilakukan secara optimal sehingga kinerja model dalam melakukan klasifikasi sentimen dapat diukur secara lebih objektif.

Pada tahap selanjutnya dilakukan proses pembobotan kata untuk menentukan tingkat kepentingan setiap kata dalam dokumen. Dalam penelitian ini, metode pembobotan yang digunakan adalah *Term Frequency–Inverse Document Frequency* (TF-IDF) [15]. Metode TF-IDF merupakan teknik pembobotan yang menggabungkan nilai *term frequency* (TF) dan *inverse document frequency* (IDF) untuk mengukur tingkat kepentingan suatu kata dalam kumpulan dokumen.

Term frequency digunakan untuk menunjukkan seberapa sering suatu kata muncul dalam sebuah dokumen, sedangkan *inverse document frequency* digunakan untuk mengukur tingkat keunikan kata tersebut dalam keseluruhan dokumen. Melalui metode TF-IDF, kata yang sering muncul dalam suatu dokumen tetapi jarang muncul pada dokumen lainnya akan memiliki bobot yang lebih tinggi. Proses pembobotan ini bertujuan untuk merepresentasikan data teks dalam bentuk numerik sehingga dapat digunakan sebagai fitur dalam proses klasifikasi menggunakan algoritma pembelajaran mesin.

Tahap selanjutnya adalah proses pembangunan model klasifikasi menggunakan algoritma Naïve Bayes. Metode ini diterapkan untuk melakukan analisis sentimen terhadap ulasan pengguna game *Honor of Kings* yang diperoleh dari Google Play Store. Algoritma Naïve Bayes merupakan metode klasifikasi berbasis probabilitas yang memanfaatkan prinsip Teorema Bayes untuk menentukan kemungkinan suatu data termasuk ke dalam kelas tertentu.

Melalui proses klasifikasi ini, setiap ulasan pengguna dianalisis berdasarkan fitur teks yang telah diekstraksi sebelumnya, kemudian dikelompokkan ke dalam kategori sentimen positif dan negatif. Model yang telah dilatih menggunakan data latih selanjutnya digunakan untuk memprediksi sentimen pada data uji sehingga dapat diketahui kinerja model dalam mengklasifikasikan opini pengguna terhadap game *Honor of Kings*.

Evaluasi merupakan tahap penting dalam mengukur kinerja model klasifikasi yang telah dibangun menggunakan algoritma Naïve Bayes. Pada penelitian ini, proses evaluasi dilakukan menggunakan metode *confusion matrix*. *Confusion matrix* merupakan teknik evaluasi yang digunakan untuk membandingkan hasil prediksi model dengan data aktual sehingga dapat diketahui tingkat ketepatan klasifikasi yang dihasilkan.

Melalui *confusion matrix*, beberapa metrik evaluasi dapat dihitung untuk menilai performa model, di antaranya akurasi, precision, recall, dan F1-score. Metrik-metrik tersebut digunakan untuk mengukur sejauh mana model mampu mengklasifikasikan data ulasan ke dalam kategori sentimen positif dan negatif secara tepat. Dengan menggunakan evaluasi ini, kinerja model dapat dianalisis secara lebih komprehensif dalam mengidentifikasi sentimen pada ulasan pengguna game *Honor of Kings*.

3. Hasil dan Diskusi

Hasil penelitian ini bertujuan untuk mengetahui tingkat akurasi algoritma Naïve Bayes dalam melakukan klasifikasi sentimen terhadap ulasan pengguna game *Honor of Kings* pada platform Google Play Store. Proses pengolahan, pengurutan, serta pengujian data dilakukan menggunakan perangkat lunak Google Colaboratory yang mendukung proses analisis data berbasis bahasa pemrograman Python. Melalui tahapan tersebut, diperoleh hasil yang menunjukkan prosedur serta tahapan perhitungan yang digunakan dalam model Naïve Bayes untuk mengklasifikasikan sentimen ulasan pengguna. Hasil analisis ini memberikan gambaran mengenai kemampuan model dalam mengidentifikasi dan mengelompokkan sentimen ulasan secara efektif ke dalam kategori positif dan negatif.

3.1 Pengumpulan Data

Proses pengambilan data berupa ulasan pengguna game *Honor of Kings* diperoleh melalui platform Google Play Store dengan menerapkan teknik *web scraping*. *Web scraping* merupakan metode yang digunakan untuk mengekstraksi informasi dari halaman website secara otomatis dan sistematis menggunakan bantuan skrip pemrograman. Melalui teknik ini, data ulasan pengguna dapat dikumpulkan dalam jumlah besar secara efisien. Gambar 2 berikut hasil web scraping data.

