RABIT : Jurnal Teknologi dan Sistem Informasi Univrab
Volume 11 No. 1 | Januari 2026 : Hal : 847-868

ISSN CETAK : 2477-2062
ISSN ONLINE : 2502-891X

PENERAPAN INDOBERT DAN BERTOPIC DALAM ABSA UNTUK EVALUASI KUALITAS APLIKASI E-GOVERNMENT INDONESIA

[1] Nathanael Abel Adrielvino [2] Afifah Trista Ayunda
[1,2] Sistem Informasi, Sains dan Teknologi, Universitas Pradita
[1,2] Jl. Gading Serpong Boulevard No.1 Tower 1 - Kabupaten Tangerang - Indonesia
E-mail : *nathanael.abel@student.pradita.ac.id, afifah.trista@pradita.ac.id*

ABSTRAK
Penelitian ini mengevaluasi respons pengguna terhadap aplikasi layanan publik pemerintah Indonesia dengan menganalisis ulasan Google Play Store dengan jumlah dataset akhir 7.000 ulasan yang dikumpulkan pada periode 1 Januari–31 Oktober 2025 dan diberi label secara manual. Tujuan utamanya adalah menerapkan *Aspect-Based Sentiment Analysis* (ABSA) berbasis kerangka E-S-QUAL untuk menilai sentimen pengguna pada empat dimensi utama, yaitu efisiensi, pemenuhan layanan, ketersediaan sistem, dan privasi. Dalam proses klasifikasi, model IndoBERT yang telah melalui *fine-tuning* menunjukkan performa terbaik dibandingkan model tradisional (SVM dan *Naïve Bayes*) mencapai akurasi 91%, sehingga dijadikan model utama dalam penelitian ini. Hasil prediksi model kemudian digunakan untuk mengarahkan analisis BERTopic guna mengidentifikasi isu-isu spesifik yang muncul dalam ulasan pengguna. Temuan penelitian menunjukkan bahwa sentimen umum pengguna terhadap aplikasi e-government cenderung negatif (62,71%). Melalui pemetaan E-S-QUAL dan *topic modeling*, ditemukan bahwa dimensi ketersediaan sistem (58,24% negatif) dan pemenuhan layanan (78,32% negatif) menjadi sumber utama ketidakpuasan pengguna, terutama terkait gangguan sistem, layanan yang tidak stabil, serta aspek privasi (91,67% negatif) yang diresahkan pengguna. Sebaliknya, aspek efisiensi didominasi sentimen positif (58.96%). Secara keseluruhan, hasil ini menegaskan perlunya pemerintah untuk memprioritaskan peningkatan stabilitas sistem dan memastikan konsistensi kualitas layanan agar manfaat transformasi digital dapat dirasakan secara optimal oleh masyarakat.

Kata Kunci: IndoBERT, SVM *(Support Vector Machine)*, *Naïve Bayes*, BERTopic, E-S-QUAL.

ABSTRACT
This research examines user responses concerning public service applications from the Indonesian government. The information collected from the Google Play Store between January 1, 2025, and October 31, 2025, final dataset consists of 7,000 reviews that were labeled manually. The primary goal is to apply Aspect-Based Sentiment Analysis (ABSA) via E-S-QUAL to assess sentiment across four dimensions (efficiency, fulfillment, system availability, and privacy). In the classification process, IndoBERT Fine-Tuned model's efficacy was compared to traditional models (SVM and Naïve Bayes, showing superiority by reaching an overall accuracy of 91%, positioning it as a benchmark for all sentiment analysis. The model then directed the BERTopic analysis to pinpoint and explain the specific issues raised by users. The general results indicate that user sentiment towards the applications is predominantly Negative (62.71%). Through the application of topic modeling and E-S-QUAL mapping, it was found that System Availability (58.24% Negative) and Fulfillment (78,32% Negative) are the key contributors to dissatisfaction, arising from particular grievances such as system outages, service disruptions, and Privacy aspect (91.67% Negative) that concerns users. Conversely, the Efficiency aspect emerges as a crucial service strength, garnering a dominant positive reaction (58.96%). Therefore, these findings clearly suggest that the government must prioritize immediate action to improve system stability and ensure consistent delivery of services, thereby realizing the full potential of its digital transformation efforts.

Keyword: *IndoBERT, SVM (Support Vector Machine), Naïve Bayes, BERTopic, E-S-QUAL*

PENDAHULUAN

Transformasi digital telah menjadi fenomena global yang mendorong perubahan signifikan dalam cara pemerintah menyelenggarakan pelayanan publik di berbagai negara [1], [2]. *Electronic government* atau *e-government* dipandang sebagai langkah yang dapat meningkatkan transparansi, efisiensi, aksesibilitas, serta akuntabilitas pelayanan bagi masyarakat yang lebih luas tanpa terbatas ruang dan waktu [3], [4], [5]. Melalui penerapan *e-government*, pemerintah juga dapat mempercepat proses

DOI : https://doi.org/10.36341/rabit.v11i1.7143
Some rights reserved BY-NC-SA 4.0 International License

administrasi (*service delivery*), dan menyediakan layanan yang lebih inklusif serta terbuka dengan warga [6], [7].

