

ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI BSI MOBILE MENGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER

Bagus Suseto Gunawan Wicaksono¹, Evanita², Aditya Akbar Riadi³

^{1,2,3}Universitas Muria Kudus

Email: ¹201951161@std.umk.ac.id, ²Evanita@umk.ac.id, ³aditya.akbar@umk.ac.id

(Naskah masuk: 10 April 2024, diterima untuk diterbitkan: 25 Mei 2024)

Abstrak

Penelitian ini bertujuan untuk mengevaluasi dan meningkatkan kualitas produk M-Banking milik Bank Syariah Indonesia (BSI) yaitu aplikasi BSI Mobile serta menganalisa sentimen ulasan positif dan negatif terhadap aplikasi tersebut menggunakan algoritma Naïve Bayes Classifier. Metode yang digunakan pada penelitian ini meliputi Scrapping data, text preprocessing, pembobotan TF-IDF, dan modelling dengan algoritma Naïve Bayes Classifier. Data mentah yang digunakan adalah ulasan pelanggan dari laman BSI Mobile di Google Play Store yang kemudian diproses melalui beberapa tahap preprocessing seperti casefolding, tokenization, stopword removal, dan stemming. Output dari penelitian ini adalah pemisahan antara komentar positif dan negatif terkait aplikasi BSI Mobile berdasarkan ulasan pengguna yang dapat digunakan sebagai bahan evaluasi pengembangan aplikasi BSI Mobile. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes Classifier memiliki akurasi sebesar 89,67% dengan nilai polaritas positif sebesar 89%, negatif 7%, dan netral 4%.

Kata kunci: Sentimen Analisis, BSI Mobil, Naïve Bayes Classifier, Google Play Store

Sentiment Analysis of BSI Mobile App Reviews Using Naïve Bayes Classifier Algorithm

Abstract

This research aims to evaluate and improve the quality of Bank Syariah Indonesia's (BSI) M-Banking product, namely the BSI Mobile application, and analyze the positive and negative sentiment reviews of the application using the Naïve Bayes Classifier algorithm. The method used in this research includes data Scrapping, text preprocessing, TF-IDF weighting, and modeling with the Naïve Bayes Classifier algorithm. The raw data used is customer reviews from the BSI Mobile page on the Google Play Store, which are then processed through several preprocessing stages such as case folding, tokenization, stopword removal, and stemming. The Output of this research is the separation of positive and negative comments related to the BSI Mobile application based on user reviews, which can be used as a basis for evaluating the development of the BSI Mobile application. The results show that the Naïve Bayes Classifier algorithm has an accuracy of 89,67% with a positive polarity value of 89%, negative 7%, and neutral 4%.

Keywords: *Sentiment Analysis, BSI Mobile, Naïve Bayes Classifier, Google Play Store*

1. PENDAHULUAN

Bank merupakan institusi yang memainkan peran penting dalam sektor ekonomi suatu negara, tidak hanya sebagai tempat penyimpanan uang seperti deposito, giro, dan tabungan (Marimin et al., 2015). Bank juga memiliki fungsi-fungsi lain seperti pembayaran tagihan listrik, belanja *Modelling*, internet, dan pembelian saham, serta berbagai produk digital seperti *Mobile Banking (M-Banking)*.

M-Banking merupakan salah satu produk digital yang disediakan oleh bank kepada nasabah untuk memudahkan transaksi keuangan (Yusmad, 2018). Setiap bank memiliki aplikasi *M-Banking* masing-masing, termasuk Bank Syariah Indonesia (BSI) yang memiliki fitur yang berbeda dengan bank-bank konvensional, karena menggunakan konsep syariah, seperti fitur pembayaran zakat, amal, jam waktu sholat, dan layanan islami lainnya. BSI memiliki aplikasi *M-Banking* bernama *BSI Mobile*, yang telah diunduh lebih dari 5 juta kali oleh pengguna di

Google Play Store, dengan rating 4.4/5.0 dan 114 ribu ulasan.

Karena antusiasme pengguna dalam memberikan ulasan untuk aplikasi BSI *Mobile* di *Google Play Store*, penulis tertarik untuk menganalisis dan menggali lebih dalam tentang pendapat dan komentar pengguna terhadap aplikasi *Mobile Banking* milik BSI ini. *Google Play Store* sendiri adalah layanan penyedia konten digital dari *Google* yang menawarkan berbagai produk daring seperti aplikasi, permainan, *film*, musik, dan buku dalam berbagai genre (Herlinawati et al., 2020).

