

# ANALISIS SENTIMEN PENGGUNA TERHADAP APLIKASI DUOLINGO DENGAN NAÏVE BAYES CLASSIFIER

Annisa Saninah<sup>1\*</sup>, Willy Prihartono<sup>2</sup>, Cep Lukman Rohmat<sup>3</sup>

<sup>1,2,3</sup>Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135; Telp. (0231) 490480

Received: 13 Desember 2024  
Accepted: 14 Januari 2025  
Published: 20 Januari 2025

## Keywords:

*Naïve Bayes Classifier*;  
*Duolingo*;  
*Google Play Store*;

## Correspondent Email:

saninahannisa@gmail.com

**Abstrak.** Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi *Duolingo* di *Google Play Store* menggunakan algoritma *Naïve Bayes Classifier*. Sebanyak 2.000 ulasan pengguna dianalisis dan diklasifikasikan menjadi sentimen positif dan negatif. Hasil penelitian menunjukkan bahwa algoritma ini memiliki performa yang sangat baik, dengan tingkat *accuracy* mencapai 88%, *precision* 95%, *recall* 92% dan *F1-Score* sebesar 93%. Mayoritas ulasan mencerminkan sentimen positif, terutama mengapresiasi fitur interaktif dan kemudahan penggunaan aplikasi. Namun, ulasan negatif sebagian besar menyoroti masalah teknis seperti *bug* dan *lag*. Penelitian ini membuktikan bahwa algoritma *Naïve Bayes Classifier* efektif untuk analisis sentimen berbasis teks sekaligus memberikan wawasan berharga bagi pengembang aplikasi dalam meningkatkan kualitas layanan. Rekomendasi dari penelitian ini meliputi pengumpulan data ulasan yang lebih banyak dengan cakupan waktu yang lebih luas guna memperoleh hasil analisis yang lebih representatif. Selain itu, membandingkan performa algoritma ini dengan metode lain, seperti *SVM*, *Random Forest*, atau *LSTM*, dapat membantu menemukan pendekatan yang lebih optimal dalam menganalisis ulasan pengguna yang kompleks. Penelitian ini juga mengungkap potensi aplikasi *Duolingo* dalam mendukung pembelajaran bahasa asing bagi anak-anak berkebutuhan khusus di sekolah. Dengan fitur interaktif dan fleksibilitasnya, aplikasi ini memungkinkan siswa belajar sesuai dengan kebutuhan individu.

**Abstract.** This study aims to analyze user sentiment towards the *Duolingo* application on the *Google Play Store* using the *Naïve Bayes Classifier* algorithm. A total of 2,000 user reviews were analyzed and classified into positive and negative sentiments. The results showed that the algorithm performed very well, with an accuracy rate of 88%, precision of 95%, recall of 92%, and *F1-Score* of 93%. The majority of reviews reflected positive sentiments, mainly appreciating the interactive features and ease of use of the app. However, negative reviews mostly highlighted technical issues such as bugs and lag. This research proves that the *Naïve Bayes Classifier* algorithm is effective for text-based sentiment analysis while providing valuable insights for app developers in improving service quality. Recommendations from this research include collecting more review data with a wider time coverage to obtain more representative analysis results. In addition, comparing the performance of this algorithm with other methods, such as *SVM*, *Random*

*Forest, or LSTM, can help find a more optimal approach in analyzing complex user reviews. This research also reveals the potential of the Duolingo app in supporting foreign language learning for children with special needs in schools. With its interactive features and flexibility, this app allows students to learn according to individual needs.*

---

## 1. PENDAHULUAN

Di era digital, pembelajaran bahasa menjadi lebih mudah, salah satunya melalui aplikasi *Duolingo* yang telah diunduh lebih dari 500 juta kali. Aplikasi ini menawarkan kursus bahasa gratis untuk berbagai tingkat kemampuan dengan pengalaman belajar yang interaktif melalui fitur permainan dan kuis. Sebagai contoh keberhasilan penerapan kecerdasan buatan (AI), *Duolingo* menggunakan AI untuk memahami dan menyesuaikan kebutuhan pengguna, sehingga mendukung pembelajaran mandiri dengan lebih efektif [1]. Aplikasi *Duolingo* memiliki sejumlah kelebihan dan kekurangan. Untuk mengetahui harapan pengguna, penting untuk menganalisis cara mereka berinteraksi dengan aplikasi ini. Ulasan pengguna memberikan informasi yang berguna bagi pengembang untuk meningkatkan layanan, sementara media sosial berfungsi sebagai platform bagi pengguna untuk mengungkapkan kepuasan atau ketidakpuasan serta berbagi pandangan tentang aplikasi. Data ini dapat digunakan untuk mengevaluasi perspektif pengguna terhadap *Duolingo*. *Duolingo* adalah platform pembelajaran bahasa gratis yang dikembangkan oleh *Luis Von Ahn* dan *Severin Hacker* pada November 2011. Aplikasi ini menyediakan berbagai pilihan bahasa, seperti *Inggris*, *Arab*, *Spanyol*, dan lainnya, dengan slogan "Pendidikan bahasa gratis untuk dunia." *Duolingo* menjadi contoh bagaimana inovasi teknologi dapat merevolusi pembelajaran bahasa. Berdasarkan data *Google Play Store* hingga Februari, aplikasi ini telah diunduh lebih dari 500 juta kali, memiliki rating 4,8, dan mendapatkan lebih dari 21 juta ulasan. Ulasan tersebut mencakup komentar positif, kritik, keluhan, serta saran, yang memberikan wawasan lebih dalam mengenai pengalaman pengguna [2]. Masalah terkait *bug* dan fitur baru yang ditemukan setelah pembaruan sering kali membuat pengguna *Duolingo* merasa tidak puas dan terganggu saat menggunakan aplikasi. Ulasan yang ada menunjukkan bahwa beberapa diantaranya berisi komentar negatif. Oleh

