

Analisis Sentimen Ulasan Pengguna Aplikasi BCA Mobile dan *myBCA* Menggunakan Metode *BiLSTM* dan SVM

Wina Elsa Wardana ¹, Anis Zubair ^{2*}

¹ Sistem Informasi, Universitas Merdeka Malang

² Sistem Informasi, Universitas Merdeka Malang

Alamat: Jalan Terusan Raya Dieng No. 62-64 Malang, Jawa Timur, Indonesia

Email : 22083000185@student.unmer.ac.id ¹, anis.zubair@unmer.ac.id ³

* Penulis Korespondensi : Anis Zubair

Abstract: The evolution from conventional services to mobile banking has led to increased usage of digital banking applications, including BCA Mobile and myBCA. The large volume of user reviews on the Google Play Store serves as a valuable source of information regarding user experiences and perceptions of application service quality. This study aims to analyze the sentiment distribution of BCA Mobile and myBCA user reviews, compare the performance of the Bidirectional Long Short-Term Memory (BiLSTM) and Support Vector Machine (SVM) methods, and determine the most optimal method for sentiment analysis. The dataset consists of 5,000 user reviews for BCA Mobile and 5,000 for myBCA, collected from the Google Play Store. The research process involved data collection, sentiment labeling, and data preprocessing comprising cleaning, case folding, normalization, tokenization, stopword removal, and stemming followed by modeling using BiLSTM and SVM with a linear kernel. Model evaluation was conducted using a confusion matrix based on four metrics accuracy, precision, recall, and F1-score along with McNemar's test to determine the significance of performance differences between the two methods. The results indicate that negative sentiment dominates both applications: 56.5% (2,741 reviews) for BCA Mobile and 57.6% (2,825 reviews) for myBCA. Model comparisons show that SVM achieved an accuracy of 86% for BCA Mobile and 87% for myBCA, whereas BiLSTM achieved 84.5% and 85.6%, respectively. McNemar's test yielded p-values of 2.60×10^{-54} for BCA Mobile and 5.84×10^{-63} for myBCA; since both p-values were below 0.05, it is concluded that there is a significant difference in performance between the two methods. Based on these findings, SVM is the most optimal method for analyzing user reviews of BCA Mobile and myBCA in this study, and the results can serve as a basis for evaluating and improving application service quality and features.

Keywords: Sentiment Analysis; Mobile Banking Application; BCA Mobile; myBCA; BiLSTM; SVM

Abstrak: Perkembangan layanan konvensional menjadi layanan berbasis mobile banking mengakibatkan meningkatnya penggunaan aplikasi perbankan digital, termasuk BCA Mobile dan myBCA. Banyaknya ulasan pengguna pada Google Play Store dapat menjadi sumber informasi untuk mengetahui pengalaman serta persepsi pengguna terhadap kualitas layanan aplikasi. Penelitian ini bertujuan untuk menganalisis distribusi sentimen ulasan pengguna BCA Mobile dan myBCA, membandingkan kinerja metode Bidirectional Long Short-Term Memory (BiLSTM) dan Support Vector Machine (SVM), serta menentukan metode yang paling optimal dalam melakukan analisis sentimen. Data penelitian berupa 5.000 ulasan pengguna BCA Mobile dan 5.000 ulasan pengguna myBCA yang diperoleh melalui platform Google Play Store. Tahapan pada penelitian ini meliputi pengumpulan data, pelabelan sentimen, preprocessing data yang terdiri dari cleaning, case folding, normalization, tokenizing, stopword removal, dan stemming, kemudian dilakukan pemodelan menggunakan BiLSTM dan SVM dengan kernel linear. Evaluasi model dilakukan menggunakan confusion matrix dengan menggunakan empat metrik evaluasi yaitu accuracy, precision, recall, dan F1-Score, serta uji McNemar untuk mengetahui signifikansi perbedaan kinerja kedua metode. Hasil penelitian menunjukkan bahwa distribusi sentimen pada kedua aplikasi didominasi oleh sentimen negatif, yaitu BCA Mobile sebesar 56,5% atau 2.741 ulasan dan myBCA sebesar 57,6% atau 2.825 ulasan. Hasil perbandingan model menunjukkan bahwa SVM memperoleh accuracy sebesar 86% pada BCA Mobile dan 87% pada myBCA, sedangkan BiLSTM memperoleh accuracy sebesar 84,5% dan 85,6%. Hasil uji McNemar menunjukkan p-value sebesar $2,60 \times 10^{-54}$ pada BCA Mobile dan $5,84 \times 10^{-63}$ pada myBCA, dimana pada kedua aplikasi menghasilkan p-value $<0,05$ sehingga dapat disimpulkan bahwa terdapat perbedaan kinerja yang signifikan antara kedua metode. Berdasarkan hasil

Masuk: 06 Juli 2026

Revisi: 12 Agustus 2026

Diterima: 30 September 2026

Tersedia: 30 September 2026

Terbit: 30 September 2026



Hak cipta: © 2026 oleh penulis. Diserahkan untuk kemungkinan publikasi akses terbuka berdasarkan syarat dan ketentuan lisensi Creative Commons Attribution (CC BY SA) (<https://creativecommons.org/licenses/by-sa/4.0/>)

tersebut, SVM merupakan metode yang paling optimal dalam analisis sentimen ulasan pengguna BCA Mobile dan myBCA pada penelitian ini serta dapat digunakan sebagai bahan evaluasi untuk meningkatkan kualitas layanan dan fitur aplikasi.

Kata kunci: Analisis Sentimen; Aplikasi Mobile Banking; BCA Mobile; myBCA; BiLSTM; SVM

1. Pendahuluan

Era digitalisasi yang semakin berkembang secara signifikan telah mendorong transformasi digital dalam berbagai sektor, salah satunya yaitu sektor perbankan. Transformasi ini ditandai dengan adanya pergeseran besar dari layanan konvensional menjadi platform berbasis mobile banking. Mobile banking didefinisikan sebagai platform yang digunakan untuk melakukan suatu transaksi finansial maupun non finansial secara praktis melalui perangkat seluler (smartphone) yang menawarkan pengalaman personal perbankan [1]. Kemajuan financial technology (fintech) pada sektor perbankan dimanfaatkan oleh lembaga keuangan untuk meningkatkan kualitas layanan dan efisiensi operasional dengan cepat, aman dan responsif, seiring dengan meningkatnya preferensi masyarakat terhadap transaksi berbasis aplikasi [2].

Salah satu lembaga keuangan yang memanfaatkan hal ini yaitu PT Bank Central Asia Tbk (BCA). PT Bank Central Asia Tbk (BCA) merupakan salah satu bank swasta terbesar di Indonesia yang ada sejak tahun 1957. BCA sendiri juga memiliki dua aplikasi mobile banking dengan fitur berbeda yaitu BCA Mobile dan myBCA [3]. Kedua aplikasi ini secara signifikan dirancang untuk memberikan pengalaman dan pendekatan yang berbeda dalam memenuhi berbagai kebutuhan nasabah [4]. Bersumber dari platform google play store, BCA Mobile dikenal kalangan masyarakat sebagai aplikasi yang telah lama beroperasi dan berfokus pada layanan perbankan dasar, dengan sistem autentikasi berbasis mPIN dengan jumlah pengguna sebanyak 50 juta. Sedangkan myBCA hadir sebagai platform terbaru dan lebih modern dengan sistem autentikasi biometrik dengan jumlah pengguna sekitar 10 juta. Adanya persaingan dari kedua aplikasi ini menimbulkan tumpang tindih fungsi dan fitur yang membuat pengguna merasa kebingungan. Sehingga di tahun 2026 telah mencapai lebih dari 1,38 juta dan 98,8 ribu ulasan pengguna aplikasi BCA Mobile dan myBCA menyampaikan keluhan kesahnya melalui platform google play store maupun app store berdasarkan pengalaman pengguna. Permasalahan ini semakin krisis karena jumlah ulasan yang bertambah setiap harinya menjadikan proses evaluasi untuk meningkatkan loyalitas dan pengalaman pengguna secara manual dinilai kurang efisien. Oleh karena itu kebutuhan otomatisasi diperlukan untuk dapat mengklasifikasi ulasan pengguna secara otomatis dan lebih efisien.