Oleh karena itu dibutuhkan sistem yang dapat melakukan sentimen analisis terhadap aplikasi BSI *Mobile* agar dapat menganalisis dan menyimpulkan ulasan pengguna aplikasi tersebut serta memberikan rekomendasi masukan terhadap pihak pengembang aplikasi agar dapat berbenah dan mengevaluasi kinerja aplikasi BSI *Mobile*.

Beberapa studi sebelumnya telah dilakukan mengenai analisis sentimen pada ulasan aplikasi *Mobile*, salah satunya adalah penelitian yang berjudul "Analisis Sentimen Pada Ulasan Aplikasi *Mobile* Menggunakan *Naïve Bayes Classifier* dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA *Mobile*)". Penelitian ini melibatkan beberapa tahapan, yaitu *text preprocessing*, normalisasi kata menggunakan metode Levenshtein Distance, dan klasifikasi menggunakan algoritma *Naïve Bayes*. Dalam hasil penelitian tersebut, ditemukan bahwa penggunaan algoritma Levenshtein Distance untuk normalisasi kata yang salah penulisannya dapat meningkatkan akurasi klasifikasi hingga 96,9% (Gunawan et al., 2017).

Salah satu penelitian lainnya yang dilakukan adalah "Analisis Sentimen pada Layanan Ojek *Modelling* Menggunakan Metode *Naïve Bayes*". Proses penelitian dimulai dengan tahap *preprocessing* teks, dilanjutkan dengan tahap pelatihan *Naïve Bayes*, kemudian dilakukan klasifikasi sentiment dengan perhitungan *VMAP*, dan dilakukan proses pengujian. Hasil penelitian menunjukkan bahwa sistem analisis sentimen mampu mencapai akurasi sebesar 80% (Garbian Nugroho et al., n.d.).

Salah satu penelitian lainnya yang dilakukan adalah "Analisis Sentimen Kualitas Layanan Teknologi Pembayaran Elektronik pada *Twitter* (Studi Kasus Ovo dan Dana)". Pada penelitian ini, terdapat beberapa tahapan seperti akuisisi data, *preprocessing* teks, klasifikasi, dan evaluasi. Metode yang digunakan dalam

penelitian ini adalah *Naïve Bayes Classifier*, *Support Vector Machine*, dan *Naïve Bayes*. Hasil penelitian menunjukkan bahwa aplikasi Ovo mendapatkan 78,92% sentimen negatif dan 21,08% sentimen positif, sedangkan aplikasi Dana mendapatkan 86,20% sentimen negatif dan 13,80% sentimen positif. Tujuan dari penelitian ini adalah untuk membandingkan kualitas layanan antara kedua aplikasi dompet digital, yaitu Dana dan Ovo (Ogi et al., n.d.).

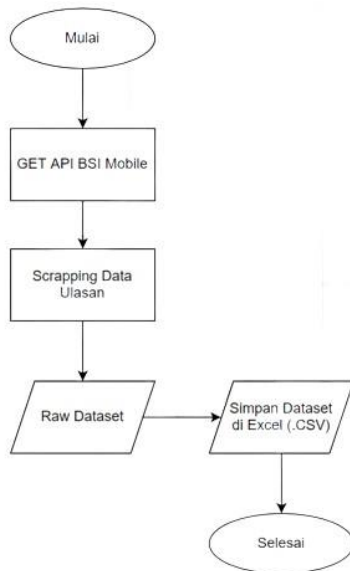
Penelitian terkait berikutnya membahas tentang "Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN". Penelitian ini menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dengan beberapa tahapan seperti *business understanding*, *data understanding*, *data collection*, *data preparation*, *modelling*, dan *evaluation*. Hasil dari penelitian menunjukkan bahwa pada ulasan aplikasi Bareksa, sentimen negatif terkait kurangnya metode pembayaran yang tersedia, kesulitan pada saat registrasi, dan aplikasi yang di perbaharui ke versi terbaru kurang bermanfaat. Sedangkan sentimen positif terkait kemudahan penggunaan aplikasi dan registrasi investasi yang lengkap. Pada aplikasi Bibit, ulasan negatif terkait pencairan dana dan registrasi yang memakan waktu lama, sementara ulasan positif terkait kepuasan pengguna terhadap kemudahan menggunakan robo advisor yang lengkap dan fitur yang disediakan. (Dwiki et al., 2021).