karena itu, penting untuk melakukan analisis sentimen sebagai pertimbangan bagi pengguna dalam memilih platform belajar online. Di sisi lain, hasil analisis ini juga bisa menjadi bahan evaluasi bagi pengembang dalam upaya meningkatkan aplikasi yang mereka buat. *Duolingo* merupakan aplikasi pembelajaran bahasa gratis yang dapat diakses di ponsel, tablet, dan komputer. Kursus bahasa *Jepang* yang ditawarkannya terbukti membantu pengguna belajar dengan cara yang lebih santai dan menyenangkan, serta meningkatkan motivasi dan pemahaman mereka. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi *Duolingo* dari *Google Play Store* dengan memanfaatkan algoritma *Naïve Bayes Classifier*. Penelitian ini berfokus pada pengelompokan ulasan menjadi sentimen positif dan negatif, menilai performa algoritma, serta mengungkap pola-pola sentimen utama. Temuan penelitian ini diharapkan dapat membantu pengembang *Duolingo* dalam meningkatkan kualitas aplikasi sekaligus memberikan kontribusi pada pengembangan teknologi analisis sentimen untuk penggunaan di aplikasi lain. Penelitian ini melakukan analisis sentimen terhadap ulasan aplikasi *Duolingo* menggunakan metode *Naïve Bayes* yang dipadukan dengan teknik *SMOTE*. Hasil penelitian menunjukkan bahwa penggunaan *SMOTE* pada *Naïve Bayes* menghasilkan *accuracy* 91,95% dan AUC 0,740, sementara tanpa *SMOTE*, *accuracy* hanya mencapai 77,14% dengan AUC 0,526. Hal ini menunjukkan bahwa *SMOTE* berperan efektif dalam meningkatkan kinerja klasifikasi sentimen [3]. Penelitian ini memanfaatkan algoritma *Naïve Bayes* untuk menganalisis sentimen ulasan aplikasi BRImo di *Google Play Store*. Data diproses menggunakan metode *KDD* yang meliputi *scraping*, pembersihan, dan transformasi menggunakan *TF-IDF*. Hasil analisis menunjukkan tingkat *accuracy* sebesar 89,58%, *precision* 85,94%, *recall* 89,58% dan *f1-score* 86,90%, dengan perhatian khusus diperlukan untuk meningkatkan *recall* pada

sentimen negatif [4]. Penelitian ini melakukan analisis sentimen terhadap ulasan pengguna aplikasi *e-commerce* Tokopedia, *Shopee* dan *Lazada* menggunakan algoritma *Naïve Bayes*. Metode yang diterapkan adalah *KDD* dan hasil analisis menunjukkan bahwa aplikasi *Shopee* memperoleh *accuracy* 92%, *precision* 92,13%, *recall* 98,8% dan *f1-score* 95,35% dengan data pelatihan 80% dan data pengujian 20% [5]. Tujuan dari penelitian ini adalah untuk menerapkan algoritma *Naive Bayes Classifier* dalam mengklasifikasikan ulasan pengguna secara otomatis berdasarkan sentimen dengan menganalisis ulasan dari platform seperti *Google Play Store* untuk menentukan sentimen positif atau negatif. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi kinerja algoritma *Naive Bayes* dengan mengukur empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score* untuk menilai sejauh mana algoritma ini dapat mengklasifikasikan sentimen ulasan pengguna secara akurat dan seimbang. Penelitian ini menggunakan algoritma *Multinomial Naïve Bayes* untuk menganalisis sentimen pengguna aplikasi *My F&B ID* berdasarkan ulasan di *Google Play Store*. Hasil pengujian dengan pembagian dataset 80:20, 70:30 dan 60:40 menunjukkan *accuracy* tertinggi sebesar 85% pada pembagian 80:20 sementara *accuracy* terendah sebesar 83% pada pembagian 60:40 yang mengindikasikan bahwa proporsi data latih yang lebih besar meningkatkan akurasi klasifikasi [6]. Penelitian ini mengkaji sentimen pelanggan *JD.ID* di *Twitter* menggunakan *Naïve Bayes Classifier* dengan tambahan fitur konversi ikon emosi dan pembobotan *TF-IDF*. Dari total 900 *tweet* (300 untuk masing-masing sentimen: positif, netral, dan negatif), data diproses melalui tahap *preprocessing* dan pembobotan sebelum diklasifikasikan. Hasil menunjukkan peningkatan *accuracy* dari 96,44% tanpa fitur tambahan menjadi 98% dengan penerapan konversi ikon emosi dan *TF-IDF* [7].

## 2. TINJAUAN PUSTAKA

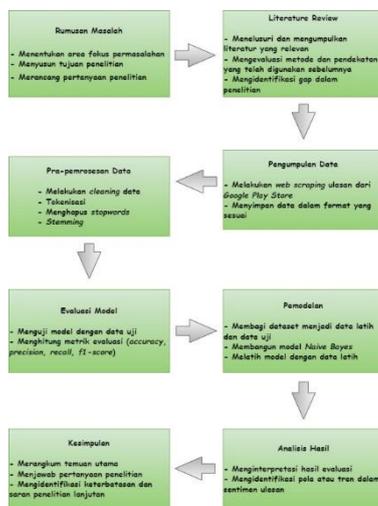
Analisis sentimen merupakan bidang yang semakin berkembang dalam penelitian teks. Tujuan utamanya adalah untuk mengidentifikasi dan memahami opini atau perasaan yang terkandung dalam suatu teks. Secara spesifik, analisis sentimen berfokus pada