A	B	C	D
username	content	score	date
Muhammad Adnan Setiawan	min saran saya tampilan mact nya di upgrade lagi soal nya saya lat lat tampilan nya ngebosenin, gada inovasi lain kah? coba bikin ta	5	12/30/2025 23:26
Arga Putra	halo level infinite sejujurnya game ini udh bagus kok, dari grafis, effect, map, jaringan yg stabil, sistem match yg adil, skin setara deng	4	1/1/2026 15:15
Naya Putra	sebenarnya ini game bagus banget, bagus banget malah, tapi minusnya delay,dan grafik yang di turun untuk beberapa HP jadi tolong	3	1/1/2026 10:05
Erika Dewi Puspita	Aku kasih bintang 3, sebelumnya untuk keseluruhan game sudah bagus, banyak event, sedikit mudah kalo mau dapetin hero ataupun	3	12/27/2025 17:05
Gusti-roof Mudzakir	ntah kenapa semenjak update muncul bug, ada satu Hero yang tiba tiba skill nya jadi warna ungu, terus sekarang grafisnya jadi buruk	1	1/2/2026 16:44
Septi Ariyanti	Game ini punya potensi besar dari segi visual dan konsep, tapi sistem pertandingannya buruk. Pemain sering dipasangkan dengan tir	3	1/1/2026 8:17
Ulin	Dari segi game Sangat bagus Tapi tolong Perbaiki Bug nya Saya main Analang gerak sendiri gecer map sendiri tolong perbaiki Dan jug	1	1/1/2026 1:24
Muhammad Syauqi	dear, untuk Dev hok/tim pengembang untuk gameplay bagus,event menarik dan tidak membosankan, sistem rank menarik, hanya sz	3	12/31/2025 5:44
Praja Muhamads	tolong perbaiki grafik yang di optimalisasi, saya sangat tidak menikmati resolusi yang sekarang karna resolusi yang dulu lebih menari	1	12/29/2025 0:55
Santri wati	hmmm untuk developer atau timi tolong di perbaiki lagi soalnya aku main mau naik eh ti tim nya beban semua tolong hilang kan drag	3	12/28/2025 21:46
Khalid Mubarak	agak gimana ya, masa saya main di rank bawah teman setim lainnya computer/boot yang hanya bisa jalan jalan tanpa menyerang dan r	4	1/1/2026 12:45
Bilal Iisa	aku kesal banget, entah kenapa kalau main tiba-tiba jaringannya jelek. padahal buat buka aplikasi lainnya masih okay banget. jujur in	1	12/30/2025 11:53
Ibnu Muhajir	Tolong masalah jaringan di perbaiki lah, masa di tengah war tiba-tiba ping jadi 300, padahal koneksi lancar buat main game aov, youtubu	2	1/2/2026 14:11
Celintal Nikuzo	hampir perfect si, sv lumayan aman skrng ga kayak du yg lag brgt, bnyk ngasi skin gratis, gimmick per season nya bagus, eventnya ji	4	12/28/2025 15:01
Suryoadi Puro	min tolong untuk sinyal nya diperbaiki, padahal dulu pas awal HOK rlin global gw main nggak ngelag, tapi pas sekarang udah main lag	4	1/1/2026 26:31
umar sidqi	ayo lah masalah grafik di benerin... ya kali ulti angka kek ldi sapu... bagian hok yg dulu. tolong kembalikan pengaturan grafik sepe	1	12/30/2025 19:22
Ramadeya Pandya	dari keseluruhan game udh bagus, eventnya menarik hanya saja matching pertimnya terlalu tidak adil. sekiranya main rank/game cej	1	12/20/2025 5:45
LANG	pliss benerin sistem match nya, udah 2 minggu aku masih belum ada progress, kalian ga mau ngasih kompensasi apapun? timnya	1	1/1/2026 15:56
Duta Pwitra Taruna	tolong untuk developer nya setelah update pembaruan banyak terjadi crash dan patch-patch terutama pada saat login masuk intervi	3	12/28/2025 20:46
Cahaya Perak	Ini cuman aku yang error atau gimana? Setiap kali login selalu "Kesalahan server klien" ???? EdR : oke, sekarang ini sudah bisa login	2	12/30/2025 5:15
Muhammad arifhan Ardan	aku kasih bintang satu dulu soalnya gue udah agak lama main, tapi tiba tiba pas masuk login gak bisa masuk, malah ada tulisan "tesa	1	12/30/2025 2:15