Namun, keberhasilan *e-government* tidak hanya diukur dari kualitas atau kecanggihan sistem yang dibangun, tetapi juga dari sisi penerimaan dan penggunaannya secara aktif dan masif oleh masyarakat [5], [8]. Berdasarkan survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia, pengguna internet di Indonesia berada di angka 80,66% pada tahun 2025, dengan mayoritas pengguna aktif berasal dari kalangan usia produktif yang cenderung melek teknologi. Kondisi ini menunjukan bahwa potensi penerimaan terhadap layanan publik digital sangat besar. Meski demikian, tingginya penetrasi internet tidak otomatis menjamin adopsi *e-government yang optimal*. Penerimaan masyarakat terhadap layanan digital pemerintah sangat dipengaruhi oleh kualitas layanan digital *(e-service quality)* seperti kemudahan penggunaan, keamanan, perlindungan data pribadi [6] keandalan sistem, kepercayaan (*trust*) [4], serta bagaimana pengalaman pengguna dalam mengakses portal layanan publik [5]. Terdapat 4 macam dimensi yang memiliki pengaruh terhadap *e-service quality* berdasarkan E-S-QUAL, yaitu *efficiency, availability, fulfillment,* dan *privacy.* Ekspektasi masyarakat akan keandalan dan performa aplikasi pemerintah juga memiliki pengaruh yang sangat besar mengingat digitalisasi layanan yang dilakukan pemerintah menjadi jalan satu-satunya masyarakat untuk mengakses layanan publik [1], [2].

Dalam konteks ini, *e-service quality* menjadi indikator utama dalam keberhasilan *e-government*. Penilaian akan *e-service quality* tidak hanya diukur melalui sisi teknis saja tetapi juga dari pengalaman nyata pengguna layanan. Pelaksanaan *e-government* di lapangan menunjukkan bahwa tidak sedikit pengguna yang mengeluhkan buruknya desain maupun pengalaman penggunaan layanan digital pemerintah terutama pada portal layanan digital pemerintahan daerah [7]. Kritik terhadap sistem yang lambat, tampilan yang membingungkan, hingga sulitnya menemukan informasi yang dibutuhkan muncul di *platform* seperti Google Play Store dimana pengguna aplikasi pemerintah dapat memberikan ulasan langsung terhadap pengalaman mereka dalam menggunakan aplikasi pemerintah. Evaluasi terhadap kualitas layanan digital pemerintah dari sudut pandang pengguna menjadi langkah yang perlu dilakukan untuk mengetahui bagaimana persepsi pengguna terhadap kualitas layanan pemerintah untuk menjadi bahan pengembangan yang menyesuaikan dengan kebutuhan pengguna [9].

Pendekatan *Aspect-Based Sentiment Analysis* (ABSA) dalam penelitian ini digunakan untuk mengolah opini publik secara lebih kontekstual [10]. ABSA memungkinkan adanya pemetaan sentimen berdasarkan aspek tertentu, misalnya tampilan, navigasi, kecepatan, atau kejelasan informasi, sehingga analisis tidak hanya menentukan polaritas positif, negatif, atau netral saja [11]. Pemodelan aspek dilakukan melalui proses *aspect extraction* dengan BERTopic yang merupakan model berbasis *transformer* yang dapat menentukan aspek dalam teks ulasan secara kontekstual dan lebih mendalam [12].

Dalam menentukan sentimen pengguna, penelitian ini memanfaatkan IndoBERT sebagai model utama. IndoBERT merupakan model berbasis arsitektur *transformer* yang secara khusus dilatih untuk Bahasa Indonesia. Penelitian terdahulu yang dilakukan oleh Cut Fiarni dan Christell James Cellose menunjukkan performa IndoBERT yang sangat baik dalam mengklasifikasikan sentimen pada ulasan Google Play Store