Ulasan pengguna pada platform tersebut menjadi suatu informasi dan evaluasi penting bagi para developer aplikasi maupun pihak manajemen. Hal ini dikarenakan ulasan tersebut mengandung berbagai macam sentimen baik positif, negatif, maupun netral. Sentimen tersebut menggambarkan tingkat kepuasan pelanggan, kualitas layanan, keluhan serta saran terhadap aplikasi tersebut [5]. Oleh karena itu, pendekatan analisis sentimen menjadi salah satu metode yang relevan dalam mengevaluasi persepsi secara sistematis.

Analisis sentimen adalah teknik pemrosesan bahasa alami atau Natural Language Processing (NLP) yang digunakan untuk mengklasifikasikan teks berdasarkan pola suatu sentimen [6]. Seiring berjalannya waktu metode machine learning dan deep learning dikembangkan untuk memaksimalkan akurasi analisis sentimen, diantaranya metode Bidirectional Long Short-Term Memory (BiLSTM) dan Support Vector Machine (SVM). Kedua metode ini sering digunakan dalam penelitian analisis sentimen karena menghasilkan performa cukup baik, namun perbandingan akurasi kedua metode tersebut pada ulasan pengguna aplikasi mobile banking perlu dikaji lebih mendalam.

Deep learning merupakan bagian dari machine learning yang menggunakan jaringan saraf tiruan yang dapat mempelajari data besar secara efisien melalui lapisan tersembunyi. RNN (Recurrent Neural Network) adalah jenis jaringan saraf yang umum digunakan untuk memproses data deret waktu. Namun, RNN menghadapi masalah seperti ketergantungan memori jangka panjang dan vanishing gradient. Untuk mengatasi ini, Long Short-Term Memory (LSTM) dikembangkan untuk mengelola informasi berurutan dalam jangka panjang. Namun, LSTM memiliki keterbatasan karena hanya memproses data dalam satu arah. Maka

dari itu agar meningkatnya akurasi, Bidirectional Long Short-Term Memory (BiLSTM) dikembangkan, yang memungkinkan pemrosesan data dalam dua arah secara simultan, memperdalam pembelajaran model [7]. Kelebihan ini sangat sesuai dengan proses bahasa alami seperti analisis sentimen. Sementara itu, Support Vector Machine (SVM) adalah suatu algoritma supervised learning, dimana algoritma ini hanya menangani klasifikasi biner (dua kelas) seperti ya atau tidak, 1 dan 0, True dan False dan sebagainya [8]. SVM diketahui mempunyai performa yang baik dalam mengklasifikasi suatu teks serta tidak membutuhkan data latih dalam jumlah yang besar.

Beberapa penelitian sebelumnya telah menganalisis sentimen ulasan aplikasi BCA Mobile. Bhatarra dan Suryono membandingkan Naïve Bayes dan SVM pada ulasan BCA Mobile dan memperoleh akurasi SVM sebesar 85% [8]. Gamma dkk. menggunakan Naïve Bayes, sedangkan Yoman dkk. menerapkan pendekatan lexicon untuk membandingkan BCA Mobile dan myBCA [4][2]. Namun, penelitian-penelitian tersebut masih terbatas pada satu algoritma machine learning klasik atau pendekatan lexicon saja. Penelitian-penelitian tersebut belum membandingkan secara langsung kinerja deep learning (BiLSTM) dengan machine learning (SVM) pada dua aplikasi BCA sekaligus, serta belum dilengkapi uji signifikansi statistik (McNemar) untuk memastikan perbedaan kinerja. Selain itu, karakteristik dataset ulasan berbahasa Indonesia yang relatif terbatas dan mengandung banyak slang, negasi, serta istilah teknis perbankan belum dievaluasi secara mendalam terhadap kedua pendekatan tersebut.

Oleh karena itu, penelitian ini mengisi gap tersebut dengan: (1) membandingkan kinerja BiLSTM dan SVM (kernel linear) pada dataset ulasan BCA Mobile dan myBCA secara terpisah; (2) menggunakan uji McNemar untuk menguji signifikansi perbedaan prediksi kedua model; dan (3) menganalisis distribusi sentimen serta fitur penting (TF-IDF) sebagai dasar evaluasi kualitas layanan aplikasi. Kontribusi ilmiah penelitian ini adalah memberikan bukti empiris mengenai metode yang lebih optimal untuk analisis sentimen ulasan mobile banking berbahasa Indonesia pada skala data menengah, sekaligus memberikan rekomendasi praktis bagi pengembang aplikasi.

2. Tinjauan Literatur

2.1. Analisis Sentimen

Analisis sentimen adalah pendekatan *Natural Language Processing* (NLP) yang bertujuan untuk mengekstraksi, mengklasifikasi ulasan secara mendalam dan menganalisis pendapat atau opini seseorang terhadap suatu topik maupun isu tertentu yang terdapat dalam teks maupun kalimat tidak terstruktur [9][10].

2.2 Web Scraping

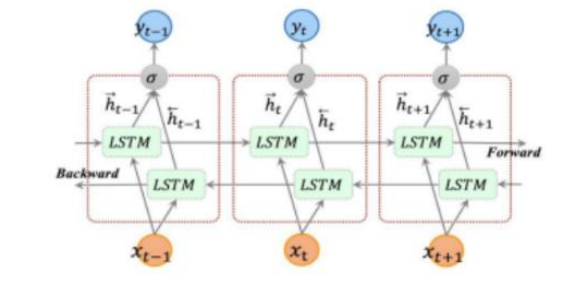
Web Scraping adalah prosedur mengekstraksi data otomatis dari situs web menggunakan perangkat lunak yang bertujuan untuk mengumpulkan data secara *real time* [11]. Teknik ini jauh lebih cepat dan efektif karena dapat digunakan untuk mengumpulkan dan menyusun data dari ribuan, atau lebih dari jutaan, halaman untuk diproses dan diambil informasinya.

2.3 Preprocessing Data

Preprocessing data merupakan pemrosesan bahasa alami untuk mengklasifikasikan data khususnya pada ulasan sistem *e-commerce* atau media sosial [12]. Tahapan ini bertujuan untuk membersihkan data agar terlihat lebih konsisten pada analisis lanjutan dengan melalui beberapa tahapan, diantaranya *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming* [8].

2.4 Bidirectional Long Short-Term Memory (BiLSTM)

Studi [7] menjelaskan bahwa Bidirectional Long Short-Term Memory merupakan perkembangan dari model LSTM dimana terdapat 2 lapisan yang saling berkebalikan arah. Pada lapisan bagian bawah bergerak maju (*forward layer*) yaitu memproses kata pertama menuju ke kata terakhir, sedangkan pada lapisan atas bergerak mundur (*backward layer*) yaitu memproses kata terakhir menuju ke kata pertama.



Gambar 1. Arsitektur BiLSTM

Secara umum, rumus dari BiLSTM sama dengan rumus LSTM, yang membedakan BiLSTM dengan LSTM yaitu dalam BiLSTM memiliki dua buah keluaran yaitu keluaran LSTM forward yang dapat dilihat pada persamaan 2.1 dan keluaran LSTM backward yang dapat dilihat pada persamaan 2.2. Keluaran dari dua buah LSTM tersebut kemudian digabungkan menjadi keluaran BiLSTM. Penggabungan biasanya dilakukan dengan mengkonkatenasi keluaran forward dan backward LSTM. Keluaran BiLSTM dapat dilihat pada persamaan dibawah ini.

