

Analisis Sentimen Pengguna Aplikasi CapCut Pada Ulasan di Play Store Menggunakan Metode Naïve Bayes

Meliyawati, Firman Noor Hasan*

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

Email: ¹yawatimelly104@gmail.com, ^{2,*}firman.noorhasan@uhamka.ac.id

Email Penulis Korespondensi: firman.noorhasan@uhamka.ac.id

Abstrak—Minat untuk berbagi pengalaman yang ditampilkan dalam visualisasi video semakin meningkat sehingga menciptakan tuntutan akan kebutuhan alat pengeditan yang efisien dan sederhana. CapCut merupakan *platform* digital kreatif *all-in-one* yang memungkinkan pengeditan video di browser, desktop, dan *mobile*. Aplikasi CapCut menjadi salah satu aplikasi yang banyak diunduh di Play Store dengan jumlah 500 juta unduhan dan tersedia gratis. Aplikasi CapCut sangat cocok untuk editor pemula karena memiliki tampilan yang sederhana dengan berbagai fitur menarik seperti *template* yang mudah dioperasikan tanpa memerlukan *software* tambahan. Namun hal tersebut tidak dapat menjamin kepuasan para penggunanya. Beragam pengalaman yang dirasakan berpengaruh terhadap penilaian yang diberikan oleh pengguna. Analisis sentimen penting dilakukan untuk mengetahui tingkat kepuasan pengguna yang hasilnya dapat dijadikan sebagai acuan untuk peningkatan kualitas aplikasi. Untuk mengetahui ulasan pengguna terhadap aplikasi CapCut dilakukan analisis sentimen menggunakan metode Naïve Bayes dengan tujuan untuk mengetahui jumlah sentimen positif dan negatif dari ulasan pengguna. Data yang digunakan diambil dari kolom ulasan yang tersedia di Play Store menggunakan teknik *web scrapping* dengan bantuan Google Colab sebanyak 880 data ulasan pengguna. Data dibagi menjadi 60% data latih yaitu 528 ulasan dan 40% data uji yaitu 352 ulasan. Analisis menghasilkan 30 lebih banyak sentimen negatif daripada sentimen positif dengan jumlah sentimen negatif sebanyak 455 ulasan dan jumlah sentimen positif sebanyak 425 ulasan. Berdasarkan evaluasi menggunakan *confusion matrix*, didapatkan hasil *accuracy* sebesar 84.09%, *precision* sebesar 91.91%, dan *recall* sebesar 73.53%.

Kata Kunci: Analisis Sentimen; CapCut; Naïve Bayes; *Confusion Matrix*; Play Store

Abstract—There is an increasing interest in sharing experiences displayed in video visualizations, creating a demand for efficient and simple editing tools. CapCut is an all-in-one creative digital platform that enables video editing on browser, desktop and mobile. The CapCut app is one of the most downloaded apps on the Play Store with 500 million downloads and is available for free. CapCut app is perfect for beginner editors as it has a simple interface with various interesting features such as templates that are easy to operate without the need for additional software. However, this cannot guarantee the satisfaction of its users. Various experiences that are felt affect the assessment given by users. Sentiment analysis is important to determine the level of user satisfaction, the results of which can be used as a reference for improving the quality of the application. To find out user reviews of the CapCut application, sentiment analysis is carried out using the Naïve Bayes method with the aim of knowing the number of positive and negative sentiments from user reviews. The data used is taken from the review column available on the Play Store using web scrapping techniques with the help of Google Colab as much as 880 user review data. The data is divided into 60% training data which is 528 reviews and 40% test data which is 352 reviews. The analysis resulted in 30 more negative sentiments than positive sentiments with the number of negative sentiments totaling 455 reviews and the number of positive sentiments totaling 425 reviews. Based on the evaluation using confusion matrix, the accuracy result is 84.09%, precision is 91.91%, and recall is 73.53%.

Keywords: Sentiment Analysis; CapCut; Naïve Bayes; Confusion Matrix; Play Store

1. PENDAHULUAN

Kemajuan teknologi terjadi di berbagai bidang, terutama dalam bidang internet di era digital saat ini. Terhitung ada 210,03 juta masyarakat Indonesia yang menggunakan internet pada tahun 2021-2022 dan meningkat sebanyak 2,67% pada tahun 2022-2023 menjadi 215,63 juta [1]. Keberadaan internet melahirkan berbagai *platform* digital yang memudahkan individu untuk saling bertukar informasi dan berbagi pengalaman. Media sosial menjadi sarana untuk bisa terhubung dengan berbagai hal [2]. Dengan memanfaatkan media sosial pengguna internet dapat berbagi pengalaman melalui berbagai media seperti foto, video, dan sebagainya. Khususnya minat untuk berbagi pengalaman yang ditampilkan dalam visualisasi video semakin meningkat sehingga menciptakan tuntutan akan kebutuhan alat pengeditan yang efisien dan sederhana dalam penggunaannya [3]. CapCut merupakan *platform* digital kreatif *all-in-one* yang memungkinkan pengeditan video di browser, desktop atau laptop, dan Android. Aplikasi CapCut menjadi salah satu aplikasi yang banyak diunduh di Play Store. Popularitasnya meningkat pesat dengan 500 juta unduhan dan ada ratusan juta pengguna yang aktif menggunakan aplikasi CapCut setiap bulannya [4]. Aplikasi CapCut disebut sebagai saudara kandung TikTok. Sama seperti TikTok, aplikasi CapCut dibuat dan dikembangkan oleh ByteDance yaitu perusahaan dari China yang bergerak di bidang teknologi. Dirilis pada tahun 2019 yang dikenal dengan sebutan “Viamaker” sebelum akhirnya pada tahun 2020 namanya diubah menjadi CapCut [5].