pengelompokan pernyataan menjadi sentimen positif atau negatif berdasarkan pemahaman terhadap isi teks tersebut [8]. Algoritma *Naïve Bayes* adalah salah satu teknik klasifikasi yang sering digunakan dalam *data mining* dan *text mining*. *Naive Bayes Classifier* terdiri dari dua tahap, yaitu pelatihan dan pengujian. Pada tahap pelatihan, model dilatih menggunakan dataset, lalu pada tahap pengujian, model diuji berdasarkan probabilitas yang diperoleh dari dataset pelatihan. Metode ini efektif dalam mengklasifikasikan opini ke dalam kategori positif, netral, atau negatif mengenai suatu produk atau isu [9]. *Naive Bayes* adalah metode klasifikasi yang sederhana, yang menghitung kemungkinan dengan mengkombinasikan kombinasi dan frekuensi nilai yang ada pada data yang diperoleh [10]. Belajar bahasa dengan *Duolingo* memberikan pengalaman yang khas berkat sistem yang memotivasi yang meliputi pelacakan kemajuan, umpan balik, dan pengembangan kosakata. *Duolingo* memanfaatkan elemen permainan untuk menjaga semangat belajar dimana pengguna harus menyelesaikan level untuk mengakses pelajaran berikutnya. Latihan berulang, seperti soal pilihan ganda, menulis dan berbicara digunakan untuk memperdalam pemahaman [11].

## 3. METODE PENELITIAN

Penelitian ini menerapkan metode kuantitatif, yang mencakup pengumpulan, pengolahan, analisis, dan penyajian data secara objektif untuk menyelesaikan masalah atau menguji hipotesis dalam rangka mengembangkan prinsip-prinsip umum [12]. Penelitian ini menerapkan pendekatan kuantitatif dengan analisis deskriptif untuk menjelaskan data numerik secara rinci. Proses penelitian meliputi pengumpulan, pengolahan, analisis dan penyajian data berdasarkan frekuensi atau jumlahnya guna menguji hipotesis dan mengidentifikasi pola. Data yang digunakan berasal dari sumber sekunder dalam bentuk angka [13]. Tujuan penelitian ini adalah untuk menganalisis dan mengevaluasi sejauh mana akurasi algoritma *Naive Bayes Classifier* dalam mengklasifikasikan sentimen positif atau negatif pada ulasan aplikasi *Duolingo* dengan memanfaatkan probabilitas dan pola kata untuk menghasilkan analisis yang akurat dan efisien. Penelitian ini mengadopsi metodologi

*Knowledge Discovery in Database (KDD)*, yang mencakup enam langkah utama, yaitu *data selection*, pra-pemrosesan, transformasi, *data mining*, evaluasi dan interpretasi. Metode ini bertujuan untuk mengidentifikasi pola-pola penting dan informasi tersembunyi yang dapat digunakan dalam pengambilan keputusan. Fokus penelitian ini adalah menganalisis sentimen ulasan pengguna aplikasi *Duolingo* di *Play Store* dengan menggunakan algoritma *Naïve Bayes Classifier*. Tahapan-tahapan metode penelitian ditampilkan pada Gambar 1 :



Gambar 1. Tahapan Metode Penelitian

### 3.1. Data Selection

Data yang relevan, yaitu ulasan pengguna *Duolingo* di *Play Store*, dipilih untuk dianalisis meliputi *ID* pengguna, isi ulasan, dan skor ulasan sebagai penanda sentimen positif atau negatif.

### 3.2. Pra-pemrosesan Data

Data yang dipilih disiapkan untuk analisis dengan membersihkan duplikat, menangani nilai kosong, melakukan tokenisasi, penghapusan *stopwords* dan *stemming* agar hanya informasi yang relevan yang diproses.

### 3.3. Data Transformation

Teks yang telah diproses diubah menjadi vektor fitur dengan menggunakan teknik seperti *TF-IDF* atau *word embeddings* sehingga algoritma dapat mengenali pola sentimen dalam ulasan.

### 3.4. Data Mining

Algoritma *Naïve Bayes* diterapkan pada data yang telah diproses untuk mengklasifikasikan sentimen, memanfaatkan kemampuannya dalam mengelola data teks besar dan keakuratannya dalam mendeteksi pola sentimen positif dan negatif.

### 3.5. Evaluation

Setelah model *Naïve Bayes* dibuat, tahap evaluasi dilakukan untuk menilai kinerjanya dengan menggunakan metrik, seperti *accuracy*, *precision*, *recall* dan *F1-score* untuk memastikan akurasi dalam klasifikasi sentimen.

### 3.6. Interpretation

Hasil analisis sentimen menggunakan model *Naïve Bayes* memberikan pemahaman tentang pandangan pengguna terhadap aplikasi *Duolingo* yang bermanfaat bagi pengembang atau pemasar untuk memahami kebutuhan dan persepsi pengguna. Wawasan ini dapat divisualisasikan untuk mendukung pengambilan keputusan berbasis data. Penggunaan metodologi *KDD* memungkinkan analisis yang lebih terstruktur dan mendalam, memberikan manfaat praktis untuk pengembangan aplikasi.

