

## ANALISIS SENTIMEN ULASAN PENGGUNA TERHADAP APLIKASI K24KLIK DI GOOGLE PLAY STORE DENGAN ALGORITMA NAÏVE BAYES

Fitri Damayanti, Abdul Rahim, Naufal Azmi Verdikha

Teknik Informatika, Universitas Muhammadiyah Kalimantan Timur

Jl. Ir. H. Juanda No.15, Sidodadi, Kec. Samarinda Ulu, Kota Samarinda, Kalimantan Timur 75124

2111102441114@umkt.ac.id

### ABSTRAK

Aplikasi K24Klik adalah platform yang menyediakan layanan kesehatan dan farmasi sehingga memudahkan siapa saja mendapatkan informasi obat dan melakukan pemesanan tanpa harus mengunjungi apotek secara langsung. Meskipun bermanfaat, pengguna masih mengeluhkan aplikasi K24Klik, seperti keterlambatan pengiriman, pembatalan tanpa persetujuan pengguna, dan sering terjadi gangguan pada sistem. Oleh karena itu, penelitian bertujuan untuk mengetahui efektifitas algoritma *Naive Bayes* yang diterapkan dalam menganalisis sentimen ulasan pengguna terhadap aplikasi K24Klik di Google Play Store. Pada kajian ini, metode yang diterapkan dalam menganalisis sentimen ulasan pengguna meliputi, *Pre-processing* data seperti *case folding*, *tokenizing*, *stopword removal* dan *stemming*, serta Algoritma *Naive Bayes* diimplementasikan dalam mengelompokkan sentimen ulasan kedalam kategori positif, netral dan negatif. Dalam penelitian ini dilakukan dengan 3 perbandingan pembagian data latih dan data uji, seperti 90:10, 80:20, dan 70:30, untuk menentukan rasio yang efektif dalam melakukan analisis sentimen. Dari penelitian ini didapatkan hasil dimana rasio 80:20 menunjukkan hasil terbaik dengan akurasi 80%, presisi 83%, recall 96%, dan F1-score 89%. nilai tersebut membuktikan kemampuan algoritma *Naive Bayes* dalam menganalisis sentimen ulasan pengguna aplikasi dengan sangat baik. Hasil pengujian ini diharapkan dapat menambah masukan bagi pengelola aplikasi dalam memperbaiki kualitas layanan di masa mendatang.

**Kata kunci :** K24Klik, Sentimen, Naive Bayes, Akurasi, Google Play Store

### 1. PENDAHULUAN

Kecanggihan sistem informasi telah memberikan pengaruh besar terhadap beragam bidang kehidupan, termasuk bidang kesehatan. Bentuk nyata dari perkembangan ini adalah munculnya aplikasi K24Klik di Google Play Store. Aplikasi K24Klik dikembangkan oleh Apotek K-24 bekerja sama dengan K24Klik.com dengan tujuan untuk memudahkan setiap orang dalam memperoleh informasi tentang obat, memesan obat, obat generik maupun resep, serta menerima konsultasi kesehatan 24 jam setiap hari dalam seminggu tanpa harus mengunjungi apotek secara langsung [1].

Kepuasan pengguna dan kinerja bisnis jangka panjang berkorelasi langsung dengan keandalan dan kualitas layanan K24Klik [2].

Seperti aplikasi lainnya, K24Klik juga mempunyai kesulitan dan permasalahan tersendiri. Banyak keluhan telah disampaikan pada ulasan di Google Play Store mengenai yang ditawarkan serta masih terdapat beberapa kekurangan dalam penggunaannya, seperti keterlambatan pengiriman, dan dibatalkan tanpa konfirmasi pengguna serta seringkali terjadi error sistem pada aplikasi [3].

Mengingat betapa pentingnya kepuasan pengguna terhadap kinerja aplikasi, penelitian menyeluruh terhadap ulasan pengguna menjadi sangat penting. Analisis sentimen dapat memberikan lebih banyak gambaran tentang bagaimana pandangan pengguna terhadap layanan yang ditawarkan K24Klik [4].

Penelitian ini berupaya menganalisis sentimen dalam ulasan pengguna baik positif maupun negatif

menggunakan algoritma *Naive Bayes* sehingga dapat menjadi landasan bagi pertumbuhan dan peningkatan layanan di masa depan [5].