Gambar 2. Hasil web scraping data

3.2 Preprocessing

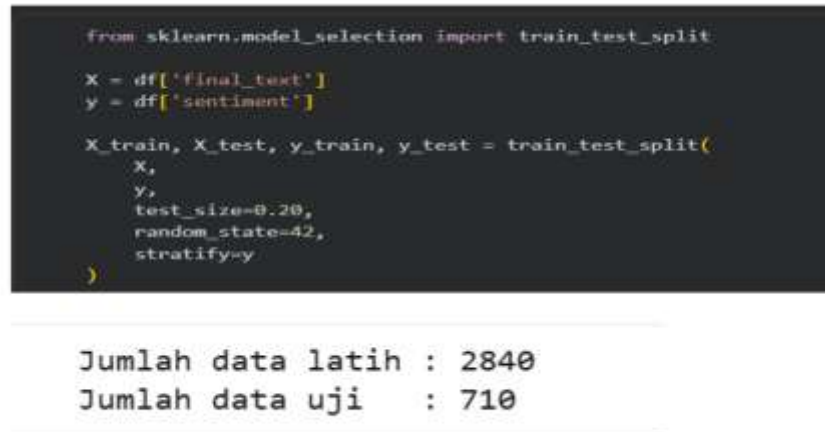
Proses *preprocessing* dalam penelitian ini meliputi beberapa tahapan, yaitu *cleaning* untuk menghapus karakter atau simbol yang tidak diperlukan, *case folding* untuk menyeragamkan huruf menjadi huruf kecil, *tokenizing* untuk memecah teks menjadi kata-kata, *stopword removal* untuk menghilangkan kata-kata umum yang tidak memiliki makna penting, serta *stemming* untuk mengubah kata menjadi bentuk dasar. Selain itu, dilakukan proses *labeling* untuk menentukan kategori sentimen pada data ulasan, serta *data splitting* untuk membagi dataset menjadi data latih dan data uji yang akan digunakan dalam proses klasifikasi. Gambar 3 berikut merupakan hasil dari seluruh alur Preprocessing.