## 4. HASIL DAN PEMBAHASAN

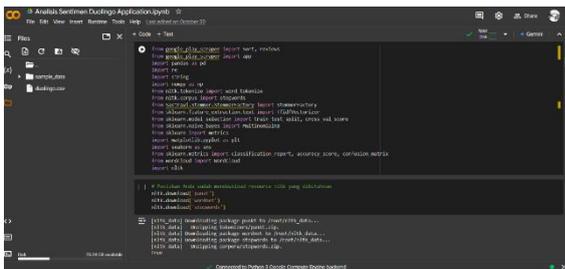
### 4.1. Hasil

#### 4.1.1. Identifikasi Permasalahan

##### Data

Penelitian ini mengidentifikasi berbagai kendala dalam analisis sentimen ulasan pengguna aplikasi *Duolingo* menggunakan algoritma *Naïve Bayes Classifier*. Beberapa di antaranya adalah ketidakseimbangan data yang lebih banyak didominasi ulasan positif, keterbatasan data yang hanya mencakup satu tahun dan fokus pada satu algoritma tanpa membandingkannya dengan metode lain. Meskipun algoritma ini memiliki *accuracy* 88%, masih diperlukan peningkatan melalui teknik optimisasi atau metode seperti *SMOTE*. Selain itu, ulasan negatif yang mencakup keluhan terkait *bug* dan *lag* menyoroti kebutuhan mendesak untuk meningkatkan aspek teknis aplikasi. Penelitian ini merekomendasikan pengumpulan data yang

lebih beragam dan analisis algoritma alternatif untuk menyempurnakan hasil sentimen dan kualitas layanan *Duolingo*. Penelitian ini menggunakan dataset yang terdiri dari 2.000 *record* dengan 9 atribut, yang diperoleh dari aplikasi *Duolingo* di *Google Play Store* (ID: 'com.duolingo') pada tahun 2024. Data tersebut disajikan dalam format file CSV dengan contoh



20 *record* ditampilkan dari total 2.000 *record* yang ada. Hasil *crawling data* ditampilkan pada Gambar 2 dan output *crawling data* pada Tabel

userName	Ulasan	rating
Najwa Rahmatia	Sejak update-an baru, saya paham kalau kalian memang ingin aplikasi kalian terlihat bagus dengan interface yang menarik, tapi tolong itu bikin lag. Saya biasanya klo tekan cepat selalu responsif tapi sejak penambahan interface baru, klo mau ngetik langsung lag. Pas nampilin hasil xp, hasil waktu mengerjakan bikin ngelag dan bikin hp jadi cepat panas. Tolong diperbaiki ya Duolingo, interface kalian sebelumnya saja sudah bagus kok, kan tidak mungkin kan kalian merelakan experience demi estetika ?	5

1: Gambar 2. Hasil *crawling data*

Tabel 1. Output *crawling data*

#### 4.1.2. Data Selection

Tahap ini adalah Pemilihan Data. Dataset yang digunakan dalam penelitian ini berasal dari ulasan pengguna aplikasi *Duolingo* dan merupakan data sekunder. Dataset ini memiliki beberapa kolom, seperti *Unnamed:0*, *reviewId*, *userImage*, *thumbsUpCount*, *reviewCreatedVersion*, *at*, *replyContent*, *repliedAt* dan *appVersion* yang akan digunakan untuk mengklasifikasikan sentimen ulasan sebagai positif atau negatif menggunakan algoritma *Naïve Bayes Classifier*.

#### 4.1.3. Preprocessing

*Preprocessing* merupakan tahap krusial sebelum klasifikasi dalam analisis sentimen, yang bertujuan untuk meningkatkan kualitas

data dan akurasi algoritma *Naïve Bayes*. Data yang diperoleh melalui *scraping* akan menjalani proses seperti pembersihan teks, normalisasi, penghapusan kata-kata yang tidak relevan dan *stemming* untuk mempersiapkannya agar siap dianalisis. Tahapan ini menyederhanakan dan menata data sehingga lebih mudah diproses oleh model *machine learning* untuk klasifikasi yang lebih akurat.

### 1. Mengubah Nama Atribut

Pada tahap awal *preprocessing*, langkah kunci adalah mengganti nama atribut dalam *DataFrame* untuk meningkatkan pemahaman dan kesesuaian dengan konteks data. Proses ini mengurangi kemungkinan kesalahan, mempermudah interpretasi, dan meningkatkan efisiensi dalam transformasi serta klasifikasi data selama analisis sentimen.

### 2. Cleaning Data

Pada tahap pembersihan data, normalisasi teks berperan penting untuk mengurangi perbedaan dalam data, sehingga memudahkan analisis menggunakan algoritma *Naïve Bayes Classifier*. Proses ini menyamakan format teks dengan mengubah kata atau karakter yang memiliki makna serupa menjadi bentuk yang konsisten, memungkinkan

Cleaning Data Ulasan	
baru diupdate karena sebelumnya ngelag terus, setelah diupdate malah tambah parah. padahal udah nyelesaiin pelajarannya karena ngebug jadi nggak dihitung dan akhirnya ngulang lagi, yang harusnya misal tiga pelajaran jadi atau bahkan lebih. jadi harus keluar apk hapus dari riwayat baru masuk lagi, dan ini bukan cuma sekali dua kali padahal koneksi internetnya bagus lancar lancar aja. padahal baru kemarin diupdate tapi performanya malah menurun.	pembelajarannya tersusun sekali, dari yang basic sampe yang rumit, bahasa yang diajarkan juga bahasa sehari-hari, untuk belajar serius atau hanya dijadikan game juga cocok, tapi jika ingin lebih cepat mengerti dan fasih belajar bahasa tidak bisa hanya mengandalkan duolingo saja, saran juga sebaiknya grammar juga lebih dijelaskan! secara keseluruhan bagus sekali, aku bukan pengguna premium tapi suka diberi premium gratis walaupun cuma beberapa hari, ada uji coba gratis juga beberapa bulan! bagus!

