

Analisis Sentimen Pengguna Terhadap Aplikasi Bing Chat di Google Play Store dengan Metode Naïve Bayes

Dimas Cahyo Ramadhan, Faldy Irwiensyah*

Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof.Dr.Hamka, Jakarta, Indonesia

Email: ¹dimascahyo2711@gmail.com, ^{2,*}faldy@uhamka.ac.id

Email Penulis Korespondensi: faldy@uhamka.ac.id

Abstrak—Perkembangan teknologi yang terjadi pada saat ini sudah semakin pesat, sehingga dapat dikatakan sebagai era revolusi teknologi dimana saat ini hampir seluruh kegiatan di masyarakat sudah menggunakan teknologi. Salah satu teknologi yang muncul pada era perkembangan teknologi pada saat ini adalah teknologi kecerdasan buatan (AI). Kecerdasan buatan merujuk pada kemampuan komputer dalam belajar, beradaptasi, dan mengambil keputusan berdasarkan data. Saat ini telah banyak teknologi kecerdasan buatan dalam bentuk aplikasi yang dapat dengan mudah diunduh secara gratis di *Google Play Store*, salah satunya adalah aplikasi hasil kemitraan antara *Microsoft* dengan *OpenAI* yaitu *Bing Chat*. Kehadiran *Bing Chat* sebagai salah satu aplikasi kecerdasan buatan di *Google Play Store* memunculkan berbagai ulasan pengguna selama menggunakan teknologi kecerdasan buatan tersebut. Berdasarkan hal tersebut dibutuhkan sebuah metode untuk menganalisis berbagai ulasan pada aplikasi *Bing Chat* tersebut. Penelitian ini bertujuan untuk menganalisis ulasan sentimen pengguna terhadap aplikasi *Bing Chat* di *Google Play Store* dengan metode *Naïve Bayes*. Sebanyak 2000 data ulasan sentimen pengguna aplikasi *Bing Chat* di *Google Play Store* dalam rentang waktu Januari sampai Februari 2024 dikumpulkan menggunakan metode *web scrapping*. Setelah melalui proses analisis didapatkan data sentimen sebanyak 1877 dengan 1653 data sentimen positif dan 224 data sentimen negatif. Hasil evaluasi dari penelitian ini terhadap sentimen aplikasi *Bing Chat* di *Google Play Store* dengan metode algoritma *Naïve Bayes* mendapatkan hasil nilai akurasi 67.16%, presisi 93.53%, dan *recall* 67.39%.

Kata Kunci: Kecerdasan Buatan; Bing Chat; *Google Play Store*; Sentimen; *Naïve Bayes*

Abstract—The development of technology that occurs at this time is increasingly rapid, so it can be said to be an era of technological revolution where at this time almost all activities in society have used technology. One of the technologies that emerged in the current era of technological development is artificial intelligence(AI) technology. Artificial intelligence refers to the ability of computers to learn, adapt, and make decisions based on data. Currently, there are many artificial intelligence technologies in the form of applications that can be easily downloaded for free on the *Google Play Store*, one of which is the application resulting from the partnership between *Microsoft* and *OpenAI*, namely *Bing Chat*. The presence of *Bing Chat* as one of the artificial intelligence applications on the *Google Play Store* raises various user reviews while using the artificial intelligence technology. Based on this, a method is needed to analyze the various reviews on the *Bing Chat* application. This research aims to analyze user sentiment reviews of the *Bing Chat* application on the *Google Play Store* with the *Naïve Bayes* method. A total of 2000 user sentiment review data for the *Bing Chat* application on the *Google Play Store* in the January to February 2024 timeframe were collected using the web scrapping method. After going through the analysis process, 1877 sentiment data were obtained with 1653 positive sentiment data and 224 negative sentiment data. The evaluation results of this research on the sentiment of the *Bing Chat* application on the *Google Play Store* with the *Naïve Bayes* algorithm method get the results of the accuracy value of 67.16%, precision 93.53%, and recall 67.39%.

Keywords: Artificial Intelligence; Bing Chat; *Google Play Store*; Sentiment; *Naïve Bayes*

1. PENDAHULUAN

Perkembangan teknologi yang terjadi pada saat ini sudah semakin pesat, sehingga dapat dikatakan sebagai era revolusi teknologi dimana saat ini hampir seluruh kegiatan di masyarakat sudah menggunakan teknologi [1]. Salah satu teknologi yang muncul pada era perkembangan teknologi pada saat ini adalah teknologi kecerdasan buatan (AI). Kecerdasan buatan merujuk pada kemampuan komputer dalam belajar, beradaptasi, dan mengambil keputusan berdasarkan data [2]. Berkembangnya teknologi kecerdasan buatan pada saat ini digunakan untuk memudahkan pekerjaan manusia [3]. Pada saat ini banyak aplikasi kecerdasan buatan yang dapat diunduh secara mudah dan gratis di *Google Play Store*, salah satunya adalah aplikasi hasil kemitraan antara *Microsoft* dengan *OpenAI* yaitu *Bing chat*. Aplikasi *Bing Chat* merupakan layanan percakapan berbasis AI yang menggunakan teknologi *GPT* [4].