Aplikasi CapCut dapat diunduh melalui Play Store dan tersedia gratis. Aplikasi CapCut memiliki tampilan yang sederhana dengan beragam fitur menarik yang mudah untuk dioperasikan tanpa memerlukan *software* tambahan sehingga sangat cocok untuk editor pemula [6]. Beragam fitur menarik yang dapat digunakan dalam pengeditan video untuk memberikan sentuhan kreatif diantaranya menambahkan musik, teks, stiker, efek dan filter untuk mempercantik, *remove background*, *green screen*, mengubah ukuran dan durasi, *text-to-speech*, dan *speech-to-text* yang dapat mengenali berbagai bahasa. Selain melakukan pengeditan video, pengguna juga dapat menggunakan banyak *template* yang tersedia atau membuat *template* untuk dapat digunakan orang lain. Hasil pengeditan video dapat langsung dihubungkan ke TikTok sehingga banyak pengguna TikTok yang memilih aplikasi CapCut sebagai alat pengeditan [7]. Menyediakan berbagai

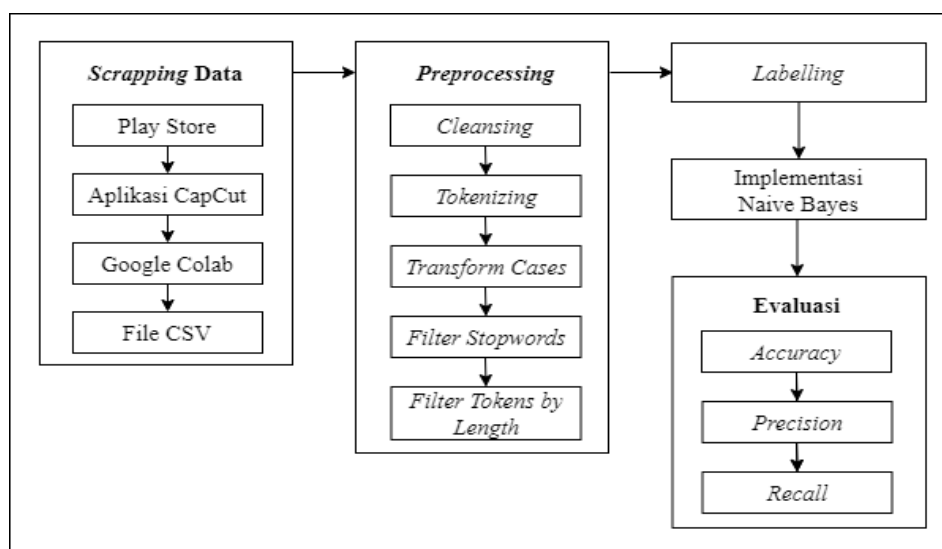
fitur menarik yang tersedia dan kemudahan dalam pengoperasiannya, aplikasi CapCut tidak dapat menjamin kepuasan para penggunanya. Karena setiap aplikasi tentunya memiliki keunggulan dan keterbatasan sehingga menjadikan pengguna mengalami pengalaman dan memiliki sudut pandang berbeda. Hal tersebut dapat dilihat berdasarkan ulasan pengguna pada kolom ulasan yang tersedia di Play Store [8]. Play Store adalah aplikasi resmi yang diproduksi oleh Google untuk dioperasikan pada perangkat *mobile* berbasis Android. Ulasan pengguna pada aplikasi CapCut yang ada di Play Store digunakan sebagai data untuk melakukan analisis sentimen [9].

Analisis sentimen merupakan teknik pemrosesan data yang digunakan untuk mendeteksi data dalam bentuk teks. Teks yang dideteksi dari ulasan pengguna dapat berlabel positif ataupun negatif. Hasil analisis sentimen akan memberikan gambaran umum mengenai opini pengguna dalam menggunakan aplikasi CapCut berdasarkan pengalamannya [10]. Naïve Bayes digunakan sebagai metode untuk mengklasifikasikan ulasan pengguna ke dalam sentimen positif dan sentimen negatif karena metode ini sederhana, cepat, dan memiliki *accuracy* tingkat tinggi [11]. Metode Naïve Bayes sudah banyak digunakan dalam melakukan analisis sentimen seperti pada penelitian sebelumnya mengenai ulasan aplikasi BRImo menghasilkan 1012 sentimen positif dan 894 sentimen negatif dengan *accuracy* sebesar 84,52%, *precision* sebesar 82,51%, dan *recall* sebesar 87,62% [12]. Peneliti Mahardika Tania Nitami dan Herny Februriyanti melakukan penelitian mengenai ulasan aplikasi Ekspedisi J&T Express menghasilkan *accuracy* sebesar 87% dengan sentimen cenderung negatif [13]. Penelitian lain yang membahas mengenai ulasan produk Shopee menghasilkan *accuracy* sebesar 85% [14]. Penelitian mengenai kinerja Dewan Perwakilan Rakyat (DPR) berdasarkan *tweet* menghasilkan 95 positif, 758 negatif, dan 693 netral, serta mendapatkan *accuracy* sebesar 80% [15]. Peneliti Adhithia Erfina dan Muhammad Fani Al-shufi melakukan penelitian mengenai aplikasi Jasa Kurir menghasilkan *accuracy* tinggi sebesar 90% [16].