Algoritma *Naive Bayes* digunakan dalam penelitian ini karena memiliki banyak keunggulan, salah satunya ialah kesederhanaan dan efisiensi algoritma ini menjadikannya salah satu metode yang paling sering digunakan untuk analisis teks [6].

Terlepas dari kesederhanaannya, metode *Naive Bayes* unggul dalam analisis sentimen, terutama pada dataset yang besar [7].

Selain itu, *Naive Bayes* terkenal karena kemampuannya memberikan hasil yang tepat dengan jumlah data latih yang relatif kecil [8].

Algoritma ini merupakan pilihan yang cocok untuk analisis sentimen ulasan pengguna di Google Play Store karena juga mudah diterapkan dan mudah dibaca. Namun, masalah muncul ketika ulasan pengguna memiliki konteks yang beragam, sehingga menghasilkan akurasi klasifikasi yang berbeda [9].

Adapun yang menjadi tujuan dari penelitian ini ialah untuk mengetahui efektifitas Algoritma *Naive Bayes* dalam menganalisis sentimen pengguna menggunakan data ulasan pada aplikasi K24Klik di Google Play Store. Langkah-langkah dalam penelitian ini mencakup pengumpulan data ulasan, *pre-processing* data, pembobotan kata dengan TF-IDF, dan analisis sentimen menggunakan algoritma *Naive Bayes*. Dengan menggunakan metode ini, diharapkan pengguna mampu mendapatkan pengetahuan yang lebih baik mengenai opini terhadap aplikasi K24Klik dari segi positif, netral dan negatif. Oleh karena itu, penelitian

ini diharapkan dapat memberikan informasi pada pengembang aplikasi, dalam menilai ulasan pengguna dan memberikan pemahaman pentingnya analisis sentimen dalam pengembangan serta meningkatkan kualitas aplikasi K24Klik.

## 2. TINJAUAN PUSTAKA

### 2.1. Penelitian Terdahulu

Berikut ini adalah penelitian sebelumnya digunakan sebagai referensi dalam kajian ini. Penelitian yang dilakukan oleh Muhammad Tirta Nugraha, dkk, dengan judul “Analisis Sentimen Ulasan Aplikasi Satu Sehat Pada Google Play Store Menggunakan *Naive Bayes Classifier*” dengan dataset sebanyak 16.245 data terdapat 12.551 data sentimen negatif dan 3.694 data sentimen positif berdasarkan rating bintang pengguna, dengan menggunakan algoritma *Naive Bayes Classifier*. Hasil yang diperoleh dengan skenario pembagian data training sebesar 90% dan data testing 10% dengan nilai *accuracy* sebesar 93,00%, *precision* sebesar 93,00% dan *recall* sebesar 93,00% [10].

Penelitian Selanjutnya yang dilakukan oleh Sendi Alpin Rizaldi yang berjudul “Analisis Sentimen Pengguna Aplikasi Jmo (Jamsostekmobile) Pada Google Play Store Menggunakan Metode *Naive Bayes*”. Data dikumpulkan dengan metode *scraped* pada bulan September 2018 sampai Februari 2023 sebanyak 5000 ulasan dan di labeli ulasan positif 1472 data dan ulasan negatif 3528 data dengan hasil *accuracy* 95%, *precision* 91%, dan *recall* 90% [11].

### 2.2. K24Klik

K24Klik merupakan apotek online pertama dan terlengkap di Indonesia, tersedia 24 jam sehari dalam seminggu. Hal ini memungkinkan masyarakat Indonesia untuk mengakses obat dengan lebih mudah setiap saat. Apotek terdekat akan mengantarkan obat ke lokasi pasien. Aplikasi ini secara otomatis akan mencari apotek terdekat untuk mengantarkan obat ke rumah, tempat kerja, atau lokasi pasien [12].

### 2.3. Analisis Sentimen

Analisis sentimen adalah interpretasi otomatis dan manipulasi data tekstual untuk ekstraksi informasi. Analisis sentimen mengidentifikasi sikap tentang orang, organisasi, atau item dalam koleksi data. Analisis sentimen memiliki dampak dan keuntungan yang signifikan, yang mengarah pada lonjakan penelitian dan aplikasi [13].

### 2.4. Pre-Processing Data

Tahap selanjutnya adalah *pre-processing* data. Pada tahap ini melakukan pembersihan bagian data yang kurang relevan atau tidak terstruktur menjadi lebih terstruktur [14].