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

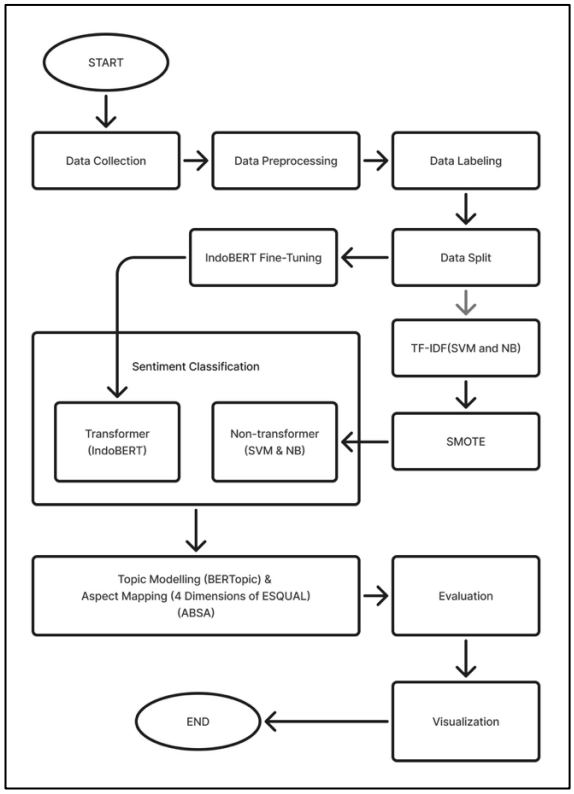
mencapai tingkat *accuracy* 95%, *precision* 94%, *recall* 94% dan F1 *score* sebesar 94% [11]. Penelitian serupa yang dilakukan oleh Andi Rifan Ghaffar Herdany dan Tanty Oktavia, mendapatkan IndoBERT menjadi model yang memiliki nilai akurasi terbaik di 92% dibandingkan dengan 3 model lainnya yang diuji, yaitu *Support Vector Machine* (SVM), *Naive Bayes*, (NB), dan *Random Forest* (RF) [13]. Berdasarkan temuan tersebut, IndoBERT terbukti atas akurasinya yang tinggi dalam memahami konteks semantik Bahasa Indonesia. Namun, untuk memastikan validitas hasil klasifikasi sentimen, penelitian ini juga menggunakan metode klasifikasi tradisional *non-transformer* yang populer di kalangan peneliti seperti *Naive Bayes* dan *Support Vector Machine* (SVM) sebagai *baseline* pembanding [14]. *Naive Bayes* (NB) dipilih karena efisiensinya dalam mengolah jumlah data yang besar dan performanya yang baik [15]. Sedangkan SVM atau *Support Vector Machine* dikenal memiliki stabilitas performa akurasi yang lebih baik [16]. Dengan menjadikan kedua model tersebut sebagai baseline, penelitian ini tidak hanya mengukur performa dari IndoBERT tetapi juga sejauh mana kemampuan model berbasis transformer dengan model *machine learning* tradisional.

Melalui pendekatan ini, penelitian tidak hanya berfokuskan untuk melihat polaritas sentimen pengguna, tetapi juga faktor atau aspek yang mempengaruhi kualitas aplikasi pemerintah Indonesia berdasarkan perspektif pengguna melalui penerapan BERTopic dan dimensi E-S-QUAL. Model IndoBERT digunakan untuk menggali makna dan emosi yang tersirat dalam ulasan pengguna, sementara hasil analisisnya divalidasi menggunakan model pembanding seperti SVM dan *Naive Bayes*. Berbeda dengan penelitian-penelitian sebelumnya yang menilai sentiment dan evaluasi kualitas layanan secara terpisah, penelitian ini mengintegrasikan beberapa metode mencakup ekstraksi topik, *aspect modeling*, klasifikasi sentimen, hingga pemetaannya terhadap E-S-QUAL untuk dapat memberikan hasil yang komprehensif. Pendekatan ini diharapkan dapat memberikan gambaran yang menyeluruh mengenai faktor-faktor yang berperan dalam membentuk *e-service quality* pada implementasi layanan digital pemerintah dari sudut pandang masyarakat pengguna dan hasil analisis ini diharapkan dapat memberikan rekomendasi yang berguna untuk pengambilan keputusan dalam pembuatan kebijakan oleh pemerintah untuk meningkatkan kematangan dan kualitas layanan aplikasi *e-government*.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan menggunakan *natural language processing* (NLP) dalam dalam menganalisa sentimen masyarakat terhadap kualitas *e-service* pada aplikasi layanan digital pemerintah Indonesia. Dengan tujuan utama untuk mengetahui aspek apa saja yang menjadi pengaruh dan sentimen masyarakat dalam menilai *e-service quality* aplikasi pemerintah berdasarkan dimensi E-S-QUAL dari sudut pandang pengguna, penelitian ini menggunakan dua macam pendekatan *machine learning* yaitu dengan berbasis *transformer* (IndoBERT) dan *non-transformer (Support Vector Machine* dan *Naive Bayes)* untuk menganalisa sentimen pengguna. Untuk dapat menghasilkan hasil yang komprehensif, dilakukan serangkaian alur penelitian mulai dari tahap pengumpulan data hingga visualisasi data seperti pada Gambar 1