Berdasarkan penelitian terdahulu, peneliti melakukan penelitian analisis sentimen dengan menggunakan metode Naïve Bayes yang bertujuan untuk mengetahui jumlah sentimen positif dan negatif dari ulasan pengguna aplikasi CapCut, mengindikasikan *accuracy*, *precision*, dan *recall* yang didapatkan dari *confusion matrix*, serta mengetahui keakuratan penggunaan Naïve Bayes dalam pengklasifikasian dengan harapan penelitian ini mendapatkan nilai *accuracy* tinggi. Pengalaman yang dirasakan pengguna tentunya mempengaruhi penilaian yang diberikan terhadap aplikasi CapCut. Penting untuk mengetahui penilaian dan opini pengguna terhadap aplikasi CapCut sehingga nantinya dapat dijadikan sebagai acuan oleh pengembang dalam meningkatkan kualitas aplikasi.

2. METODOLOGI PENELITIAN

Tahap-tahap yang dilakukan dalam penelitian analisis sentimen mengenai ulasan pengguna pada aplikasi CapCut di kolom ulasan Play Store dengan menggunakan metode Naïve Bayes ditampilkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Penelitian dimulai dengan melakukan *scrapping* data ulasan pengguna aplikasi CapCut di Play Store menggunakan bantuan Google Colab. Data yang didapatkan dari hasil *scrapping* kemudian akan memasuki tahap *preprocessing* atau pembersihan seperti *cleansing*, *tokenizing*, *transform cases*, *filter stopwords*, dan *filter tokens by length*. Setelah dilakukan *preprocessing* kemudian dilanjutkan ke *labelling* atau pelabelan. Data dibagi menjadi 2 yaitu data latih dan data uji. Dilakukan pelabelan manual pada data latih yang digunakan untuk melatih model sebelum melakukan pengklasifikasian. Sementara data uji digunakan pada tahap implementasi Naïve Bayes. Pada tahap implementasi Naïve Bayes, data akan diprediksi dan diklasifikasikan ke dalam sentimen positif ataupun sentimen negatif berdasarkan probabilitas kemiripan pengalaman sebelumnya. Tahap akhir dari penelitian adalah evaluasi dari hasil menggunakan *confusion matrix* untuk mengetahui *accuracy*, *precision*, dan *recall*. Berikut persamaan yang digunakan dalam *confusion matrix*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

3. HASIL DAN PEMBAHASAN

3.1 Scrapping Data

Data diambil pada 03 Oktober 2023 menggunakan teknik *web scrapping*. *Web scrapping* adalah teknik yang digunakan untuk mengumpulkan informasi berupa data ulasan dari Play Store dengan tujuan untuk mengekstrak sebuah informasi [17]. Gambar 2 menunjukkan proses ekstraksi data yang dimulai dengan mengakses Play Store, kemudian menuju aplikasi CapCut untuk mendapatkan *link* yang dimasukkan ke dalam Google Colab dan data hasil *scrapping* disimpan dalam format file *Comma Separated Value* (CSV).



Gambar 2. Proses Ekstraksi Data

Gambar 3 menunjukkan proses *scrapping* data ulasan aplikasi CapCut. *Scrapping* data dilakukan dengan bantuan Google Colab dengan *link* tujuan aplikasi CapCut yaitu ‘com.lemon.lvoverseas’ dan menggunakan 2 *library* yaitu *numpy* dan *pandas* [18]. Sebanyak 880 data berupa ulasan pengguna didapatkan dari hasil *scrapping* berdasarkan kategori terbaru (*newest*) berisi atribut nama pengguna (*username*), rating (*score*), tanggal (*at*), dan ulasan (*content*).

```

[ ] from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.lemon.lvoverseas',
    lang='id',
    country='id',
    sort=Sort.NEWEST,
    count=880,
    filter_score_with=None
  )
  
```

Gambar 3. Proses Scrapping Data Ulasan Aplikasi CapCut

Gambar 4 menunjukkan proses menyimpan file. Setelah proses *scrapping* berhasil, data hasil *scrapping* kemudian disimpan dalam format *Comma Separated Value* (CSV) dan disimpan dengan nama “*Scrapping CapCut*”.

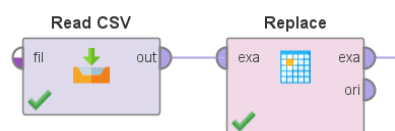
```

[ ] my_df.to_csv("Scrapping CapCut.csv", index = False)
  
```

Gambar 4. Proses Menyimpan File

3.2 Preprocessing

Proses pembersihan atau *preprocessing* berfungsi untuk membersihkan data. Pada tahapan ini data hasil *scrapping* akan diolah dengan memilih data yang akan digunakan dan menghapus data yang tidak berguna sehingga data menjadi lebih terstruktur [19]. Tahap *preprocessing* terdiri dari *cleansing*, *tokenizing*, *transform cases*, *filter stopwords*, dan *filter tokens by length*. *Tools* yang digunakan dalam *preprocessing* adalah RapidMiner. Gambar 5 menunjukkan proses *cleansing* yang dimana proses dimulai dengan membaca keseluruhan data ulasan menggunakan operator *read CSV*. Kemudian dihubungkan dengan operator *replace* untuk menghapus karakter-karakter khusus dan tanda baca pada setiap ulasan.



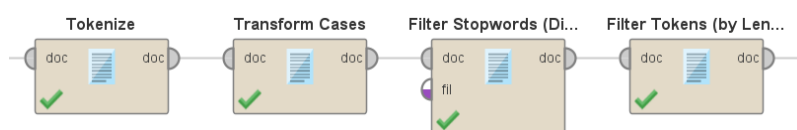
Gambar 5. Proses Cleansing

Tahap pertama adalah *cleansing*. *Cleansing* dilakukan untuk menyaring data seperti menghapus karakter-karakter khusus dan tanda baca pada data ulasan [20].