Google Play Store merupakan sebuah *platform* layanan digital yang digunakan oleh pengguna *smartphone* berbasis *android* untuk mengunduh berbagai aplikasi dan produk – produk lainnya baik secara gratis maupun berbayar [5]. Pada *Google Play Store*, pengguna dapat memberikan ulasan mengenai pengalaman mereka selama menggunakan suatu aplikasi atau produk, ulasan yang diberikan mencakup ulasan positif, negatif, atau netral. Ulasan dapat diartikan sebagai pemberian umpan balik, pendapat, atau evaluasi terhadap suatu aplikasi, konsep ulasan juga melibatkan analisis atau ringkasan dari berbagai sumber seperti buku, berita, film, aplikasi, dan lainnya [6]. Ulasan yang diberikan oleh setiap pengguna di *Google Play Store* dapat memengaruhi calon pengguna sebagai faktor yang dipertimbangkan dalam memutuskan untuk menggunakan suatu aplikasi [7]. Kehadiran *Bing Chat* sebagai salah satu aplikasi kecerdasan buatan di *Google Play Store* memunculkan berbagai ulasan pengguna selama menggunakan teknologi kecerdasan buatan tersebut. Berdasarkan hal tersebut dibutuhkan sebuah metode untuk menganalisis berbagai ulasan pada aplikasi *Bing Chat* tersebut.

Analisis sentimen atau *opinion mining* merupakan cabang ilmu dari *data mining* yang biasanya digunakan untuk menganalisis data tekstual berupa sebuah opini yang mengandung polaritas sehingga nantinya menghasilkan sebuah informasi yang memiliki nilai positif, negatif, atau netral [8]. Analisis sentimen memiliki banyak kegunaan diantaranya melakukan analisis terhadap ulasan produk dan pemantauan popularitas suatu merek [9]. Algoritma *Naïve Bayes*

merupakan suatu metode klasifikasi statistik yang mampu meramalkan probabilitas keanggotaan suatu kelas data sesuai dengan prinsip probabilitas [10]. Algoritma *Naïve Bayes* dipilih karena memiliki kelebihan mudah digunakan, dan hanya membutuhkan lebih sedikit data pelatihan [11].

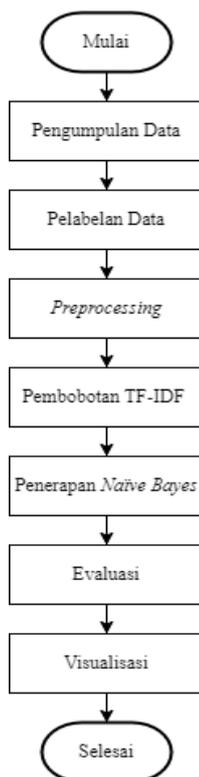
Penelitian serupa mengenai analisis sentimen yang telah dilakukan oleh peneliti – peneliti sebelumnya diantaranya penelitian analisis sentimen terhadap kenaikan harga bahan pokok menggunakan metode *naive bayes classifier* yang dilakukan oleh Muhammad Muslimin dan Veronica Lusiana. Pada penelitian ini dari 2070 *data tweet*, menunjukkan bahwa sentimen negatif lebih sering muncul dibanding sentimen positif dengan nilai persentase sentimen positif 2,8% dan sentimen negatif 97,2%. Setelah dilakukan pengujian didapatkan hasil *accuracy* tertinggi 94,38%, *precision* 59,67%, *recall* 67,93%, dan *F-measure* 62,32% [12]. Penelitian analisis sentimen lainnya dilakukan oleh Indah Novitasari, Tri Basuki Kurniawan, Deshinta Arrova Dewi, dan Misinem. Penelitian tersebut membahas mengenai analisis sentimen masyarakat terhadap *tweet* ruang guru menggunakan algoritma *naive bayes classifier*, penelitian tersebut menunjukkan hasil prediksi lebih dominan positif dan dari pengujian metode *naive bayes classifier* didapatkan hasil presisi 71%, *recall* 69% dan *F1-score* 69% dengan nilai akurasi adalah 69% [13].

Sebagai pembanding untuk mengukur seberapa efektif metode klasifikasi *naive bayes classifier* untuk analisis sentimen terhadap suatu aplikasi kecerdasan buatan diperlukan sebuah penelitian mengenai analisis sentimen dengan metode yang berbeda, seperti penelitian yang dilakukan oleh Muhammad Nanda Fahriza, dan Noviana Riza. Penelitian ini melakukan analisis sentimen terhadap aplikasi *Chat GPT* menggunakan metode klasifikasi *K-Nearest Neighbor*(KNN) dan menghasilkan nilai akurasi 96.6%, jika dibandingkan dengan metode *Naïve Bayes Classifier* yang menghasilkan nilai akurasi sebesar 67.16%, hasil akurasi dari kedua metode klasifikasi analisis sentimen tersebut membuktikan bahwa efektifitas metode *K-Nearest Neighbor*(KNN) dalam menganalisis sentimen lebih tinggi dibanding metode *Naïve Bayes Classifier*. Pembuktian berdasarkan penelitian pembanding yang menganalisis sentimen aplikasi *Chat GPT* dengan menggunakan metode klasifikasi *K-Nearest Neighbor*(KNN). Setelah dilakukan evaluasi, didapatkan nilai akurasi 0.9662921348314607, presisi: 0.48314606741573035, *recall*: 0.5, *F1 score*: 0.4914285714285714 [14].

Berdasarkan uraian yang telah dijelaskan diatas, yang membedakan penelitian ini dengan penelitian sebelumnya yaitu pada penelitian ini peneliti melakukan analisis sentimen dengan menggunakan metode *Naïve Bayes* terhadap aplikasi kecerdasan buatan yang berbeda yaitu aplikasi *Bing Chat*. Penelitian ini bertujuan untuk menganalisis ulasan sentimen pengguna terhadap aplikasi *Bing Chat* di *Google Play Store* dengan metode *Naïve Bayes* dan melihat kecenderungan sentimen pengguna terhadap aplikasi kecerdasan buatan tersebut apakah cenderung positif atau negatif.