Penelitian yang berjudul "Analisis Sentimen Aplikasi Streaming Film dengan Algoritma *Support Vector Machine* di Play Store" melalui empat tahapan, yaitu pengambilan data, *preprocessing*, pemodelan data, dan *cross validation*. Algoritma yang digunakan adalah *Support Vector Machine*. Hasil dari penelitian menunjukkan bahwa aplikasi iflix adalah yang terbaik dengan skor akurasi 92,67%, namun sentimen negatifnya jauh melebihi sentimen positif, sehingga skor akurasi bukanlah satu-satunya metrik untuk menentukan aplikasi terbaik. Hal ini disebabkan oleh proses pembaruan aplikasi yang tidak sesuai dengan harapan pengguna. Selain itu, aplikasi Netflix mengalami masalah dalam metode pembayaran yang jarang digunakan oleh masyarakat Indonesia, dan banyak ulasan pengguna yang memberikan penilaian negatif terhadap layanan yang ditawarkan oleh aplikasi dengan akurasi yang rendah. Beberapa masalah yang dihadapi oleh pengguna termasuk konten yang tidak lengkap, tidak tersedia terjemahan bahasa Indonesia, dan bug yang banyak di aplikasi (Fani Al-shufi & Erfina, 2021).

2. METODOLOGI PENELITIAN

2.1 Data Acquisition

Tahapan pertama dalam penelitian ini adalah Data Acquisition, yang merupakan proses pengumpulan data. Data yang diambil berasal dari ulasan pengguna pada aplikasi BSI *Mobile*, dan metode yang digunakan adalah *Scrapping* dengan menggunakan bantuan *library Google Play Scrapper*. *Scrapping* adalah metode pengumpulan data melalui halaman *web internet* dengan bantuan API (*Application Programming Interface*), dan data yang berhasil diambil kemudian disimpan dalam *spreadsheet* menggunakan aplikasi seperti *Microsoft Excel* atau *Google Spreadsheet* (Kesuma & Iskandar, 2022). Diagram alir yang menjelaskan tahapan *Scrapping* data dapat dilihat pada Gambar 1



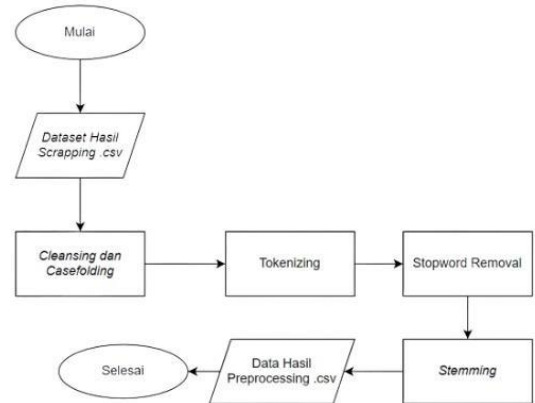
Gambar 1. Proses *Scrapping* data

2.2 Preprocessing Data

Pada tahap ini data yang telah didapatkan dari proses sebelumnya kemudian dibersihkan dan melalui beberapa tahapan diantaranya *cleansing*, *casefolding*, *tokenizing*, *stopword removal*, dan *stemming*. Pada tahap *cleansing* data dibersihkan dari unsur simbol, tautan, dan emoji, setelah itu dilakukan tahapan *casefolding* dimana setiap kata akan yang memiliki huruf besar akan ditransformasikan menjadi huruf kecil, lalu dilakukan *tokenizing* yang bertujuan untuk memecah setiap kata menjadi beberapa bagian serta spasi akan dihapus dan dipisahkan dengan tanda koma (Prihatini, 2016).

Tahapan berikutnya adalah *stopword removal* dimana proses ini bertujuan untuk menghilangkan kata yang memiliki relevansi rendah, biasanya kata yang ada dalam kamus *stopword* adalah kata yang

memiliki kemunculan paling tinggi dalam sebuah dokumen (Prihatini, 2016). Terakhir adalah proses *stemming* dimana semua data yang telah melalui proses *stopword removal* akan diubah menjadi kata dasar (Gusriani et al., n.d.). Pada tahap *preprocessing*, terdapat beberapa tahapan yang dapat dilihat pada Gambar 2.