clean_text	case_folding	stopword_removed	tokens	stemmed
min saran saya tampilan r min saran saya tam; min saran tampilan mact ny	min saran saya tam; min saran tampilan mact ny	min saran tampilan mact ny	['min', 'saran', 'tampilan', 'mact', 'min', 'saran', 'tampil', 'mact', 'nya', 'upgi	
halo level infinite sejujurn halo level infinite sej halo level infinite sejujurny	halo level infinite sejujurn halo level infinite sej halo level infinite sejujurny	halo level infinite sejujurny	['halo', 'level', 'infinite', 'sejujur', 'halo', 'level', 'infinite', 'jujur', 'game', 'ud	
ntah kenapa semenjak up ntah kenapa semenj ntah semenjak update mun	ntah kenapa semenjak up ntah kenapa semenj ntah semenjak update m	ntah kenapa semenjak update m	['ntah', 'semenjak', 'update', 'm', 'ntah', 'semenjak', 'update', 'muncul', 'b	
dari segi game sangat bag dari segi game sanga tolong perli	dari segi game sangat bag dari segi game sanga tolong perli	dari segi game sangat tolong perli	['segi', 'game', 'bagus', 'tolong', 'segi', 'game', 'bagus', 'tolong', 'baik', 'bu	
tolong perbaiki grafik yang tolong perbaiki grafil tolong perbaiki grafik optim	tolong perbaiki grafik yang tolong perbaiki grafil tolong perbaiki grafik optim	tolong perbaiki grafik optim	['tolong', 'perbaiki', 'grafik', 'opi', 'tolong', 'baik', 'grafik', 'optimalisasi', 'nik	
agak gimana ya masa saya; agak gimana ya masi; gimana ya main rank teman	agak gimana ya masa saya; agak gimana ya masi; gimana ya main rank teman	agak gimana ya masi; gimana ya main rank teman	['gimana', 'ya', 'main', 'rank', 'te', 'gimana', 'ya', 'main', 'rank', 'teman', 'tim	
aku kesal banget entah ke aku kesal banget ent kesal banget main tibataba j	aku kesal banget entah ke aku kesal banget ent kesal banget main tibataba j	aku kesal banget main tibataba j	['kesal', 'banget', 'main', 'tibatit', 'kesal', 'banget', 'main', 'tibatiba', 'jaring	
tolong masalah jaringan d tolong masalah jarin tolong jaringan perbaiki wai	tolong masalah jaringan d tolong masalah jarin tolong jaringan perbaiki wai	tolong masalah jaringan perbaiki wai	['tolong', 'jaringan', 'perbaiki', 'i', 'tolong', 'jaring', 'baik', 'war', 'ping', 'kon	
hampir perfect si sv lumay hampir perfect si sv perfect si sv lumayan aman	hampir perfect si sv lumay hampir perfect si sv perfect si sv lumayan aman	hampir perfect si sv perfect si sv lumayan aman	['perfect', 'si', 'sv', 'lumayan', 'a', 'perfect', 'si', 'sv', 'lumayan', 'aman', 'skri	
min tolong untuk sinyal ny min tolong untuk sin min tolong sinyal nya diperl	min tolong untuk sinyal ny min tolong untuk sin min tolong sinyal nya diperl	min tolong untuk sinyal nya diperl	['min', 'tolong', 'sinyal', 'nya', 'd', 'min', 'tolong', 'sinyal', 'nya', 'baik', 'pas',	
ayo lah masalah grafik di layo lah masalah graayo grafik benerin ya kali ul	ayo lah masalah grafik di layo lah masalah graayo grafik benerin ya kali ul	ayo lah masalah graayo grafik benerin ya kali ul	['ayo', 'grafik', 'benerin', 'ya', 'ki', 'ayo', 'grafik', 'benerin', 'ya', 'kali', 'ulti', 'i	
dari keseluruhan game ud dari keseluruhan gar game udh bagus eventnya r	dari keseluruhan game ud dari keseluruhan gar game udh bagus eventnya r	dari keseluruhan game udh bagus eventnya r	['game', 'udh', 'bagus', 'eventny', 'game', 'udh', 'bagus', 'eventnya', 'tarik',	
pliss benerin sistem matcl pliss benerin sistem pliss benerin sistem match i	pliss benerin sistem matcl pliss benerin sistem pliss benerin sistem match i	pliss benerin sistem match i	['pliss', 'benerin', 'sistem', 'matr', 'pliss', 'benerin', 'sistem', 'match', 'nya',	
ini cuman aku yang error i ini cuman aku yang cuman error atqu gimana ki	ini cuman aku yang error i ini cuman aku yang cuman error atqu gimana ki	ini cuman aku yang cuman error atqu gimana ki	['cuman', 'error', 'atqu', 'giman', 'cuman', 'error', 'atqu', 'gimana', 'kali', 'l	
game yg sangat menyenar game yg sangat men game yg menyenangkangr	game yg sangat menyenar game yg sangat men game yg menyenangkangr	game yg sangat menyenangkangr	['game', 'yg', 'menyenangkangr', 'game', 'yg', 'menyenangkangr', 'game', 'yg', 'menyenangkangr', 'yg',	
mengapa sekarang tak bis mengapa sekarang t memainkan grafik max efek	mengapa sekarang tak bis mengapa sekarang t memainkan grafik max efek	mengapa sekarang t memainkan grafik max efek	['memainkan', 'grafik', 'max', 'e', 'main', 'grafik', 'max', 'efek', 'nya', 'buruk	
jujur saya suka sama gami jujur saya suka samajujuju suka game nya kalo m	jujur saya suka sama gami jujur saya suka samajujuju suka game nya kalo m	jujur saya suka samajujuju suka game nya kalo m	['jujur', 'suka', 'game', 'nya', 'ka', 'jujur', 'suka', 'game', 'nya', 'kalo', 'main',	
kenapa sih habis update g kenapa sih habis uprsih habis update gamenya c	kenapa sih habis update g kenapa sih habis uprsih habis update gamenya c	kenapa sih habis update gamenya c	['sih', 'habis', 'update', 'gameny', 'sih', 'habis', 'update', 'gamenya', 'delay',	
tolong dong perbaiki lagi t tolong dong perbaiki tolong perbaiki sistem repo	tolong dong perbaiki lagi t tolong dong perbaiki tolong perbaiki sistem repo	tolong dong perbaiki tolong perbaiki sistem repo	['tolong', 'perbaiki', 'sistem', 're', 'tolong', 'baik', 'sistem', 'reportnya', 'kad	
maaf saya ada masukan t maaf saya ada masu maaf masukan terkait peng	maaf saya ada masukan t maaf saya ada masu maaf masukan terkait peng	maaf saya ada masu maaf masukan terkait peng	['maaf', 'masukan', 'terkait', 'pe', 'maaf', 'masuk', 'kait', 'atur', 'grafik', 'bar	
game enggak terstruktur i game enggak terstru game terstruktur hero farm	game enggak terstruktur i game enggak terstru game terstruktur hero farm	game enggak terstru game terstruktur hero farm	['game', 'terstruktur', 'hero', 'fa', 'game', 'struktur', 'hero', 'farm', 'lane', 'd	

