

Knowledge Discovery in Sharia Mobile Banking Reviews Using Aspect-Based Sentiment Analysis and Machine Learning

Muthia Nashiroh Ramadhani ¹, Ken Ditha Tania ^{2*}, Mira Afrina ³

^{1,2,3}Sistem Informasi, Universitas Sriwijaya

muthianashiroh@gmail.com ¹, kenya.tania@gmail.com ², mirafrina81@gmail.com ³

Article Info

Article history:

Received 2025-11-12

Revised 2025-12-25

Accepted 2026-01-07

Keyword:

ABSA,
BYOND by BSI,
Classification,
Knowledge Discovery,
Sentiment Analysis.

ABSTRACT

User reviews provide important insights into the quality of digital banking applications; however, their large volume makes manual analysis inefficient. This study applies Aspect-Based Sentiment Analysis (ABSA) to examine user perceptions of the BYOND by BSI application based on three aspects: interface, features and performance, and services. Three classification algorithms were compared: Naïve Bayes, Support Vector Machine (SVM), and Random Forest, evaluated with accuracy, precision, recall, *F1-score*, and ROC-AUC. The results indicate that SVM and Naïve Bayes achieved the best performance, with an accuracy of 0.95 and an *F1-score* of 0.92, whereas Random Forest exhibited slightly lower performance with an *F1-score* of 0.89. Furthermore, sentiment analysis reveals the features and performance aspect exhibits the highest proportion of negative sentiment (39.6%), primarily associated with system reliability issues, login problems, transaction failures, and application instability. These findings demonstrate that ABSA can serve as an effective knowledge discovery approach for identifying critical functional issues and supporting data-driven prioritization in improving digital banking services, particularly within the context of sharia banking applications.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Transformasi digital secara signifikan telah membawa perubahan dalam sektor perbankan di Indonesia. Perubahan ini mendorong pengembangan layanan perbankan yang lebih efisien dan fleksibel melalui penerapan *mobile banking* [1]. *Mobile banking* hadir sebagai layanan keuangan berbasis digital yang mempermudah nasabah dalam melakukan berbagai transaksi secara *online* melalui koneksi internet dan perangkat seluler [2]. Kemudahan dalam mengelola keuangan, menjadikan *mobile banking* sebagai layanan yang sangat dibutuhkan oleh masyarakat [3].

PT Bank Syariah Indonesia Tbk (BSI) secara resmi berdiri pada tanggal 1 Februari 2021, merupakan bank syariah terbesar di Indonesia [4]. BSI meluncurkan BYOND by BSI pada November 2024 sebagai *superapp* layanan perbankan syariah yang komprehensif dan inovatif. Aplikasi ini merupakan pengembangan dari BSI Mobile sebagai bagian dari upaya transformasi digital untuk meningkatkan kualitas

pengalaman transaksi pengguna melalui layanan yang lebih modern, aman, dan efisien [5]. Seiring meningkatnya penggunaan *mobile banking* di Indonesia, diperlukan evaluasi yang berkelanjutan terhadap kualitas layanan digital perbankan, termasuk BYOND by BSI [6]. Ulasan pengguna pada platform seperti Google Play Store menjadi sumber data berharga dalam memahami tingkat kepuasan, keluhan, dan ekspektasi pengguna terhadap kinerja aplikasi perbankan [7]. Dalam hal ini, metode analisis berbasis teks seperti *sentiment analysis* dapat dimanfaatkan untuk mengekstraksi informasi penting dari ribuan ulasan yang bersifat subjektif [8].

Analisis sentimen adalah bagian dari *Natural Language Processing* (NLP) yang bertujuan dalam melakukan identifikasi, penilaian, dan klasifikasi opini dalam bentuk teks [9]. Pendekatan ini digunakan untuk memahami persepsi pengguna terhadap suatu produk maupun layanan tertentu [10]. Analisis sentimen memiliki peran penting dalam mengekstrak informasi dari ulasan pengguna, karena

dapat dijadikan sebagai landasan dalam proses pengambilan keputusan [11]. Secara umum, analisis sentimen mengelompokkan opini menjadi tiga jenis kategori, sentimen positif, negatif, dan netral [9].

Aspect-Based Sentiment Analysis (ABSA) merupakan bagian dari analisis sentimen yang berkembang pesat dan semakin mendapat perhatian dalam beberapa tahun terakhir [12]. Teknik ini memungkinkan identifikasi sentimen secara spesifik terhadap aspek-aspek tertentu dalam suatu aplikasi [13]. Penelitian sebelumnya menerapkan ABSA untuk mengevaluasi opini pengguna aplikasi Flip dari Google Play dengan menerapkan *Support Vector Machine* (SVM) sebagai algoritma dalam klasifikasi sentimen. Tiga aspek utama yang menjadi fokus dalam penelitian tersebut meliputi kecepatan, keamanan, dan biaya. Penelitian ini mengindikasikan bahwa model SVM dengan kernel linear dan nilai parameter C sebesar 1.0 memberikan akurasi terbaik secara keseluruhan, dengan akurasi tertinggi pada aspek biaya [14].

Sebuah studi juga menerapkan *Aspect-Based Sentiment Analysis* (ABSA) menggunakan *Naïve Bayes* pada ulasan aplikasi JConnect Mobile, dengan rata-rata akurasi sebesar 93,1% dan mengidentifikasi tiga aspek utama, yaitu tampilan, fitur dan performa, serta layanan. Aspek tampilan dan layanan didominasi oleh sentimen positif, sedangkan aspek fitur dan performa lebih banyak mengandung sentimen negatif [13]. Penelitian serupa juga dilakukan terhadap aplikasi OVO dengan menerapkan ABSA menggunakan *topic modelling* dari LDA dan SVM untuk mengidentifikasi empat aspek utama, yaitu *Account Access and Registration*, *User Experience*, *Transaction Efficiency*, serta *Balance and Charges*. Model terbaik diperoleh pada SVM Linear dengan nilai *F1-score* sebesar 0,804 [15].

Penelitian selanjutnya menggunakan algoritma SVM untuk mengevaluasi sentimen dalam ulasan pengguna terhadap aplikasi layanan perbankan, seperti Livin' by Mandiri, BRI Mobile, dan BCA Mobile. Dari hasil evaluasi akurasi tertinggi diperoleh dari dataset BRI Mobile sebesar 92.97%, disusul oleh dataset gabungan dengan akurasi 90.05%, kemudian BCA Mobile sebesar 89.73%, dan Livin' by Mandiri sebesar 87.46% [16]. Penelitian lain melakukan perbandingan terhadap efektivitas algoritma *Random Forest*, *Naïve Bayes*, dan SVM dalam menganalisis opini terhadap aplikasi Ajaib. Data ulasan diperoleh sebanyak 2.000 dengan teknik *web scraping*. Tahap *preprocessing* mencakup *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Proses pelabelan dilakukan dengan memberi label sentimen positif (skor >3), netral (skor =3), dan negatif (skor <3), lalu diolah menggunakan teknik pembobotan TF-IDF dan *feature extraction* dengan CountVectorizer. Berdasarkan hasil evaluasi, SVM memperoleh akurasi tertinggi sebesar 91%, diikuti Random Forest 85%, dan Naïve Bayes 83% [17].

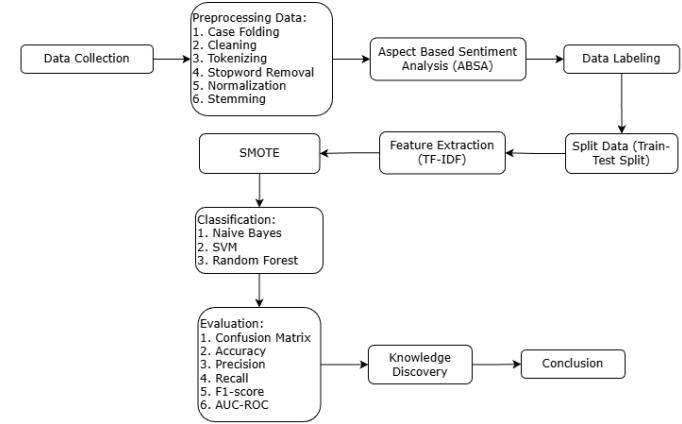
Berdasarkan berbagai penelitian sebelumnya, *Aspect-Based Sentiment Analysis* telah banyak digunakan untuk mengevaluasi persepsi pengguna terhadap layanan

perbankan digital secara lebih mendalam melalui analisis sentimen berbasis aspek. Namun, penelitian-penelitian tersebut umumnya masih berfokus pada aplikasi perbankan konvensional atau layanan finansial umum, serta belum secara khusus mengkaji penerapan ABSA pada domain perbankan syariah khususnya BYOND by BSI. Selain itu, kajian yang mengombinasikan pendekatan pelabelan sentimen secara otomatis berbasis model bahasa dengan evaluasi komparatif algoritma *machine learning* dalam kerangka ABSA masih relatif terbatas.