algoritma untuk mengenali pola sentimen dengan lebih tepat. Hasil *cleaning data* ditampilkan pada Tabel 2 :

Tabel 2. Cleaning Data Ulasan

### 3. Tokenizing

Pada tahap tokenisasi, teks dibagi menjadi kata atau token yang memudahkan analisis teks dengan cara yang lebih terorganisir. Proses ini membantu algoritma dalam mengenali pola sentimen secara lebih akurat dan menjadi dasar untuk tahap selanjutnya, seperti penghapusan *stopwords* dan *stemming* yang meningkatkan ketepatan analisis sentimen. Hasil *tokenizing* ditampilkan pada Tabel 3 :

Tabel 3. Ouput *Tokenizing*

<i>Tokenizing</i>
[ 'pembelajarannya', 'tersusun', 'sekali', ',', 'dari', 'yang', 'basic', 'sampe', 'yang', 'rumit', ',', 'bahasa', 'yang', 'diajarkan', 'juga', 'bahasa', 'sehari-hari', ',', 'untuk', 'belajar', 'serius', 'atau', 'hanya', 'dijadikan', 'game', 'juga', 'cocok', ',', 'tapi', 'jika', 'ingin', 'lebih', 'cepat', 'mengerti', 'dan', 'fasih', 'belajar', 'bahasa', 'tidak', 'bisa', 'hanya', 'mengandalkan', 'duolingo', 'saja', ',', 'saran', 'juga', 'sebaiknya', 'grammar', 'juga', 'lebih', 'dijelaskan', '!', 'secara', 'keseluruhan', 'bagus', 'sekali', ',', 'aku', 'bukan', 'pengguna', 'premium', 'tapi', 'suka', 'diberi', 'premium', 'gratis', 'walaupun', 'cuma', 'beberapa', 'hari', ',', 'ada', 'uji', 'coba', 'gratis', 'juga', 'beberapa', 'bulan', '!', 'bagus', '!' ]

### 4. Stopword Removal

Tahap *Stopword Removal* merupakan proses penting dalam pembersihan data teks untuk menghapus kata-kata umum yang tidak memberikan kontribusi pada analisis, seperti "ini," "dan," atau "yang." Penghapusan *stopwords* ini bertujuan untuk mengurangi kompleksitas data dan lebih menekankan pada kata-kata yang relevan untuk analisis sentimen. Hasil *stopword removal* ditampilkan pada Tabel 4 :

Tabel 4. Ouput *stopword removal*

<i>Stopword Removal</i>
[ 'pembelajarannya', 'tersusun', ',', 'basic', 'sampe', 'rumit', ',', 'bahasa', 'diajarkan', 'bahasa', 'sehari-hari', ',', 'belajar', 'serius', 'dijadikan', 'game', 'cocok', ',', 'cepat', 'mengerti', 'fasih', 'belajar', 'bahasa', 'mengandalkan', 'duolingo', ',', 'saran', 'grammar', '!', 'bagus', ',', 'pengguna', 'premium', 'suka', 'premium', 'gratis', ',', 'uji', 'coba', 'gratis', '!', 'bagus', '!' ]

### 5. Stemming

Tahap *stemming* dalam analisis sentimen berfungsi untuk menyatukan kata-kata yang memiliki akar kata sama namun bentuk yang berbeda, seperti "bermain" dan "dimainkan" menjadi "main." Proses ini mengurangi pengulangan, meningkatkan konsistensi data, dan memudahkan algoritma dalam mengenali pola sentimen dengan tepat. *Stemming* memastikan analisis sentimen lebih

efektif dengan memfokuskan pada kata-kata yang relevan. Hasil *stemming* ditampilkan pada Tabel 5 :

Tabel 5. Ouput *stemming*

<i>Stemming</i>
[ 'diupdate', 'ngelag', ',', 'diupdate', 'parah', ',', 'udah', 'nyelesaiin', 'ajar', 'ngebug', 'nggak', 'hitung', 'ngulang', ',', 'ajar', ',', 'apk', 'hapus', 'riwayat', ',', 'kali', 'koneksi', 'internetnya', 'bagus', 'lancar', 'lancar', 'aja', ',', 'kemarin', 'diupdate', 'performa', 'turun', ']' ]

### 6. Slangwords

Penanganan *slangwords* dalam analisis sentimen sangat penting untuk menjaga akurasi klasifikasi sentimen. *Slang* yang tidak ditangani dengan baik dapat menyebabkan kesalahan interpretasi karena kata-kata tersebut sering memiliki makna tertentu. Proses ini meliputi identifikasi *slang* dan penyesuaian makna yang tepat sehingga analisis sentimen dapat lebih akurat mencerminkan ekspresi dalam ulasan pengguna. Hasil *slangwords* ditampilkan pada Tabel 6 :

Tabel 6. Ouput *slangwords*

<i>Slangwords</i>
[ 'seruu', 'banget', 'soalnya', 'gk', 'perlu', 'kursus/les', 'lagi', 'dulu', 'gk', 'bisa', 'bahasa', 'inggris', 'jadi', 'bisa', 'sedikit', ',', 'sedikit', 'demi', 'sedikit', 'jadi', 'lumayan', 'lancar', 'pakai', 'bahasa', 'inggris', 'tapi', 'setiap', 'salah', 'pasti', 'di', 'kasih', 'kesempatan', 'terus', 'pas', 'ngejawab', 'nya', 'gk', 'terasa', 'banget', 'soal', 'nya', 'juga', 'gak', 'terlalu', 'susah', 'pas', 'pertama', 'kali', 'coba', 'biasa', 'aja', 'tapi', 'makin', 'lama', 'jadi', 'lebih', 'seru', 'dan', 'gak', 'seru', 'kalau', 'cuma', 'coba', 'sekali', 'jadi', 'seruu', 'banget', 'banyak', 'teman', 'ada', 'berbagai', 'bahasa', 'mandarin', ',', 'arab', ',', 'inggris', 'dan', 'masih', 'banyak', 'lagi', ',', 'inti', 'nya', 'si', 'seru', 'ya' ]