2. METODOLOGI PENELITIAN

Pada penelitian analisis sentimen terhadap aplikasi *Bing Chat* dengan menggunakan metode *Naïve Bayes* ini, mencakup beberapa tahapan yang saling berkaitan dalam proses pembuatan penelitian mulai dari pengumpulan data hingga selesai di proses visualisasi. Gambar 1 berikut merupakan alur tahapan pada penelitian ini.



Gambar 1. Alur Tahapan Penelitian

Pada gambar 1 merupakan alur tahapan penelitian dimulai dari tahap pengumpulan data, pelabelan data, *preprocessing*, pembobotan TF-IDF, penerapan *Naïve Bayes*, evaluasi, dan terakhir tahap visualisasi.

2.1 Pengumpulan Data

Pada penelitian ini proses pengumpulan data ulasan terhadap aplikasi Bing Chat pada *Google Play Store* menggunakan metode *web scrapping* dengan bantuan *tools Google Collaboratory*. Metode *web scrapping* adalah tahapan pengambilan data dari dokumen setengah terstruktur di internet, umumnya berupa halaman *web* [15]. Data yang akan diproses pada penelitian ini yaitu data ulasan dan penilaian dari para pengguna aplikasi Bing Chat. Dalam proses ini, data – data yang telah dikumpulkan dianalisis sesuai jumlah data yang dibutuhkan dan kemudian disimpan dalam bentuk *file csv*.

Tabel 1. Contoh Data Yang Digunakan

| <i>Comment</i> | Label |
|---------------------------|---------|
| aplikasi jelek banyak bug | negatif |
| bagus sekaliiii | positif |
| kasih bintang 4 | positif |
| sangat bagus buat foto | positif |
| kurang optimal saat ini | negatif |

Pada tabel 1 merupakan contoh data ulasan sentimen yang akan digunakan dalam penelitian ini.

2.2 Pelabelan Data

Data yang telah dikumpulkan selanjutnya masuk ke tahap proses pelabelan data. Proses ini bertujuan memisahkan data menjadi tiga sentimen positif, negatif, dan netral [16]. Pada tahap ini data ulasan akan diberikan label berdasarkan ulasan dan penilaian yang telah dikumpulkan. Proses pelabelan akan dilakukan secara manual menggunakan *microsoft excel*.

2.3 Preprocessing

Data yang sudah diberi label kemudian masuk ke tahap *preprocessing*. *Preprocessing* merupakan proses yang dilakukan dengan tujuan untuk mempersiapkan teks menjadi data yang akan diolah pada proses selanjutnya [17]. Terdapat beberapa tahapan pada proses ini seperti *cleansing*, *case folding*, *tokenizing*, *filtering stopword* dan *filter tokens by length* dalam bahasa Indonesia [18]. Berikut penjelasan tahapan – tahapan tersebut.

- Cleansing* merupakan langkah untuk menghapus semua karakter yang tidak diinginkan dalam data.
- Case Folding* merupakan langkah dimana semua huruf kapital diubah menjadi huruf kecil secara keseluruhan.
- Tokenizing* merupakan proses memisahkan kalimat menjadi kata per kata agar mendapatkan kata yang memiliki nilai.
- Filtering Stopword* adalah proses penghapusan kata – kata yang tidak memiliki makna.
- Filter Tokens By Length* adalah proses menghilangkan kata berdasarkan panjang karakter.

2.4 Pembobotan TF – IDF

Setelah tahap *preprocessing*, selanjutnya masuk ke tahap pembobotan TF – IDF. Pembobotan TF – IDF (Term Frequency-Inverse Document Frequency) merupakan metode untuk memberikan bobot pada setiap kata berdasarkan tingkat kemunculan kata tersebut dalam sebuah dokumen [19]. Pada proses ini nilai pembobotan akan semakin tinggi apabila jumlah kata serupa yang sering muncul semakin banyak.

2.5 Penerapan Naïve Bayes

Data yang telah dilakukan pembobotan TF – IDF, selanjutnya masuk ke tahap penerapan *naïve bayes*. *Naïve Bayes* mengadopsi asumsi independensi, yang membuatnya unggul dalam kinerja dibandingkan dengan algoritma lain karena memerlukan volume data pelatihan yang lebih kecil namun tetap memberikan hasil yang baik, terutama dalam konteks data kategorikal [20]. Pada tahap ini algoritma *naïve bayes* diterapkan untuk melakukan klasifikasi pada setiap data sentimen yang telah memiliki bobot. Berikut merupakan rumus perhitungan metode *Naïve Bayes* [21].

$$P(c|d) = \frac{P(c).P(d|c)}{P(d)} \quad (1)$$

Dengan keterangan sebagai berikut :

$P(c|d)$: merupakan nilai probabilitas c yang dipengaruhi oleh d (*posterior probability*)

$P(d|c)$: merupakan nilai probabilitas d yang dipengaruhi oleh c (*likelihood*)

Adapun tahapan – tahapan dalam metode *Naïve Bayes* sebagai berikut [22].

- Menghitung jumlah label yang ada
- Menghitung jumlah kasus dari masing – masing label
- Mengalikan semua variabel

2.6 Evaluasi

Pada tahap evaluasi akan dilakukan penghitungan dengan menggunakan metode *confusion matrix*. *Confusion matrix* merupakan metode prediksi perhitungan *matrix* pada akurasi *data mining* [23]. Dengan menggunakan metode ini dapat memungkinkan peneliti untuk melihat seberapa efektif model dalam mengklasifikasi sentimen.