Penelitian ini bertujuan untuk mengisi celah tersebut dengan melakukan *knowledge discovery* melalui penerapan *Aspect-Based Sentiment Analysis* dalam menganalisis ulasan pengguna terhadap aplikasi BYOND by BSI sebagai aplikasi perbankan syariah generasi terbaru. Penelitian ini juga mengevaluasi dan membandingkan efektivitas tiga algoritma *machine learning*, yaitu *Naïve Bayes*, *Support Vector Machine* (SVM), dan *Random Forest*, dalam klasifikasi sentimen berbasis aspek. Diharapkan, hasil penelitian ini dapat memberikan wawasan empiris yang bermanfaat dalam memahami persepsi pengguna serta berkontribusi terhadap peningkatan kualitas layanan digital pada perbankan syariah.

II. METODE

Penelitian ini dilaksanakan melalui tahapan yang tersusun secara sistematis dan berurutan dari tahap awal hingga tahap akhir. Alur dari penelitian digambarkan melalui diagram alur (*flowchart*) seperti pada gambar 1.



Gambar 1. Alur Penelitian

A. Data Collection

Data dikumpulkan melalui pemanfaatan teknik *web scraping* dari ulasan aplikasi BYOND by BSI yang tersedia di Google Play Store. Pengambilan data dilakukan pada rentang waktu 3 Juni hingga 10 Juli 2025 menggunakan library *Google Play Scraper* di platform Google Colaboratory (Colab). Data awal yang discraping berjumlah 5.000 ulasan, kemudian data yang bersifat duplikat dan kosong dihapus, sehingga menyisakan 4.910 data yang

digunakan untuk analisis sentimen. Data yang dikumpulkan mencakup *review text*, rating (skala 1-5), dan *date*, hasil *scraping* kemudian disimpan dalam format CSV untuk mempermudah analisis selanjutnya. Tabel 1 menyajikan contoh ulasan yang berhasil dikumpulkan.

TABEL I
DATA ULASAN PENGGUNA

Review	Rating	Date
Sistem BUSUK. Sering error, apalagi kalau saat penting, seperti pembayaran yang harus dibayar waktu itu juga.	1	6/13/2025 9:04
Terimakasih bsi aplikasih memudahkan transaksi dan gak ribet mudah untuk proses apa pun lancar gak ada masalah sama sekali kerreeeeeeeeeen 😊	5	6/3/2025 11:52
sangat tidak nyaman, setiap mau buka byond bsi hampir setiap hari selalu tidak bisa di buka padahal mau cek transaksi dan lain2.. tolong pihak bsi segera di perbaiki agar tidak mengecewakan nasabah	2	7/8/2025 5:30
Suka banget sama aplikasinya! Bikin hidup jadi lebih praktis. Terima kasih BYON by BSI	5	6/15/2025 4:14
BYOND sering gangguan dan untuk yg jauh lokasi dengan cabang sangat susah. karena call center pun tidak memberikan solusi dan diarahkan ke kantor cabang.	1	6/13/2025 4:07

B. Preprocessing Data

Data ulasan yang dikumpulkan pada tahap sebelumnya selanjutnya diproses melalui tahapan *preprocessing* sebelum dilakukan analisis sentimen. Tahap *preprocessing* memiliki peran penting karena bertujuan untuk meningkatkan kualitas data sebelum dilakukan pemodelan [18]. Proses ini mencakup penghilangan noise atau gangguan pada data, penyesuaian format data agar lebih efektif untuk dianalisis, pengurangan potensi bias atau kesalahan dengan memperhatikan *missing value*, serta penyiapan data agar sesuai dengan model [19].

Tahapan pertama dalam proses *preprocessing* adalah *case folding*, yaitu proses standarisasi bentuk huruf dengan cara mengubah teks secara keseluruhan menjadi huruf kecil [20]. Selanjutnya dilakukan *cleaning*, yaitu pembersihan data dengan menghapus elemen yang tidak penting, seperti URL, karakter khusus, tag HTML, angka, emoji, serta duplikasi data [21]. Tahap selanjutnya *tokenizing*, yaitu pemrosesan teks dengan cara membagi kalimat menjadi token atau bentuk yang lebih kecil, yang dipisahkan berdasarkan spasi [22]. Setelah memisahkan kalimat menjadi token, kemudian dilakukan *stopword removal*, yaitu proses penyaringan kata-

kata umum yang tidak memiliki makna, seperti konjungsi dan preposisi contohnya seperti kata “di”, “untuk”, “ke” “dan”, serta “yang” [9]. Kemudian dilakukan normalisasi dengan cara mengubah teks tidak baku, seperti singkatan, bahasa gaul atau bentuk tulisan informal lainnya, menjadi format standar sehingga dapat diproses secara optimal oleh sistem [9]. Tahap terakhir adalah *stemming*, yaitu proses mengembalikan kata ke bentuk aslinya, dengan menghapus afiks seperti awalan, akhiran, dan imbuhan lainnya. Hal ini dilakukan menggunakan algoritma berbahasa Indonesia seperti Sastrawi [23]. Hasil *preprocessing data* dapat dilihat pada tabel 2.

TABEL II
HASIL PREPROCESSING DATA

Preprocessing	Text
Dataset	Sistem BUSUK. Sering error, apalagi kalau saat penting, seperti pembayaran yang harus dibayar waktu itu juga.
Case Folding	sistem busuk. sering error, apalagi kalau saat penting, seperti pembayaran yang harus dibayar waktu itu juga.
Cleaning	sistem busuk sering error apalagi kalau saat penting seperti pembayaran yang harus dibayar waktu itu juga
Tokenizing	['sistem', 'busuk', 'sering', 'error', 'apalagi', 'kalau', 'saat', 'penting', 'seperti', 'pembayaran', 'yang', 'harus', 'dibayar', 'waktu', 'itu', 'juga']
Stopword Removal	['sistem', 'busuk', 'error', 'pembayaran', 'dibayar']
Normalization	['sistem', 'busuk', 'error', 'pembayaran', 'dibayar']
Stemming	['sistem', 'busuk', 'error', 'bayar', 'bayar']

C. Aspect Based Sentiment Analysis (ABSA)

Aspect Based Sentiment Analysis (ABSA) merupakan pendekatan analisis sentimen yang bersifat lebih mendalam, dengan tujuan untuk mengidentifikasi aspek-aspek tertentu dalam sebuah ulasan dan menentukan polaritas sentimen terhadap masing-masing aspek tersebut [24][25]. Pada penelitian ini, pemilihan aspek didasarkan pada studi literatur terhadap penelitian sebelumnya yang menganalisis ulasan aplikasi layanan keuangan digital menggunakan ABSA. Berdasarkan referensi [26][13], penggunaan aspek yang dianalisis mencakup tiga kategori utama, yaitu aspek tampilan yang berkaitan dengan desain antarmuka dan kemudahan navigasi; aspek fitur dan performa yang mencakup keandalan sistem serta fitur-fitur yang tersedia; serta aspek layanan yang berfokus pada pengalaman pengguna secara keseluruhan, termasuk tingkat kepuasan serta persepsi emosional mereka terhadap layanan yang diberikan oleh aplikasi. Ketiga aspek tersebut umumnya digunakan dan dianggap representatif dalam menggambarkan pengalaman pengguna aplikasi *mobile banking*.