*Pre-processing data* dilakukan dengan 4 tahap yaitu *case folding*, *stopword removal*, *tokenizing* dan *stemming*, yang akan dijelaskan sebagai berikut:

#### a. Case Folding

*Case folding* adalah tahapan mengubah seluruh huruf pada teks ulasan, diubah dari huruf besar ke huruf kecil sehingga menghasilkan format penulisan yang konsisten [15].

#### b. Tokenizing

*Tokenizing* adalah tahapan untuk membagi kalimat menjadi beberapa bagian atau kata-kata individual. Pada tahap ini menghasilkan kata yang dikenal sebagai token [16].

#### c. Stopword Removal

*Stopwords Removal* pada tahapan ini dilakukan untuk menghapus kata-kata yang tidak relevan seperti kata sambung yaitu “ke”, “di”, “dari” dan “ini” [17].

#### d. Stemming

*Stemming* adalah tahapan akhir dari *pre-processing*. Tahap ini akan mengubah kata menjadi bentuk dasar dengan cara menghilangkan awalan, imbuhan atau akhiran. Hal ini bertujuan agar kata-kata menjadi lebih sederhana dan mudah untuk dianalisis [18].

### 2.5. Ekstraksi Fitur TF-IDF

Ekstraksi fitur TF-IDF adalah tahap pembobotan kata yang berfungsi mengubah teks menjadi numerik pada data. TF-IDF bertujuan untuk menemukan kata yang penting dalam sebuah dokumen atau kumpulan dokumen. Perhitungan atau rumus TF-IDF dibagi menjadi dua yaitu, TF (*Term Frekuensi*) dan IDF (*Inverse Document Frequency*), yang akan digabungkan pada tahap akhir perhitungan setelah masing-masing dihitung menggunakan rumus dan teknik yang berbeda [19].

### 2.6. Naive Bayes

Penelitian ini menggunakan algoritma *Naive Bayes* dengan probabilitas sederhana berdasarkan pendekatan Teorema Bayes (*Bayes Theorem*) dimana hasil yang diperoleh dari setiap kelas bersifat independen, karena tidak saling berhubungan sehingga hasil yang di dapatkan murni hanya diproses dari dokumen itu sendiri [20]. Dibawah ini adalah rumus dari Teorema Bayes (*Bayes Theorem*).

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \quad (1)$$

Keterangan:

A = Hipotesis data B adalah kelas khusus

B = Data dengan dengan kelas tidak dikenal

$P(A|B)$  = Probabilitas hipotesis A didasarkan pada kondisi B

$P(B|A)$  = Probabilitas hipotesis B didasarkan pada kondisi A

$P(A)$  = Probabilitas hipotesis A

$P(B)$  = Probabilitas hipotesis B

### 2.7. Evaluasi

Evaluasi yang dilakukan dapat mengukur performa algoritma *Naive Bayes* analisis sentimen

pada ulasan pengguna aplikasi K24Klik mencari nilai *Recall*, *Precision*, *Accuraci* dan *F1-Score*.

Nilai-nilai perhitungan tersebut diambil dari data *confusion matrix*, seperti terlihat pada Tabel 1. Data ini dihasilkan setelah pengujian model dengan dataset[21].

Tabel 1. *Confusion matrix*

| Predicted Class | True Class         |                     |                     |
|-----------------|--------------------|---------------------|---------------------|
|                 | Netral             | Positif             | Negatif             |
| Netral          | True Neutral (TNt) | False Positive (FP) | False Negative (TN) |
| Positif         | False Neural (FNt) | True Positive (TP)  | False Negative (TN) |
| Negatif         | False Neural (FNt) | False Positive (FP) | True Negative (TN)  |

Ketika nilai dari keempat term tersebut telah diperoleh, maka nilai *accuracy*, *precision*, *recall* dan *f1-score* dapat diketahui dengan menghitung rumus berikut ini :