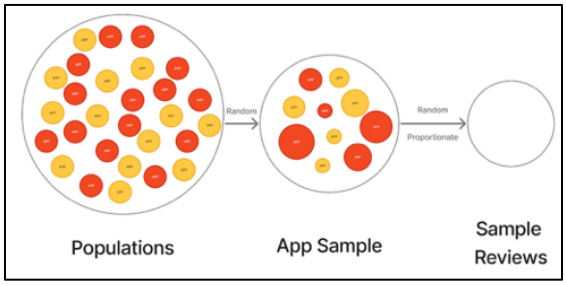
DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino



Gambar 1. Alur Proses Penelitian

Pengumpulan Data dan Persiapan Data

Proses pengumpulan data dalam penelitian ini memanfaatkan scraping data ulasan aplikasi dari Google Play Store. Pemilihan aplikasi dilakukan dengan teknik *cluster sampling* yang memungkinkan adanya dua tahapan dalam menentukan sampel.



Gambar 2. Teknik *Cluster Random Sampling*

Tahap pertama, seperti pada Gambar 2, dilakukan penentuan sampel aplikasi yang akan diteliti dari sebuah populasi dengan jumlah aplikasi yang sama dari masing-masing *cluster*. Kemudian dari masing-masing aplikasi yang terpilih akan diambil data reviewnya secara random dengan pendekatan proporsional agar memberikan hasil yang lebih representatif dan adil dari setiap aplikasi [13].

Cluster aplikasi pada penelitian ini dibagi menjadi dua, yaitu *cluster* aplikasi pemerintah daerah dan pemerintah pusat. Dikarenakan adanya keterbatasan sumber daya dan waktu, penelitian ini dibatasi untuk meneliti 10 aplikasi pemerintah yang terdiri dari 5 aplikasi dari masing masing cluster. Jumlah sampel *review* pun dibatasi pada 7.000 *review dataset* akhir dikarenakan akan dilakukannya proses *manual labeling* yang memakan waktu. Berikut merupakan detail sampel yang digunakan.

Tabel 1. Distribusi Jumlah Sampel Ulasan Proposional

No	Nama Aplikasi	*Cluster*	Jumlah Ulasan pada Google Play Store	Jumlah Ulasan yang Digunakan	Proporsi
1	Sapa Warga- Jabar Super Apps	Pemerintah Daerah	12.100	95	1.36%
2	Salaman Kota Bandung	Pemerintah Daerah	9.980	78	1.12%
3	Tangerang Live	Pemerintah Daerah	14.600	115	1.64%
4	Alpukat Betawi	Pemerintah Daerah	13.000	102	1.46%

DOI : https://doi.org/10.36341/rabit.v11i1.7143
Corresponding Author : Nathanael Abel Adrielvino

5	Serba Digi	Pemerintah Daerah	9.340	74	1.05%
6	Info BMKG	Pemerintah Pusat	82.400	648	9.26%
7	KAI Access	Pemerintah Pusat	399.000	3.138	44.83%
8	Cek Bansos	Pemerintah Pusat	74.400	585	8.36%
9	SIGNAL	Pemerintah Pusat	271.000	2.131	30.45%
10	Siap Kerja	Pemerintah Pusat	4.320	34	0.49%

Data yang telah dikumpulkan kemudian akan melewati tahap *pre-processing data.* Pada tahap ini, data teks ulasan akan melalui proses seperti *case folding,* pembersihan dari karakter non-alfabet, angka, tanda baca dan juga *emoji.* Normalisasi teks juga dilakukan untuk mendapatkan hasil yang lebih akurat dan terstandar. Tokenisasi juga dilakukan untuk keperluan melatih model *non-transformer* seperti SVM dan NB sedangkan model IndoBERT akan menggunakan *tokenizer* bawaan dari IndoBERT.

Tabel 2. Contoh Proses *Data Pre-processing*

Tahapan	Hasil
Data teks mentah	untuk kereta api oke lah , tp untuk Krl 🤦 ribet dan gak jelas !
Case folding	untuk kereta api oke lah , tp untuk krl 🤦 ribet dan gak jelas !
Cleaning	untuk kereta api oke lah tp untuk krl ribet dan gak jelas
Normalization	untuk kereta api oke lah tapi untuk krl ribet dan tidak jelas
Tokenization	["untuk","kereta","api","oke", "lah","tapi","untuk","krl","ribet","dan","tidak","jelas"]
Stopword Removal	["kereta","api","oke","krl","ribet","jelas"]

Setelah diperoleh data yang bersih, *manual labeling* dilakukan pada 7.000 dengan memberi kode angka seperti 0 untuk sentimen negatif, 1 untuk sentimen netral, dan 2 untuk sentimen positif. Kemudian dilakukan proses *data splitting* dengan rasio 70% untuk *training model*, 20% untuk *testing*, dan 10% untuk data validasi model. Pada tahap *splitting* ini juga pembagian dilakukan secara stratifikasi untuk menjaga keseimbangan jumlah sampel di tiap sentimennya.