Proses ekstraksi aspek pada penelitian ini dilakukan menggunakan pendekatan berbasis kata kunci (*keyword-based aspect extraction*). Setiap aspek direpresentasikan oleh sekumpulan kata kunci yang ditentukan secara manual berdasarkan studi literatur dan karakteristik domain aplikasi perbankan. Ulasan pengguna kemudian dikelompokkan ke dalam aspek tampilan, fitur dan performa, serta layanan dengan mencocokkan kata kunci yang muncul pada teks ulasan. Pendekatan ini bersifat *rule-based* dan memungkinkan proses ekstraksi aspek yang terstruktur. Daftar kata kunci yang digunakan untuk masing-masing aspek disajikan pada Tabel 3.

TABEL III
KATA KUNCI ASPEK ULASAN

No	Aspek	Kata Kunci
1	Tampilan	'antarmuka', 'bersih', 'clear', 'dark mode', 'desain', 'font', 'halaman', 'ikon', 'interface', 'jelas', 'keren', 'layout', 'menarik', 'menu', 'modern', 'mudah', 'mudah digunakan', 'navigasi', 'praktis', 'rapi', 'ribet', 'simpel', 'simple', 'tampilan', 'tata letak', 'theme', 'ui', 'uiux', 'user friendly', 'ux', 'warna'
2	Fitur dan Performa	'aktivasi', 'akun', 'atm', 'bayar', 'bug', 'buka', 'cepat', 'crash', 'digital', 'down', 'error', 'fitur', 'fiturnya', 'fungsi', 'gagal', 'gangguan', 'hang', 'keluar', 'kendala', 'koneksi', 'lag', 'lambat', 'lancar', 'lemot', 'limit', 'loading', 'login', 'masuk', 'mutasi', 'notifikasi', 'otp', 'qr', 'respon lambat', 'saldo', 'server', 'sistem', 'stuck', 'system', 'tarik', 'topup', 'transaksi', 'transfer', 'uang', 'update', 'verif', 'verifikasi', 'versi', 'virtual account', 'wallet', 'wifi'
3	Layanan	'admin', 'akses', 'akun', 'aman', 'amanah', 'bagus', 'baik', 'balas', 'bantu', 'bantuin', 'batal', 'buruk', 'cabang', 'call center', 'cepat', 'cs', 'customer', 'daftar', 'help', 'hubungi', 'kecewa', 'keluhan', 'komplain', 'layanan', 'mantap', 'memuaskan', 'nasabah', 'nyaman', 'ok', 'oke', 'pelayanan',

D. Data Labeling

Secara umum pelabelan data dilakukan untuk mengelompokkan ulasan menjadi tiga kategori sentimen: positif, negatif, dan netral [27]. Namun, dalam penelitian ini analisis dibatasi hanya pada label positif dan negatif, karena ulasan netral umumnya tidak memuat evaluasi yang eksplisit dan tidak secara jelas mencerminkan preferensi pengguna [28]. Proses pelabelan sentimen dilakukan secara otomatis menggunakan model *pre-trained* IndoBERT yang tersedia pada pustaka Hugging Face Transformers. Dalam penelitian ini, IndoBERT digunakan sebagai *pseudo-labeler* untuk menghasilkan label sentimen terhadap teks ulasan berbahasa Indonesia pada masing-masing aspek yang telah diidentifikasi.

Karena satu ulasan dapat mengandung lebih dari satu aspek, setiap ulasan kemudian dipecah menjadi beberapa baris data, di mana setiap baris mewakili satu aspek dan satu label sentimen. Pendekatan ini memastikan bahwa setiap pasangan aspek dan sentimen dianalisis secara terpisah sesuai dengan prinsip ABSA. Dalam proses ini, penentuan label juga mempertimbangkan nilai rating ulasan sebagai pertimbangan tambahan, di mana ulasan dengan rating ≤ 2 diklasifikasikan sebagai sentimen negatif dan ulasan dengan rating ≥ 4 diklasifikasikan sebagai sentimen positif. Selain itu, dilakukan validasi manual terhadap 300 data yang dipilih secara acak untuk mengevaluasi kualitas hasil pelabelan otomatis. Hasil validasi menunjukkan tingkat kesesuaian label sebesar 0,94, yang mengindikasikan bahwa pendekatan pelabelan yang digunakan memiliki tingkat akurasi yang tinggi dan dapat diandalkan.

TABEL IV
HASIL PELABELAN

Review	Rating	Label Sentimen Aspek		
		Fitur dan Performa	Layanan	Tampilan
sistem busuk sering error apalagi kalau saat penting seperti pembayaran yang harus dibayar waktu itu juga	1	Negatif	None	None
terimakasih bsi aplikasi memudahkan transaksi dan gak ribet mudah untuk proses apa pun lancar gak ada masalah sama sekali keran	5	Positif	None	Positif
sangat tidak nyaman setiap mau buka byond bsi hampir setiap hari selalu tidak bisa di buka padahal mau cek transaksi dan lain tolong pihak bsi segera di perbaiki agar tidak	2	Negatif	Negatif	None

mengecewakan nasabah				
suka banget sama aplikasinya bikin hidup jadi lebih praktis terima kasih byond by bsi	5	None	Positif	Positif
byond sering gangguan dan untuk yg jauh lokasi dengan cabang sangat susah karena call center pun tidak memberikan solusi dan diarahkan ke kantor cabang	1	None	Negatif	None

E. Split Data

Untuk memastikan bahwa model dapat melakukan generalisasi terhadap data baru, dilakukan pembagian data menjadi data pelatihan dan data pengujian. Pada penelitian ini, 80% data dibagi untuk pelatihan (*training*) dan 20% digunakan untuk pengujian (*testing*). Pembagian ini bertujuan agar model dapat belajar mengenali pola dari data pelatihan dan kemudian diuji menggunakan data baru yang belum pernah dilihat sebelumnya. Untuk menghindari data *leakage*, pembagian dilakukan berdasarkan *Review_ID* unik bukan berdasarkan baris data. Dengan demikian, seluruh pasangan aspek dan sentimen yang berasal dari satu ulasan yang sama ditempatkan pada subset data yang sama, sehingga tidak terjadi kebocoran informasi antara data pelatihan dan pengujian.

TABEL V
PEMBAGIAN DATA

Rasio Perbandingan	Data Pelatihan	Data Pengujian
80:20	5949	1488

F. Feature Extraction

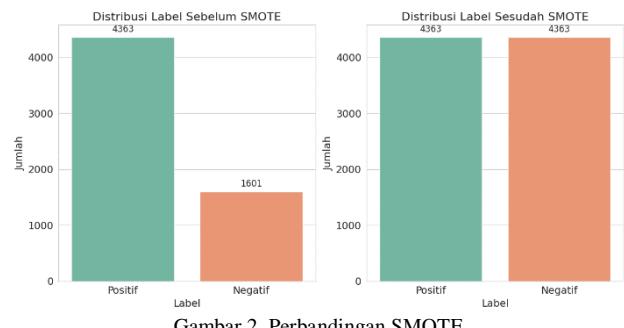
Metode pembobotan kata dalam dokumen, *Term Frequency-Inverse Document Frequency* (TF-IDF) dihitung berdasarkan hasil perkalian antara TF dan IDF [13]. TF-IDF digunakan untuk menilai relevansi suatu kata dengan mempertimbangkan frekuensi kemunculannya dalam dokumen tertentu (*Term Frequency* atau TF) dan frekuensi kemunculannya di seluruh kumpulan dokumen (*Document Frequency* atau DF). Melalui perhitungan IDF, kata yang terlalu banyak muncul dalam dokumen dianggap kurang signifikan dan akan diberikan bobot yang lebih rendah [29]. Hasil vektorisasi TF-IDF kemudian digunakan sebagai masukan untuk tahap SMOTE, yang bertujuan

menyeimbangkan jumlah data antar kelas sebelum proses pelatihan model dilakukan.