- Accuracy* yang menilai seberapa efektif model melakukan klasifikasi secara menyeluruh, pada umumnya ditemukan pada masalah klasifikasi yang seimbang. Rumus ini yang digunakan untuk menentukan akurasi :

$$Akurasi = \frac{TP+TN+TNt}{TP+TN+TNt+FP+FN+FNt} \times 100\% \quad (2)$$

- Precision* adalah metrik yang menghitung tingkat ketepatan prediksi positif suatu model adalah benar. Jika nilai presisi lebih tinggi, semakin baik model dalam mencegah prediksi positif yang salah. Rumus ini yang dapat digunakan untuk menentukan Presisi :

$$Precision = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

- Recall* adalah parameter yang menentukan sejauh mana model dapat menganalisis seluruh kemungkinan positif yang sebenarnya. Rumus ini yang dapat digunakan untuk menentukan Recall:

$$Recall = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

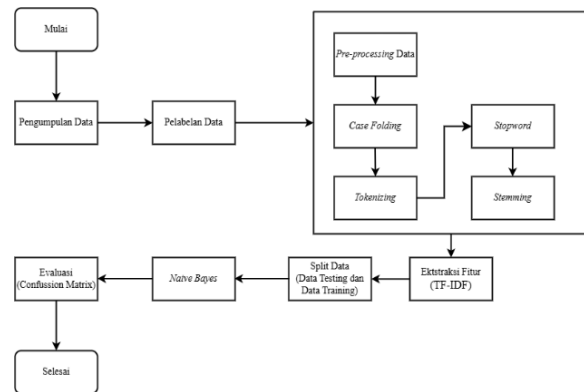
- F1-score* digunakan untuk menyatukan dua metrik, *precision* dan *recall*, dengan memproses nilai keduanya menjadi satu angka tunggal. Rumus ini yang dapat digunakan untuk menentukan Recall :

$$F1 - score = \frac{(2 \times Recall \times Precision)}{Recall + Precision} \quad (5)$$

### 3. METODE PENELITIAN

#### 3.1. Alur Penelitian

Pada penelitian ini akan membahas tahap-tahap yang dilakukan dalam proses riset ini.



Gambar 1. Alur penelitian

Gambar diatas adalah proses-proses yang akan dilakukan dalam Analisis Sentimen Ulasan Pengguna tentang Aplikasi k24klik di Google Play Store dengan Algoritma *Naive Bayes*, pada alur diatas akan dijelaskan pada komponen-komponen di bawah ini.

#### 3.2. Pengumpulan Data

Langkah awal dari penelitian ini adalah mengumpulkan data. Data yang diperoleh berupa ulasan pengguna mengenai aplikasi K24Klik di Google Play Store menggunakan teknik *scraping* melalui Google Colab dengan bahasa pemrograman *python* dari rating 1 sampai dengan rating 5, kemudian akan disimpan dalam format csv.

#### 3.3. Pelabelan Data

Proses pelabelan data dilakukan setelah data dikumpulkan untuk mempermudah analisis sentimen. Pada pelabelan data ini terbagi dalam tiga kelas yaitu positif, netral dan negatif secara manual berdasarkan *score*. Jika *score* bernilai empat dan lima maka ulasan akan diklasifikasikan sebagai positif, jika *score* bernilai tiga maka ulasan akan di kategorikan sebagai netral sedangkan untuk *score* bernilai satu dan dua akan dikategorikan sebagai negatif [22].

#### 3.4. Pre-processing Data

Proses berikutnya ialah *pre-processing* data dilakukan untuk membersihkan data menggunakan teknik *case folding*, *stopword removal*, *tokenize* *stemming* dan kemudian ekstraksi fitur. Setelah melalui proses *cleansing*, dilakukan penlabelan *dataset* dengan memberikan label positif dan negatif pada data.

#### 3.5. Naive Bayes

Penelitian ini menerapkan algoritma *Naive Bayes* dengan probabilitas sederhana berdasarkan pendekatan Teorema Bayes (*Bayes Theorem*) dimana hasil yang diperoleh dari setiap kelas bersifat independen, karena tidak saling berhubungan sehingga hasil yang di dapatkan murni hanya diproses dari dokumen itu sendiri [20].

Dibawah ini adalah rumus dari *Teorema Bayes* (*Bayes Theorem*).

$$P(A|B) = \frac{P(A/B) \cdot P(A)}{P(B)} \quad (6)$$

Keterangan:

A = Hipotesis data B adalah kelas khusus

B = Data dengan dengan kelas tidak dikenal

$P(A|B)$  = Probabilitas hipotesis A didasarkan pada kondisi B

$P(B|A)$  = Probabilitas hipotesis B didasarkan pada kondisi A

$P(A)$  = Probabilitas hipotesis A

$P(B)$  = Probabilitas hipotesis B

## 4. HASIL DAN PEMBAHASAN

### 4.1. Pengumpulan Data

Tahapan pertama yang dilakukan adalah pengumpulan data yang diperoleh dengan teknik *scraped* dalam memperoleh data dari Google Play Store menggunakan Google Colab. Dari hasil proses *scraped* diperoleh sebanyak 4.444 data ulasan dari 24.000 lebih ulasan yang ada. Alasan penggunaan jumlah data tersebut karena adanya pembatasan pustaka dalam menarik data pada aplikasi K24Klik. Berikut hasil *scraping* dapat dilihat dalam Tabel 2 dibawah ini.