ABSA (Aspect-Based Sentiment Analysis)

Analisis Sentimen Berbasis Aspek (*Aspect-Based Sentiment Analysis*/ABSA) merupakan bentuk analisis sentimen yang lebih terperinci karena tidak hanya menentukan sentimen keseluruhan dari sebuah teks, tetapi juga berdasarkan aspek atau fitur tertentu dari suatu produk maupun layanan [17]. Tujuan utama ABSA adalah mengidentifikasi opini yang diekspresikan dengan lebih kontekstual, sehingga menghasilkan pemahaman yang lebih mendalam. Pendekatan ini terbukti relevan dan bermanfaat di berbagai sektor, mulai dari industri komersial hingga pemerintahan, karena memungkinkan analisis yang lebih spesifik terhadap pengalaman dan persepsi pengguna [18]. Melalui ABSA, organisasi maupun instansi pemerintah dapat memahami dengan lebih rinci aspek-aspek yang dinilai positif maupun negatif, sehingga hasilnya dapat dijadikan dasar dalam meningkatkan kualitas

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

produk, layanan, maupun kebijakan publik secara lebih tepat sasaran.

IndoBERT dan IndoBERT Fine-Tuning

IndoBERT merupakan sebuah *pre-trained model* berbasis *transformer* yang secara khusus dilatih untuk mengenali konteks semantik Bahasa Indonesia. Dilatih dengan menggunakan kosakata dari kamus WordPiece yang bersumber dari tiga sumber besar seperti Wikipedia (74 juta kata), artikel berita seperti Kompas dan lainnya (55 juta kata), dan juga dari korpus web Indonesia (90 juta kata) membuat IndoBERT dapat memahami berbagai topik bahkan kata-kata yang informal [10]. Dalam penelitian ini, IndoBERT menjadi model utama dalam menganalisa sentimen ulasan aplikasi pemerintah Indonesia. Agar model dapat lebih mengerti topik pada penelitian ini, dilakukan *fine-tuning* dengan melalui proses *hyperparameter tuning* dan didapatkan hasil terbaik sebagai berikut

Tabel 3. *Hyperparameter Tuning*

Hyperparameter	**Detail**
Learning Rate	1E-5
Batch Size	16
Max Length	128
Epoch	3
Optimizer	AdamW

Support Vector Machine dan Naive Bayes

Support Vector Machine (SVM) dan Naive Bayes dalam penelitian ini, menjadi pembanding terhadap model IndoBERT. Kedua model *machine learning* klasik ini melewati proses TF-IDF untuk mengetahui frekuensi dan relevansinya pada setiap kata. Data kemudian melalui proses SMOTE atau *Synthetic Minority Over-Sampling Technique* untuk menangani *imbalanced data* pada proses *training machine learning*-nya. Pada penelitian ini model SVM dan Naïve Bayes menggunakan parameter *standard* sebagai model pembanding umum yang sering digunakan dalam penelitian analisis sentiment dan tidak melaui proses *hyperparameter tuning*. Model SVM akan mencari *hyperplane* yang paling optimal dalam mengklasifikasikan sentimen [19]. Sedangkan model Naive Bayes menggunakan probabilistik dan statistik dalam proses menentukan sentimennya [20], [21]. Pada Naive bayes terdapat persamaan seperti berikut dalam proses klasifikasinya [22].

$$P(C|X) = \frac{P(X|C)\,x\,P(C)}{P(X)} \quad (1)$$

Keterangan:

1. P(C|X): Probabilitas hipotesis (kelas) C diberikan data X (*posterior*)
2. P(X|C): Probabilitas data X diberikan bahwa kelasnya adalah C (*likelihood*)
3. P(C): Probabilitas awal dari kelas C (*prior*)
4. P(X): Probabilitas keseluruhan dari data X (*evidence*)

BERTopic dan Penilaian E-S-QUAL

Analisis topik atau *topic modelling* dilakukan dengan metode BERTopic untuk menghasilkan analisa yang mendalam. Pada penelitian ini, digunakan model *default* BERTopic *multilingual sentence-transfrormer* yang kemudian dimensi yang ditemukan akan direduksi oleh UMAP dengan konfigurasi parameter *n_neighbors = 15, n_components = 5,* dan *min_dist = 0.0* dan proses *cluster reduce embeddings* oleh HDBSCAN dengan parameter *min_cluster_size = 5 dan min_samples = 2*, kemudian c-TF-IDF untuk dapat mengekstrak topik yang ada di dalam setiap ulasan [23]. Dan dalam penilaian sesuai dengan 4 dimensi E-S-QUAL (*efficiency, fullfillment, system availability,* dan *privacy*),

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

hasil *topic extraction* BERTopic akan dipetakan ke 4 dimensi tersebut berdasarkan keterangan item penilaian dalam setiap aspek ke 4 dimensi utama E-S-QUAL seperti berikut [24], [25].