G. SMOTE

Teknik Oversampling Minoritas Sintetis (SMOTE) adalah metode yang digunakan untuk *oversampling* dalam mengatasi masalah distribusi data yang tidak seimbang, terutama ketika jumlah data dalam satu kategori tidak sebanding dengan kategori lainnya [30]. Metode ini membuat data sintetis pada kelas minoritas menggunakan interpolasi antar titik data yang sudah ada, bukan hanya menduplikasi data seperti pada *oversampling* konvensional [31]. Tujuan penerapan SMOTE adalah untuk meningkatkan representasi kelas minoritas dengan menghasilkan sampel sintetis sehingga distribusi kelas menjadi lebih seimbang. Pendekatan ini membantu meningkatkan kinerja model dengan menyediakan lebih banyak data pelatihan bagi kelas minoritas [32].

Namun demikian, pada konteks data teks yang direpresentasikan menggunakan TF-IDF, sampel sintetis yang dihasilkan oleh SMOTE bersifat interpolasi numerik dan tidak selalu merepresentasikan struktur linguistik yang realistik. Oleh karena itu, penggunaan SMOTE dalam penelitian ini digunakan untuk meningkatkan pembelajaran model pada kelas minoritas, dengan fokus pada evaluasi kinerja klasifikasi, bukan pada interpretasi linguistik data sintetis.



Gambar 2. Perbandingan SMOTE

Gambar 2 menunjukkan perbandingan distribusi label sebelum dan sesudah penerapan SMOTE. Sebelum *oversampling*, jumlah data pada label *positif* sebanyak 4363 dan *negatif* sebanyak 1601. Setelah penerapan SMOTE, jumlah data pada kedua kelas menjadi seimbang, masing-masing sebanyak 4363 data. Proses ini dilakukan agar model dapat belajar secara lebih optimal dan tidak bias terhadap kelas mayoritas.

H. Classification

Dalam penelitian ini, digunakan beberapa algoritma *machine learning* untuk melakukan klasifikasi. Selanjutnya, dilakukan komparasi terhadap kinerja dari masing-masing algoritma tersebut. Berikut tiga algoritma yang digunakan:

- 1) *Naïve Bayes*: Merupakan metode klasifikasi yang didasarkan pada *Teorama Bayes* dengan memperkirakan peluang dari data masuk ke dalam kategori tertentu [33]. Metode ini bersifat probabilistik dengan anggapan bahwa setiap fitur bekerja secara independen, tidak bergantung satu sama lain [12].
- 2) *Support Vector Machine (SVM)*: SVM beroperasi dengan menentukan *hyperplane* pemisah terbaik antar kelas, serta mampu menangani data non-linear melalui penggunaan fungsi kernel [33].
- 3) *Random Forest*: Algoritma yang beroperasi dengan membentuk sejumlah pohon prediksi yang dilatih secara independen menggunakan subset data acak dari pelatihan. Hasil akhir ditentukan melalui proses *voting* berdasarkan pemilihan mayoritas suara dari semua pohon [34].

Seluruh algoritma klasifikasi pada penelitian ini diimplementasikan menggunakan *library* Scikit-learn dengan konfigurasi hiperparameter default. *Naïve Bayes* menggunakan MultinomialNB dengan parameter bawaan, SVM diimplementasikan menggunakan LinearSVC dengan kernel linear, serta *Random Forest* menggunakan RandomForestClassifier. Penelitian ini tidak melakukan proses *hyperparameter tuning* agar setiap model memperoleh perlakuan yang setara, sehingga perbandingan kinerja dapat dilakukan secara adil dan objektif.

TABEL VI
HIPERPARAMETER KLASIFIKASI MODEL

Model	Hiperparameter
<i>Naïve Bayes</i>	alpha=1.0, fit_prior=True
<i>Support Vector Machine</i>	kernel='linear', C=1.0, random_state=42
<i>Random Forest</i>	n_estimators=100, criterion='gini', max_depth=None, random_state=42, n_jobs=-1

I. Evaluation

Untuk mengetahui seberapa baik kinerja model dilakukan evaluasi model dengan membandingkan prediksi yang dihasilkan terhadap data sebenarnya. Hasil perbandingan ini kemudian disusun dalam bentuk *confusion matrix* untuk memberikan gambaran performa model secara keseluruhan [35]. *Confusion matrix* terdiri atas empat komponen, yaitu *True Positive* (TP) prediksi data positif yang secara benar diprediksi positif, *False Positive* (FP) prediksi positif tetapi secara benar adalah negatif, *False Negative* (FN) data sebenarnya positif tetapi diprediksi negatif, dan *True Negative* (TN) negatif yang diprediksi secara benar dengan

negatif. Nilai-nilai tersebut dapat membantu dalam mengevaluasi efektivitas suatu model melalui metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* [35]. Rumus metrik evaluasi dapat dilihat pada persamaan berikut [36].

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

Accuracy merupakan metrik yang paling umum digunakan dalam mengevaluasi kinerja algoritma pada permasalahan klasifikasi. Nilai akurasi diperoleh dengan membandingkan jumlah prediksi yang benar dari positif maupun negatif terhadap total keseluruhan data yang diuji.

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

Precision digunakan untuk mengukur tingkat ketepatan model dalam memprediksi kelas positif, yaitu seberapa banyak data positif yang berhasil diprediksi dengan benar oleh model. Nilai *precision* dihitung dengan membagi jumlah prediksi positif yang benar (TP) dengan jumlah keseluruhan prediksi positif (TP + FP).

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

Recall digunakan untuk mengukur kemampuan model dalam menemukan atau mengenali seluruh data positif yang sebenarnya. Dengan kata lain, *recall* menunjukkan seberapa banyak data positif yang berhasil diprediksi dengan benar dari seluruh data positif yang ada.

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

F1-score merupakan metrik evaluasi yang digunakan untuk mengukur keseimbangan antara *precision* dan *recall*. Metrik ini berguna ketika dibutuhkan keseimbangan antara ketepatan prediksi positif (*precision*) dan kemampuan model dalam mengenali seluruh data positif (*recall*).

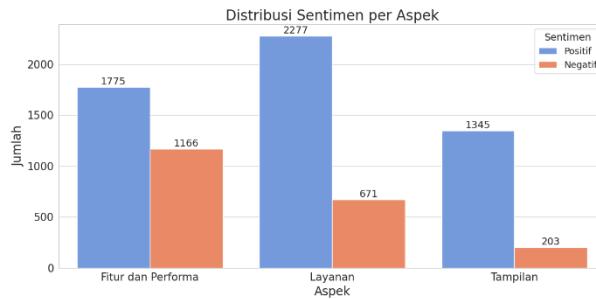
J. Knowledge Discovery

Knowledge discovery didefinisikan sebagai metode untuk menemukan pola yang tersembunyi, belum diketahui sebelumnya, dan berpotensi berguna dari sekumpulan data yang besar [37]. *Knowledge discovery* dan data mining berperan penting dalam mengekstrak informasi dan pola yang tersembunyi dari data yang luas dan kompleks, sehingga dapat mengungkap wawasan baru yang bermanfaat untuk mendukung pengambilan keputusan, inovasi layanan, hingga pengembangan strategi bisnis [38][39][40]. Dalam penelitian ini, proses *knowledge discovery* diterapkan untuk menemukan pola sentimen yang terkandung dalam ulasan pengguna aplikasi BYOND oleh BSI. Melalui pendekatan ini,

data ulasan yang bersifat mentah diolah melalui beberapa tahap menjadi pengetahuan yang lebih bermakna, seperti dominasi sentimen positif dan negatif terhadap fitur tertentu, serta kecenderungan opini pengguna terhadap kinerja aplikasi. Hasil *knowledge discovery* ini diharapkan dapat menjadi bahan evaluasi serta dasar dalam penyusunan strategi peningkatan kualitas layanan mobile banking BYOND by BSI.