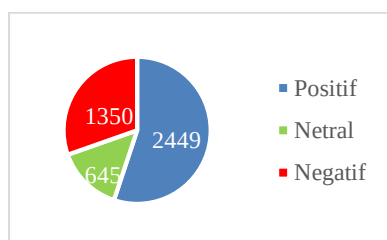
Tabel 2. Data hasil *scraping*

| Score | Ulasan                                                                                                    |
|-------|-----------------------------------------------------------------------------------------------------------|
| 1     | Pengiriman otomatis 3-5 hari, ga bisa diganti, jarak 1,2 km nunggu lebih dari 3 hari? Mending yg lain dah |
| 4     | Aplikasi sangat baik. Dan perlu ditingkatkan aplikasinya lagi                                             |
| 5     | terimakasih K24, cepet banget pelayanannya, semoga semakin sukses yaaa dan makin banyak orang terbantu    |

Pada Tabel 2 diatas menunjukkan beberapa ulasan yang diperoleh dari hasil *scraping* dengan atribut yang digunakan dalam penelitian ini, yaitu *score* dan ulasan.

### 4.2. Pelabelan Data

Setelah mengumpulkan data, tahap selanjutnya adalah memberikan label positif, netral dan negatif berdasarkan *score*. Dari jumlah data sebanyak 4.444 dengan hasil label positif sebesar 2449 data, label netral sejumlah 645 data dan hasil label negatif sebanyak 1350 data. Berikut Gambar 3.1 yang menampilkan persentase label positif, netral dan negatif dalam bentuk diagram lingkaran.



Gambar 2. Diagram pelabelan data

Pada Gambar 2 merupakan diagram lingkaran yang menampilkan jumlah label positif, label netral dan label negatif pada setiap masing-masing dataset.

### 4.3. Pre-processing Data

Langkah pertama dalam *pre-processing* data adalah membersihkan dan menyiapkan data agar lebih terorganisir, sehingga memudahkan dalam meningkatkan akurasi dan kemudahan dalam analisis. Ada beberapa tahap dalam *pre-processing* data yaitu sebagai berikut :

### 4.4. Case Folding

Dalam tahap ini, semua *font* pada ulasan diubah menjadi huruf kecil (*lowercase*) dilakukan, menghapus simbol khusus, *white space* dan emoji. Pada Tabel 3 dibawah menunjukkan hasil sebelum dan sesudah dari pemrosesan *case folding*.

Tabel 3. Hasil *case folding*

| Input                                                                                                  | Output                                                                                               |
|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| Terimakasih K24, cepet banget pelayanannya, semoga semakin sukses yaaa dan makin banyak orang terbantu | terimakasih k24 cepet banget pelayanannya semoga semakin sukses yaaa dan makin banyak orang terbantu |
| Mudah cepat sampai.. dan dapat diakses kapanpun dan di manapun saat di butuhkan.. mantapp pokoknya     | mudah cepat sampai dan dapat diakses kapanpun dan di manapun saat di butuhkan mantapp pokoknya       |
| Beli obat gak usah keluar rumah , praktis sangat membantu.                                             | beli obat gak usah keluar rumah praktis sangat membantu                                              |

### 4.5. Tokenizing

Langkah selanjutnya yaitu tokenisasi, proses ini membagi teks yang awalnya kalimat menjadi kata-kata individual. Pada Tabel 4 dibawah menunjukkan hasil sebelum dan sesudah dari proses *Tokenizing*.