Gambar 2. Proses *preprocessing* data

2.3 Feature extraction

Tahap ini melibatkan proses *labelling* dan memberikan bobot pada setiap kata yang telah melalui tahap *preprocessing* data. Metode yang digunakan adalah *Term Frequency - Inverse Document Frequency* (TF-IDF). Berikut merupakan tahapan dari *feature extraction* :

2.3.1 Labelling Data

Pada tahapan *labelling* bertujuan untuk menentukan polaritas sentimen dari data ulasan yang telah diperoleh dan diberikan label positif, negatif, atau netral. Proses *labelling* dilakukan secara otomatis dengan menggunakan bantuan *library TextBlob* yang telah disediakan oleh *Python*. Namun, penggunaan *library* tersebut belum tersedia untuk bahasa Indonesia, sehingga hal tersebut dapat mempengaruhi akurasi hasil yang akan diperoleh karena pengambilan data ulasan dilakukan dalam bahasa Indonesia.

2.3.2 Pembobotan Kata (TF-IDF)

Pada tahap berikutnya, dilakukan pembobotan kata sebelum masuk ke tahap klasifikasi. Pembobotan kata diimplementasikan dengan menggunakan metode TF-IDF (*Term Frequency - Inverse Document Frequency*) yang berfungsi untuk menghitung bobot setiap kata dan mencerminkan nilai frekuensi kata-kata yang penting atau sering muncul dalam dokumen. Bobot nilai ini akan berdampak pada hasil klasifikasi pada tahap selanjutnya. Berikut adalah tahapan dari proses pembobotan kata dengan TF-IDF :

2.3.2.1 Menghitung jumlah TF (Term Frequency)

yaitu kata tertentu yang sering muncul dalam sebuah dokumen, maka kata tersebut memiliki tingkat relevansi / kepentingan yang tinggi daripada kata yang lain yang ditunjukkan pada persamaan (1)

$$TF_{(t,d)} = \frac{f(td)}{\sum td} \quad (1)$$

Keterangan :

|D| = Jumlah dokumen / kalimat yang ada dalam corpus

df(t) = Jumlah dokumen / kalimat dimana muncul kata.

2.3.2.2 Menghitung total IDF (Inverse Document Frequency)

langkah selanjutnya merupakan menghitung IDF yang ditunjukkan pada persamaan (2)

$$IDF_{(t)} = \log \frac{|D|}{df(t)+1} \quad (2)$$

Keterangan :

TF = Hasil perhitungan dari Term Frequency

IDF = Hasil perhitungan dari Inverse Document Frequency

2.3.2.3 Menghitung Bobot Kata

Langkah terakhir adalah menghitung bobot dari setiap kata yang ada di dalam dokumen yaitu hasil TF akan dikali dengan hasil IDF, langkah ini dapat menggunakan persamaan (3)

$$W_{(t,d)} = TF_{(t,d)} \cdot IDF_{(t)} \quad (3)$$

Keterangan :

f(t,d) = Frekuensi setiap kata (t) yang muncul dalam suatu kalimat (d)

$\sum t,d$ = Total keseluruhan kata (t) yang ada di dalam suatu kalimat (d)

2.4 Naïve Bayes Classification

Pada tahap klasifikasi, dilakukan menggunakan algoritma *Naïve Bayes* untuk memproses data yang telah dilakukan *labelling* pada tahap sebelumnya. Model klasifikasi ini membaca polaritas dokumen yang telah ditransformasi menjadi nilai angka, dimana nilai polaritas yang mendekati atau sama dengan 1 akan diklasifikasikan sebagai sentimen positif, nilai polaritas yang sama dengan 0 akan diklasifikasikan sebagai sentimen netral, dan nilai

polaritas yang kurang dari 0 akan diklasifikasikan sebagai sentimen negatif. Selanjutnya, data akan dibagi menjadi dua bagian, yaitu data latih sebanyak

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

70% dan data uji sebanyak 30%. Adapun persamaan *Naïve Bayes* dapat dilihat pada persamaan (2).

Keterangan :

X = Data dengan kelas yang belum diketahui

H = Hipotesis data X merupakan suatu kelas spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori probabilitas)

P(H) = Probabilitas hipotesis H (prior probabilitas)

P(X|H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

2.5 Evaluation

Untuk menguji performa model algoritma *Naïve Bayes*, digunakan metode Confusion Matrix. Confusion matrix adalah tabel yang merepresentasikan jumlah data uji yang diklasifikasikan dengan benar dan salah. Tabel ini digunakan untuk mengevaluasi performa model machine learning. Pada Pengujian ini dipilih sebagai metode evaluasi karena data yang digunakan telah memiliki label. Metode evaluasi ini dipilih karena data yang digunakan telah diberi label sebelumnya. Pengujian ini bertujuan untuk menilai tingkat akurasi model pada data latih dengan menghitung nilai accuracy, precision, recall, dan f1-score.