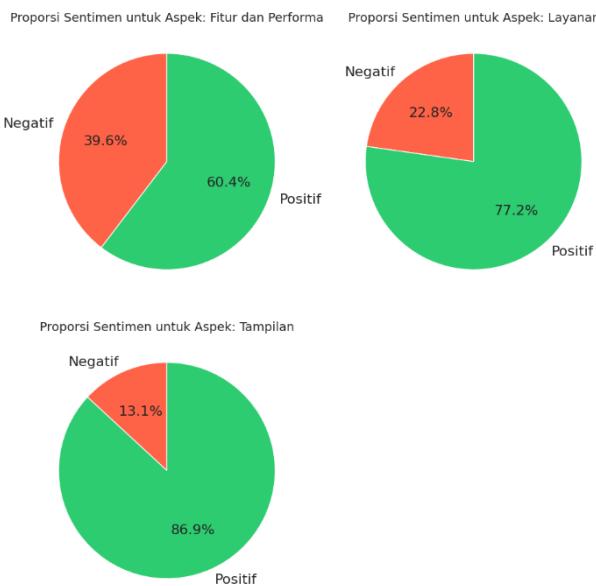
III. HASIL DAN PEMBAHASAN

A. Analisis Sentimen



Gambar 3. DIstribusi Sentimen Setiap Aspek

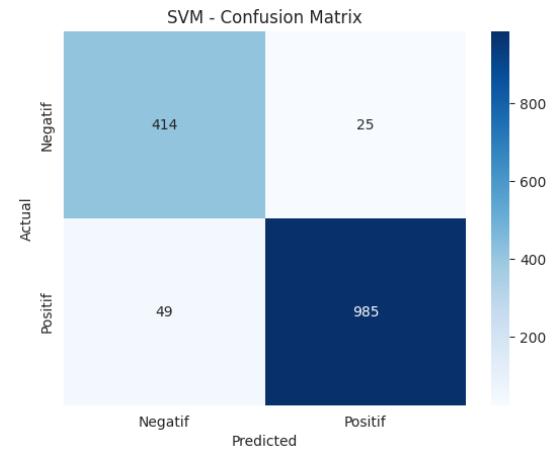
Hasil analisis menunjukkan adanya variasi distribusi sentimen pada ketiga aspek. Berdasarkan gambar 3, aspek fitur dan performa memiliki 1775 ulasan positif dan 1166 ulasan negatif. Aspek layanan mendapatkan 2277 ulasan positif dan 671 ulasan negatif. Aspek tampilan mendapatkan 1345 ulasan positif dan 203 ulasan negatif. Secara keseluruhan, terdapat 5397 ulasan positif dan 2040 ulasan negatif.



Gambar 4. Proporsi Sentimen Setiap Aspek

Gambar 4 menyajikan proporsi sentimen pada masing-masing aspek berdasarkan klasifikasi sentimen positif dan negatif. Berdasarkan hasil proporsi tersebut, isu yang paling dominan dalam ulasan pengguna berkaitan dengan aspek fitur dan performa. Hal ini ditunjukkan oleh tingginya proporsi sentimen negatif pada aspek tersebut, yaitu sebesar 39.6%, yang merupakan persentase tertinggi dibandingkan dengan aspek layanan (22.8%) dan tampilan (13.1%). Hal ini mengindikasikan bahwa ulasan pengguna paling banyak menyoroti masalah yang berkaitan dengan fitur dan performa aplikasi dibandingkan dengan dua aspek lainnya, meskipun secara keseluruhan sentimen positif tetap mendominasi pada ketiga aspek.

B. Hasil Klasifikasi Model SVM



Gambar 5. Confusion Matrix SVM

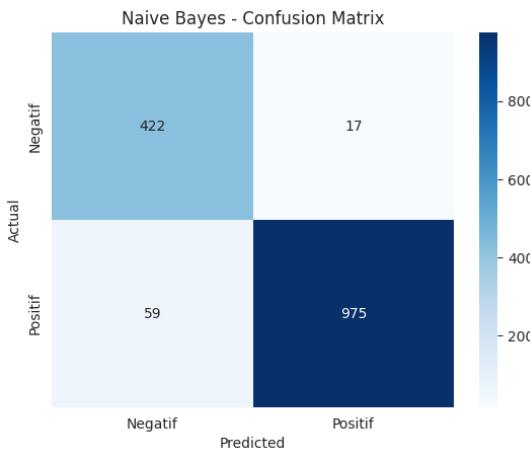
Berdasarkan hasil *confusion matrix*, model SVM berhasil mengklasifikasikan 985 data positif dan 414 data negatif secara benar (*True Positive* dan *True Negative*). Namun, masih terdapat kesalahan klasifikasi berupa 25 data negatif yang keliru diprediksi sebagai positif (*False Positive*) dan 49 data positif yang salah diprediksi sebagai negatif (*False Negative*). Hasil ini menunjukkan bahwa algoritma SVM memiliki kemampuan klasifikasi yang baik, dengan jumlah prediksi benar yang jauh lebih besar dibandingkan jumlah kesalahan klasifikasi. Hal tersebut mengindikasikan bahwa SVM mampu memisahkan kelas sentimen secara cukup optimal pada data yang digunakan.

SVM - Classification Report:				
	precision	recall	f1-score	support
Negatif	0.89	0.94	0.92	439
Positif	0.98	0.95	0.96	1034
accuracy			0.95	1473
macro avg	0.93	0.95	0.94	1473
weighted avg	0.95	0.95	0.95	1473

Gambar 6. Classification Report Model SVM

Secara Keseluruhan, akurasi model SVM mencapai 0,95, yang menunjukkan bahwa model ini cukup andal dalam mengklasifikasikan data sentimen. Kinerja terbaik terlihat pada kelas positif dengan *precision* 0,98, *recall* 0,95, dan *F1-score* 0,96, yang menandakan bahwa prediksi sentimen positif relatif konsisten dan jarang salah. Sementara itu, pada kelas negatif diperoleh *precision* 0,89, *recall* 0,94, dan *F1-score* 0,92, yang menunjukkan kemampuan model dalam mengenali ulasan negatif juga baik, meskipun nilainya sedikit lebih rendah dibanding kelas positif. Nilai *macro average F1-score* sebesar 0,94 dan *weighted average F1-score* sebesar 0,95 mengindikasikan performa model cukup seimbang pada kedua kelas. Temuan ini menegaskan bahwa SVM memiliki performa yang stabil untuk mendukung proses ABSA dalam mengekstraksi opini pengguna berdasarkan aspek, sehingga dapat digunakan sebagai dasar dalam knowledge discovery terkait peningkatan kualitas aplikasi BYOND by BSI.

C. Hasil Klasifikasi Model Naïve Bayes



Gambar 7. Confusion Matrix Naïve Bayes

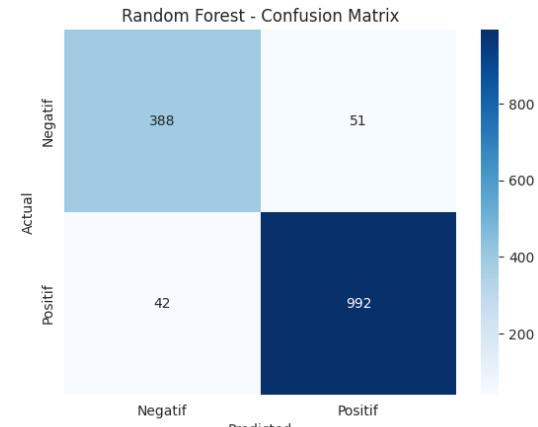
Confusion matrix pada model *Naïve Bayes*, diperoleh hasil klasifikasi yang juga tergolong baik. Model ini berhasil mengklasifikasikan 975 data positif dan 422 data negatif secara benar. Kesalahan klasifikasi yang terjadi relatif lebih rendah, yaitu 17 data negatif yang diprediksi sebagai positif (*False Positive*) dan 59 data positif yang diprediksi sebagai negatif (*False Negative*). Jika dibandingkan dengan model SVM, *Naïve Bayes* menghasilkan jumlah *False Positive* yang lebih kecil, namun memiliki jumlah *False Negative* yang lebih besar. Hal ini menunjukkan bahwa *Naïve Bayes* cenderung lebih berhati-hati dalam memprediksi kelas positif, sehingga masih melewatkannya sebagian data positif yang seharusnya terkласifikasi dengan benar.