Tabel 2. *Confusion Matrix*

| | <i>True negative</i> | <i>True positive</i> |
|---------------------------|----------------------|----------------------|
| <i>Predicted negative</i> | TN | FN |
| <i>Predicted positive</i> | FP | TP |

Pada tabel 2 menjelaskan mengenai 4 hasil nilai klasifikasi dari metode *confusion matrix* yaitu *True Negative* (TN) adalah nilai negatif yang diprediksi dengan benar. *False Negative* (FN) adalah nilai positif yang terprediksi secara negatif. *False Positive* (FP) adalah nilai negatif yang terprediksi secara positif. *True Positive* (TP) adalah nilai positif yang diprediksi dengan benar. Pada proses evaluasi diterapkan metode perhitungan, seperti akurasi, presisi, dan *recall* [24]. Berikut rumus perhitungan tersebut.

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

$$\text{Presisi} = \frac{TP}{(TP+FP)} \quad (3)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (4)$$

Rumus di atas merupakan rumus perhitungan pada tahap evaluasi dengan menggunakan metode *confusion matrix*. Pada tahap ini akan dilakukan perhitungan untuk mencari nilai akurasi, presisi, dan *recall* berdasarkan 4 hasil klasifikasi metode *confusion matrix* yaitu *True Negative* (TN), *False Negative* (FN), *False Positive* (FP), *True Positive* (TP).

2.7 Visualisasi

Setelah melalui berbagai tahapan diatas, data hasil analisis sentimen terhadap aplikasi Bing pada *Google Play Store* kemudian divisualisasikan kedalam bentuk *worldcloud* untuk menampilkan kata – kata sentimen yang paling sering muncul dalam data tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pengumpulan data ulasan pengguna terhadap aplikasi Bing Chat di *Google Play Store* dilakukan dengan menggunakan metode *web scrapping* dengan bantuan *tools Google Collaboratory*. Proses *scrapping data* menggunakan bahasa pemrograman *python*. Data yang dikumpulkan merupakan data ulasan aplikasi Bing Chat yang paling relevan dalam rentang waktu Januari sampai Februari 2024 di *Google Play Store*. Dari hasil proses *scrapping* didapatkan data sebanyak 2000 data ulasan.

```
[ ] from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.microsoft.bing',
    lang='id',
    country='id',
    sort=Sort.MOST_RELEVANT,
    count=2000,
    filter_score_with=None
)
```

Gambar 2. Proses Pengumpulan Data

Pada gambar 2 diatas merupakan proses pengumpulan data menggunakan metode *web scrapping* dengan bantuan *Google Collaboratory*. Peneliti memasukan tautan '<https://com.microsoft.bing/>' kedalam *script coding* untuk mendapatkan data ulasan aplikasi Bing Chat pada platform *Google Play Store*. Pengambilan data ulasan aplikasi Bing Chat dalam Bahasa Indonesia sebanyak 2000 data berdasarkan data ulasan yang paling relevan.

3.2 Pelabelan Data

Data yang telah dikumpulkan pada proses sebelumnya kemudian masuk proses pelabelan data. Pada proses ini peneliti melakukan pelabelan terhadap 2000 data ulasan aplikasi Bing Chat di *Google Play Store* secara manual dengan

menggunakan *microsoft excel 2019*. Proses pelabelan dilakukan dengan memberikan label sentimen positif atau negatif berdasarkan ulasan dan penilaian data yang telah dikumpulkan sebelumnya.

Tabel 3. Hasil Pelabelan Data

| No | Content | Label |
|----|---------------------|---------|
| 1 | Aplikasi luar biasa | Positif |
| 2 | Sangat memuaskan | Positif |
| 3 | Kenapa sering eror | Negatif |
| 4 | Semakin ribet | Negatif |
| 5 | Mantap | Positif |

Tabel 3 merupakan hasil dari proses pelabelan terhadap 2000 data ulasan aplikasi Bing Chat di Google Play Store yang dilakukan secara manual. Dari 2000 data ulasan, didapatkan 1877 data ulasan yang terdiri dari 1653 data sentimen positif dan 224 data sentimen negatif.

3.3 Preprocessing

Proses selanjutnya masuk ke tahap *preprocessing* dengan menggunakan RapidMiner. Data yang telah diberi label kemudian dipersiapkan dengan tahapan – tahapan *preprocessing* berikut.

- a. Tahap pertama yaitu *Cleansing*, pada tahap ini dilakukan pembersihan data dari semua karakter yang tidak diinginkan seperti simbol, emoji, dan lainnya.



Gambar 3. Proses *Cleansing*

Gambar 3 diatas merupakan proses *Cleansing* menggunakan operator *replace* pada RapidMiner.

Tabel 4. Hasil Proses *Cleansing*

| Sebelum <i>Cleansing</i> | Setelah <i>Cleansing</i> |
|-------------------------------|-------------------------------|
| Plus AI jadi lebih praktis | Plus AI jadi lebih praktis |
| sangat memuaskan ðŸŒ°ðŸŒ°ðŸŒ° | sangat memuaskan |
| Sangat membantu | Sangat membantu |
| ini sangat keren ðŸŽ | ini sangat keren |
| Mantap AI nya membantu sekali | Mantap AI nya membantu sekali |

Tabel 4 menampilkan hasil proses *Cleansing*, terlihat karakter yang tidak diinginkan pada data telah dibersihkan setelah proses *Cleansing*.

- b. Selanjutnya adalah tahap *Case Folding*, pada tahap ini dilakukan perubahan semua huruf kapital menjadi huruf kecil secara keseluruhan.