Gambar 3. Hasil Preprocessing data

3.3 Splitting Data

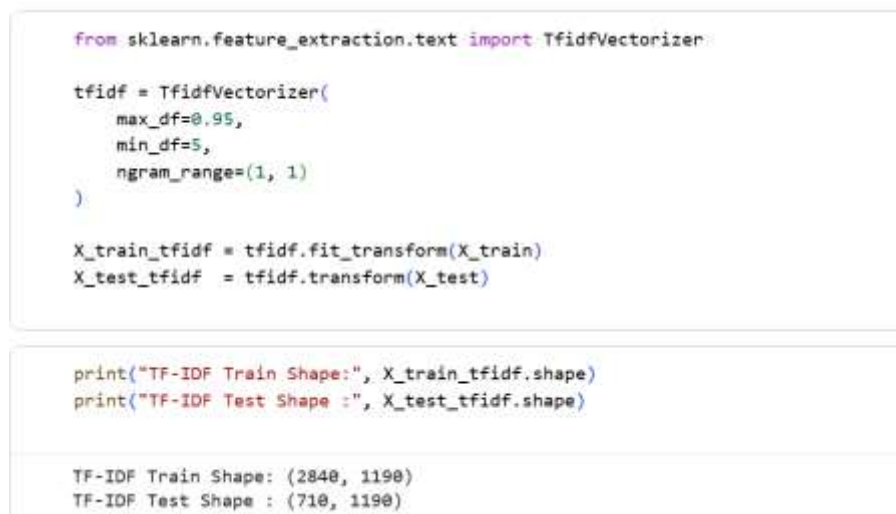
Tahapan berikutnya adalah *data splitting*. Pada tahap ini dataset dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan perbandingan 80% : 20%. Data latih digunakan untuk membangun dan melatih model klasifikasi menggunakan algoritma Naïve Bayes sehingga model dapat mempelajari pola dari data yang tersedia. Sementara itu, data uji digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya. Proses pembagian data ini bertujuan untuk mengukur kemampuan model dalam melakukan prediksi sentimen secara lebih objektif. Gambar 4 berikut kode dan hasil splitting data



Gambar 4. Kode dan Hasil splitting data

3.4 Pembobotan TF-IDF

Pembobotan TF-IDF merupakan proses mengubah data teks hasil *preprocessing* menjadi representasi numerik melalui teknik ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency*. Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen serta tingkat kepentingannya terhadap keseluruhan dokumen dalam kumpulan data. Dengan menggunakan metode TF-IDF, kata yang sering muncul dalam suatu dokumen namun jarang muncul pada dokumen lain akan memiliki bobot yang lebih tinggi, sehingga dapat merepresentasikan informasi yang lebih relevan dalam proses analisis sentimen. Representasi numerik yang dihasilkan kemudian digunakan sebagai fitur dalam proses klasifikasi menggunakan algoritma Naïve Bayes. Gambar 5 berikut Kode dan hasil TF-IDF.



Gambar 5. Kode dan Hasil TF-IDF

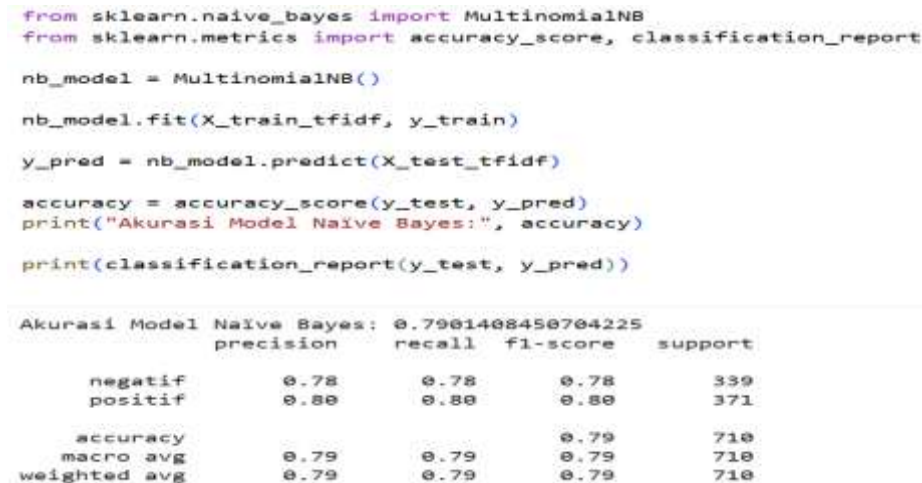
3.5 Pembuatan Model dan Pengujian Naïve Bayes