Tabel 1. Hasil Cleansing

Sebelum	Sesudah
Saya suka banget pake capcut, Pekerjaan Editing semakin cepat.	Saya suka banget pake capcut Pekerjaan Editing semakin cepat
Aplikasi ini sangat membantu buat saya yang sering bikin konten atau edit vidio...	Aplikasi ini sangat membantu buat saya yang sering bikin konten atau edit vidio
Sangat membantu dalam membuat video	Sangat membantu dalam membuat video
Keren banget, suka sekali dengan aplikasi ini	Keren banget suka sekali dengan aplikasi ini
Bagus aplikasi nya aku suka	Bagus aplikasi nya aku suka

Gambar 6 menunjukkan bagian operator *preprocessing*. Setelah melalui proses *cleansing* sebagai tahap awal, proses dilanjutkan dengan beberapa tahapan yaitu *tokenizing*, *transform cases*, *filter stopwords*, dan *filter tokens by length*.



Gambar 6. Operator Preprocessing

Tahap kedua adalah *tokenizing*. *Tokenizing* dilakukan untuk memisahkan potongan-potongan kata pada kalimat [21]. Potongan-potongan kata pada setiap ulasan dipisahkan menggunakan operator *tokenize*.

Tabel 2. Hasil Tokenizing

Sebelum	Sesudah
Saya suka banget pake capcut Pekerjaan Editing semakin cepat	Saya, suka, banget, pake, capcut, Pekerjaan, Editing, semakin, cepat
Aplikasi ini sangat membantu buat saya yang sering bikin konten atau edit vidio	Aplikasi, ini, sangat, membantu, buat, saya, yang, sering, bikin, konten, atau, edit, vidio
Sangat membantu dalam membuat video	Sangat, membantu, dalam, membuat, video
Keren banget suka sekali dengan aplikasi ini	Keren, banget, suka, sekali, dengan, aplikasi, ini
Bagus aplikasi nya aku suka	Bagus, aplikasi, nya, aku, suka

Tahap ketiga adalah *transform cases*. *Transform cases* dilakukan untuk mengubah semua data ulasan jadi huruf kecil dengan menggunakan opsi *lower case* [22]. Proses perubahan ulasan jadi huruf kecil secara keseluruhan menggunakan operator *transform cases*.

Tabel 3. Hasil Transform Cases

Sebelum	Sesudah
Saya, suka, banget, pake, capcut, Pekerjaan, Editing, semakin, cepat	saya, suka, banget, pake, capcut, pekerjaan, editing, semakin, cepat
Aplikasi, ini, sangat, membantu, buat, saya, yang, sering, bikin, konten, atau, edit, vidio	aplikasi, ini, sangat, membantu, buat, saya, yang, sering, bikin, konten, atau, edit, vidio
Sangat, membantu, dalam, membuat, video	sangat, membantu, dalam, membuat, video
Keren, banget, suka, sekali, dengan, aplikasi, ini	keren, banget, suka, sekali, dengan, aplikasi, ini
Bagus, aplikasi, nya, aku, suka	bagus, aplikasi, nya, aku, suka

Tahap keempat adalah *filter stopwords*. *Filter stopwords* dilakukan untuk menghapus kata yang masuk dalam kategori tidak penting berdasarkan kamus *stopwords* Bahasa Indonesia [23]. Kamus *stopwords* Bahasa Indonesia diunduh dari www.kaggle.com. Penghapusan kata tidak penting berdasarkan kamus menggunakan operator *filter stopwords* (*dictionary*).

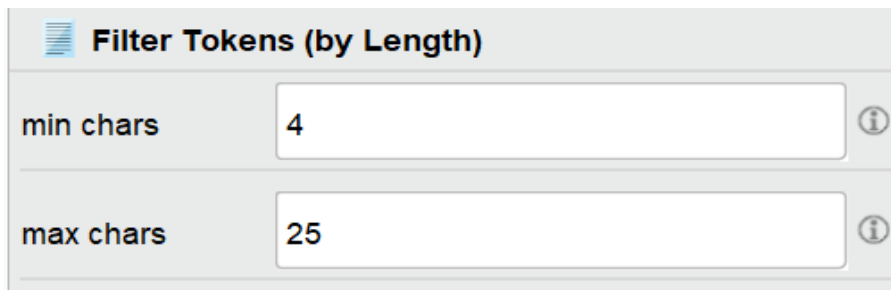
Tabel 4. Hasil Filter Stopwords

Sebelum	Sesudah
saya, suka, banget, pake, capcut, pekerjaan, editing, semakin, cepat	suka, banget, pake, capcut, pekerjaan, editing, cepat
aplikasi, ini, sangat, membantu, buat, saya, yang, sering, bikin, konten, atau, edit, vidio	aplikasi, membantu, bikin, konten, edit, video

sangat, membantu, dalam, membuat, video keren, banget, suka, sekali, dengan, aplikasi, ini bagus, aplikasi, nya, aku, suka

membantu, video keren, banget, suka, aplikasi bagus, aplikasi, nya, suka

Tahap kelima adalah *filter tokens by length*. *Filter tokens by length* dilakukan untuk membatasi karakter [24]. Pada proses ini batas minimum dan maksimum adalah 4 sampai 25 karakter. Gambar 7 menunjukkan proses pembatasan karakter. Pembatasan karakter menggunakan operator *filter tokens (by length)*.