Tabel 5. Hasil Proses *Case Folding*

| Sebelum <i>Case Folding</i> | Setelah <i>Case Folding</i> |
|-------------------------------|-------------------------------|
| Plus AI jadi lebih praktis | plus ai jadi lebih praktis |
| sangat memuaskan | sangat memuaskan |
| Sangat membantu | sangat membantu |
| ini sangat keren | ini sangat keren |
| Mantap AI nya membantu sekali | mantap ai nya membantu sekali |

Tabel 5 menampilkan hasil dari proses *Case Folding* dimana semua huruf kapital pada data telah diubah menjadi huruf kecil setelah dilakukan proses tersebut.

- c. Setelah tahap *case folding* selanjutnya masuk ke tahap *Tokenizing*, pada tahap ini dilakukan pemisahan kalimat sentimen menjadi kata per kata yang lebih kecil.

Tabel 6. Hasil Proses *Tokenizing*

| Sebelum <i>Tokenizing</i> | Setelah <i>Tokenizing</i> |
|-------------------------------|-----------------------------------|
| plus ai jadi lebih praktis | plus, ai, jadi, lebih, praktis |
| sangat memuaskan | sangat, memuaskan |
| sangat membantu | sangat, membantu |
| ini sangat keren | ini, sangat, keren |
| mantap ai nya membantu sekali | mantap, ai, nya, membantu, sekali |

Tabel 6 menampilkan hasil dari proses *Tokenizing*, dapat dilihat sebelum dilakukan proses *Tokenizing* data sentimen masih dalam bentuk kalimat dan setelah dilakukan proses *Tokenizing* data sentimen telah berubah menjadi kata per kata yang lebih kecil.

d. Selanjutnya adalah tahap *Filtering Stopword*, tahap ini menghapus setiap kata yang tidak memiliki makna.

Tabel 7. Hasil Proses *Filtering Stopword*

| Sebelum <i>Filtering Stopword</i> | Setelah <i>Filtering Stopword</i> |
|-----------------------------------|-----------------------------------|
| plus, ai, jadi, lebih, praktis | plus, ai, praktis |
| sangat, memuaskan | memuaskan |
| sangat, membantu | membantu |
| ini, sangat, keren | keren |
| mantap, ai, nya, membantu, sekali | mantap, ai, membantu |

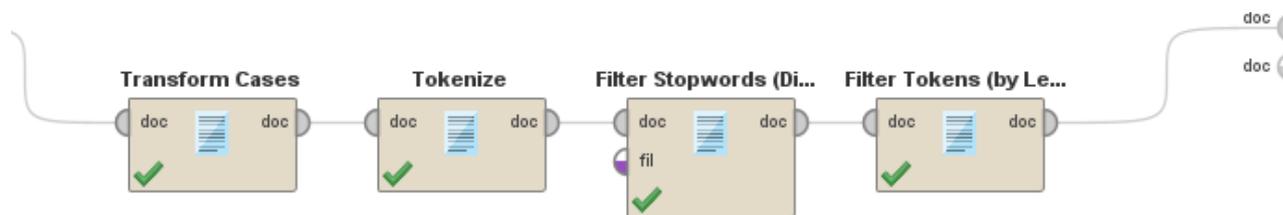
Pada tabel 7 menampilkan hasil dari proses *Filtering Stopword*, peneliti menggunakan dataset *stopword* Bahasa Indonesia yang didapatkan dari *website kaggle.com* yang dimasukkan ke dalam operator *Filter Stopword* pada RapidMiner. Proses ini menghilangkan kata yang tidak memiliki makna yang terdapat dalam data.

e. Selanjutnya adalah tahap *Filter Tokens by Length*, tahap dilakukan proses menghilangkan kata – kata pada data berdasarkan panjangnya karakter.

Tabel 8. Hasil Proses *Filter Tokens by Length*

| Sebelum <i>Filtering Tokens by Length</i> | Setelah <i>Filtering Tokens by Length</i> |
|---|---|
| plus, ai, praktis | plus, praktis |
| memuaskan | memuaskan |
| membantu | membantu |
| keren | keren, |
| mantap, ai, membantu | mantap, membantu |

Pada tabel 8 menampilkan hasil dari proses *Filter Tokens by Length*. Dapat dilihat pada proses ini dengan menggunakan operator *Filter Tokens by Length*, kata yang memiliki panjang kurang dari 4 karakter dan melebihi 25 karakter sudah dihilangkan dari data.



Gambar 4. Proses *Preprocessing*

Pada gambar 4 diatas merupakan operator yang digunakan pada proses *preprocessing*.

3.4 Pembobotan TF – IDF

Data yang telah diposes pada *preprocessing*, selanjutnya dilakukan pembobotan dengan metode TF – IDF. Proses pembobotan TF – IDF dilakukan dengan menggunakan RapidMiner.