Persamaan *Forward* :

$$\vec{h}_t = \mathcal{H}(W_{xh}x_t + W_{hh}\vec{h}_{t-1} + b_h) \quad (1)$$

Keterangan :

x_t : Input pada waktu/urutan ke- t

$\vec{h}_t$: *Hidden state* pada waktu ke- t dari arah *forward*

$\vec{h}_{t-1}$: *Hidden state* sebelumnya dari arah forward

W_{xh} : Matriks yang menghubungkan input x_t dengan *hidden state*

W_{hh} : Matriks yang menghubungkan *hidden state* dengan *hidden state* saat ini

b_h : Bias pada *hidden layer*

$\mathcal{H}$: Fungsi aktivasi yang digunakan untuk menghasilkan *hidden state*

Persamaan *Backward* :

$$\overleftarrow{h}_t = \mathcal{H}(W_{xh}x_t + W_{hh}\overleftarrow{h}_{t+1} + b_h) \quad (2)$$

Keterangan :

x_t : Input pada waktu/urutan ke- t

$\overleftarrow{h}_t$: *Hidden state* pada waktu ke- t dari arah *backward*

$\overleftarrow{h}_{t+1}$: *Hidden state* sebelumnya dari arah *backward*

W_{xh} : Matriks yang menghubungkan input x_t dengan *hidden state*

W_{hh} : Matriks yang menghubungkan *hidden state* dengan *hidden state* saat ini

b_h : Bias pada *hidden layer*

$\mathcal{H}$: Fungsi aktivasi yang digunakan untuk menghasilkan *hidden state*

Persamaan *Output* :

$$y_t = W_{hy} h_t + W_{hy} h_t + b_y \quad (3)$$

Keterangan :

y_t : Output yang dihasilkan pada waktu/urutan ke- t

$\vec{h}_t$: *Hidden state* dari arah *forward*

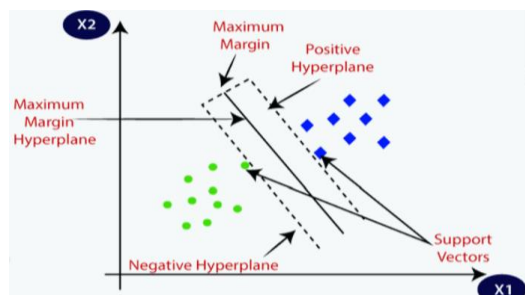
$\overleftarrow{h}_t$: *Hidden state* dari arah *backward*

W_{hy} : Matriks bobot yang menghubungkan *hidden state* dengan output

b_y : Bias pada *layer output*

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu metode *machine learning* dengan prinsip *Structural Risk Minimization* (SRM) yang dikenalkan pertama kali oleh Vapnik tahun 1992 dan bertujuan untuk menemukan hyperplane terbaik dengan memisahkan dua buah *class* pada *input space* [13]. SVM juga didefinisikan sebagai sistem pembelajaran yang menerapkan ruang hipotesis berupa fungsi-fungsi linear didalam sebuah fitur yang memiliki dimensi tinggi dan dilatih menggunakan algoritma pembelajaran yang berdasarkan teori optimasi [14][15].



Gambar 2. *Support Vector Machine*

Berdasarkan gambar diatas, SVM bertugas mencari *hyperplane* terbaik antara kedua kelas data yang memisahkan kedua data agar hyperplane dapat direpresentasikan pada ruang berdimensi tinggi dengan menggunakan persamaan :

$$f(x) = w \cdot x + b \quad (4)$$

Keterangan :

w : vektor bobot penentu arah

x : vektor fitur dari titik data
 b : bias, penentu posisi hyperplane

2.6 Confusion Matrix

Konsep *Confusion Matrix* pertama kali ditemukan oleh seorang ahli biologi dan statistik Inggris bernama Karl Pearson pada tahun 1904. *Confusion matrix* sendiri didefinisikan sebagai tabel evaluasi yang digunakan pada model *machine learning* dengan tujuan untuk pengukuran kinerja dalam mengklasifikasi setiap pengujian metode [16][17].

		gold standard labels		
		gold positive	gold negative	
system output labels	system positive	true positive	false positive	precision = $\frac{tp}{tp+fp}$
	system negative	false negative	true negative	
		recall = $\frac{tp}{tp+fn}$		accuracy = $\frac{tp+tn}{tp+fp+tn+fn}$

Gambar 3. *Confusion Matrix*

Berdasarkan gambar tabel *confusion matrix* diatas dijelaskan bahwa komponen utama pada matrix ini berupa *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), *True Negative* (TN). Dari keempat komponen tersebut dihasilkan empat matrix evaluasi yang digunakan sebagai pembandingan [18], dengan persamaan :

a. Accuracy

Mengukur proporsi berdasarkan jumlah seluruh prediksi benar dari total data uji dengan persamaan :

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

b. Precision

Mengukur seberapa tepat model dalam memprediksi kelas yang dapat diartikan dari semua data yang diprediksi berapa yang benar positif. Hasil *precision* didapatkan dari persamaan :

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

c. Recall

Mengukur seberapa lengkap model dalam mendeteksi suatu kelas yang dapat dihitung menggunakan persamaan :

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

d. F1-Score

Untuk *f1-score* sendiri digunakan sebagai penyeimbang antara nilai *precision* dan nilai *recall* melalui persamaan :

$$f1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

2.7 Uji McNemar Test

Uji *McNemar* adalah alat statistik yang dimanfaatkan untuk membandingkan frekuensi dari dua variable biner yang diukur pada subjek yang sama [19] [20]. Pada penelitian terkait uji *McNemar* umumnya menggunakan data yang dikumpulkan dari subjek yang sama sebelum dan sesudah perlakuan diberikan. Persamaan uji *McNemar* dijabarkan sebagai berikut :

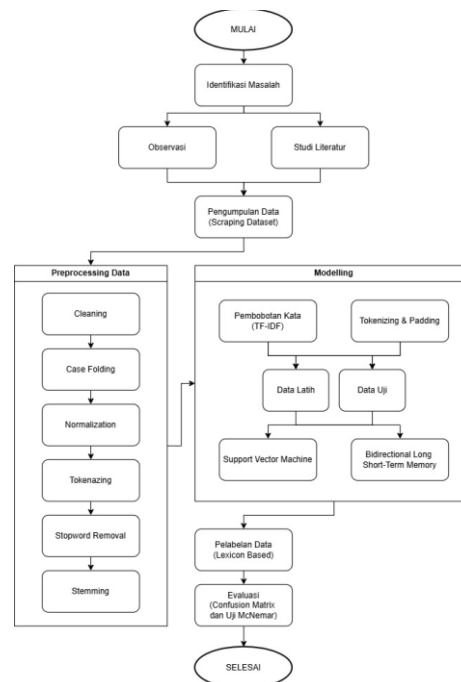
$$\chi^2 = \frac{(|b - c| - 1)^2}{(b + c)} \quad (9)$$

Keterangan :

χ^2 : Nilai Chi-Square

b dan c : Jumlah subjek yang mengalami perubahan

3. Metode



Gambar 4. Alur Penelitian

Penelitian ini menggunakan data ulasan aplikasi *BCA Mobile* dan *myBCA* sebagai objek penelitian. Tujuan dari penelitian ini mengetahui metode dengan kinerja terbaik dalam masalah ini, serta dapat memberikan kontribusi bagi pengembang aplikasi agar lebih baik dalam melakukan pelayanan di industri perbankan Indonesia berdasarkan distribusi ulasan.