Gambar 7. Proses Pembatasan Karakter

Tabel 5. Hasil Filter Tokens by Length

Sebelum	Sesudah
suka, banget, pake, capcut, pekerjaan, editing, cepat aplikasi, membantu, bikin, konten, edit, vidio membantu, video keren, banget, suka, aplikasi bagus, aplikasi, nya, suka	suka, banget, pake, capcut, pekerjaan, editing, cepat aplikasi, membantu, bikin, konten, edit, vidio membantu, video keren, banget, suka, aplikasi bagus, aplikasi, suka

3.3 Labelling

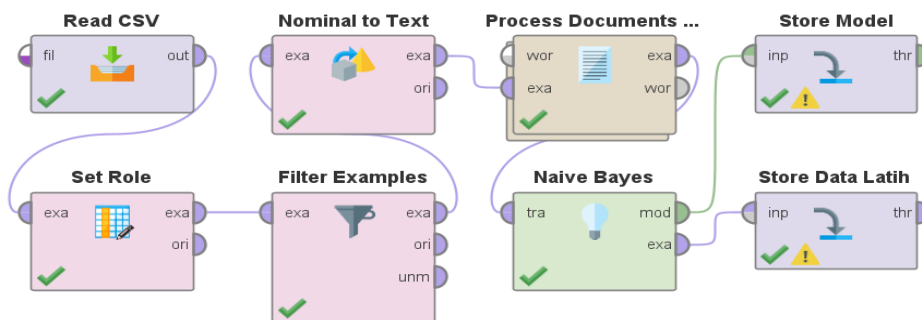
Setelah *preprocessing*, tahapan selanjutnya adalah *labelling* atau pelabelan. Pelabelan dilakukan secara manual dengan pembagian data latih dan data uji. Dari data hasil *scrapping* sebanyak 880 ulasan pengguna yang sudah melalui tahap *preprocessing* diantaranya 528 digunakan untuk data latih dan 352 digunakan untuk data uji. Data latih sebanyak 528 ulasan pengguna diberi label positif atau negatif secara manual berdasarkan ulasan yang dimana ulasan yang diberi label positif adalah ulasan yang berisi pujian atau kepuasan pengguna. Sementara ulasan yang diberi label negatif adalah ulasan yang berisi ketidakpuasan pengguna. Sisanya sebanyak 352 ulasan pengguna tidak dilabeli untuk dijadikan sebagai data uji dalam klasifikasi Naïve Bayes [25].

Tabel 6. Hasil Labelling

Sentimen	Ulasan
Positif	bagus
Negatif	fitur pencarian mengalami kendala
Positif	bagus aplikasinya banyak fitur
Negatif	banyak bug
Positif	praktis mudah digunakan

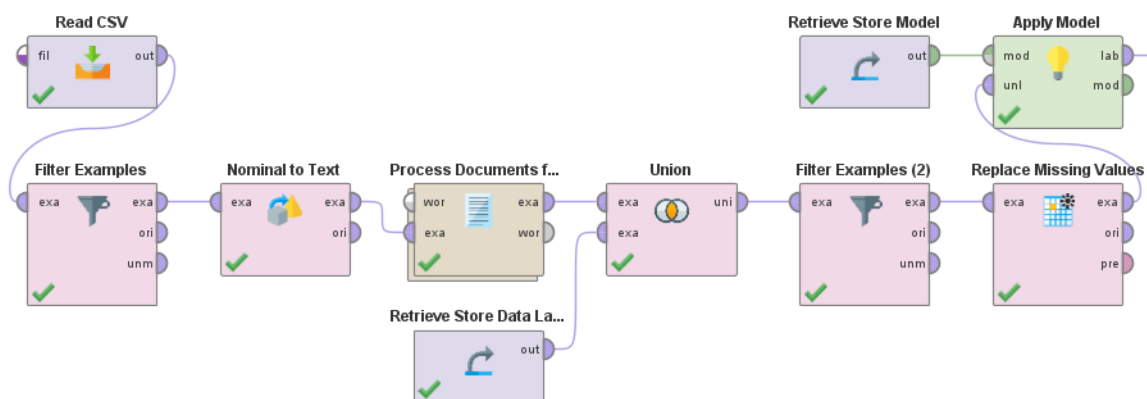
3.4 Implementasi Naïve Bayes

Naïve Bayes mengklasifikasikan berdasarkan prinsip probabilitas dan statistik yang termasuk dalam salah satu metode pada *machine learning*. Metode Naïve Bayes bekerja dengan cara memprediksi probabilitas kemiripan pengalaman sebelumnya dengan probabilitas di masa depan. Naïve Bayes menjadi metode yang terbukti memiliki *accuracy* yang tinggi dalam pengaplikasiannya pada data dengan jumlah yang besar [26]. Pada tahap ini dimulai dengan pembuatan data latih. Gambar 8 menunjukkan proses pembuatan data latih. File yang berisi data ulasan pengguna yang sudah dilabeli dan tidak dilabeli dimasukkan pada operator *read CSV*. Kemudian operator *read CSV* dihubungkan pada operator *set role* untuk menjadikan kolom sentimen dibaca sebagai label. Selanjutnya dihubungkan dengan operator *filter examples* untuk menyaring data yang tidak hilang (*is not missing*) atau data yang sudah dilabeli secara manual. Operator *nominal to text* digunakan agar dapat melanjutkan ke operator berikutnya dan bertujuan untuk mengubah *nominal ke text*. Operator *process document* berisi operator *tokenize*, *transform cases*, *filter stopwords*, dan *filter tokens by length*. Kemudian digunakan operator Naïve Bayes untuk melakukan analisis pada data yang akan disimpan pada *store* bernama *store model* dan *store data latih*. Hasil analisis akan digunakan pada tahap selanjutnya di gambar 9.