3. HASIL DAN PEMBAHASAN

Berdasarkan penelitian yang telah dilakukan data yang dikumpulkan sebanyak 6000 data untuk dilakukan proses sentimen analisis, pada data tersebut nantinya akan dilakukan pembagian dimana 80% digunakan untuk data latih dan 20% digunakan untuk data uji, adapun data hasil *Scrapping* dapat dilihat pada Tabel 1.

Tabel 1. Hasil Scrapping Data

<u>UserName</u>	<u>Ulasan</u>
Nurchiah Herliana putri	Makin nyaman pakai BSI sekarang, aplikasi ringan, fitur lengkap dan menarik pendaftaran mudah. Bertransaksi jadi lebih gampang ga ribet. Simpan uang lebih aman di BSI. BSI Keren
Syahrul Ramadani	Aplikasi ini sangat membantu sekali dalam hal keuangan saya mudah dan cepat serta tidak ribet.....

Vira Agresia	Aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi,,,fitur nya juga lengkap l,,,semoga sukses terus kedepan nya😊 📱
Dian Magfiroh	Good job BSI E seneng banget apk nya mudah dipahami dan gak beribet Semoga sllu lebih di tingkatkan LG ya. Biar kita nyaman,semua seneng dgn aplikasinya.
SRI HERLIANAWATI	

keuangan saya mudah dan cepat serta tidak ribet....	saya mudah dan cepat serta tidak ribet
Aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi,,,fitur nya juga lengkap l,,,semoga sukses terus kedepan nya😊 📱	aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi fitur nya juga lengkap lsemoga sukses terus kedepan nya
Good job BSI E seneng banget apk nya mudah dipahami dan gak beribet Semoga sllu lebih di tingkatkan LG ya. Biar kita nyaman,semua seneng dgn aplikasinya.	good job bsi seneng banget apk nya mudah dipahami dan gak beribet semoga sllu lebih di tingkatkan lg ya biar kita nyaman semua seneng dgn aplikasinya

Pada tahap selanjutnya data ulasan tersebut dibersihkan melalui beberapa tahapan diantaranya :

3.1 Cleansing

Pada tahapan ini seluruh data akan dilakukan *cleansing* , tujuan dalam proses pembersihan ini adalah untuk menghapus angka, simbol, tanda baca, hastag, mention. Hasil *cleansing* tersaji pada Tabel 2.

Tabel 2. Hasil Cleansing

Ulasan	Cleansing
Makin nyaman pakai BSI sekarang, aplikasi ringan, fitur lengkap dan menarik pendaftaran mudah. Bertransaksi jadi lebih gampang ga ribet. Simpan uang lebih aman di BSI BSI Keren	Makin nyaman pakai BSI sekarang aplikasi ringan fitur lengkap dan menarik pendaftaran mudah. Bertransaksi jadi lebih gampang ga ribet Simpan uang lebih aman di BSI BSI Keren
Aplikasi ini sangat membantu sekali dalam hal keuangan saya mudah dan cepat serta tidak ribet....	Aplikasi ini sangat membantu sekali dalam hal keuangan saya mudah dan cepat serta tidak ribet
Aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi,,,fitur nya juga lengkap l,,,semoga sukses terus kedepan nya😊 📱	Aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi fitur nya juga lengkap lsemoga sukses terus kedepan nya
Good job B S I E seneng banget apk nya mudah dipahami dan gak beribet Semoga sllu lebih di tingkatkan LG ya. Biar kita nyaman,semua seneng dgn aplikasinya.	Good job BSI seneng banget apk nya mudah dipahami dan gak beribet Semoga sllu lebih di tingkatkan LG ya Biar kita nyaman semua seneng dgn aplikasinya

3.2 Casefolding

Pada tahapan *casefolding* data yang sudah dibersihkan dari proses sebelumnya kemudian semua kata yang memiliki huruf kapital akan ditransformasi menjadi huruf kecil. Hasil yang diperoleh pada proses ini dapat dilihat pada Tabel 3.