Penelitian ini memiliki beberapa tahapan analisis mulai dari pengumpulan data, *preprocessing* data, *modelling*, dan evaluasi. Tahapan dilakukan secara runtut dan sistematis yang bertujuan untuk memastikan proses penelitian berjalan terarah sehingga mendapatkan hasil yang valid.

3.1. Pengumpulan Data

Data dikumpulkan melalui teknik web scraping dengan menggunakan *library google-play-scraper* pada bahasa pemrograman *Python* dan *Google Colaboratory* sebagai *tools* pendukung. Dataset ulasan pengguna aplikasi *BCA Mobile* dan *myBCA* yang ditarik memiliki batasan berbahasa Indonesia.

3.2. Eksplorasi Data

Eksplorasi dataset dilakukan dengan tujuan memaparkan dataset agar sesuai dengan kebutuhan analisis. Dataset yang dibutuhkan pada penelitian ini mencakup user, tanggal, ulasan pengguna, dan rating. Selain itu proses pembersihan data kosong dan duplikat juga dilakukan agar data siap diolah pada tahap selanjutnya.

3.3. Preprocessing Data

Preprocessing data merupakan tahap awal dalam bidang data dengan tujuan membersihkan, mentransformasi, dan mengubah data mentah menjadi lebih terstruktur dan siap diolah pada analisis lanjutan. Tahap ini memiliki beberapa proses, diantaranya :

a. *Cleaning*

Proses membersihkan data dengan menghilangkan suatu elemen yang tidak relevan agar nantinya tidak berdampak pada tahap analisis lanjutan.

b. *Case Folding*

Proses mengubah semua huruf menjadi huruf kecil agar terlihat konsisten pada saat proses analisis lanjutan.

c. *Normalization*

Proses mengubah kata yang tidak baku menjadi kata baku dengan menggunakan kamus lexicon.

d. *Tokenizing*

Tokenizing merupakan proses memecah teks menjadi unit-unit yang lebih kecil seperti kata atau frasa.

e. *Stopword Removal*

Proses menghilangkan atau menghapus kata-kata umum yang tidak relevan untuk analisis lanjutan. Proses ini menggunakan daftar stopwords Bahasa Indonesia yang dimodifikasi. Kata negasi penting seperti “tidak”, “bukan”, “jangan”, “belum” tidak dihapus agar informasi polaritas tetap terjaga. Hal ini dilakukan setelah observasi bahwa penghapusan kata “tidak” dapat membalikkan sentimen. Sebagai contoh “tidak lemot” menjadi “lemot”.

f. *Stemming*

Proses menghilangkan kata imbuhan berupa awalan, akhiran dan menjadikannya sebagai kata dasar dengan menggunakan *library* Sastrawi.

Dampak preprocessing terhadap distribusi polaritas diverifikasi secara manual pada 100 sampel acak. Dari hasil verifikasi ini tidak ditemukan perubahan label yang signifikan akibat penghapusan stopwords.

3.4. Pelabelan

Proses pelabelan bertujuan memberikan label sentimen pada setiap ulasan pengguna aplikasi BCA Mobile dan myBCA. Penelitian ini menggunakan pendekatan lexicon-based dengan kamus InSet (*Indonesian Sentiment Lexicon*). Setiap token hasil preprocessing dihitung polaritasnya: kata yang terdapat pada daftar InSet positif diberi skor +1, kata pada daftar InSet negatif diberi skor -1, dan kata yang tidak terdapat pada kamus diberi skor 0. Skor polaritas total suatu ulasan diperoleh dari penjumlahan skor seluruh token. Aturan klasifikasi yang digunakan adalah:

- Skor polaritas > 0 diklasifikasikan sebagai positif
- Skor polaritas ≤ 0 diklasifikasikan sebagai negatif

Ulasan dengan rating 3 atau netral digolongkan ke dalam kelas negatif. Keputusan ini didasarkan pada observasi bahwa sebagian besar ulasan rating 3 mengandung keluhan tersirat atau konteks yang lebih dekat dengan sentimen negatif. Untuk memastikan kualitas label, sebanyak 200 sampel diverifikasi secara manual. 200 sampel berasal dari 100 sampel tiap-tiap aplikasi. Tingkat kesepakatan antar-annotator diukur menggunakan Cohen's Kappa.

3.5. Modelling

Pada tahap modelling, data yang telah melalui proses preprocessing dan pelabelan digunakan untuk membangun dua model klasifikasi sentimen. Dua model itu Bidirectional Long Short-Term Memory (BiLSTM) dan Support Vector Machine (SVM). Dataset dibagi menjadi data latih dan data uji dengan rasio 80:20. Pembagian data dilakukan untuk memperoleh data yang digunakan dalam proses pembelajaran model serta data independen untuk mengevaluasi kemampuan generalisasi model.

Model BiLSTM digunakan untuk mempelajari pola dan hubungan kata berdasarkan urutan teks. Data teks terlebih dahulu direpresentasikan dalam bentuk sequence dan selanjutnya diproses menggunakan embedding layer sebelum diteruskan ke lapisan BiLSTM. Arsitektur model terdiri atas embedding layer, dua lapisan Bidirectional LSTM, dropout layer, dense layer, dan output layer dengan dua kelas klasifikasi. Masing-masing positif dan negatif.

4.2. Eksplorasi Data

Eksplorasi data penting dilakukan untuk melihat bahwa data yang berhasil diambil dan dikumpulkan pada tahap sebelumnya sebanyak 5.000 ulasan pengguna *BCA Mobile* dan 5.000 ulasan pengguna *myBCA*. Pada tahap ini ada beberapa hal yang dilakukan, diantaranya :

- Memastikan kolom ulasan pengguna sudah benar berbahasa Indonesia, yang menghasilkan data bersih sebanyak 4.997 ulasan *BCA Mobile* dan 4.997 ulasan *myBCA*.
- Menyeleksi kolom yang tidak dibutuhkan, seperti kolom “review Id”, “user image”, “thumbsup”, “review create version”, “reply content”, “replied at”, “app version”.
- Merubah nama kolom yang sudah sesuai dengan kebutuhan, seperti kolom “username” menjadi “user”, “content” menjadi “ulasan pengguna”, “score” menjadi “rating”, dan “at” menjadi “tanggal”. Dan menambahkan kolom “no”.
- Menghapus data duplikasi dan data dengan nilai kosong.

Dari tahap eksplorasi data yang telah dilakukan didapatkan sebanyak 4.997 dan 4.997 data bersih dan siap diolah untuk tahap selanjutnya, dimana data ulasan pertama pada aplikasi *BCA Mobile* tanggal 21 November 2025 hingga data terbaru pada tanggal 19 April 2026. Sedangkan untuk aplikasi *myBCA* data ulasan pertama pada tanggal 01 Desember 2025 hingga data terbaru pada tanggal 19 April 2026.

4.3. Preprocessing Data

Sebelum dilakukan analisis lebih lanjut data perlu melalui beberapa proses pada tahap preprocessing data, diantaranya seperti *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, *stemming*. Hal ini bertujuan untuk modifikasi data.