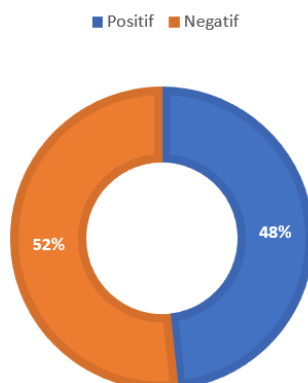
Gambar 8. Proses Pembuatan Data Latih

Setelah proses pembuatan data latih, tahap selanjutnya adalah implementasi Naïve Bayes. Gambar 9 menunjukkan proses implementasi Naïve Bayes. Proses implementasi dimulai dengan membaca data menggunakan operator *read CSV* yang dihubungkan dengan operator *filter examples* untuk menyaring data yang hilang (*is missing*) atau yang belum dilabeli. Penjelasan mengenai *nominal to text* dan operator *process document* sama dengan yang terlihat pada gambar 8. Selanjutnya data hasil analisis pada gambar 8 yang bernama *store data latih* diunionkan dengan hasil dari operator *process document* dan dihubungkan dengan operator *filter examples*, kemudian dihubungkan lagi dengan operator *replace missing values*. Sementara data hasil analisis yang bernama *store model* langsung dihubungkan dengan operator *apply model* dan digabungkan dengan data hasil analisis sebelumnya (*store data latih*).



Gambar 9. Proses Implementasi Naïve Bayes

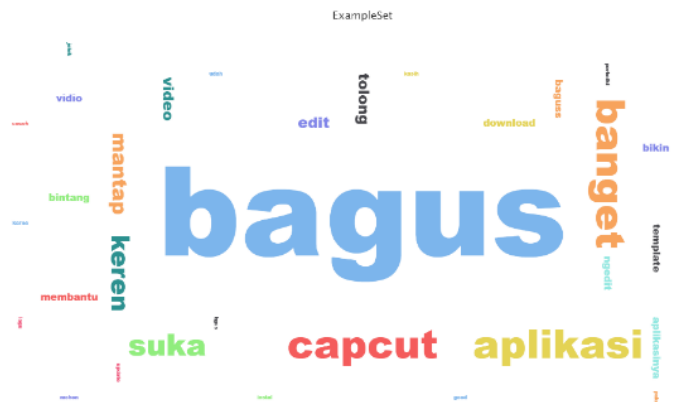
Setelah melalui tahap *preprocessing*, *labelling*, dan implementasi Naïve Bayes, didapatkan hasil prediksi. Gambar 10 menunjukkan diagram pie dari hasil implementasi Naïve Bayes. Dari jumlah 880 data didapatkan sebanyak 455 ulasan yang masuk dalam sentimen negatif dan sebanyak 425 ulasan yang masuk dalam sentimen positif. Berdasarkan hasil prediksi, dapat disimpulkan bahwa sentimen pengguna terhadap aplikasi CapCut cenderung negatif yang dimana ada 30 lebih banyak sentimen negatif daripada sentimen positif. Hal tersebut disebabkan oleh banyaknya pengguna yang menyayangkan adanya versi *update* pada aplikasi CapCut. Pengguna merasa tidak puas dan mengalami beberapa kendala setelah melakukan *update* pada aplikasi CapCut.



Gambar 10. Diagram Pie

Selain diagram pie, hasil implementasi Naïve Bayes juga divisualisasikan dalam bentuk *wordcloud* untuk melihat kata apa yang paling banyak muncul dan yang paling sering digunakan dalam 880 data ulasan pengguna yang digunakan

dalam penelitian [27]. Gambar 11 menunjukkan visualisasi *wordcloud* yang menampilkan kata dengan *range* dari 1 sampai 35 setelah data diurutkan secara *descending* (dari besar ke kecil). Pada *wordcloud* terlihat kata ‘bagus’ adalah kata yang paling banyak digunakan dengan jumlah kemunculan sebanyak 296 kali. Kemudian diurutkan kedua yang memiliki jumlah kemunculan 105 kali adalah kata ‘banget’. Dan diurutkan ketiga yang muncul sebanyak 102 kali adalah kata ‘aplikasi’.



Gambar 11. Visualisasi Wordcloud

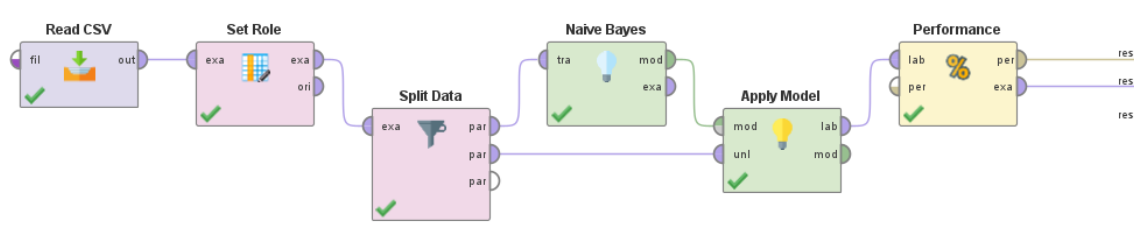
3.5 Evaluasi

Tahap akhir dalam proses analisis sentimen yaitu evaluasi. Evaluasi ini bertujuan untuk mengetahui hasil kinerja suatu model. Yang digunakan untuk mengetahui hasil kinerja pada penelitian ini adalah *confusion matrix*. *Confusion matrix* digunakan untuk menghitung tingkat *accuracy*, *precision*, dan *recall* [28]. Tabel 7 menunjukkan istilah-istilah yang digunakan dalam *confusion matrix*. *True Positive* (TP) pada saat *actual class* dan *predicted class*-nya positif. *False Positive* (FP) pada saat *actual class*-nya negatif dan *predicted class*-nya positif. *False Negative* (FN) pada saat *actual class*-nya positif dan *predicted class*-nya negatif. Dan *True Negative* (TN) pada saat *actual class* dan *predicted class*-nya negatif.