Tabel 4. Hasil *tokenizing*

| Input                                                                                                | Output                                                                                                              |
|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| terimakasih k24 cepet banget pelayanannya semoga semakin sukses yaaa dan makin banyak orang terbantu | [terimakasih, k24, cepet, banget, pelayanannya, semoga, semakin, sukses, yaaa, dan, makin, banyak, orang, terbantu] |
| mudah cepat sampai dan dapat diakses kapanpun dan di manapun saat di butuhkan mantapp pokoknya       | [mudah,cepat,sampai, dan,dapat,diakses, kapanpun,dan,di, manapun,saat, di, butuhkan,mantapp, pokoknya]              |
| beli obat gak usah keluar rumah praktis sangat membantu                                              | [beli,obat, gak, usah, keluar, rumah, praktis, sangat, membantu]                                                    |

### 4.6. Stopword Removal

Dalam tahapan ini, proses menghilangkan kata-kata yang kurang bermakna penting, misalnya kata "semakin", "dan", "makin" dan "banyak" yang terlihat pada ulasan pertama. Pada Tabel 5 dibawah

menunjukkan hasil sebelum dan sesudah dari proses *stopword removal*.

Tabel 5. Hasil *stopword removal*

| Input                                                                                                 | Output                                                                       |
|-------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| [terimakasih,k24,cepat,banget,pelayanannya,semoga,semakin,sukses,yaa,dan,makin,banyak,orang,terbantu] | [terimakasih,k24,cepat,banget,pelayanannya,semoga,sukses,yaa,orang,terbantu] |
| [mudah,cepat,sampai,dan,dapat,diakses,kapanpun,dan,di,mana pun,saat,di,memerlukan,mantapp,pokoknya]   | [mudah,cepat,diakses,manapun,memerlukan,mantapp,pokoknya]                    |
| [beli,obat,gak,usah,keluar,rumah,praktis,sangat,membantu]                                             | [beli,obat,gak,rumah,praktis,membantu]                                       |

#### 4.7. Stemming

Langkah akhir dari *pre-processing* yaitu *stemming*. Pada proses ini menghilangkan imbuhan dan akhiran menjadi kata dasar. Seperti “nya”, “di”, “kan” dan lain sebagainya. Dapat dilihat pada Tabel 6 dibawah hasil dari sebelum dan sesudah dilakukan proses *Stemming* pada data.

Tabel 6 Hasil *stemming*

| Input                                                                         | Output                                                                     |
|-------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| [terimakasih,k24,cepat,banget,pelayanannya,semoga,sukses,yaaa,orang,terbantu] | [terimakasih, k24, cepat, banget, layan, moga, sukses, yaaa, orang, bantu] |
| [mudah,cepat,diakses,manapun,memerlukan,mantapp,pokoknya]                     | [mudah, cepat, akses, mana, butuh, mantapp, pokok]                         |
| [beli,obat,gak,rumah,praktis,membantu]                                        | [beli, obat, gak, rumah, praktis, bantu]                                   |

#### 4.8. Split Data

Pada penelitian ini digunakan tiga perbandingan rasio data yakni 90:10, 80:20 dan 70:30. Perbandingan data latih dan data uji terlihat di Tabel 7 dibawah.

Tabel 7. *Split data*

| Pembagian  | Rasio 90:10 | Rasio 80:20 | Rasio 70:30 |
|------------|-------------|-------------|-------------|
| Data Latih | 3999        | 3555        | 3110        |
| Data Uji   | 445         | 889         | 1334        |

#### 4.9. Ekstraksi Fitur TF-IDF

Langkah selanjutnya adalah ekstraksi fitur untuk pembobotan kata dengan menggunakan *Term Frequency - Inverse Document Frequency* (TF-IDF) untuk data latih dan data uji. Pada Gambar 3 terlihat hasil dari pemrosesan TF-IDF.

|           |                     |
|-----------|---------------------|
| (0, 638)  | 0.20350151769589386 |
| (0, 655)  | 0.24128342430932478 |
| (0, 1173) | 0.2670779932375243  |
| (0, 1298) | 0.1318228283288169  |
| (0, 2636) | 0.17266918785033572 |
| (0, 3387) | 0.3064491366618625  |
| (0, 4293) | 0.25644155309721645 |
| (0, 4444) | 0.249547541576139   |
| (0, 5286) | 0.32566238561504424 |
| (0, 5300) | 0.25209347435142987 |
| (0, 5598) | 0.25701118629412084 |
| (0, 5765) | 0.2958756858387115  |
| (0, 5804) | 0.2417100708308916  |
| (0, 6260) | 0.42222212465285275 |

Gambar 3. Hasil TF-IDF

Berdasarkan hasil ekstraksi diatas, menunjukkan bobot masing-masing kata pada setiap dokumen. Nilai pada matriks mencerminkan seberapa penting suatu term pada dokumen relatif pada keseluruhan dari kumpulan dokumen. Pada output (0, 6260) dapat dilihat bahwa kata dengan indeks ke-6260 memiliki nilai 0. 4222 di dokumen ke-0 menunjukkan bahwa *term* memiliki peran penting dalam dokumen. Pada output (0, 1298) berarti kata dengan indeks ke-1298 memiliki bobot 0.1318 pada dokumen ke-0 mempunyai nilai yang lebih rendah menandakan bahwa *term* kurang penting dalam dokumen.