Setelah tahap *preprocessing* data selesai dilakukan, langkah selanjutnya adalah membangun model klasifikasi menggunakan algoritma Naïve Bayes. Data yang telah melalui tahapan *cleaning*, *case folding*, *stopword removal*, *tokenizing*, *stemming*, *labeling*, *data splitting*, serta pembobotan TF-IDF kemudian digunakan sebagai masukan (*input*) dalam proses pelatihan model *machine learning*. Pada tahap ini, model Naïve Bayes dilatih menggunakan data latih untuk mempelajari pola yang terdapat dalam dataset sehingga mampu mengklasifikasikan sentimen ulasan pengguna ke dalam kategori positif dan negatif secara lebih akurat.

Setelah model Multinomial Naïve Bayes dilatih menggunakan data latih, tahap selanjutnya adalah melakukan evaluasi kinerja model menggunakan data uji. Proses pengujian ini bertujuan untuk mengukur kemampuan model dalam mengklasifikasikan sentimen pada data baru yang belum pernah digunakan pada tahap pelatihan, sekaligus menjawab rumusan masalah yang telah ditetapkan dalam penelitian.

Hasil evaluasi menunjukkan bahwa model yang dibangun mampu memberikan performa yang cukup baik dengan tingkat akurasi mencapai 79%. Hal ini menunjukkan bahwa model mampu memprediksi sentimen ulasan pengguna dengan benar pada sebagian besar data uji, sehingga algoritma Naïve Bayes dinilai cukup efektif dalam melakukan klasifikasi sentimen terhadap ulasan pengguna game *Honor of Kings* pada platform Google Play Store.

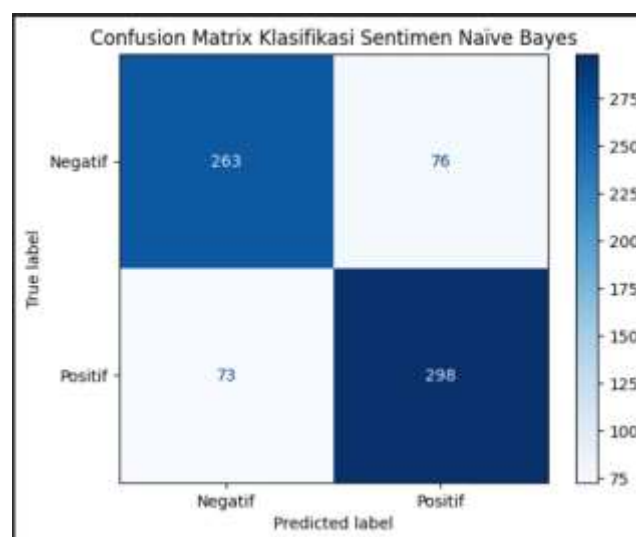
Gambar 6 berikut Kode dan Hasil Pengujian Naïve Bayes.



Gambar 6. Kode dan Hasil Pengujian Naive Bayes

3.6 Evaluasi Confusion Matrix

Untuk melakukan analisis yang lebih mendalam terhadap kinerja model, digunakan confusion matrix yang ditampilkan pada Gambar 7. *Confusion matrix* digunakan untuk memvisualisasikan hasil klasifikasi model dengan membandingkan antara nilai prediksi yang dihasilkan oleh model dan nilai aktual dari data uji. Melalui visualisasi ini, dapat diketahui jumlah prediksi yang benar maupun yang salah pada masing-masing kelas sentimen, yaitu positif dan negatif, sehingga performa model klasifikasi dapat dianalisis secara lebih jelas dan terperinci.



Gambar 7. Evaluasi Confusion Matrix

Berdasarkan hasil evaluasi model menggunakan *confusion matrix*, diperoleh informasi mengenai performa klasifikasi sentimen pada data uji. Hasil pengujian menunjukkan bahwa model Multinomial Naïve Bayes memiliki kemampuan yang cukup baik dalam membedakan sentimen positif dan negatif pada ulasan game *Honor of Kings*. Dari total 710 data uji, sebagian besar data berhasil diklasifikasikan dengan benar oleh model.

Pada kategori sentimen negatif, model berhasil mengklasifikasikan 263 data sebagai sentimen negatif dengan benar (*True Negative*), namun terdapat 76 data negatif yang salah diklasifikasikan sebagai sentimen positif (*False Positive*). Hal ini menunjukkan bahwa beberapa ulasan negatif memiliki karakteristik kata yang mirip dengan ulasan positif sehingga menyebabkan kesalahan klasifikasi.