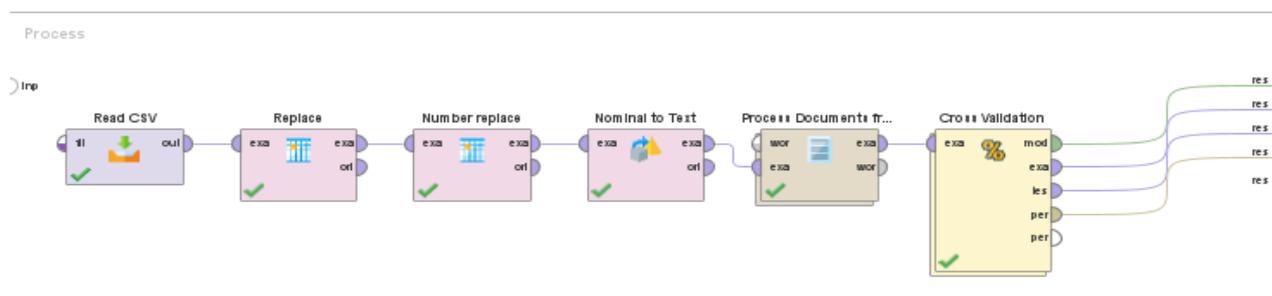
| text | aamiin | abis | abiss | acakanmoh... | aces | account |
|--|--------|------|-------|--------------|------|---------|
| aplikasi bagus membantu menemukan disertai link su... | 0 | 0 | 0 | 0 | 0 | 0 |
| aplikasi bagus bermanfaat berita trending creat desain ... | 0 | 0 | 0 | 0 | 0 | 0 |
| kredit unlimited | 0 | 0 | 0 | 0 | 0 | 0 |
| aplikasinya bagus | 0 | 0 | 0 | 0 | 0 | 0 |
| alhamdulillah mudahmudahan manfaat barokah aamiin | 0.501 | 0 | 0 | 0 | 0 | 0 |
| bintang batasi sekarang | 0 | 0 | 0 | 0 | 0 | 0 |
| please kalo login pake akun google pake microsoft | 0 | 0 | 0 | 0 | 0 | 0 |

Gambar 5. Hasil Proses TF – IDF

Gambar 5 merupakan hasil proses pembobotan TF – IDF dengan operator *Process Documents from Data* pada RapidMiner. Pemberian bobot berdasarkan kata yang tingkat kemunculannya tinggi.

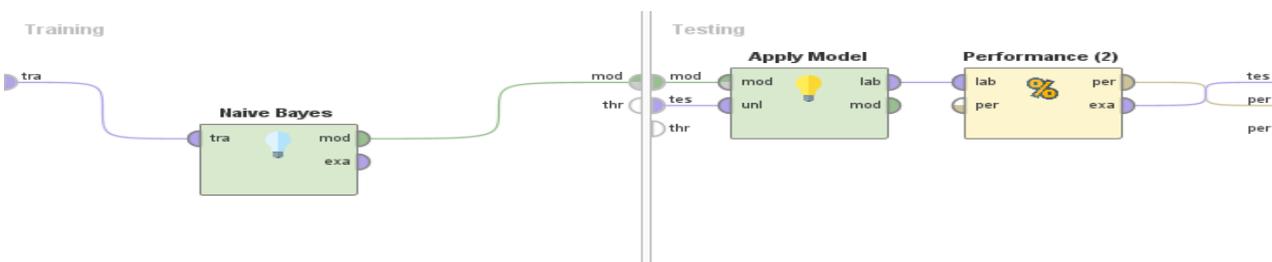
3.5 Penerapan Naïve Bayes

Data yang telah melewati proses pembobotan TF – IDF, kemudian masuk ke proses penerapan *Naïve Bayes* menggunakan operator *Cross Validation*. Pada proses ini dilakukan pengujian mengenai seberapa akurat kinerja algoritma *Naïve Bayes* dalam melakukan klasifikasi terhadap data sentimen yang telah dilakukan pembobotan.



Gambar 6. Proses Analisis Sentimen Menggunakan RapidMiner

Gambar 6 merupakan seluruh operator yang digunakan pada proses analisis sentimen menggunakan RapidMiner, dimulai dari membaca data hasil proses *scrapping* dan pelabelan yang telah disimpan dalam bentuk *csv*. Kemudian masuk ke tahap *Cleansing* oleh operator *Replace* dan *Number Replace*. Kemudian dilakukan perubahan data dari bentuk nominal ke bentuk *text* dan pembobotan TF – IDF oleh operator *Process Documents from Data*. Data kemudian diproses oleh operator *Cross Validation* untuk dilakukan penerapan *Naïve Bayes* dan untuk memunculkan hasil evaluasi dari data sentimen yang telah diproses.



Gambar 7. Penerapan Algoritma Naïve Bayes

Gambar 7 merupakan proses penerapan algoritma *Naïve Bayes*. Pada proses ini dilakukan klasifikasi data menggunakan operator *Naïve Bayes*, kemudian data yang telah diklasifikasi diproses oleh operator *Apply Model* untuk mendapatkan hasil prediksi terhadap data klasifikasi tersebut. Kemudian operator *Performance* melakukan evaluasi kinerja terhadap data klasifikasi algoritma *Naïve Bayes*.

3.6 Evaluasi

Data yang telah selesai pembobotan kemudian dilakukan evaluasi dengan metode *confusion matrix* dengan RapidMiner. Hasil dari evaluasi algoritman *Naïve Bayes* didapatkan nilai akurasi 67.16%, dengan perolehan nilai *confusion matrix* seperti yang terdapat pada gambar dibawah ini.

accuracy: 67.16% +/- 4.77% (micro average: 67.16%)

| | true positif | true Negatif | class precision |
|---------------|--------------|--------------|-----------------|
| pred. positif | 1114 | 77 | 93.53% |
| pred. Negatif | 539 | 146 | 21.31% |
| class recall | 67.39% | 65.47% | |

Gambar 8. Hasil Evaluasi Naïve Bayes

Gambar 8 merupakan hasil dari evaluasi Naïve Bayes dengan menggunakan RapidMiner.