#### 4.1.4. Transformation

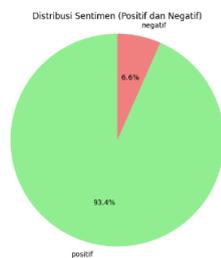
Tahap *transformation* mengonversi data yang telah diproses ke dalam format yang dapat diolah oleh algoritma, termasuk pelabelan sentimen dan penggunaan metode *TF-IDF* untuk merepresentasikan teks dalam bentuk numerik. Proses ini mempersiapkan data untuk analisis lebih lanjut oleh model *machine learning*.

#### Pelabelan

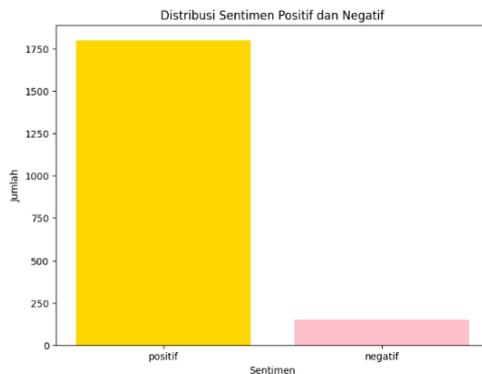
Data ulasan dilabeli sebagai positif atau negatif untuk mendukung analisis lanjutan menggunakan algoritma *Naïve Bayes Classifier*. Hasil pelabelan ditampilkan pada Tabel 7 :

Tabel 7. Proses Menampilkan Pelabelan

Reviews	Labelling
sangat bagus dan pembelajarannya itu tidak langsung dengan pertanyaan yang banyak, sedikit demi sedikit jadi pembelajaran ini membuat kita itu mudah paham dan mudah mengerti juga dan belajarnya itu asik karena sambil bermain dan santai juga terimakasih <i>Duolingo</i> , <i>Duolingo</i> ini sangat membantu 🙏🙏	positif
baru diupdate karena sebelumnya ngelag terus, setelah diupdate malah tambah parah. Padahal udah nyelesaiin pelajarannya karena ngebug jadi nggak dihitung dan akhirnya ngulang lagi yang harusnya misal tiga pelajaran jadi 4 atau bahkan lebih. Jadi harus keluar aplikasj hapus dari riwayat baru masuk lagi, dan ini bukan cuma sekali dua kali padahal koneksi internetnya bagus lancar lancar aja. Padahal baru kemarin diupdate tapi performanya malah menurun.	negatif



Gambar 3. Output Pelabelan



Gambar 4. Output Diagram Batang



Gambar 5. Output Split Ratio

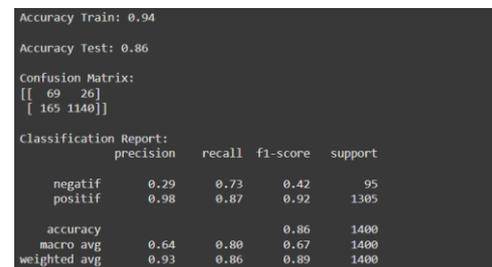
#### 4.1.5. Data Mining

Proses ini membagi data ulasan menjadi data latih dan uji dengan perbandingan

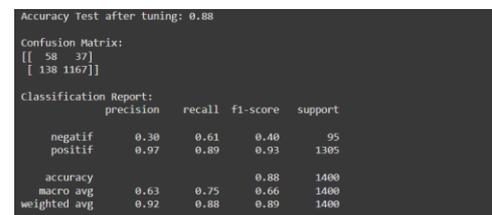
70:30, lalu menggunakan *'TfidfVectorizer'* untuk mengubah teks menjadi format numerik, fokus pada 200 kata yang paling relevan. Data latih diproses menggunakan *fit\_transform*, sementara data uji diubah dengan transform untuk persiapan pelatihan dan pengujian model.

#### 4.1.6. Evaluation / Interpretation

Pada tahap Evaluasi, data dibagi dengan proporsi 70:30 dengan 70% untuk pelatihan dan 30% untuk pengujian. Model *Naïve Bayes* dievaluasi menggunakan metrik seperti *accuracy*, *precision*, *recall*, *F1-score* dan *confusion matrix* untuk menilai kinerjanya dalam analisis sentimen. Evaluasi ini bertujuan untuk mengukur seberapa baik model dapat menggeneralisasi dan memprediksi data yang belum pernah dilihat sebelumnya. Pemilihan rasio 70:30 bertujuan untuk menciptakan keseimbangan antara data pelatihan dan pengujian, mencegah *overfitting* dan memberikan gambaran yang lebih akurat mengenai kinerja model. Hasil laporan klasifikasi ditampilkan pada Gambar 6 :



Gambar 6. Proses Menampilkan Hasil Laporan Akurasi Pemodelan *Naïve Bayes*



Gambar 7. Proses Menampilkan Pemodelan *Naïve Bayes Classifier* dengan menggunakan metode *Grid Search* untuk mengoptimalkan Akurasi



Gambar 8. Output Proses Menampilkan *Cross Validation*

## 4.2 Pembahasan

Penelitian ini akan menilai kinerja algoritma *Naïve Bayes Classifier* dalam menganalisis sentimen ulasan aplikasi *Duolingo*. Evaluasi dilakukan dengan menghitung metrik seperti *accuracy*, *precision*, *recall*, *F1-score* dan *confusion matrix* untuk mengukur efektivitas algoritma. Metrik-metrik tersebut digunakan untuk menilai kemampuan algoritma dalam mengklasifikasikan sentimen pengguna dimana akurasi menunjukkan persentase prediksi yang benar, *precision* dan *recall* fokus pada pengenalan sentimen spesifik, *F1-score* menyeimbangkan *precision* dan *recall* serta *confusion matrix* menggambarkan distribusi prediksi benar dan salah untuk mengidentifikasi kesalahan dalam klasifikasi.