Naive Bayes - Classification Report:				
	precision	recall	f1-score	support
Negatif	0.88	0.96	0.92	439
Positif	0.98	0.94	0.96	1034
accuracy			0.95	1473
macro avg	0.93	0.95	0.94	1473
weighted avg	0.95	0.95	0.95	1473

Gambar 8. Classification Report Model Naïve Bayes

Model *Naïve Bayes* juga menunjukkan performa yang baik dengan akurasi keseluruhan 0,95. Pada kelas positif, model menghasilkan *precision* 0,98, *recall* 0,94, dan *F1-score* 0,96, yang menunjukkan bahwa model mampu mengidentifikasi sentimen positif dengan baik. Untuk kelas negatif, nilai *precision* 0,88, *recall* 0,96, dan *F1-score* 0,92 menunjukkan bahwa *Naïve Bayes* cenderung baik dalam menangkap ulasan negatif, meskipun ketepatan prediksi negatifnya sedikit lebih rendah, yang mengindikasikan masih ada sebagian prediksi negatif yang tidak sepenuhnya tepat. Nilai *macro average F1-score* 0,94 dan *weighted average F1-score* 0,95 menunjukkan performa yang relatif seimbang. Jika dibandingkan dengan SVM, performa *Naïve Bayes* mendekati sama dengan SVM, terutama pada *F1-score* kelas positif yang sama-sama tinggi, sehingga model ini dapat menjadi alternatif dalam klasifikasi sentimen berbasis aspek.

D. Hasil Klasifikasi Model Random Forest



Gambar 9. Confusion Matrix Random Forest

Berdasarkan *confusion matrix* pada model *Random Forest*, diketahui bahwa model berhasil mengklasifikasikan 992 data positif dan 388 data negatif secara benar. Namun, kesalahan klasifikasi masih ditemukan, yaitu 51 data negatif yang diprediksi sebagai positif (*False Positive*) dan 42 data positif yang diprediksi sebagai negatif (*False Negative*). Dibandingkan dengan model *Naïve Bayes* dan SVM, *Random Forest* menunjukkan jumlah *True Positive* tertinggi, tetapi juga menghasilkan jumlah *False Positive*

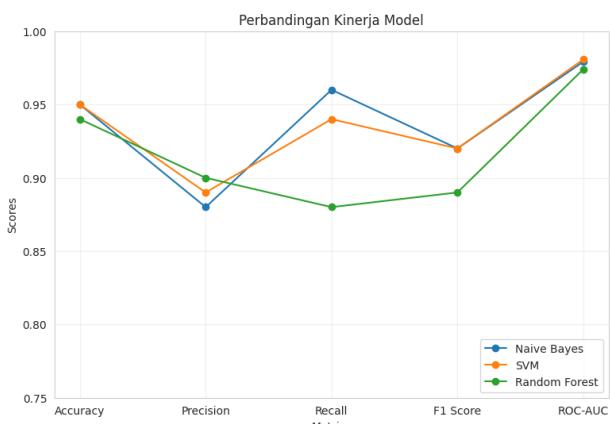
yang lebih besar. Hal ini menunjukkan bahwa meskipun *Random Forest* sangat baik dalam mengenali kelas positif, model ini cenderung lebih sering salah mengklasifikasikan data negatif sebagai positif.

Random Forest - Classification Report:				
	precision	recall	f1-score	support
Negatif	0.90	0.88	0.89	439
Positif	0.95	0.96	0.96	1034
accuracy			0.94	1473
macro avg	0.93	0.92	0.92	1473
weighted avg	0.94	0.94	0.94	1473

Gambar 10. Classification Report Model Random Forest

Random Forest menghasilkan akurasi keseluruhan yang sedikit lebih rendah, yaitu 0,94. Pada kelas positif, model menunjukkan performa yang sangat baik dengan *precision* 0,95, *recall* 0,96, dan *F1-score* 0,96, menandakan bahwa model efektif dalam mengenali sentimen positif. Namun, kelemahan utama terlihat pada kelas negatif, dengan *precision* 0,90, *recall* 0,88, dan *F1-score* 0,89, yang menunjukkan bahwa masih terdapat ulasan negatif yang tidak berhasil teridentifikasi dengan baik dibandingkan dua model lainnya. Nilai *macro average F1-score* sebesar 0,92 dan *weighted average F1-score* sebesar 0,94 mengindikasikan bahwa performa antar kelas pada *Random Forest* kurang seimbang, terutama karena penurunan pada metrik kelas negatif. Hal ini menunjukkan bahwa *Random Forest* cenderung lebih optimal pada kelas dominan (positif), tetapi kurang sensitif dalam menangkap keluhan atau sentimen negatif.

E. Perbandingan Kinerja Model



Gambar 11. Perbandingan Kinerja Model

Berdasarkan perbandingan metrik evaluasi, SVM menunjukkan kinerja terbaik dengan ROC-AUC tertinggi (0.9807) serta nilai akurasi dan *F1-score* sebesar 0.95 dan 0.92, yang menandakan kemampuan serta kestabilan model

yang baik. *Naïve Bayes* memiliki performa yang hampir setara dengan SVM, dengan *recall* tertinggi (0.96) dan ROC-AUC 0.9792, namun *precision* yang sedikit lebih rendah menunjukkan kecenderungan kesalahan pada sebagian prediksi positif. Sedangkan *Random Forest* menunjukkan performa relatif lebih rendah dengan ROC-AUC 0.9740 serta penurunan pada *recall* dan *F1-score*. Dengan demikian, meskipun SVM dan *Naïve Bayes* menunjukkan performa yang relatif setara berdasarkan metrik evaluasi, SVM dipilih sebagai model utama dalam klasifikasi sentimen berbasis aspek karena menunjukkan ketebalahan performa yang lebih konsisten serta nilai ROC-AUC tertinggi dibandingkan model lainnya..

F. Knowledge Discovery

Knowledge discovery dalam penelitian ini bertujuan untuk mengekstrak *insight* dari hasil analisis sentimen berbasis aspek (ABSA) pada ulasan pengguna aplikasi BYOND by BSI. Temuan penelitian menunjukkan adanya pola sentimen yang berbeda antar aspek. Aspek fitur dan performa memiliki sentimen negatif paling tinggi dibandingkan dengan dua aspek lainnya dengan total 1166 data, yang menandakan bahwa permasalahan utama pengguna BYOND by BSI berkaitan langsung dengan keandalan sistem dan kelancaran proses transaksi, sebagai fungsi kritis dalam layanan perbankan digital. Untuk mengidentifikasi keluhan yang paling dominan, dilakukan visualisasi *word cloud* berdasarkan ulasan negatif.

Gambar 8 menampilkan kata-kata dominan seperti “lama”, “gagal”, “transaksi”, “masuk (*login*)”, “susah”, dan “error”. Contoh ulasan pengguna yang menegaskan hal ini antara lain: “sistem busuk sering error apalagi kalau saat penting seperti pembayaran yang harus dibayar waktu itu juga”, “bermasalah saat mau login selalu keluar”, serta “susah di buka aplikasi nya sangat lemot transaksi jadi lama kadang gak bisa di buka sama sekali”. Berdasarkan komentar dan kata-kata dominan tersebut, sebagian besar keluhan pengguna berkaitan dengan gangguan pada stabilitas sistem, responsivitas aplikasi, serta mekanisme autentifikasi pengguna.



Gambar 12. *Word Cloud* Ulasan Negatif Pada Aspek Fitur dan Performa

Temuan ini sejalan dengan literatur kualitas layanan digital yang menekankan bahwa *reliability* dan *system*

performance merupakan determinan utama kepuasan pengguna pada aplikasi perbankan digital [41]. Gangguan pada aspek ini berpotensi menurunkan kepercayaan pengguna dan meningkatkan risiko *churn*, meskipun aspek tampilan dan layanan secara umum memperoleh sentimen positif.

Hasil dari *knowledge discovery* ini dapat diterjemahkan secara operasional menjadi prioritas perbaikan sistem, di mana pengembang perlu memfokuskan roadmap pengembangan pada peningkatan stabilitas transaksi, optimasi proses login, serta penanganan error yang bersifat kritis dan berulang. Dengan demikian, ABSA tidak hanya berfungsi sebagai alat analisis sentimen, tetapi juga sebagai mekanisme *early warning system* yang membantu pengembang mengidentifikasi area teknis paling berdampak terhadap pengalaman pengguna.