Tabel 7. Confusion Matrix

Evaluasi		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Gambar 12 menunjukkan proses *confusion matrix*. Proses *confusion matrix* dimulai dengan memasukkan file berisi data hasil prediksi menggunakan metode Naïve Bayes ke dalam operator *read CSV* yang kemudian dihubungkan dengan operator *set role* untuk menjadikan kolom *prediction*(*sentiment*) dibaca sebagai label. Operator *split* data digunakan untuk membagi data dengan ratio 60% banding 40% yang kemudian dihubungkan dengan operator *apply model*. Lalu operator *apply model* dihubungkan dengan operator *performance* untuk melihat hasil *accuracy*.



Gambar 12. Proses Confusion Matrix

Hasil dari proses *confusion matrix* menghasilkan nilai *accuracy* sebesar 84.09% dengan nilai *True Positive* (TP) sebesar 125, *False Negative* (FN) sebesar 45, *False Positive* (FP) sebesar 11, dan *True Negative* (TN) sebesar 171. Tabel 8 menunjukkan hasil *accuracy* klasifikasi.

Tabel 8. Hasil Accuracy Klasifikasi

Accuracy: 84.09%			
	True Positif	True Negatif	Class Precision
Pred. Positif	125	11	91.91%
Pred. Negatif	45	171	79.17%
Class Recall	73.53%	93.96%	

Berikut hasil perhitungan dari *accuracy*, *precision*, dan *recall* menggunakan persamaan yang ada pada *confusion matrix* berdasarkan hasil data.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{125+171}{125+171+11+45} = \frac{296}{352} = 0.8409 = 84.09\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{125}{125+11} = \frac{125}{136} = 0.9191 = 91.91\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{125}{125+45} = \frac{125}{170} = 0.7353 = 73.53\%$$

Berdasarkan perhitungan, hasil evaluasi klasifikasi data ulasan pengguna aplikasi CapCut didapatkan sebesar 84.09% untuk nilai *accuracy*, sebesar 91.91% untuk nilai *precision*, dan sebesar 73.53% untuk nilai *recall*.

4. KESIMPULAN

Data sebanyak 880 ulasan pengguna dari kolom ulasan di Play Store diambil menggunakan teknik *web scrapping* dengan *link* tujuan aplikasi CapCut yaitu 'com.lemon.lvoverseas' yang kemudian dimasukkan ke Google Colab dan data hasil *scrapping* disimpan dalam format file *Comma Separated Value (CSV)*. Setelah melalui tahap *preprocessing* seperti *cleansing*, *tokenizing*, *transform cases*, *filter stopwords*, dan *filter tokens by length*, tahap selanjutnya adalah *labelling*. Pada tahap *labelling* dilakukan pembagian data untuk data latih dan data uji diantaranya 528 digunakan untuk data latih dan 352 digunakan untuk data uji. Setelah dibagi, data latih sebanyak 528 dilabeli secara manual. Dilanjutkan dengan implementasi Naïve Bayes yang mendapatkan hasil 30 lebih banyak sentimen negatif daripada sentimen positif. Ada 455 ulasan yang termasuk sentimen negatif dan 425 ulasan yang termasuk sentimen positif. Hal tersebut disebabkan oleh banyaknya pengguna yang menyangkan adanya versi *update* pada aplikasi CapCut. Pengguna merasa tidak puas dan mengalami beberapa kendala setelah melakukan *update*. Analisis sentimen menggunakan metode Naïve Bayes dan *confusion matrix* di tahap evaluasi pada penelitian ini menghasilkan sebesar 84.09% untuk nilai *accuracy*, sebesar 91.91% untuk nilai *precision*, dan sebesar 73.53% untuk nilai *recall*. Berdasarkan hasil *accuracy* yang didapatkan, dapat disimpulkan bahwa metode Naïve Bayes memiliki keakuratan tinggi dalam pengklasifikasian data. Saran untuk penelitian selanjutnya adalah menggunakan lebih banyak data untuk mendapatkan hasil *accuracy* di atas 90% dan membuat perbandingan antara satu metode dengan metode lainnya.