Tabel 1. Preprocessing Data *BCA Mobile*

Proses	Hasil
ulasan	parah transfer qris uangnya ke potong buktinya ga ada, ke orang nya ga ke kirim emang m-BCA ga jelas
<i>cleaning</i>	parah transfer qris uangnya ke potong buktinya ga ada, ke orang nya ga ke kirim emang m BCA ga jelas
<i>case folding</i>	parah transfer qris uangnya ke potong buktinya ga ada ke orang nya ga ke kirim emang m bca ga jelas
<i>normalization</i>	parah transfer qris uangnya ke potong buktinya tidak ada ke orang nya tidak ke kirim emang m bca tidak jelas
<i>tokenizing</i>	['parah', 'transfer', 'qris', 'uangnya', 'ke', 'potong', 'buktinya', 'tidak', 'ada', 'ke', 'orang', 'nya', 'tidak', 'ke', 'kirim', 'memang', 'sama', 'bca', 'tidak', 'jelas']
<i>stopword removal</i>	['parah', 'transfer', 'qris', 'uangnya', 'potong', 'buktinya', 'tidak', 'orang', 'tidak', 'kirim', 'bca', 'tidak']
<i>stemming</i>	['parah', 'transfer', 'qris', 'uang', 'potong', 'bukti', 'orang', 'kirim', 'bca']

Tabel 2. Preprocessing Data *myBCA*

Proses	Hasil
ulasan	Tidak scan galeri,Harus Poto pake hp lain .
<i>cleaning</i>	Tidak scan galeri Harus Poto pake hp lain
<i>case folding</i>	tidak scan galeri harus poto pake hp lain
<i>normalization</i>	tidak scan galeri harus foto pakai hp lain
<i>tokenizing</i>	['tidak', 'scan', 'galeri', 'harus', 'foto', 'pakai', 'hp', 'lain']

Proses	Hasil
<i>stopword removal</i>	['tidak', 'scan', 'galeri', 'foto', 'pakai', 'hp']
<i>stemming</i>	['tidak', 'scan', 'galeri', 'foto', 'pakai', 'hp']

4.4. Pelabelan Data

Tahap ini merupakan proses pelabelan terhadap data ulasan pengguna aplikasi dengan menentukan kelas dan label dari sentimen, baik positif, negatif, maupun netral. Namun, pada penelitian ini kelas akan dibagi menjadi dua yaitu kelas positif dan kelas negatif. Sehingga, ulasan dengan rating 1, 2, kelas netral atau ulasan dengan rating 3 akan dikelompokkan ke dalam label negatif karena ulasan tersebut memiliki konteks yang mengarah ke ulasan yang bersifat negatif atau keluhan pengguna terhadap aplikasi *BCA Mobile* dan *myBCA*. Sedangkan ulasan dengan rating 4 dan 5 akan dikelompokkan pada kelas positif yang berisikan tentang kepuasan pengguna terhadap aplikasi mobile banking BCA tersebut.

Pelabelan ini akan menggunakan pendekatan *lexicon-based*. Kamus ini yang akan digunakan untuk perbandingan data yang ada. Apabila kata termasuk InSet positif maka nilai polaritas akan bertambah +1. Namun sebaliknya apabila kata termasuk InSet negatif maka nilai polaritas akan berkurang -1. Sehingga didapatkan bahwa pada aplikasi *BCA Mobile* yang berlabel negatif sebanyak 2.741 data dan label positif sebanyak 2.114 data. Sedangkan untuk aplikasi *myBCA* ulasan yang berlabel negatif sebanyak 2.825 data dan label positif sebanyak 2.082.

Tabel 3. Distribusi Data BCA Mobile

Label	Distribusi	Jumlah Data
Positif	43.5 %	2.114
Negatif	56.5 %	2.741

Tabel 4. Distribusi Data myBCA

Label	Distribusi	Jumlah Data
Positif	42.4 %	2.082
Negatif	57.6 %	2.825

Hasil pelabelan data ulasan aplikasi *BCA Mobile* dan *myBCA* pada penelitian ini dapat dilihat pada Tabel 3 dan Tabel 4, dimana hasil dari kedua distribusi aplikasi menunjukkan bahwa mayoritas ulasan terklasifikasi pada ulasan negatif. Pada aplikasi *BCA Mobile* data ulasan negatif yang tertera sebanyak 2.741 dengan presentase sebesar 56.5 %, sedangkan data ulasan negatif pada aplikasi *myBCA* sebanyak 2.825 dengan presentase sebesar 57.6 %.

4.5. Modelling

Pada tahap ini akan dilakukan proses penerapan algoritma menggunakan data yang sudah siap diolah. Penelitian ini memiliki 2 model penelitian yaitu *Bidirectional Long Short-Term Memory* (BiLSTM) dan *Support Vector Machine* (SVM). Data yang akan digunakan untuk pemodelan akan dibagi menjadi data latih dan data uji dengan rasio 80:20. Data latih digunakan untuk melatih model algoritma agar dapat mengenai pola, sedangkan data uji digunakan untuk menguji akurasi, performa, serta kemampuan prediksi model.

Tabel 5. Pembagian Data Latih dan Data Uji

Aplikasi	Data Latih	Data Uji
<i>BCA Mobile</i>	3.884	971
<i>myBCA</i>	3.925	982

a. *Bidirectional Long Short-Term Memory (BiLSTM)*Tabel 6. Arsitektur BiLSTM *BCA Mobile*

<i>Layer (Type)</i>	<i>Output Shape</i>	<i>Parameter</i>
<i>Embedding</i>	(None, 100, 128)	377.600
<i>Bidirectional</i>	(None, 100, 128)	98.816
<i>Bidirectional_1</i>	(None, 64)	41.216
<i>Dropout</i>	(None, 64)	0
<i>Dense</i>	(None, 64)	4.160
<i>Dense_1</i>	(None, 2)	130

Total *Params* : 521.922Trainable *Params* : 521.922Tabel 7. Arsitektur BiLSTM *myBCA*

<i>Layer (Type)</i>	<i>Output Shape</i>	<i>Parameter</i>
<i>Embedding</i>	(None, 100, 128)	388.992
<i>Bidirectional</i>	(None, 100, 128)	98.816
<i>Bidirectional_1</i>	(None, 64)	41.216
<i>Dropout</i>	(None, 64)	0
<i>Dense</i>	(None, 64)	4.160
<i>Dense_1</i>	(None, 2)	130

Total *Params* : 533.314Trainable *Params* : 533.314

Berdasarkan ringkasan model BiLSTM pada Tabel 6 dan Tabel 7, kedua aplikasi memiliki struktur lapisan dan menghasilkan output yang sama. Perbedaan antara kedua model terdapat pada jumlah parameter. Model BiLSTM pada aplikasi *BCA Mobile* memiliki 521.922 parameter, sedangkan untuk aplikasi *myBCA* memiliki 533.314 parameter. Perbedaan jumlah parameter terutama terdapat pada embedding layer, yang menunjukkan adanya perbedaan ukuran vocabulary yang digunakan pada kedua dataset.

b. *Support Vector Machine (SVM)*Tabel 8. *Term Rata-Rata TF-IDF Tertinggi BCA Mobile*

Term	Index	TF	DF	IDF	TF-IDF
Aplikasi	156	927	787	2.82	37.81
Lemot	1610	426	389	3.52	31.42
Merah	1854	560	508	3.26	28.08
Update	3182	409	352	3.62	22.24
Buka	470	417	364	3.59	21.47
Pakai	2167	475	401	3.49	18.26
Qris	2409	329	295	3.79	18.01
Hijau	1063	307	275	3.86	15.59
Saldo	2558	322	255	3.94	13.45
Bagus	219	710	689	2.95	8.26

Tabel 9. *Term Rata-Rata TF-IDF Tertinggi myBCA*

Term	Index	TF	DF	IDF	TF-IDF
Update	3291	1055	870	2.72	480.23
Buka	480	561	486	3.31	273.56
Aplikasi	148	1451	1141	2.46	51.35
MyBCA	1978	655	559	3.17	26.58
BCA	281	734	608	3.08	26.34
Baik	227	438	417	3.46	24.25
Transaksi	3197	475	426	3.44	23.61
Masuk	1786	481	392	3.52	20.56
Mudah	1951	230	220	4.10	20.27
Tolong	3175	469	443	3.40	19.46

Berdasarkan hasil pembobotan kata (TF-IDF) pada Tabel 8 dan Tabel 9 dapat disimpulkan bahwa ulasan pengguna aplikasi *BCA Mobile* lebih mengarah pada aspek kinerja aplikasi, diantaranya pada kata *lemot*, *update*, *merah*, *hijau*. Sedangkan ulasan pengguna aplikasi *myBCA* lebih menyoroti dalam hal kemudahan penggunaan.