G. Kontribusi Penelitian

Temuan ABSA menunjukkan bahwa aspek fitur dan performa menjadi area prioritas karena memiliki sentimen negatif tertinggi, dengan kata dominan seperti “error”, “gagal”, “lama”, “login”, dan “transaksi”. Berdasarkan temuan tersebut, pengembang BYOND by BSI dapat menyusun prioritas perbaikan sebagai berikut:

- Mengoptimasi alur login/OTP dan penanganan kegagalan masuk
- Meningkatkan mekanisme transaksi agar tidak mudah gagal/timeout, serta penyajian status transaksi yang transparan (*pending/success/failed*) untuk mengurangi keluhan “gagal” dan ketidakpastian pengguna
- Mengurangi *loading time*, peningkatan respons UI, dan perbaikan *crash* atau *force close* yang tercermin dari kata “lama” dan “error”
- Mengintegrasikan hasil ABSA sebagai *monitoring* berkala sebagai landasan perbaikan ketika frekuensi kata keluhan tertentu meningkat, sehingga ABSA berperan sebagai *early warning* untuk tim pengembang

Pendekatan ini memungkinkan hasil ABSA dimanfaatkan secara langsung sebagai dasar penyusunan prioritas pengembangan aplikasi, dengan mengacu pada pola keluhan pengguna pada aspek yang paling bermasalah.

IV. KESIMPULAN

Hasil analisis sentimen berbasis aspek pada ulasan pengguna BYOND by BSI menunjukkan bahwa secara umum sentimen positif masih mendominasi, namun aspek fitur dan performa menjadi sumber keluhan utama, ditandai oleh kata-kata seperti “lama”, “gagal”, “transaksi”, “masuk (login)”, “susah”, dan “error”. Temuan ini menegaskan bahwa keandalan sistem dan kinerja teknis merupakan faktor krusial yang paling memengaruhi pengalaman pengguna aplikasi perbankan syariah digital. Dari sisi performa model, *Support Vector Machine* (SVM) menunjukkan kinerja paling

konsisten dengan *Naïve Bayes* memiliki performa yang relatif setara, sedangkan Random Forest cenderung lebih rendah dalam stabilitas klasifikasi. Hasil penelitian ini memberikan implikasi praktis berupa prioritas perbaikan pada stabilitas login dan kelancaran transaksi, serta menunjukkan bahwa pendekatan ABSA berbasis *knowledge discovery* dapat menjadi alat pendukung pengambilan keputusan yang efektif dalam peningkatan kualitas layanan digital BYOND by BSI. Meskipun demikian, penelitian ini masih terbatas pada jumlah aspek dan metode ekstraksi berbasis kata kunci, sehingga penelitian selanjutnya disarankan untuk memperluas aspek analisis dan mengeksplorasi pendekatan otomatis berbasis *deep learning* untuk memperoleh insight yang lebih mendalam.

DAFTAR PUSTAKA

- [1] M. Rahman, H. P. Yee, M. A. K. Masud, and M. U. H. Uzir, “Examining the dynamics of mobile banking app. Adoption during the COVID-19 pandemic: A digital shift in the crisis,” *Digit. Bus.*, vol. 4, no. 2, p. 100088, 2024, doi: 10.1016/j.digbus.2024.100088.
- [2] I. A. Firmansyah, R. Yasirandi, and R. G. Utomo, “The influence of efficacy, credibility, and normative pressure to M-banking adoption level in Indonesia,” *Procedia Comput. Sci.*, vol. 197, no. 2021, pp. 51–60, 2021, doi: 10.1016/j.procs.2021.12.117.
- [3] N. S. Kirani, D. Widyaningtyas, E. Handayani, and I. Randikaparsa, “The Influence of Perceived of Usefulness, Ease, Risk and Trust on Continuous Intention in Using BRI Mobile Banking (Study of BRI Mobile Banking Users in Banyumas Regency),” *Asian J. Manag. Anal.*, vol. 4, no. 1, pp. 215–230, 2025, doi: 10.5592/ajma.v4i1.13459.
- [4] Bank Syariah Indonesia, “Tentang Kami.” Accessed: Jul. 05, 2025. [Online]. Available: <https://www.bankbsi.co.id/company-information/tentang-kami>
- [5] Bank Syariah Indonesia, “SuperApp BYOND by BSI Resmi Diluncurkan! Hadirkan Layanan Komprehensif yang Semakin Nyaman & Aman Diakses.” Accessed: Jul. 05, 2025. [Online]. Available: <https://www.bankbsi.co.id/index.php/news-update/berita/superapp-byond-by-bsi-resmi-diluncurkan-hadirkan-layanan-komprehensif-yang-semakin-nyaman-aman-diakses>
- [6] A. Habib, E. Pramana, H. Junaedi, and E. Ronando, “Extending the Expectation Confirmation Model to Examine Continuous Use Mobile Banking: Security, Trust, and Convenience,” *INTENSIF J. Ilm. Penelit. dan Penerapan Teknol. Sist. Inf.*, vol. 9, no. 1, pp. 76–96, 2025, doi: 10.29407/intensif.v9i1.23751.
- [7] C. Rosanti, F. A. Artanto, R. E. Saputra, E. Syariah, U. Muhammadiyah, and P. Pekalongan, “Analisis Sentiment Pengguna Aplikasi Mobile Banking Pada Bank Syariah Dengan Support Vector Regression Sentiment Analysis of Mobile Banking Application User in Sharia Bank Using Support Vector Regression,” vol. 4, no. 8, pp. 341–347, 2024.
- [8] C. Rosanti, F. A. Artanto, and R. E. Saputra, “Perception of user opinions towards Sharia mobile banking applications in Indonesia based on comments on Google Play Store,” *SERAMBI J. Ekon. Manaj. dan Bisnis Islam*, vol. 7, no. 1, pp. 97–108, 2025, doi: 10.36407/serambi.v7i1.1486.
- [9] M. Kumar, L. Khan, and H. T. Chang, “Evolving techniques in sentiment analysis: a comprehensive review,” *PeerJ Comput. Sci.*, vol. 11, pp. 1–61, 2025, doi: 10.7717/PEERJ-CS.2592.
- [10] Y. T. Bau, T. E. Leong, and C. Le Goh, “Sentiment Analysis of E-Wallet Companies: Exploring Customer Ratings and Perceptions,” *J. Logist. Informatics Serv. Sci.*, vol. 10, no. 4, pp.