Tabel 4. Item Penilaian Dimensi E-S-Qual

Dimensi	Keterangan Item
Efficiency	EFF1: Halamanya mudah untuk menemukan apa yang dibutuhkan customer.
	EFF2: Halamannya memudahkan pelanggan untuk berpindah kemana saja.
	EFF3: Memungkinkan pelanggan untuk menyelesaikan transaksi dengan cepat.
	EFF4: Informasi yang tertera tersusun dengan baik.
	EFF5: Dapat memuat halaman dengan cepat.
	EFF6: Mudah untuk digunakan.
	EFF7: Memungkinkan pelanggan untuk mengakses dengan cepat.
	EFF8: Halaman tertata dengan baik.
Fulfillment	FUL1: Mengirimkan pesanan sesuai dengan yang dijanjikan.
	FUL2: Penyediaan layanan tepat pada periode yang sesuai.
	FUL3: Mengirimkan pesanan dengan cepat.
	FUL4: Mengirimkan barang yang dipesan.
	FUL5: Memiliki ketersediaan stok yang sesuai dengan klaim perusahaan.
	FUL6: Jujur terhadap penawaran yang diberikan.
	FUL7: Memberikan janji pengiriman produk secara akurat.
System Availability	SYS1: Layanan selalu tersedia.
	SYS2: Layanan dapat diluncurkan dan berjalan dengan segera.
	SYS3:Layanan tidak mengalami *crash*.
	SYS4: Layanan tidak mengalami freeze dalam penggunaan layanan.
Privacy	PR1: Melindungi informasi pelanggan dalam perilaku kegiatan onlinenya.
	PR2: Tidak memberikan informasi pribadi pelanggan ke layanan lain.
	PR3: Melindungi informasi pembayaran pelanggan.

Evaluasi dan Visualisasi

Evaluasi dilakukan dengan menggunakan penilaian yang meliputi nilai *accuracy, precision, recall,* dan *juga* F1 *Score*. Berikut merupakan persamaan-persamaan yang digunakan dalam mengukur nilai-nilai tersebut.

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

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

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

$$F1\ Score = \frac{2\ x\ Precision\ x\ Recall}{Precision+Recall} \quad (5)$$

Keterangan:

1. *True Positive (TP)*
2. *False Positive (FP)*
3. *True Negative (TN)*
4. *False Negative (FN)*

Selain penilaian numerik, hasil penelitian juga ada divisualisasikan dengan beberapa metode untuk memberikan gambaran yang lebih komprehensif dalam penilaiannya. *Confusion matrix* digunakan untuk melihat hasil performa masing-masing model dalam memprediksi sentimen di setiap kelasnya. Sedangkan *word*

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

cloud dimanfaatkan untuk melihat kata-kata yang paling dominan muncul di setiap kelas sentimennya. Pemanfaatan *data visualization* dengan *dashboard* juga dilakukan menggunakan Metabase untuk melihat hasil data yang lebih akurat dan menyeluruh.

HASIL DAN PEMBAHASAN

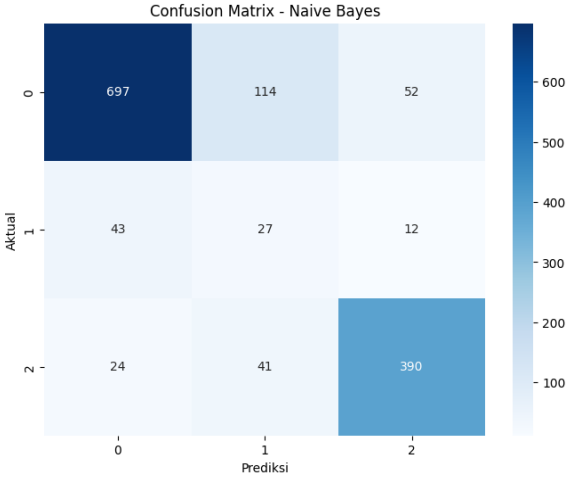
Hasil Klasifikasi Sentimen Naive Bayes

Klasifikasi dengan model *non-transformer* seperti *Naive Bayes* dan SVM akan melewati tahap TF-IDF dan juga SMOTE (*Synthetic Minority Over-Sampling Technique)* sebagai *best practice* dari proses klasifikasi data ulasan. Model diuji dengan 1.400 data yang sama yang digunakan juga untuk menguji model SVM maupun IndoBERT. Secara keseluruhan, model *Naive Bayes* berhasil meraih akurasi hingga 80% yang tergolong baik. Tabel 5 menampilkan hasil evaluasi model *Naive Bayes* mulai dari nilai *precision, recall*, hingga F1-*Score*nya.