### 4.2.1. Accuracy

Pada model analisis sentimen dengan *Naïve Bayes Classifier*, akurasi mengukur sejauh mana model dapat membuat prediksi yang benar secara keseluruhan. Metrik ini krusial untuk menilai kemampuan model dalam membedakan sentimen positif dan negatif pada ulasan aplikasi *Duolingo*. Akurasi yang tinggi menandakan bahwa model mampu mengklasifikasikan ulasan dengan akurat.

Model analisis sentimen yang menggunakan *Naïve Bayes Classifier* memperoleh *accuracy* 88% yang menandakan kinerjanya yang baik dalam mengklasifikasikan sentimen ulasan *Duolingo*. Model ini berhasil membedakan ulasan positif dan negatif dengan tepat pada 88% data uji yang menunjukkan kemampuannya dalam mengidentifikasi sentimen pengguna dengan tingkat kesalahan yang rendah.

### 4.2.2. Precision

Metrik *precision* mengukur sejauh mana model dapat dengan tepat mengidentifikasi ulasan positif di antara semua prediksi positif yang dihasilkan. *Precision* sangat penting untuk memastikan bahwa ulasan yang diberi label positif benar-benar menggambarkan kepuasan pengguna. *Precision* yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali ulasan positif dengan sedikit kesalahan. Hasil perhitungan menunjukkan bahwa *precision* sebesar 95% menggambarkan akurasi model *Naïve Bayes Classifier* dalam

mengklasifikasikan ulasan positif. Artinya, 95% dari ulasan yang dikategorikan positif oleh model memang benar-benar positif. Namun, terdapat 63 *false positives* yang mengindikasikan bahwa model perlu diperbaiki untuk mengurangi kesalahan dalam mengidentifikasi ulasan negatif sebagai positif sehingga dapat meningkatkan akurasi klasifikasi.

### 4.2.3. Recall

Metrik *recall* pada model *Naïve Bayes Classifier* berfungsi untuk mengukur sejauh mana model dapat menemukan seluruh ulasan positif. *Recall* sangat penting untuk memastikan tidak ada ulasan positif yang terlewat. Dengan demikian, *recall* mengevaluasi kemampuan model dalam mengidentifikasi sentimen positif dalam dataset.

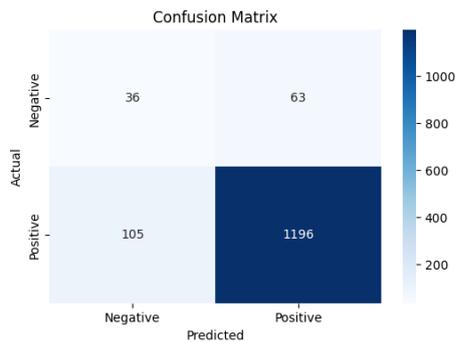
### 4.2.4. F1-Score

Metrik *F1-score* menghitung rata-rata harmonis antara *precision* dan *recall* yang mencerminkan keseimbangan di antara keduanya. *F1-score* sangat penting ketika perlu memperhatikan baik ketepatan dalam mengidentifikasi kelas positif (*precision*) maupun kemampuan model dalam menemukan semua contoh kelas positif (*recall*). *F1-score* memberikan satu angka yang menggambarkan kinerja model secara keseluruhan.

Nilai *F1-Score* sebesar 0.93 menunjukkan keseimbangan yang baik antara *precision* dan *recall* yang mengindikasikan bahwa model efektif dalam mengklasifikasikan ulasan positif dengan akurat.

### 4.2.5. Confusion Matrix

Matriks kebingungan (*confusion matrix*) menyajikan hasil klasifikasi yang benar dan salah dari model serta memperlihatkan distribusi klasifikasi untuk setiap kategori. Ini sangat berguna untuk mengevaluasi cara model mengklasifikasikan ulasan positif dan negatif sekaligus mengidentifikasi kesalahan yang dapat memengaruhi akurasi penilaian terhadap kepuasan pengguna. Dengan menganalisis matriks kebingungan, maka kita dapat meningkatkan kinerja model dalam membedakan sentimen positif dan negatif. *Heatmap Confusion Matrix* pada Gambar 9 :



Gambar 9. Output Proses Menampilkan Heatmap Confusion Matrix

#### 4.2.6. Word Cloud

Word Cloud adalah tampilan visual dari teks yang dimana kata-kata yang sering muncul ditampilkan dalam ukuran besar sedangkan kata-kata yang jarang muncul lebih kecil. Teknik ini digunakan dalam analisis sentimen ulasan aplikasi Duolingo untuk menampilkan kata-kata positif dan negatif. Word Cloud ditampilkan pada Gambar 10 :



Gambar 10. Output Word Cloud Positif



Gambar 11. Output Word Cloud Negatif

### 5. KESIMPULAN

Berdasarkan penelitian penerapan algoritma Naive Bayes Classifier dalam analisis sentimen ulasan pengguna aplikasi Duolingo di Google Play Store, maka dapat disimpulkan, diantaranya :