#### 4.10. Hasil

Dibawah ini adalah tabel hasil dari nilai akurasi, presisi, *recall* dan *f1-score* dari 3 pembagian rasio data.

Tabel 8. Hasil pembagian 3 rasio data

|          | 90:10 | 80:20 | 70:30 |
|----------|-------|-------|-------|
| Akurasi  | 78%   | 80%   | 78%   |
| Presisi  | 80%   | 83%   | 81%   |
| Recall   | 96%   | 96%   | 96%   |
| F1-score | 87%   | 89%   | 88%   |

Hasil pembagian rasio pada Tabel 8 menunjukkan bahwa rasio 80:20 menghasilkan performa tertinggi dengan akurasi 80%, presisi 83%, dan *F1-score* 89%, dibandingkan dengan rasio 90:10 dan 70:30 yang menghasilkan akurasi dan *F1-score* yang lebih rendah (78% dan 87-88%). *Recall* secara konsisten tinggi pada 96% di semua rasio, yang menunjukkan kemampuan model yang baik dalam mendeteksi kelas positif. Dengan demikian, rasio 80:20 memiliki keunggulan disemua metrik perhitungan.

## 5. KESIMPULAN DAN SARAN

Pada penelitian yang dilakukan, dapat disimpulkan bahwa Algoritma *Naive Bayes* terbukti efektif dalam menganalisis sentimen ulasan pengguna terhadap aplikasi K24klik di Google Play Store. Efektifitas ini diukur dari perbandingan tiga rasio yakni 90:10, 80:20 dan 70:30, yang menghasilkan model yang dilatih dengan perbandingan data latih dan data uji 80:20 memiliki performa yang paling baik, dengan akurasi 80%, presisi 83%, *recall* 96%, dan *F1-*

score 89%. Berdasarkan hasil tersebut, model mampu membedakan sentimen positif, netral dan negatif secara akurat, terutama pada data uji yang mewakili seluruh bagian dari populasi. Nilai *recall* model yang secara signifikan tinggi yaitu 96% pada semua rasio, membuktikan kemampuan model dalam mendeteksi ulasan dengan sentimen positif, netral dan negatif dengan baik, namun akurasi mengalami penurunan pada rasio 90:10 dan 70:30. Berdasarkan hasil ini, dapat diketahui bahwa model ini efektif dalam mengkategorikan sentimen ulasan pengguna, dengan rasio 80:20 yang menghasilkan gabungan data latih dan data uji yang sesuai.

Namun, diperlukan lebih banyak penelitian untuk mengidentifikasi aspek-aspek yang dapat menyebabkan kesalahan dan meningkatkan kinerja model secara keseluruhan. Selain itu, penelitian selanjutnya diharapkan dapat mempertimbangkan penggunaan algoritma alternatif, seperti Support Vector Machine atau Deep Learning, untuk membandingkan kinerjanya dengan Naive Bayes. Penelitian ini juga dapat diperluas dengan menggunakan analisis sentimen multikategori seperti, positif, netral dan negatif untuk memperoleh pemahaman yang lebih baik tentang perspektif pengguna