Tabel 5. Hasil Evaluasi Model *Naïve Bayes*

	Precision	***Recall***	***F1-Score***	***Support***
Negatif (0)	0.91	0.81	0.86	863
Netral (1)	0.15	0.33	0.20	82
Positif (2)	0.86	0.86	0.86	455
Macro AVG	0.64	0.66	0.64	1.400
Weighted AVG	0.85	0.80	0.81	1.400

Confusion matrix menampilkan bagaimana performa *Naive Bayes* yang sesungguhnya dalam melakukan klasifikasi teks berbahasa Indonesia. Dalam mengolah data dengan sentimen negatif seperti pada Gambar 3, *Naive Bayes* berhasil mendapatkan 697 ulasan dengan tepat. Dalam mengklasifikasikan sentimen netral, model *Naive Bayes* masih mengalami kesulitan dalam keakuratannya. Sedangkan dalam kelas sentimen positif, *Naive Bayes* dapat menentukan sentimen positif dengan baik dan mendapatkan 390 prediksi yang benar.



Gambar 3. *Confusion Matrix* Evaluasi Model *Naive Bayes*

Hasil Klasifikasi Sentimen Support Vector Machine (SVM)

Klasifikasi yang dilakukan dengan menggunakan model SVM menghasilkan klasifikasi lebih akurat jika dibandingkan dengan model *Naive Bayes* dengan akurasi keseluruhan di angka 84%. Berikut ini merupakan tabel hasil evaluasi model SVM.

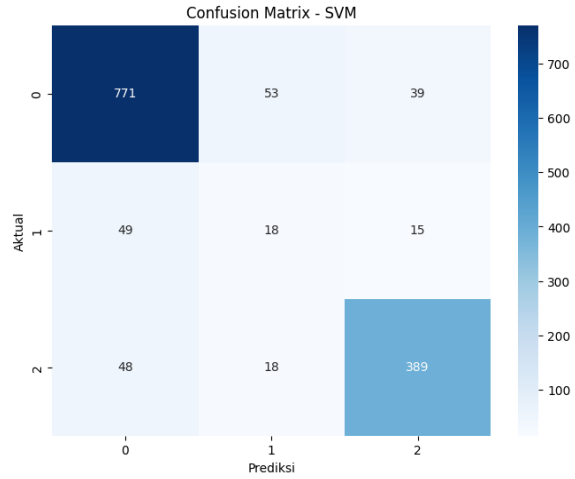
Tabel 6. Hasil Evaluasi Model SVM

	Precision	***Recall***	***F1-Score***	***Support***
Negatif (0)	0.89	0.89	0.89	863
Netral (1)	0.20	0.22	0.21	82
Positif (2)	0.88	0.85	0.87	455
Macro AVG	0.66	0.66	0.66	1.400
Weighted AVG	0.84	0.84	0.84	1.400

Dengan memperoleh rata-rata nilai yang lebih unggul dibandingkan dengan model *Naive Bayes*, hal ini juga tercerminkan pada *confusion matrix* seperti pada Gambar 4 yang

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

menunjukkan jumlah ketepatan prediksi yang lebih tinggi.



Gambar 4. *Confusion Matrix* Evaluasi Model SVM

Penggambaran *confusion matrix* model SVM ini menunjukkan bahwa model SVM dapat menentukan klasifikasi sentimen yang lebih baik pada kelas sentimen positif dan negatif jika dibandingkan dengan model NB. Namun kedua model juga masih mengalami kesulitan dalam menentukan sentimen netralnya.

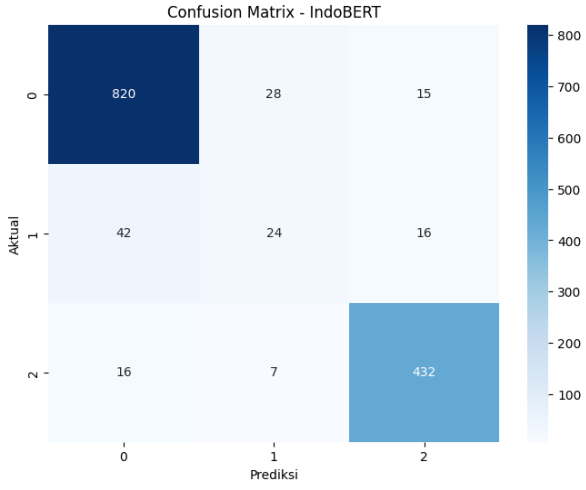
IndoBERT Fine-Tuning dan Hasil Klasifikasi IndoBERT

Pada model IndoBERT, proses *fine-tuning* dilakukan dengan melalui proses *hyperparameter tuning* dengan konfigurasi terbaik seperti yang terdapat pada Tabel 3, dengan *validation accuracy* di angka 91%. Setelah melalui proses *fine-tuning*, model kemudian diuji pada data uji yang sama seperti yang digunakan untuk menguji model NB dan SVM. Berikut merupakan hasil klasifikasi sentimen dengan model IndoBERT.