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(1114+146)}{(1114+146+77+539)} = \frac{1260}{1876} = 0,6716$$

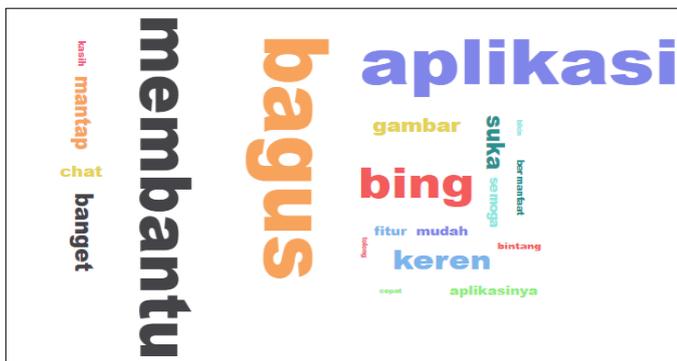
$$\text{Presisi} = \frac{TP}{(TP+FP)} = \frac{1114}{(1114+77)} = \frac{1114}{1191} = 0,9353$$

$$\text{Recall} = \frac{TP}{(TP+FN)} = \frac{1114}{(1114+539)} = \frac{1114}{1653} = 0,6739$$

Dari hasil perhitungan evaluasi algoritman Naïve Bayes diatas, didapatkan nilai akurasi 67.16%, presisi 93.53%, dan recall 67.39%.

3.7 Visualisasi

Hasil dari analisis sentimen pengguna terhadap aplikasi Bing Chat di Google Play Store dengan metode Naïve Bayes. Kemudian divisualisasikan dalam bentuk wordcloud untuk menampilkan hasil berupa kata yang paling sering muncul dalam data sentimen tersebut



Gambar 9. Hasil Visualisasi Wordcloud

Gambar 9 merupakan hasil dari visualisasi data sentimen dalam bentuk wordcloud.

4. KESIMPULAN

Berdasarkan hasil pengumpulan data sentimen yang telah dilakukan dengan metode *web scrapping* terhadap aplikasi Bing Chat di Google Play Store, didapatkan 2000 data sentimen pada rentang waktu Januari sampai Februari 2024. Setelah dilakukan analisis terhadap 2000 data sentimen tersebut, didapatkan 1877 data ulasan yang terdiri dari 1653 data sentimen positif dan 224 data sentimen negatif. Dari hasil analisis data didapatkan kesimpulan bahwa sentimen pengguna terhadap aplikasi Bing Chat di Google Play Store cenderung lebih ke sentimen positif. Hasil evaluasi data sentimen dengan menggunakan metode algoritma Naïve Bayes menunjukkan nilai akurasi 67.16%, presisi 93.53%, dan recall 67.39%. Saran terhadap pengembangan penelitian ini adalah penelitian mengenai aplikasi kecerdasan buatan dapat dikembangkan lagi menggunakan metode pengklasifikasi sentimen lainnya, sehingga hasil dari penelitian ini dapat dijadikan sebagai pembandingan hasil kinerja dan evaluasi pada penelitian selanjutnya.

REFERENCES

- [1] S. Suginam, "Transformasi Digital di Masa Pandemi Covid 19: Studi Fenomenologi Pada UKM Kota Medan," Journal of Business and Economics Research (JBE), vol. 3, no. 2, pp. 296–299, 2022, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejournal.seminar-id.com/index.php/jbe/index>
- [2] U. Muzakir, B. Baharuddin, A. Manuhutu, and H. Widoyo, "Penerapan Kecerdasan Buatan Dalam Sistem Informasi: Tinjauan Literatur Tentang Aplikasi, Etika, dan Dampak Sosial," Jurnal Review Pendidikan dan Pengajaran (JRPP), vol. 6, no. 4, pp. 1163–1169, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://journal.universitaspahlawan.ac.id/index.php/jrpp/index>