#### DAFTAR PUSTAKA

- [1] A. Yulandari, A. R. Satria, D. Setyawan, and N. A. Budiadi, "Efek Penolakan Perubahan Terhadap Niat Menggunakan Aplikasi Apotek K24klik," *Jurnal Bisnis, Manajemen, dan Ekonomi*, vol. 5, no. 1, pp. 22–37, 2024.
- [2] S. Sriani and S. Suhardi, "ANALISIS SENTIMEN PENGGUNA APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER DAN C4. 5," *JOURNAL OF SCIENCE AND SOCIAL RESEARCH*, vol. 7, no. 2, pp. 555–563, 2024.
- [3] F. Alifiana, M. F. Asnawi, I. A. Ihsannudin, M. A. M. Baihaqy, and D. Asmarajati, "ANALISIS SENTIMEN APLIKASI DUOLINGO MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT MACHINE LEARNING," *Device*, vol. 13, no. 2, pp. 223–230, 2023.
- [4] L. R. Dharmawan, I. Arwani, and D. E. Ratnawati, "Analisis Sentimen pada Sosial Media Twitter Terhadap Layanan Sistem Informasi Akademik Mahasiswa Universitas Brawijaya dengan Metode K-Nearest Neighbor," *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, vol. 4, no. 3, pp. 959–965, 2020.
- [5] M. F. El Firdaus, N. Nurfaizah, and S. Sarmini, "Analisis Sentimen Tokopedia Pada Ulasan di Google Playstore Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 5, pp. 1329–1336, 2022.
- [6] S. H. Alizadeh, A. Hediehloo, and N. S. Harzevili, "Multi independent latent component extension of naive Bayes classifier," *Knowl Based Syst*, vol. 213, p. 106646, 2021, doi: <https://doi.org/10.1016/j.knosys.2020.106646>.
- [7] N. Boyko and K. Boksho, "Application of the Naive Bayesian Classifier in Work on Sentimental Analysis of Medical Data.," *IDDM*, vol. 2020, p. 3rd, 2020.
- [8] H. Chen, S. Hu, R. Hua, and X. Zhao, "Improved naive Bayes classification algorithm for traffic risk management," *EURASIP J Adv Signal Process*, vol. 2021, no. 1, p. 30, 2021.
- [9] A. T. Rizkya, R. Rianto, and A. I. Gufroni, "Implementation of the Naive Bayes Classifier for Sentiment Analysis of Shopee E-Commerce Application Review Data on the Google Play Store," *International Journal of Applied Information Systems and Informatics (JAISI)*, vol. 1, no. 1, 2023.
- [10] M. T. Nugraha, N. N. Sulistiyowati, and U. U. Enri, "ANALISIS SENTIMEN ULASAN APLIKASI SATU SEHAT PADA GOOGLE PLAY STORE MENGGUNAKAN NAÏVE BAYES CLASSIFIER," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 5, pp. 3593–3601, 2023.
- [11] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 109–117, 2023.
- [12] Hestia, "K24Klik.com, Apotek Online Paling Komplit - 24 Menit Sampai!," K24Klik.com.
- [13] C. F. Hasri and D. Alita, "Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," *Jurnal informatika dan rekayasa perangkat lunak*, vol. 3, no. 2, pp. 145–160, 2022.
- [14] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM): Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 375–384, 2024.
- [15] S. Alam and M. I. Sulistyo, "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, 2023.
- [16] R. A. Saputra, D. P. Ray, and F. Irwiensyah, "Analisis Sentimen Aplikasi Tokocrypto Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Naïve Bayes," *KLIK:*

- Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 4, pp. 2028–2036, 2024.
- [17] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *Jurnal KomtekInfo*, pp. 1–7, 2023.
- [18] S. Syafrizal, M. Afdal, and R. Novita, “Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor: Sentiment Analysis of PLN Mobile Application Review Using Naïve Bayes Classifier and K-Nearest Neighbor Algorithm,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 10–19, 2024.
- [19] E. R. N. Mustaqim, U. Pagalay, and C. Crysdian, “Prediksi tingkat kepercayaan masyarakat terhadap PILPRES 2024 menggunakan TF-IDF dan Bow menggunakan metode SVM,” *Jurnal Cahaya Mandalika ISSN 2721-4796 (online)*, vol. 5, no. 1, pp. 515–530, 2024.
- [20] R. Rahmadani, A. Rahim, and R. Rudiman, “ANALISIS SENTIMEN ULASAN ‘OJOL THE GAME’ DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES DAN MODEL EKSTRAKSI FITUR TF-IDF UNTUK MENINGKATKAN KUALITAS GAME,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4988.
- [21] R. D. Yahya, S. A. Wibowo, and N. Vendyansyah, “ANALISIS SENTIMEN UNTUK DETEKSI UJARAN KEBENCIAN PADA MEDIA SOSIAL TERKAIT PEMILU 2024 MENGGUNAKAN METODE SUPPORT VECTOR MACHINE,” 2024.
- [22] F. A. Irawan, A. R. Atmadja, and A. Wahana, “Analisis Sentimen Ulasan Aplikasi Bank Digital Menggunakan Algoritma Naïve Bayes,” *Explorer (Hayward)*, vol. 4, no. 2, pp. 60–68, 2024.