Sementara itu, pada kategori sentimen positif, model mampu mengklasifikasikan 298 data positif dengan benar (*True Positive*), sedangkan 73 data positif lainnya salah diklasifikasikan sebagai sentimen negatif (*False Negative*). Kesalahan ini menunjukkan adanya kemiripan kosakata antara ulasan positif dan negatif yang menyebabkan model mengalami kesulitan dalam membedakan kedua kelas tersebut.

Secara keseluruhan, jumlah prediksi yang benar lebih dominan dibandingkan prediksi yang salah, sehingga model memperoleh tingkat akurasi sebesar 79%. Hasil ini menunjukkan bahwa algoritma Naïve Bayes cukup efektif dalam melakukan klasifikasi sentimen ulasan pengguna game *Honor of Kings*, meskipun masih terdapat keterbatasan dalam menangani ulasan yang bersifat ambigu atau memiliki kemiripan kata antar kelas sentimen.

Berikut adalah perhitungan dari confusion matrix:

1. Perhitungan untuk Kelas Negatif

a. Precision Negatif

Rumus 1:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Precision_{Negatif} = \frac{263}{263 + 73} = \frac{263}{336} = 0,783$$

b. Recall Negatif

Rumus 2:

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Recall_{Negatif} = \frac{263}{263 + 76} = \frac{263}{339} = 0,776$$

c. F1-Score Negatif

Rumus 3:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

$$F1_{Negatif} = \frac{2 \times 0,783 \times 0,776}{0,783 + 0,776} = 0,779$$

2. Perhitungan untuk Kelas Positif

a. Precision Positif

$$Precision_{Positif} = \frac{298}{298 + 76} = \frac{298}{374} = 0,797$$

b. Recall Positif

$$Recall_{Positif} = \frac{298}{298 + 73} = \frac{298}{371} = 0,803$$

c. F1-Score Positif

$$F1_{Positif} = \frac{2 \times 0,797 \times 0,803}{0,797 + 0,803} = 0,800$$

3. Perhitungan Akurasi Model

Rumus Akurasi:

$$Accuracy = \frac{TP+TN}{Total\ Data} \quad (4)$$

$$Accuracy = \frac{263 + 298}{263 + 76 + 73 + 298} = \frac{561}{710} = 0,79 \text{ (79\%)}$$

Berdasarkan perhitungan confusion matrix, model Naïve Bayes memperoleh nilai akurasi sebesar 79%. Untuk kelas sentimen negatif, diperoleh nilai precision sebesar 78,3%, recall sebesar 77,6%, dan F1-score sebesar 77,9%. Sementara itu, pada kelas sentimen positif diperoleh nilai precision sebesar 79,7%, recall sebesar 80,3%, dan F1-score sebesar 80,0%. Hasil ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang cukup seimbang pada kedua kelas sentimen.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Naïve Bayes mampu melakukan klasifikasi sentimen terhadap ulasan pengguna game *Honor of Kings* pada platform Google Play Store dengan performa yang cukup baik. Model yang dibangun berhasil mengklasifikasikan ulasan ke dalam kategori sentimen positif dan negatif dengan tingkat akurasi sebesar 79%, yang menunjukkan bahwa sebagian besar data uji dapat diprediksi dengan benar oleh model. Hasil evaluasi menunjukkan bahwa pada kelas sentimen positif diperoleh nilai precision sebesar 79,7%, recall sebesar 80,3%, dan F1-score sebesar 80,0%. Sementara itu, pada kelas sentimen negatif diperoleh nilai precision sebesar 78,3%, recall sebesar 77,6%, dan F1-score sebesar 77,9%. Nilai-nilai tersebut menunjukkan bahwa algoritma Naïve Bayes memiliki kemampuan yang cukup seimbang dalam membedakan sentimen positif dan negatif, meskipun masih terdapat beberapa kesalahan klasifikasi yang disebabkan oleh kemiripan kosakata antar kelas sentimen. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa algoritma Naïve Bayes merupakan metode yang cukup efektif dan efisien dalam melakukan analisis sentimen berbasis teks. Model yang dihasilkan mampu mengidentifikasi pola sentimen pada ulasan pengguna secara baik, sehingga dapat memberikan gambaran mengenai persepsi pengguna terhadap kualitas game *Honor of Kings*. Hasil penelitian ini juga diharapkan dapat menjadi referensi bagi pengembang dalam melakukan evaluasi dan peningkatan kualitas layanan aplikasi.