4.6. Evaluasi

Tahap terakhir pada penelitian ini adalah evaluasi. Evaluasi berperan untuk membandingkan performa kedua model yaitu *Bidirectional Long Short-Term Memory* (BiLSTM) dan *Support Vector Machine* (SVM). Pada tahap ini dibagi menjadi 2 bagian yaitu *Confusion Matrix* dan Uji *McNemar Test*.

a. *Confusion Matrix*

Analisis Error

Beberapa contoh kesalahan klasifikasi yang sering terjadi:

- Negasi – “aplikasinya tidak *lemot*” (seharusnya positif, diprediksi negatif karena kata “*lemot*”).
- Sarkasme / ironi – “bagus banget sering error” (seharusnya negatif).
- Slang & typo – “app nya *lemot bgt*” atau “qris gagal trs”.
- Istilah teknis perbankan – “saldo tidak *update*”, “verifikasi gagal”, “mPIN salah terus”.

Kesalahan ini menunjukkan bahwa model masih kesulitan menangani konteks negasi dan sarkasme, terutama BiLSTM yang bergantung pada urutan kata, sementara SVM lebih robust terhadap noise pada data berukuran menengah. Maka dari itu, *Confusion Matrix* dimanfaatkan sebagai tolak ukur suatu model dengan mengukur kinerjanya. Parameter yang digunakan berupa empat metrik evaluasi, diantaranya *accuracy*, *precision*, *recall*, dan *F1-Score* seperti yang tertera pada Tabel 10 dan Tabel 11.

Tabel 10. Metrik *Confusion Matrix* BiLSTM

Aplikasi	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>BCA Mobile</i>	84.5 %	85 %	84 %	84 %
<i>myBCA</i>	85.6 %	86 %	86 %	86 %

Hasil *confusion matrix* pada model BiLSTM menunjukkan bahwa aplikasi *myBCA* memiliki *accuracy* tertinggi yaitu sebesar 85.6% dibandingkan dengan aplikasi *BCA Mobile* yang hanya memiliki *accuracy* 84.5%. Namun, tingkat klasifikasi pada kedua model sangat baik karena jumlah prediksi benar yang jauh lebih tinggi dibandingkan dengan prediksi salah.

Tabel 11. Metrik *Confusion Matrix* SVM

Aplikasi	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>BCA Mobile</i>	86 %	87 %	86 %	86 %
<i>myBCA</i>	87 %	88 %	87 %	87 %

Untuk hasil *confusion matrix* pada model SVM menunjukkan bahwa aplikasi *myBCA* juga memiliki *accuracy* tertinggi yaitu sebesar 87% dibandingkan dengan aplikasi *BCA Mobile* yang hanya memiliki *accuracy* 86%. Sehingga dapat disimpulkan bahwa pola sentimen pada ulasan pengguna *myBCA* lebih mudah dikenali oleh model SVM.

Hasil penelitian diatas memperlihatkan bahwa model Support Vector Machine (SVM) menghasilkan kinerja yang lebih baik dibandingkan dengan model Bidirectional Long Short-Term Memory (BiLSTM) pada analisis sentimen ulasan pengguna aplikasi BCA Mobile dan myBCA. Kondisi ini dipengaruhi oleh beberapa faktor yang terkait. Pertama, dataset yang terbatas atau relatif kecil lebih dominan pada kinerja metode SVM karena metode ini memiliki prinsip Structural Risk Minimization (SRM) yang bertujuan untuk menemukan hyperplane terbaik dan dinilai efektif pada data dengan skala kecil sampai menengah, sementara metode LSTM membutuhkan data yang berskala besar atau kompleks untuk menghasilkan kinerja terbaiknya. Kedua, pelabelan data dengan pendekatan lexicon-based dapat mempengaruhi kinerja metode karena dinilai kurang konsisten pada kata yang mengandung sarkasme atau konteks ambigu. Ketidakkonsistenan ini dapat mempengaruhi metode BiLSTM yang kinerjanya mengutamakan konteks dan urutan kata, dibandingkan dengan kinerja SVM yang tidak bergantung pada urutan kata karena mengandalkan pembobotan TF-IDF. Faktor-faktor inilah yang menghasilkan akurasi performa metode SVM lebih baik dibandingkan dengan metode BiLSTM.

b. Uji *McNemar Test*

Dikarenakan penelitian ini termasuk dalam penelitian komparatif, maka perlu dilakukan uji statistik *McNemar Test*. Pengujian ini dimanfaatkan untuk memastikan perbedaan kinerja kedua model signifikan secara statistik. Langkah awal pada pengujian *McNemar test* adalah menentukan prediksi benar atau salah pada aplikasi *BCA Mobile* dan *myBCA* seperti Tabel 12 dan Tabel 13.

Tabel 12. Prediksi Benar atau Salah BiLSTM dan SVM Aplikasi *BCA Mobile*

Label Aktual	Prediksi BiLSTM	Prediksi SVM	BiLSTM	SVM
Positif	Negatif	Positif	Salah	Benar
Negatif	Positif	Negatif	Salah	Benar
Positif	Positif	Negatif	Benar	Salah
Negatif	Negatif	Negatif	Benar	Benar
Negatif	Positif	Negatif	Salah	Benar

Tabel 13. Prediksi Benar atau Salah BiLSTM dan SVM Aplikasi *myBCA*

Label Aktual	Prediksi BiLSTM	Prediksi SVM	BiLSTM	SVM
Positif	Positif	Positif	Benar	Benar
Negatif	Negatif	Negatif	Benar	Benar
Negatif	Negatif	Positif	Benar	Salah
Positif	Positif	Negatif	Benar	Salah
Negatif	Negatif	Positif	Benar	Salah

Setelah mengetahui hasil prediksi benar atau salah dari kedua aplikasi *mobile banking*, langkah selanjutnya adalah melakukan perhitungan uji McNemar dengan data yang telah didapat pada Tabel 14 dan Tabel 15.

Tabel 14. Uji McNemar *BCA Mobile*

	SVM (Benar)	SVM (Salah)
BiLSTM (Benar)	434	65
BiLSTM (Salah)	401	71

Berdasarkan Tabel 14 dapat dijabarkan bahwa 434 merupakan prediksi benar BiLSTM dan SVM, 65 merupakan prediksi BiLSTM benar tetapi SVM salah, 401 merupakan prediksi BiLSTM salah, tetapi SVM benar, dan 71 kedua model sama-sama melakukan kesalahan.