Tabel 3. Hasil Casefolding

Cleansing	Casefolding
Makin nyaman pakai BSI sekarang, aplikasi ringan, fitur lengkap dan menarik pendaftaran mudah. Bertransaksi jadi lebih gampang ga ribet. Simpan uang lebih aman di BSI BSI Keren	makin nyaman pakai bsi sekarang aplikasi ringan fitur lengkap dan menarik pendaftaran mudah. bertransaksi jadi lebih gampang ga ribet simpan uang lebih aman di bsi bsi keren
Aplikasi ini sangat membantu sekali dalam hal	aplikasi ini sangat membantu sekali dalam hal keuangan

3.3 Tokenizing

Tahap selanjutnya adalah *tokenizing*, bertujuan untuk memecahkan kalimat menjadi kata-kata yang dipisahkan oleh spasi dan tanda petik. Berikut merupakan hasil dari proses *tokenizing* yang dapat dilihat pada Tabel 4.

Tabel 4. Hasil Tokenizing

Casefolding	Tokenizing
makin nyaman pakai bsi sekarang aplikasi ringan fitur lengkap dan menarik pendaftaran mudah. bertransaksi jadi lebih gampang ga ribet simpan uang lebih aman di bsi bsi keren	[makin,nyaman,pakai,bsi,sekarang,aplikasi,ringan,fitur,lengkap,dan,menarik,pendaftaran,mudah,bertransaksi,jadi,lebih,gampang,ga,ribet,simpan,uang,lebih,aman,di,bsi,bsi,keren]
aplikasi ini sangat membantu sekali dalam hal keuangan saya mudah dan cepat serta tidak ribet	[aplikasi,ini,sangat,membantu,sekali,dalam,hall,keuangan,saya,mudah,dan,cepat,serta,tidak,ribet]
aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi fitur nya juga lengkap lsemoga sukses terus kedepan nya	[aplikasi,nya,sangat,membantu,dan,mempermudah,kita,dalam,bertransaksifitur,nya,juga,lengkap,lsemoga,sukses,terus,kedepan,nya]
good job bsi seneng banget apk nya mudah dipahami dan gak beribet	[good,job,bsi,seneng,banget,apk,nya,mudah,dipahami,dan,gak,beribet]
semoga sllu lebih di tingkatkan lg ya biar kita nyaman semua seneng dgn aplikasinya	[semoga,slu,lebih,di,tingkatkan,lg,ya,biar,kita,nyaman,semua,seneng,dgn,aplikasinya]

3.4 Stopword removal

Proses pembersihan *stopword* pada dataset yang sudah dilakukan tokenize dilakukan untuk membuang kata-kata yang tidak memiliki pengaruh terhadap proses klasifikasi. Berikut merupakan hasil dari proses *stopword* removal yang dapat dilihat pada tabel 5.

Tabel 5. Stopword Removal

Tokenizing	Stopword Removal
[makin,nyaman,pakai,bsi,sekarang,aplikasi,ringan,fitur,lengkap,dan,menarik,pendaftaran,mudah,bertransaksi,jadi,lebih,gampang,ga,ribet,simpan,uang,lebih,aman,di,bsi,bsi,keren]	nyaman,pakai,ringan,fitur,lengkap,menarik,pendaftaran,mudah,bertransaksi,gampang,ribet,simpan,uang,aman,keren

[aplikasi, ini, sangat, membantu, sekali, dalam, hal, keuangan, saya, mudah, dan, cepat, serta, tidak, ribet]	membantu, keuangan, mudah, cepat, ribet	Aplikasi ini sangat membantu sekali dalam hal keuangan saya mudah dan cepat serta tidak ribet....	bantu, uang, mudah, cepat, ribet	Positif
[aplikasi, nya, sangat, membantu, dan, mempermudah, kita, dalam, bertransaksi, fitur, nya, juga, lengkap, semoga, sukses, terus, kedepan, nya]	membantu, mempermudah, bertransaksi, fitur, lengkap, semoga, sukses, kedepan	Aplikasi nya sangat membantu dan mempermudah kita dalam bertransaksi, ..., fitur nya juga lengkap l, ..., semoga sukses terus kedepan nya 🙏	bantu, mudah, transaksi, lengkap, semoga, sukses, depan	Positif
[good, job, bsi, seneng, banget, apk, nya, mudah, dipahami, dan, gak, beribet]	good, job, seneng, banget, apk, mudah, dipahami, gak, beribet	Good job B S I 🙏 seneng banget apk nya mudah dipahami dan gak beribet	good, job, seneng, banget, apk, mudah, paham, gak, ribet	Positif
[semoga, sllu, lebih, di, tingkatkan, lg, ya, biar, kita, nyaman, semua, seneng, dgn, aplikasi, nya]	semoga, tingkatkan, biar, nyaman, semua, seneng, aplikasinya	Semoga sllu lebih di tingkatkan LG ya. Biar kita nyaman, semua seneng dgn aplikasinya.	semoga, tingkat, biar, nyaman, semua, seneng, aplikasi	Positif