Referensi

- [1] J. Feng and X. Zhang, "Analysis of Mobile Game Intellectual Property Marketing Strategy: A Case Study of Honor of Kings," *Applied Economics and Finance*, vol. 9, no. 4, p. 1, Sep. 2022, doi: 10.11114/aef.v9i4.5707.
- [2] M. Haikal and U. Hayati, "ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI GAME ONLINE PUBG MOBILE MENGGUNAKAN ALGORITMA NAIVE BAYES," 2023.
- [3] M. Tirta Nugraha, N. Sulistiyowati, and U. Enri, "ANALISIS SENTIMEN ULASAN APLIKASI SATU SEHAT PADA GOOGLE PLAY STORE MENGGUNAKAN NAIVE BAYES CLASSIFIER," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 5, Oct. 2023.
- [4] M. Adhit and D. Yuda, "Implementasi Web Scraping Untuk Ekstraksi Data Penjual dan Produk Panel Surya Di E-Commerce," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 5, no. 1, Jun. 2025.
- [5] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, Jan. 2022, doi: 10.23960/jitet.v10i1.2262.
- [6] R. Andrea Arifa, N. Suarna, A. Bahtiar, N. Rahaningsih, and W. Prihartono, "MODEL MACHINE LEARNING UNTUK PREDIKSI RISIKO PENYAKIT LIVER DENGAN RANDOM FOREST TEROPTIMASI," *JURNAL SISTEM INFORMASI DAN TEKNOLOGI (SINTEK)*, vol. VI, no. 1, Jan. 2026, Accessed: Jul. 05, 2026. [Online]. Available: <https://www.sintek.stmikku.ac.id/index.php/home/article/view/204>

- [7] R. Zulfiquri, B. N. Sari, and T. N. Padilah, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI MEDIA SOSIAL INSTAGRAM PADA SITUS GOOGLE PLAY STORE MENGGUNAKAN NAÏVE BAYES CLASSIFIER," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4995.
- [8] N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 1, pp. 47–54, Apr. 2022.
- [9] S. Pambudi, P. Setiaji, and W. A. Triyanto, "Sentiment Analysis of Fizzo Novel Application Using Support Vector Machine and Naïve Bayes Algorithm with SEMMA Framework," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 4, pp. 1861–1880, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.4875.
- [10] R. Andrea Arifa, N. Suarna, A. Bahtiar, N. Rahaningsih, and W. Prihartono, "MODEL MACHINE LEARNING UNTUK PREDIKSI RISIKO PENYAKIT LIVER DENGAN RANDOM FOREST TEROPTIMASI," *JURNAL SISTEM INFORMASI DAN TEKNOLOGI (SINTEK)*, vol. VI, no. 1, Jan. 2026, Accessed: Jul. 05, 2026. [Online]. Available: <https://www.sintek.stmikku.ac.id/index.php/home/article/view/204>
- [11] A. H. Nurdy, A. Rahim, and Arbansyah, "Analisis Sentimen Ulasan Game Stumble Guys Pada Playstore Menggunakan Algoritma Naïve Bayes," *Teknika*, vol. 13, no. 3, pp. 388–395, Sep. 2024, doi: 10.34148/teknika.v13i3.993.
- [12] G. Darmawan, S. Alam, and M. Imam Sulisty, "ANALISIS SENTIMEN BERDASARKAN ULASAN PENGGUNA APLIKASI MYPERTAMINA PADA GOOGLE PLAYSTORE MENGGUNAKAN METODE NAÏVE BAYES," *STORAGE – Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 100–108, Aug. 2023, doi: 10.55123.
- [13] A. Faros Wibawa, L. S. Marita, M. Faisal, W. Widyastuty, and A. Sumandito, "ANALISIS EFEKTIFITAS IMPLEMENTASI ALGORITMA NAÏVE BAYES PADA ROLE HERO DALAM TURNAMEN MOBILE LEGENDS," 2024.
- [14] E. Hasibuan and E. A. Heriyanto, "ANALISIS SENTIMEN PADA ULASAN APLIKASI AMAZON SHOPPING DI GOOGLE PLAY STORE MENGGUNAKAN NAIVE BAYES CLASSIFIER," *JTS*, vol. 1, no. 3.
- [15] M. Abib Allesdio, A. Irma Purnamasari, I. Ali, N. Suarna, and A. Bahtiar, "ANALISA PENGGUNAAN METODE LEXICON BASED DAN ALGORITMA NAIVE BAYES PADA SENTIMEN ULASAN APLIKASI DUOLINGO," *JURNAL SISTEM INFORMASI DAN TEKNOLOGI (SINTEK)*, vol. VI, no. 2, May 2026, Accessed: Jul. 05, 2026. [Online]. Available: <https://www.sintek.stmikku.ac.id/index.php/home/article/view/261>