Pada uji *McNemar test* ini nilai yang digunakan yaitu 65 dan 401 karena kedua model menghasilkan prediksi yang berbeda atau biasa disebut discordant pairs. Sehingga dapat dihitung seperti berikut :

$$\begin{aligned}
 \chi^2 &= \frac{(|65 - 401| - 1)^2}{(65 + 401)} \\
 &= \frac{(336 - 1)^2}{446} \\
 &= \frac{(335)^2}{446} \\
 &= \frac{112225}{446} \\
 &= 240,826 = 240,83
 \end{aligned}$$

```

1 hasil = mcnemar(table, exact=False, correction=True)
2
3 print("Statistik McNemar :", hasil.statistic)
4 print("P-value :", hasil.pvalue)

Statistik McNemar : 240.82618025751074
P-value : 2.5975422274000704e-54

```

Gambar 11. Script dan Hasil p-value *BCA Mobile*

Berdasarkan hasil perhitungan McNemar test didapatkan bahwa nilai uji statistik sebesar 240,83 dengan p-value $2,60 \times 10^{-54}$. Dengan demikian hasil p-value $< 0,05$ yang artinya H_0 ditolak dan H_1 diterima.

Tabel 15. Uji McNemar *myBCA*

	SVM (Benar)	SVM (Salah)
BiLSTM (Benar)	420	61
BiLSTM (Salah)	435	66

Sedangkan Tabel 15 dapat dijabarkan bahwa 420 merupakan prediksi benar BiLSTM dan SVM, 61 merupakan prediksi BiLSTM benar tetapi SVM salah, 435 merupakan prediksi BiLSTM salah, tetapi SVM benar, dan 66 kedua model sama-sama melakukan kesalahan. Kemudian dilakukan perhitungan *McNemar test* seperti berikut :

$$\chi^2 = \frac{(|61 - 435| - 1)^2}{(61 + 435)}$$

$$\begin{aligned}
 &= \frac{(374 - 1)^2}{496} \\
 &= \frac{(373)^2}{496} \\
 &= \frac{139129}{496} \\
 &= 280.502 = 280,50
 \end{aligned}$$

```

1 hasil = McNemar(table, exact=False, correction=True)
2
3 print("Statistik McNemar :", hasil.statistic)
4 print("P-value :", hasil.pvalue)

Statistik McNemar : 280.50201612903226
P-value : 5.837117972196774e-63

```

Gambar 12. *Script dan Hasil p-value myBCA*

Pada Gambar 12 dijelaskan bahwa hasil perhitungan McNemar test sebesar 280,50 untuk uji statistik dengan p-value $5,84 \times 10^{-63}$. Dengan demikian hasil p-value $< 0,05$ yang artinya H_0 ditolak dan H_1 diterima. Dari hasil perhitungan pada Gambar 11 dan Gambar 12 dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara metode BiLSTM dan SVM pada analisis sentimen ulasan pengguna BCA Mobile dan myBCA.

Berdasarkan hasil pengujian, menunjukkan bahwa distribusi sentimen dalam ulasan pengguna untuk aplikasi BCA Mobile dan myBCA dominan berlabel negatif. Pada aplikasi BCA Mobile dengan presentase sebesar 56,5% sentimen negatif (setara dengan 2.741 data), sedangkan 43,5% sentimen positif (total 2.114 data). Sementara itu, pada aplikasi myBCA, sentimen negatif mencapai 57,6% atau 2.825 data, sedangkan sentimen positif sebesar 42,4% atau 2.082 data. Sentimen negatif tersebut mengindikasikan masih adanya kendala yang dihadapi pengguna saat menggunakan kedua aplikasi, khususnya terkait kinerja teknis seperti aplikasi yang lambat, proses pembaruan, kegagalan transaksi, indikator, dan tahapan verifikasi. Temuan ini menunjukkan bahwa analisis sentimen dapat menjadi salah satu cara untuk lebih memahami pendapat dan keluhan pengguna, sehingga mempermudah upaya evaluasi dan peningkatan kualitas aplikasi. Selanjutnya, hasil perbandingan menunjukkan bahwa metode Support Vector Machine (SVM) mencapai akurasi 86% pada BCA Mobile dan 87% pada myBCA, sedangkan metode Bidirectional Long Short-Term Memory (BiLSTM) masing-masing mencapai akurasi 84,5% dan 85,6%. Perbedaan tersebut menunjukkan bahwa SVM memiliki kinerja yang lebih baik dibandingkan BiLSTM pada dataset yang digunakan dalam penelitian ini. Kondisi ini dapat dipengaruhi oleh karakteristik dataset yang terbatas, penggunaan pembobotan TF-IDF, serta pelabelan berbasis leksikon. SVM bekerja dengan baik dalam mengklasifikasikan data berdimensi tinggi, sementara BiLSTM memerlukan data yang lebih kompleks dan berskala besar untuk mempelajari konteks serta hubungan antar-kata secara efektif. Perbedaan hasil kinerja tersebut kemudian diperkuat melalui uji statistik McNemar, yang digunakan untuk menentukan apakah perbedaan prediksi antara kedua metode tersebut signifikan secara statistik. Pada BCA Mobile, nilai statistik uji adalah 240,83 dengan p-value sebesar $2,60 \times 10^{-54}$, sedangkan pada myBCA, nilai statistik uji adalah 280,50 dengan p-value sebesar $5,84 \times 10^{-63}$. Karena kedua nilai p-value $< 0,05$, H_0 ditolak dan H_1 diterima. Hal ini menunjukkan bahwa adanya perbedaan kinerja yang signifikan secara statistik antara metode BiLSTM dan SVM pada kedua aplikasi tersebut. Dengan demikian, keunggulan SVM dalam penelitian ini tidak hanya terlihat dari nilai accuracy, precision, recall, dan F1-Score, tetapi juga didukung oleh hasil uji statistik McNemar yang mengindikasikan adanya perbedaan kinerja yang signifikan. Temuan penelitian ini sejalan dengan penelitian sebelumnya oleh [8] mengenai evaluasi aplikasi BCA Mobile menggunakan algoritma Naïve Bayes dan Support Vector Machine. Studi mereka menunjukkan bahwa Support Vector

Machine mencapai akurasi sebesar 85%, yang lebih baik dibandingkan Naïve Bayes dengan akurasi sebesar 83%.

Dengan demikian, penelitian ini memperkuat hasil penelitian sebelumnya tentang kemampuan SVM dalam mengklasifikasikan sentimen dari ulasan aplikasi BCA, sekaligus melakukan pengembangan dengan membandingkan SVM dengan BiLSTM serta menggunakan dua objek penelitian, yaitu BCA Mobile dan myBCA. Berdasarkan hasil keseluruhan tersebut, masalah pada latar belakang yang berkaitan dengan kesulitan menganalisis banyak ulasan pengguna secara manual dapat diatasi dengan menggunakan analisis sentimen. Hasil uji perbandingan menunjukkan bahwa SVM bisa direkomendasikan sebagai metode yang memiliki kinerja terbaik dalam penelitian ini. Hasil analisis juga bisa digunakan sebagai bahan untuk mengevaluasi pengembang agar bisa meningkatkan kualitas aplikasi, terutama pada bagian kinerja teknis, stabilitas, proses pembaruan, indikator dan notifikasi, verifikasi, serta kemudahan dalam menggunakan aplikasi..

5. Kesimpulan

Berdasarkan hasil kajian mengenai analisis sentimen pada ulasan pengguna aplikasi BCA Mobile dan myBCA, distribusi sentimen pada kedua aplikasi didominasi oleh sentimen negatif. Pada aplikasi BCA Mobile diperoleh 56,5% ulasan negatif (2.741 data) dan 43,5% ulasan positif (2.114 data), sedangkan pada aplikasi myBCA diperoleh 57,6% ulasan negatif (2.825 data) dan 42,4% ulasan positif (2.082 data). Hasil analisis TF-IDF menunjukkan bahwa ulasan pengguna BCA Mobile lebih banyak membahas aspek kinerja aplikasi, seperti pembaruan, kecepatan yang lambat, dan indikator sistem, sementara ulasan myBCA lebih berfokus pada kemudahan penggunaan.