3.5 Stemming

Tahapan terakhir dari *preprocessing* adalah *stemming* dimana setiap kata akan dibersihkan imbuhan dan awalnya sebagai contoh kata “bertransaksi” diubah menjadi “transaksi”, hal ini bertujuan untuk mempermudah dalam proses *labelling* dan pembobotan kata. Hasil dari proses *stemming* dapat dilihat pada tabel 6.

Tabel 6. Hasil Stemming

Stopword Removal	Stemming
nyaman, pakai, ringan, fitur, lengkap, menarik, pendaftaran, mudah, bertransaksi, gampang, ribet, simpan, uang, aman, keren	nyaman, pakai, ringan, fitur, lengkap, menarik, daftar, mudah, bertransaksi, gampang, ribet, simpan, uang, aman, keren
membantu, keuangan, mudah, cepat, ribet	bantu, uang, mudah, cepat, ribet
membantu, mempermudah, bertransaksi, fitur, lengkap, semoga, sukses, kedepan	bantu, mudah, transaksi, lengkap, semoga, sukses, depan
good, job, seneng, banget, apk, mudah, paham, gak, beribet	good, job, seneng, banget, apk, mudah, paham, gak, ribet
semoga, tingkatkan, biar, nyaman, semua, seneng, aplikasinya	semoga, tingkat, biar, nyaman, semua, seneng, aplikasi

3.6 Feature extraction

Setelah *preprocessing* data selesai langkah selanjutnya adalah proses *labelling* data dimana proses ini bertujuan untuk memberikan label pada data, jika polaritas data bernilai 0 maka akan memiliki label netral, jika data bernilai lebih dari 0 maka memiliki label positif dan jika nilai polaritas kurang dari 0 maka memiliki label negatif. Berikut merupakan hasil *labelling* yang ditunjukkan pada tabel 7.

Tabel 7. Hasil Labelling

Ulasan	Data Bersih	Label
Makin nyaman pakai BSI sekarang, aplikasi ringan, fitur lengkap dan menarik pendaftaran mudah. Bertransaksi jadi lebih gampang ga ribet. Simpan uang lebih aman di BSI. BSI Keren	nyaman, pakai, ringan, fitur, lengkap, menarik, daftar, mudah, bertransaksi, gampang, ribet, simpan, uang, aman, keren	Positif

Setelah dilakukan *labelling* data guna menentukan polaritas langkah selanjutnya adalah pembobotan kata dimana metode yang digunakan dalam proses ini adalah TF-IDF, proses ini bertujuan untuk memerikan bobot dari setiap kata yang nantinya akan dijadikan input untuk proses klasifikasi sentimen dimana dapat dihitung menggunakan persamaan (1), (2), (3). Contoh perhitungan TF-IDF dapat dilihat pada tabel 8.

Tabel 8. Hasil Perhitungan TF-IDF

Term	TF		TF	DF	IDF	TF-IDF
	D1	D2				
data	1	1	2	2	0.176	0.352
udah	1	0	1	1	0.477	0.477
foto	1	0	1	1	0.477	0.477
teknologi	1	1	2	2	0.176	0.352
scan	1	1	2	2	0.176	0.352
wajah	1	1	2	2	0.176	0.352
mumpuni	1	1	2	2	0.176	0.352
apps	1	0	1	2	0.176	0.176
belah	1	0	1	1	0.477	0.477
pasti	1	0	1	1	0.477	0.477
mampu	1	0	1	1	0.477	0.477
mohon	1	0	1	1	0.477	0.477
baik	1	0	1	1	0.477	0.477

3.7 Naive Bayes Classification dan Evaluation

Pada tahapan tahapan ini *dataset* yang telah diberikan label dan bobot kemudian dilakukan klasifikasi, hal ini bertujuan untuk memisahkan data ulasan positif dan data ulasan negatif. Sebelumnya akan dilakukan proses pemisahan data yaitu data latih dan data uji yang dimana perbandingannya adalah 80:20. Berdasarkan hasil pengujian klasifikasi yang telah dilakukan, maka akurasi yang diperoleh sebesar 89,67%, hasil akurasi tersebut diperoleh berdasarkan tabel Confusion matrix yang ada pada tabel 9.