- 1) Algoritma Naive Bayes Classifier efektif dalam mengklasifikasikan ulasan ke dalam kategori positif dan negatif berdasarkan analisis terhadap 2.000 ulasan. Sebagian besar ulasan positif memuji fitur interaktif dan kemudahan penggunaan aplikasi sementara ulasan negatif menyoroti masalah teknis. Model ini menunjukkan akurasi yang tinggi, memberikan wawasan berharga bagi pengembang.
- 2) Model ini memperoleh accuracy 88%, precision 95%, recall 92% dan F1-score 93% menunjukkan kinerja yang sangat baik dalam mendeteksi sentimen positif. Meskipun mayoritas ulasan positif, beberapa ulasan negatif menyoroti masalah teknis yang menunjukkan bahwa Naive Bayes Classifier dapat diandalkan dalam analisis sentimen dan memberikan masukan yang berguna bagi pengembang aplikasi.

### 6. SARAN

Penelitian yang menerapkan algoritma Naive Bayes Classifier untuk analisis sentimen ulasan pengguna aplikasi Duolingo menunjukkan hasil yang baik dengan accuracy 88% yang menunjukkan bahwa algoritma ini efektif dalam mengklasifikasikan sentimen. Namun, tantangan muncul dari asumsi independensi antar fitur yang bisa menyulitkan dalam memahami konteks teks yang lebih kompleks, seperti ironi atau sarkasme. Proses preprocessing, seperti pembersihan data, stemming dan penghapusan stopwords juga memengaruhi hasil. Untuk penelitian selanjutnya, beberapa rekomendasi yang dapat dipertimbangkan, yaitu :

- 1) Mengumpulkan dataset yang lebih besar dan beragam untuk memperoleh analisis sentimen yang lebih akurat.
- 2) Membandingkan algoritma Naive Bayes dengan metode lain, seperti SVM, Random Forest atau LSTM untuk menemukan algoritma yang lebih efektif.
- 3) Mengimplementasikan aplikasi Duolingo di sekolah untuk anak-anak

berkebutuhan khusus, disesuaikan dengan kebutuhan mereka dan menyediakan pelatihan bagi guru untuk memanfaatkan aplikasi ini secara maksimal.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

### DAFTAR PUSTAKA

- [1] D. Winoto, V. Desta Aditia, C. Sorisa, R. Priskila, and V. Handrianus Pranatawijaya, "Analisis Sentimen Pada Ulasan Pengguna Terhadap Aplikasi Pembelajaran Bahasa Duolingo: Menggunakan Algoritma Naïve Bayes Dan K-Nearest Neighbor," *JATI (Jurnal Mhs. Tek. Inform., vol. 8, no. 3, pp. 3230–3236, 2024, doi: 10.36040/jati.v8i3.9647.*
- [2] S. Mauliddiyah, M. N. F. Hidayat, and F. Rizal, "Analisis sentiment ulasan aplikasi pembelajaran duolingo di play store menggunakan distilbert," vol. 7, pp. 502–511, 2024, doi: 10.37600/tekinkom.v7i1.1395.
- [3] S. Chohan, A. Nugroho, A. Maezar, B. Aji, and W. Gata, "Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naïve Bayes dan Synthetic Minority Over Sampling Technique," vol. 22, no. 2, 2020.
- [4] A. Nuraini, A. Faqih, G. Dwilestari, N. Dienwati Nuris, and R. Narasati, "Analisis Sentimen Terhadap Review Aplikasi Brimo Di Google Play Store Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform., vol. 7, no. 6, pp. 3661–3666, 2024, doi: 10.36040/jati.v7i6.8228.*
- [5] Rahel Lina Simanjuntak, Theresia Romauli Siagian, Vina Anggriani, and Arnita Arnita, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Shopee Dengan Menggunakan Algoritma Naïve Bayes," *J. Tek. Mesin, Elektro dan Ilmu Komput., vol. 3, no. 3, pp. 23–39, 2023, doi: 10.55606/teknik.v3i3.2411.*
- [6] N. Ayu Puspita, F. Tri Anggraeny, and A. Mustika Rizki, "Analisis Sentimen Dengan Algoritma Naïve Bayes Classifier Terhadap Ulasan Aplikasi My F&B Id," *JATI (Jurnal Mhs. Tek. Inform., vol. 8, no. 1, pp. 1090–1095, 2024, doi: 10.36040/jati.v8i1.8842.*
- [7] F. V. Sari, "ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD . ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI," vol. 10, no. 2, pp. 681–686, 2019.
- [8] S. Arofah *et al.*, "Analisis Sentimen Pemakaian Sistem Absensi Berbasis Web," vol. 8, no. 3, pp. 2619–2625, 2024.
- [9] A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," *J. Inform. dan Tek. Elektro Terap., vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.*
- [10] R. Rachman and R. N. Handayani, "Klasifikasi Algoritma Naïve Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," *J. Inform., vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.*
- [11] Rifidinal, "Keefektifan penggunaan duolingo dalam pembelajaran kosakata Bahasa Inggris," *J. Manaj. Pendidik. dan Ilmu Sos., vol. 2, no. 2, pp. 697–704, 2021, [Online]. Available: <https://dinastirev.org/JMPIS/article/view/627>*
- [12] N. L. Mauliddiyah, "PENGARUH BIMBINGAN BELAJAR PADA MASA PANDEMI COVID-19 TERHADAP PRESTASI BELAJAR MATEMATIKA SISWA KELAS 6 SDN KRADINAN 01," p. 6, 2021.
- [13] Heliyanti Susana, "Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet," *J. Ris. Sist. Inf. dan Teknol. Inf., vol. 4, no. 1, pp. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.*