REFERENCES

- [1] F. A. Indriyani, A. Fauzi, and S. Faisal, "Analisis Sentimen Aplikasi Tiktok Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine," *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, vol. 10, no. 2, pp. 176–184, Jul. 2023, doi: 10.37373/teknov10i2.419.
- [2] U. Yudatama et al., *Memahami Teknologi Informasi: Prinsip, Pengembangan, dan Penerapan*. Kaizen Media Publishing, 2023.
- [3] A. R. Abdullah and F. N. Hasan, "Analisis Sentimen Terhadap Kandidat Calon Presiden Berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier," *SMATIKA: STIKI Informatika Jurnal*, vol. 13, no. 1, pp. 117–130, Jul. 2023, doi: 10.32664/smatika.v13i01.750.
- [4] I. D. Rahayu, "Penerapan Aplikasi Capcut Untuk Meningkatkan Keterampilan Berbicara Bahasa Inggris Kelas IX-A SMP Negeri 2 Gudo," *Jurnal PETISI*, vol. 03, no. 02, pp. 44–50, 2022, doi: 10.36232/jurnalpetisi.v3i2.2585.
- [5] D. Nurdiansyah, P. Handayani, and F. Zabadi, "Peningkatan Skill Editing Video Karang Taruna Menggunakan Aplikasi Mobile Phone CapCut Di Desa Lengkong Kulon," *Communnity Development Journal*, vol. 4, no. 3, pp. 5572–5580, 2023, doi: 10.31004/cdj.v4i3.17082.
- [6] V. Wijaya, N. Sitompul, U. H. Mulyanto, and S. Wahyuni, "Pelatihan Editing Video Menggunakan Aplikasi Smartphone CapCut Bagi Ibu-Ibu Gabungan Organisasi Wanita Sambas," *Jurnal Pengabdian kepada Masyarakat Nusantara (JPKMN)*, vol. 4, no. 2, pp. 1343–1351, 2023, doi: 10.55338/jpkmn.v4i2.980.
- [7] K. Anam, "Perancangan Video Promosi Iklan Indomie Menggunakan Aplikasi Ibispaint X, Flipaclip, Dan Capcut," *Jurnal Sistem Informasi dan Teknologi Informasi*, vol. 1, no. 3, pp. 10–18, 2022, doi: 10.1234/siti.v1i3.123.
- [8] M. Riski, M. Fikry, and Yusra, "Klasifikasi Sentimen Ulasan Aplikasi WhatsApp di Play Store Menggunakan Metode K-Nearest Neighbor," *KLIK (Kajian Ilmiah Informatika dan Komputer)*, vol. 4, no. 1, pp. 438–444, 2023, doi: 10.30865/klik.v4i1.1050.
- [9] Y. W. S. Putra et al., *Pengantar Aplikasi Mobile*. Haura Utama, 2023.
- [10] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [11] N. Q. Rizkina and F. N. Hasan, "Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes," *Journal of Information System Research*, vol. 4, no. 4, pp. 1136–1144, 2023, doi: 10.47065/josh.v4i4.3803.
- [12] M. Khoirul, U. Hayati, and O. Nurdianwan, "Analisis Sentimen Aplikasi BRImo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, pp. 478–482, 2023, doi: 10.36040/jati.v7i1.6373.
- [13] M. T. Nitami and H. Februariyanti, "Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma Naïve Bayes," *Jurnal Manajemen Informatika & Sistem Informasi (MISI)*, vol. 5, no. 1, pp. 20–29, 2022, doi: 10.36595/misi.v5i1.396.
- [14] L. O. Sihombing, H. Hannie, and B. A. Dermawan, "Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naïve Bayes Classifier," *Edumatic: Jurnal Pendidikan Informatika*, vol. 5, no. 2, pp. 233–242, Dec. 2021, doi: 10.29408/edumatic.v5i2.4089.

- [15] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, pp. 34–40, Jan. 2022, doi: 10.23960/jitet.v10i1.2262.
- [16] A. Erfina and M. Fani Al-shufi, "Analisis Sentimen Aplikasi Jasa Kurir Di Play Store Menggunakan Algoritma Naive Bayes," *Jurnal Sistem Informasi dan Informatika (SIMIKA)*, vol. 5, no. 2, pp. 103–110, 2022, doi: 10.47080/simika.v5i2.1789.
- [17] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 9, no. 2, pp. 247–253, Jul. 2020, doi: 10.32736/sisfokom.v9i2.903.
- [18] E. Yuniar, D. Safiroh, and D. Wahyuningsih, "Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompok Digital dengan Menggunakan Algoritma Machine Learning," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, no. 1, pp. 35–42, Apr. 2022, doi: 10.25008/janitra.v2i1.145.
- [19] A. Saputra and F. Noor Hasan, "Analisis Sentimen Terhadap Aplikasi Coffee Meets Bagel Dengan Algoritma Naïve Bayes Classifier," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 2, pp. 465–474, Jan. 2023, doi: 10.54443/sibatik.v2i2.579.
- [20] A. Purnamawati, M. N. Winarto, and M. Mailasari, "Analisis Sentimen Aplikasi TikTok menggunakan Metode BM25 dan Improved K-NN Fitur Chi-Square," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 7, no. 1, pp. 97–105, May 2023, doi: 10.31603/komtika.v7i1.8938.
- [21] A. P. Astuti, S. Alam, and I. Jaelani, "Komparasi Algoritma Support Vector Machine Dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo," *Bangkit Indonesia*, vol. 11, no. 2, pp. 1–6, 2022, doi: 10.52771/bangkitindonesia.v11i2.196.
- [22] A. Wibowo, F. Noor Hasan, L. Akbar Ramadhan, R. Nurhayati, and dan Arief Wibowo, "Analisis Sentimen Opini Masyarakat Terhadap Keefektifan Pembelajaran Daring Selama Pandemi COVID-19 Menggunakan Naïve Bayes Classifier," *Jurnal Asimetri: Jurnal Ilmiah Rekayasa Dan Inovasi*, vol. 4, no. 2, pp. 239–248, 2022, doi: 10.35814/asiimetrik.v4i1.3577.
- [23] K. Yolanda, Yusra, and M. Fikry, "Klasifikasi Sentimen Ulasan Aplikasi Whatsapp Di Play Store Menggunakan Naïve Bayes Classifier," *J-INTECH (Journal of Information and Technology)*, vol. 11, no. 1, pp. 1–9, 2023, doi: 10.32664/j-intech.v11i1.867.
- [24] I. T. Julianto, D. Kurniadi, M. R. Nashrulloh, and A. Mulyani, "Comparison Of Classification Algorithm And Feature Selection In Bitcoin Sentiment Analysis," *Jurnal Teknik Informatika (JUTIF)*, vol. 3, no. 3, pp. 739–744, 2022, doi: 10.20884/1.jutif.2022.3.3.343.
- [25] A. Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.
- [26] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *TEKNO KOMPAK*, vol. 15, no. 1, pp. 131–145, 2021, doi: 10.33365/jtk.v15i1.744.
- [27] M. Raffi, A. Suharso, and I. Maulana, "Analisis Sentimen Ulasan Aplikasi Binar Pada Google Play Store Menggunakan Algoritma Naïve Bayes," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 6, no. 1, pp. 451–462, 2023, doi: 10.31539/intecom.v6i1.6117.
- [28] B. Gunawan, H. Sasty Pratiwi, and E. Esyudha Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 4, no. 2, pp. 17–29, 2018, doi: 10.26418/jp.v4i2.27526.