- [11] 189–205, 2023, doi: 10.33168/JLISS.2023.0413.
A. Alsokkar, M. Otair, H. E. Alfar, A. Y. Nasereddin, K. Aldiabat, and L. Abualigah, “Sentiment analysis for Arabic call center notes using machine learning techniques,” *J. Auton. Intell.*, vol. 7, no. 3, pp. 1–16, 2024, doi: 10.32629/jai.v7i3.940.
- [12] M. Wankhade, A. Chandra, S. Rao, and C. Kulkarni, *A survey on sentiment analysis methods, applications, and challenges / Artificial Intelligence Review*, no. 0123456789, 2022.
- [13] A. R. Putra and D. E. Ratnawati, “Analisis Sentimen Berbasis Aspek pada Aplikasi Mobile Menggunakan Naïve Bayes berdasarkan Ulasan Pengguna Playstore (Studi Kasus : Jconnect Mobile),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 2, pp. 293–300, Apr. 2025, doi: 10.25126/jtiik.2025127556.
- [14] Nurul Hidayati, Faqih Hamami, and Riska Yanu Fa’rifah, “Aspect-Based Sentiment Analysis On FLIP Application Reviews (Play Store) Using Support Vector Machine (SVM) Algorithm,” *J. Informatics Telecommun. Eng.*, vol. 7, no. 1, pp. 183–197, 2023, doi: 10.31289/jite.v7i1.9768.
- [15] E. F. Aprilia, A. A. Arifiyanti, and N. Sembilu, “Aspect-Based Sentiment Analysis on User Perceptions of OVO using Latent Dirichlet Allocation and Support Vector Machine,” *Aviat. Electron. Inf. Technol. Telecommun. Electr. Control.*, vol. 7, no. 2, p. 163, 2025, doi: 10.28989/avitec.v7i2.3035.
- [16] M. J. Prasetyo and I. M. A. Agastya, “Analisis Sentimen Ulasan Aplikasi Perbankan di Google Play Store menggunakan Algoritma Support Vector Machine Sentiment Analysis of Banking Application Reviews on Google Play Store using Support Vector Machine Algorithm,” *Sist. J. Sist. Inf.*, vol. 13, pp. 2386–2400, 2024.
- [17] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, “Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 11, no. 1, pp. 87–97, 2025, doi: 10.25077/teknosi.v11i1.2025.87-97.
- [18] K. L. Tan, C. P. Lee, and K. M. Lim, “A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research,” *Appl. Sci.*, vol. 13, no. 7, 2023, doi: 10.3390/app13074550.
- [19] L. A. Supriyono *et al.*, *Buku Ajar Big Data dan Data Mining: Konsep, Metodologi, dan Aplikasi*. PT. Sonpedia Publishing Indonesia, 2025. [Online]. Available: <https://books.google.co.id/books?id=jNF9EQAAQBAJ>
- [20] M. T. Hidayat and T. Mauritsius, “Perception of the Indonesian Society on the Performance of the Ministry of Investment/BKPM Based on Sentiment Analysis,” *Int. J. Eng. Trends Technol.*, vol. 72, no. 2, pp. 242–253, 2024, doi: 10.14445/22315381/IJETT-V72I2P125.
- [21] S. Akter *et al.*, “a Comprehensive Study of Machine Learning Approaches for Customer Sentiment Analysis in Banking Sector,” *Am. J. Eng. Technol.*, vol. 6, no. 10, pp. 100–111, 2024, doi: 10.37547/tajet/volume06issue10-11.
- [22] S. S. Aljameel *et al.*, “A sentiment analysis approach to predict an individual’s awareness of the precautionary procedures to prevent covid-19 outbreaks in Saudi Arabia,” *Int. J. Environ. Res. Public Health*, vol. 18, no. 1, pp. 1–12, 2021, doi: 10.3390/ijerph18010218.
- [23] A. Sabir, H. A. Ali, and M. A. Aljabery, “ChatGPT Tweets Sentiment Analysis Using Machine Learning and Data Classification,” *Inform.*, vol. 48, no. 7, pp. 103–112, 2024, doi: 10.31449/inf.v48i7.5535.
- [24] I. P. Ariyani, K. D. Tania, A. Wedhasmara, and A. Meiriza, “Ekstraksi Pengetahuan dari Ulasan Aplikasi CapCut Menggunakan Metode Aspect-Based Sentiment Analysis dan Klasifikasi,” vol. 16, no. April, pp. 80–92, 2025.
- [25] G. Zhao, Y. Luo, Q. Chen, and X. Qian, “Aspect-based sentiment analysis via multitask learning for online reviews,” *Knowledge-Based Syst.*, vol. 264, p. 110326, 2023, doi: 10.1016/j.knosys.2023.110326.
- [26] A. F. Inayah, K. D. Tania, A. Wedhasmara, and A. Meiriza, “Knowledge Extraction Using Aspect-Based Sentiment Analysis and Classification Method,” *JEPIN (Jurnal Edukasi dan Penelit. Inform.)*, vol. 10, no. 3, pp. 378–388, 2024.
- [27] Regina, T. H. Saragih, and D. Kartini, “Analisis Sentimen Brand Ambassador Bts Terhadap Tokopedia Menggunakan Klasifikasi Bayesian Network Dengan Ekstraksi Fitur Tf-Idf,” *J. Inform. Polinema*, vol. 9, no. 4, pp. 383–390, 2023, doi: 10.33795/jip.v9i4.1333.
- [28] I. Z. Mustaqim, H. M. Puspasari, A. T. Utami, R. Syalevi, and Y. Ruldeviyani, “Assessing public satisfaction of public service application using supervised machine learning,” *IAES Int. J. Artif. Intell.*, vol. 13, no. 2, pp. 1608–1618, 2024, doi: 10.11591/ijai.v13.i2.pp1608-1618.
- [29] F. H. Rachman and I. Imamah, “Pendekatan Data Science untuk Mengukur Empati Masyarakat terhadap Pandemi Menggunakan Analisis Sentimen dan Seleksi Fitur,” *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 3, p. 492, 2022, doi: 10.26418/jp.v8i3.56655.
- [30] M. Sulistiyo, Y. Pristyanto, S. Adi, and G. Gumelar, “Implementasi Algoritma Synthetic Minority Over-Sampling Technique untuk Menangani Ketidakseimbangan Kelas pada Dataset Klasifikasi,” *Sistemasi*, vol. 10, no. 2, p. 445, 2021, doi: 10.32520/stmsi.v10i2.1303.
- [31] W. Suci and S. Samsudin, “Algoritma K-Nearest Neighbors dan Synthetic Minority Oversampling Technique dalam Prediksi Pemesanan Tiket Pesawat,” *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1775, 2022, doi: 10.30865/mib.v6i3.4374.
- [32] V. Siahaan and R. H. Sianipar, *EMOTION PREDICTION FROM TEXT USING MACHINE LEARNING AND DEEP LEARNING WITH PYTHON GUI*. BALIGE PUBLISHING, 2023. [Online]. Available: <https://books.google.co.id/books?id=Nw1nEAAAQBAJ>
- [33] J. G. Brandão *et al.*, “Optimization of machine learning models for sentiment analysis in social media,” *Inf. Sci. (Ny.)*, vol. 694, no. November 2024, p. 121704, 2025, doi: 10.1016/j.ins.2024.121704.
- [34] Mubaqarah, A. N. Puteri, and A. Sumardin, “Comparison of Random Forest and XGBoost for Diabetes Classification with SHAP and LIME Interpretation,” *JTERA (Jurnal Teknol. Rekayasa)*, vol. 9, no. 2, pp. 121–130, 2024, doi: 10.31544/jtera.v9.i1.2024.121-130.
- [35] A. F. Alshammari, “Implementation of Model Evaluation Using Confusion Matrix in Python,” *Int. J. Comput. Appl.*, vol. 186, no. 50, pp. 42–48, 2024, doi: 10.5120/ijca2024924236.
- [36] H. Swapnarekha, J. Nayak, H. S. Behera, P. B. Dash, and D. Pelusi, “An optimistic firefly algorithm-based deep learning approach for sentiment analysis of COVID-19 tweets,” *Math. Biosci. Eng.*, vol. 20, no. 2, pp. 2382–2407, 2023, doi: 10.3934/mbe.2023112.
- [37] N. Singhal and Himanshu, “A Review on Knowledge Discovery from Databases,” *Lect. Notes Electr. Eng.*, vol. 860, no. January, pp. 457–464, 2022, doi: 10.1007/978-981-16-9488-2_43.
- [38] Z. T. Zafira, K. D. Tania, W. K. Sari, S. Informasi, and U. Sriwijaya, “Sentiment-Based Knowledge Discovery of Wondr by BNI App Reviews Using SVM, KNN, and Naive Bayes for CRM Enhancement,” vol. 9, no. 5, pp. 2498–2508, 2025.
- [39] D. K. Harahap, K. D. Tania, and P. E. Sentyuni, “Topic Mining-Based Knowledge Discovery of User Health Information Needs,” vol. 7, no. 4, pp. 641–653, 2025.
- [40] S. A. Putri, K. D. Tania, and K. Pasemah, “Knowledge Discovery Through Sentiment Analysis and Topic Modeling of BCA Mobile and MyBCA,” vol. 8, no. 2, pp. 669–682, 2025.
- [41] C. Tedjokusumo and W. R. Murhadi, “Customer satisfaction as a mediator between service quality and customer loyalty : a case study of Bank Central Asia,” vol. 27, no. 2, pp. 156–169, 2023, doi: 10.20885/jsb.vol27.iss2.art3.