- [3] Kharisma Agustya Zahra Salsabilla, Tasya Diva Fortuna Hadi, Widya Pratiwi, and Siti Mukaromah, "PENGARUH PENGGUNAAN KECERDASAN BUATAN TERHADAP MAHASISWA DI PERGURUAN TINGGI," SEMINAR NASIONAL TEKNOLOGI DAN SISTEM INFORMASI, vol. 3, no. 1, pp. 168–175, Sep. 2023, doi: <https://doi.org/10.33005/sitasi.v3i1.371>.
- [4] A. Afgiansyah, "Artificial Intelligence Neutrality: Framing Analysis of GPT Powered-Bing Chat and Google Bard," Jurnal Riset Komunikasi, vol. 6, no. 2, pp. 179–193, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://jurnalrisetkomunikasi.org/index.php/jrk/index>
- [5] P. Aditiya, U. Enri, and I. Maulana, "Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine," JURIKOM (Jurnal Riset Komputer), vol. 9, no. 4, pp. 1020–1028, 2022, Accessed: Feb. 28, 2024. [Online]. Available: <http://www.ejurnal.stmik-budidarma.ac.id/index.php/jurikom/index>
- [6] N. B. Sidauruk and N. Riza, "SENTIMEN ANALISIS DATA PENGGUNA TERHADAP KAI ACCESS," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 7, no. 2, pp. 1297–1303, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/index>
- [7] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes," JATISI (Jurnal Teknik Informatika dan Sistem Informasi), vol. 9, no. 2, pp. 785–795, 2022, Accessed: Feb. 28, 2024. [Online]. Available: <https://jurnal.mdp.ac.id/index.php/jatisi/index>
- [8] B. Laurensz and E. Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," Jurnal Nasional Teknik Elektro dan Teknologi Informasi, vol. 10, no. 2, 2021.
- [9] I. Zulfahmi, "Analisis Sentimen Aplikasi PLN Mobile Menggunakan Metode Decision Tree," Jurnal Penelitian Rumpun Ilmu Teknik, vol. 3, no. 1, pp. 11–21, 2024, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejurnal.politeknipratama.ac.id/index.php/JUPRIT/index>
- [10] U. Kusnia and F. Kurniawan, "Analisis Sentimen Review Aplikasi Media Berita Online Pada Google Play menggunakan Metode Algoritma Support Vector Machines (SVM) Dan Naive Bayes," Explore IT: Jurnal Keilmuan dan Aplikasi Teknik Informatika, vol. 14, no. 1, pp. 24–28, 2022.
- [11] M. M. Alfritri, N. Nurahman, M. Minarni, and D. Rusda, "Evaluasi Performa Algoritma Naïve Bayes Dalam Mengklasifikasi Penerima Bantuan Pangan Non Tunai," Jurnal Media Informatika Budidarma, vol. 7, no. 3, pp. 1433–1445, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/index>
- [12] M. Muslimin and V. Lusiana, "Analisis Sentimen Terhadap Kenaikan Harga Bahan Pokok Menggunakan Metode Naive Bayes Classifier," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 7, no. 3, pp. 1200–1209, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/index>
- [13] I. Novitasari, T. B. Kurniawan, and D. A. Dewi, "Analisis Sentimen Masyarakat Terhadap Tweet Ruang Guru Menggunakan Algoritma Naive Bayes Classifier (NBC)," Jurnal Mantik, vol. 6, no. 3, pp. 3308–3318, 2022, Accessed: Feb. 28, 2024. [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik/index>
- [14] M. N. Fahriza and N. Riza, "ANALISIS SENTIMEN PADA ULASAN APLIKASI CHAT GENERATIVE PRE-TRAINED TRANSFORMER GPT MENGGUNAKAN METODE KLASIFIKASI K-NEAREST NEIGHBOR (KNN)," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 7, no. 2, pp. 1351–1358, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/index>
- [15] R. T. Handayanto, H. Herlawati, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, and D. Y. Septia, "Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine," Jurnal Komtika (Komputasi dan Informatika), vol. 5, no. 2, pp. 153–163, 2021, Accessed: Feb. 28, 2024. [Online]. Available: <https://journal.unimma.ac.id/index.php/komtika/index>
- [16] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," Sistemasi: Jurnal Sistem Informasi, vol. 10, no. 1, pp. 222–238, 2021, Accessed: Feb. 28, 2024. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id/index.php/stmsi/index>
- [17] T. V. Meiyanti, M. Hatta, and A. Sevtiana, "Analisis Sentimen Mahasiswa Dengan Dosen Menggunakan Metode Naïve Bayes Classifier Pada Kuesioner Dosen," Jurnal Manajemen Sistem Informasi, vol. 1, no. 2, pp. 55–59, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://jurnal.cic.ac.id/index.php/jurminsi/index>
- [18] W. Yulita, "Analisis sentimen terhadap opini masyarakat tentang vaksin covid-19 menggunakan algoritma naïve bayes classifier," Jurnal Data Mining dan Sistem Informasi, vol. 2, no. 2, pp. 1–9, 2021, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/JDMSI/index>
- [19] M. Kholilullah, M. Martanto, and U. Hayati, "ANALISIS SENTIMEN PENGGUNA TWITTER (X) TENTANG PIALA DUNIA USIA 17 MENGGUNAKAN METODE NAIVE BAYES," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 8, no. 1, pp. 392–398, 2024, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/index>
- [20] E. Indrayuni, A. Nurhadi, and D. A. Kristiyanti, "Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc," Faktor Exacta, vol. 14, no. 2, pp. 64–71, 2021, Accessed: Feb. 28, 2024. [Online]. Available: https://journal.lppmunindra.ac.id/index.php/Faktor_Exacta/index
- [21] Z. Annisa and B. S. S. Ulama, "Analisis Sentimen Data Ulasan Pengguna Aplikasi 'PeduliLindungi' pada Google Play Store Menggunakan Metode Naïve Bayes Classifier Model Multinomial," Jurnal Sains dan Seni ITS, vol. 11, no. 6, pp. D464–D471, 2023, Accessed: Apr. 02, 2024. [Online]. Available: https://ejournal.its.ac.id/index.php/sains_seni/index
- [22] A. Putri et al., "Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir: Comparison of K-NN, Naive Bayes and SVM Algorithms for Final-Year Student Graduation Prediction," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 3, no. 1, pp. 20–26, 2023, Accessed: Apr. 02, 2024. [Online]. Available: <https://journal.irpi.or.id/index.php/malcom/index>
- [23] D. Sandi, E. Utami, and K. Kusnawi, "Analisis Sentimen Publik Terhadap Elektabilitas Ganjar Pranowo di Tahun Politik 2024 di Twitter dengan Algoritma KNN dan Naïve Bayes," Jurnal Media Informatika Budidarma, vol. 7, no. 3, pp. 1097–1108, 2023, Accessed: Feb. 28, 2024. [Online]. Available: <https://ejurnal.stmik-budidarma.ac.id/index.php/mib/index>
- [24] I. Verawati and B. S. Audit, "Algoritma Naïve Bayes Classifier Untuk Analisis Sentiment Pengguna Twitter Terhadap Provider By. u," Jurnal Media Informatika Budidarma, vol. 6, no. 3, pp. 1411–1417, 2022, Accessed: Feb. 28, 2024. [Online]. Available: <http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/index>