Perbandingan kinerja model menunjukkan bahwa Support Vector Machine (SVM) menghasilkan performa yang lebih baik dibandingkan Bidirectional Long Short-Term Memory (BiLSTM) pada kedua aplikasi. Pada BCA Mobile, BiLSTM memperoleh accuracy 84,5%, precision 85%, recall 84%, dan F1-Score 84%, sedangkan SVM memperoleh accuracy 86%, precision 87%, recall 86%, dan F1-Score 86%. Pada myBCA, BiLSTM memperoleh accuracy 85,6%, precision 86%, recall 86%, dan F1-Score 86%, sedangkan SVM memperoleh accuracy 87%, precision 88%, recall 87%, dan F1-Score 87%. Selisih kinerja tersebut diperkuat oleh uji McNemar yang menghasilkan p-value $2,60 \times 10^{-54}$ pada BCA Mobile dan $5,84 \times 10^{-63}$ pada myBCA. Karena kedua p-value $< 0,05$, perbedaan kinerja antara BiLSTM dan SVM dinyatakan signifikan secara statistik.

Dengan demikian, dalam konteks dataset, preprocessing, dan konfigurasi eksperimen pada penelitian ini, metode SVM dengan kernel linear dapat direkomendasikan sebagai metode yang lebih optimal untuk menganalisis sentimen ulasan pengguna aplikasi BCA Mobile dan myBCA. Temuan ini tidak dimaksudkan untuk digeneralisasi bahwa SVM selalu unggul dibandingkan BiLSTM pada semua kondisi data. Hasil analisis sentimen tersebut juga dapat dimanfaatkan sebagai bahan evaluasi bagi pengembang untuk meningkatkan kualitas layanan dan fitur aplikasi, khususnya terkait kinerja teknis, stabilitas, proses pembaruan, serta kemudahan penggunaan.

Referensi

- [1] A. Ramadhan and A. Priyono, "Analisis Faktor-faktor yang Mempengaruhi Konsumen M-Banking BCA untuk Terus Menggunakan Layanan," *Sel. Manaj. J. Mbs. Bisnis Manaj.*, vol. 01, no. 04, pp. 267–277, 2022.
- [2] J. P. Yoman, C. Cok, K. Laigusten, and G. Zovintho, "Analisis Sentimen Ulasan Perbedaan Aplikasi BCA Mobile dengan MyBCA di Playstore Menggunakan Metode Lexicon," *JDMIS J. Data Min. Inf. Syst.*, vol. 3, no. 2, pp. 64–76, 2025, doi: 10.54259/jdmis.v3i2.4236.
- [3] H. Natanael, A. Pinandito, and M. C. Saputra, "Evaluasi User Experience pada Aplikasi BCA Mobile dan MyBCA dengan menggunakan Metode Usability Testing dan User Experience Questionnaire (UEQ)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 1, pp. 495–504, 2023.
- [4] M. Gamma, A. Hakim, and F. Irwiensyah, "Analisis Sentimen Terhadap Ulasan Pengguna Pada Aplikasi BCA Mobile Menggunakan Metode Naïve Bayes," *J. Inf. Syst. Res.*, vol. 5, no. 4, 2024, doi: 10.47065/josh.v5i4.5343.
- [5] R. Ferdiana *et al.*, "Dataset Indonesia untuk Analisis Sentimen," *JNTETI*, vol. 8, no. 4, pp. 334–339, 2019.
- [6] S. Adi *et al.*, "Analisis Sentimen Aplikasi Halo BCA di Google Play Store Menggunakan Metode Naive Bayes, Support Vector Machine dan Random Forest," *HOAQ J. Teknol. Inf.*, vol. 15, no. c, pp. 69–79, 2024.
- [7] S. H. Permatasari, I. Manfaati Nur, and F. Fauzi, "Metode Bidirectional Long Short-Term Memory (BiLSTM) Untuk Memprediksi Harga Saham BBRI Dengan Optimasi Nesterov Adaptive Moment (Nadam)," *Pros. Semin. Nas. Unimus*, vol. 7, pp. 1151–1159, 2024.
- [8] D. W. Bhatara and R. R. Suryono, "Analisis sentimen aplikasi bca mobile menggunakan algoritma naïve bayes dan support vector machine," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 4, pp. 1907–1917, 2024.
- [9] A. Sole and A. Zubair, "Analisis Perbandingan Naïve Bayes dan Random Forest untuk Evaluasi PKKMB Berbasis Sentimen Mahasiswa Comparative Analysis of Naïve Bayes and Random Forest for Student Sentiment-Based PKKMB Evaluation," *J. Janitra Inform. dan Sist. Inf.*, vol. 1, no. 1, pp. 20–32, 2026, doi: 10.59395/nf4xmg03.
- [10] E. Budianita, E. P. Cynthia, A. Pranata, and D. Abimanyu, "Pendekatan berbasis Machine Learning dan Leksikal Pada Analisis Sentimen," *Semin. Nas. Teknol. Informasi, Komun. dan Ind.*, pp. 99–104, 2022.
- [11] M. A. Khder, "Web Scraping or Web Crawling : State of Art , Techniques , Approaches and Application," *Int. J. Adv. Soft Compu*, vol. 13, no. 3, 2021, doi: 10.15849/IJASCA.211128.11.
- [12] H. T. Duong, T. Anh, and N. Thi, "A review : preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Networks*, pp. 1–16, 2021, doi: 10.1186/s40649-020-00080-x.
- [13] M. I. Hossain, "Support Vector Machine," *Master Sci. High Integr. Syst.*, Art. no. 1340303, 2022, doi: 10.18411/trnio-12-2023-769.
- [14] N. H. Ovirianti, M. Zarlis, and H. Mawengkang, "Support Vector Machine Using A Classification Algorithm," *Sink. J. dan Penelit. Tek. Inform.*, vol. 6, no. 3, pp. 2103–2107, 2022.
- [15] D. Masruroh *et al.*, "Wirausaha Majoo Menggunakan Metode SVM," *JATI (Jurnal Mbs. Tek. Inform.*, vol. 9, no. 5, pp. 8051–8057, 2025.
- [16] E. Helmud, E. Helmud, and P. Romadiana, "Classification Comparison Performance of Supervised Machine Learning Random Forest and Decision Tree Algorithms Using Confusion Matrix," *J. SISFOKOM (Sistem Inf. dan Komputer)*, vol. 13, pp. 92–97, 2024.
- [17] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, "Visualisasi Word Cloud Hasil Analisis Sentimen Berbasis Fitur Layanan Aplikasi Gojek dengan Support Vector Machine," *J. Serina Sains, Tek. dan Kedokt.*, vol. 02, no. 01, pp. 61–70, 2024.
- [18] N. M. C. Utami, N. L. P. L. S. Setiawati, A. A. I. A. S. Komaladewi, and F. P. P. Ferdiansyah, "Jurnal Teknologi Terpadu," *J. Teknol. Terpadu*, vol. 7, no. 1, pp. 15–22, 2021, [Online]. Available: <https://media.neliti.com/media/publications/493730-water-ph-and-turbidity-control-system-in-0a553e14.pdf>
- [19] E. Trihartiningsih and D. P. Putri, "Pengaruh Pendidikan Kesehatan Terhadap Peningkatan Pengetahuan Kesehatan Reproduksi Remaja," *Indones. J. Innov. Multidisipliner Res.*, vol. 1, no. 3, pp. 385–391, 2023, doi: 10.69693/ijim.v1i3.145.
- [20] E. Lutfina, W. Andriana, S. Quamila, P. Wiratmaja, and E. Febrianti, "Metode dan Algoritma Dalam Sentimen Analisis : Systematic Literature Review," *Sci. Technol. Manag. J.*, vol. 4, no. 2, pp. 67–79, 2024.