Tabel 9. Hasil Pengujian Confusion Matrix

		Hasil Prediksi		
		Positif	Netral	Negatif
Kelas Aktual	Positif	797	0	1
	Netral	33	0	1
	Negatif	58	0	10

Berdasarkan tabel 9. Akurasi model dapat dihitung menggunakan persamaan (5)

4. KESIMPULAN

1. Dalam penelitian ini, algoritma *Naïve Bayes Classifier* mampu melakukan analisis sentimen secara otomatis pada ulasan pengguna BSI *Mobile* di *Google Play Store* dengan akurasi sebesar 89,67%. Hasil klasifikasi sentimen positif menunjukkan bahwa aplikasi BSI *Mobile* sangat membantu nasabah dalam melakukan transaksi keuangan dan mudah digunakan. Namun, hasil klasifikasi sentimen negatif menunjukkan bahwa pengguna kesulitan dalam proses aktivasi BSI *Mobile* menggunakan verifikasi foto wajah dan prosesnya memakan waktu yang lama. Penulis merekomendasikan untuk memperbaiki kinerja fitur tersebut lebih efektif dan efisien. Sistem yang telah dibuat berhasil menampilkan hasil klasifikasi dengan berbagai bentuk diagram yang mempermudah interpretasi dan pemahaman data oleh pengguna.
2. Akurasi dapat berubah seiring dengan pertambahan jumlah ulasan dan trend.

DAFTAR PUSTAKA

- Dwiki, A., Putra, A., & Juanita, S. (2021). *Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN*. 8(2). <http://jurnal.mdp.ac.id>
- Fani Al-shufi, M., & Erfina, A. (2021). *SENTIMEN ANALISIS MENGENAI APLIKASI STREAMING FILM MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DI PLAY STORE*.
- Garbian Nugroho, D., Herry Chrisnanto, Y., Wahana Jurusan Informatika, A., & Matematika dan Ilmu Pengetahuan Alam Universitas Jenderal Achmad Yani Jalan Terusan Jenderal Sudirman, F. (n.d.). *ANALISIS SENTIMEN PADA JASA OJEK ONLINE*
- MENGGUNAKAN METODE NAÏVE BAYES*.
- Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2017). *Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile)* (Vol. 1, Issue 10). <http://j-ptiik.ub.ac.id>
- Gusriani, S., Diah Kusuma Wardhani, K., & Ihsan Zul, M. (n.d.). *Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka)*.
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., Komputer STMIK Nusa Mandiri Jl Damai No, I., Jati Barat, W., & Selatan, J. (2020). *ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE* (Vol. 5, Issue 2).
- Kesuma, M. E.-K., & Iskandar, R. (2022). *Analisis Toko dan Asal Toko Fashion Pria di Shopee Menggunakan Data Scrapping dan Exploratory Data Analysis*. *Majalah Ilmiah Teknologi Elektro*, 21(1), 127. <https://doi.org/10.24843/mite.2022.v21i01.p17>
- Marimin, A., Haris Romdhoni, A., Tira, D., Fitria, N., & Surakarta, S.-A. (2015). *PERKEMBANGAN BANK SYARIAH DI INDONESIA*. *Jurnal Ilmiah Ekonomi Islam*, 01(02).
- Ogi, E., Pratiwi1, I., & Yustanti2, W. (n.d.). *Analisis Sentimen Kualitas Layanan Teknologi Pembayaran Elektronik pada Twitter (Studi Kasus Ovo dan Dana)*. *JEISBI*, 02, 2021.
- Prihatini, P. M. (2016). *IMPLEMENTASI EKSTRAKSI FITUR PADA PENGOLAHAN DOKUMEN BERBAHASA INDONESIA* The Implementation of Extraction Feature on Indonesian Documents' Processing. In *JURNAL MATRIX* (Vol. 6, Issue 3).

Yusmad, M. A. (2018). *Aspek Hukum Perbankan Syariah dari Teori ke Praktik*. Deepublish.