Tabel 7. Hasil Evaluasi Model IndoBERT

	Precision	***Recall***	***F1-Score***	***Support***
Negatif (0)	0.93	0.95	0.94	863
Netral (1)	0.41	0.29	0.34	82
Positif (2)	0.93	0.95	0.94	455
Macro AVG	0.76	0.73	0.74	1.400
Weighted AVG	0.90	0.91	0.91	1.400

Berdasarkan hasil tersebut, dapat terlihat kemampuan IndoBERT yang sangat baik dalam mengolah teks Bahasa Indonesia. Dalam pengujian ini, diperoleh *overall accuracy* hingga 91% menjadikan IndoBERT unggul jika dibandingkan dengan model NB (80%) dan SVM (84%). Hasil inipun konsisten terlihat pada performanya yang sangat baik dalam mengklasifikasikan sentimen positif dan negatif yang masing-masingnya mencapai nilai *precision* 93%, *recall* 95%, dan F1-*Score* 94%. Sedangkan, kelas sentimen netral mendapatkan skor lebih rendah daripada dua kelas sentimen lainnya seperti yang terjadi pada model SVM dan NB. Hal ini pun juga digambarkan melalui hasil *confusion matrix* yang menunjukan ketepatan yang sangat baik dalam menentukan sentimen positif dan negatif seperti pada Gambar 5.



Gambar 5. *Confusion Matrix* Evaluasi Model IndoBERT

Hasil dari *confusion matrix* menunjukkan bahwa model masih kesulitan

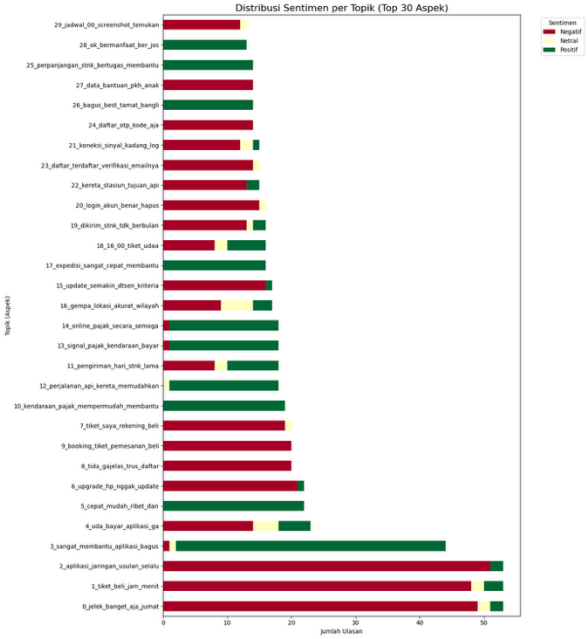
DOI : https://doi.org/10.36341/rabit.v11i1.7143
Corresponding Author : Nathanael Abel Adrielvino

dalam mengklasifikasikan ulasan berpolaritas netral yang sesuai dengan hasil evaluasi nilai *precision*, *recall*, dan F1-*Score* yang lebih rendah daripada kelas sentimen lainnya. Melalui *confusion matrix* juga tercerminkan kemampuan *fine-tuned* IndoBERT yang lebih unggul jika dibandingkan dengan model *non-transformer* seperti SVM dan NB dengan berhasil meraih lebih banyak prediksi yang tepat.

Pada kelas sentimen netral, ketiga model menunjukan performa yang belum maksimal dalam menentukan sentimen netral. Hal ini dapat dilihat dari nilai *support* pada tabel 5-7 yang menampilkan jumlah data netral yang lebih sedikit jika dibandingkan dengan kelas sentiment lainnya sehingga pola sentimen netral lebih sulit untuk dipelajari ketiga model. Faktor lain yang mempengaruhi juga terdapat pada ulasan yang mengandung kombinasi kalimat yang positif dan negatif tanpa adanya penekanan emosi yang jelas. Kondisi tersebut dapat meningkatkan kemungkinan kesalahan klasifikasi pada kelas netral.

BERTopic Extraction Evaluation

Setelah mendapatkan hasil klasifikasi polaritas sentimennya, dipilihlah IndoBERT sebagai acuan sentimen BERTopic sebagai model klasifikasi yang paling akurat dalam penelitian ini. Pemrosesan BERTopic menghasilkan total 86 topik yang teridentifikasi dari 1.400 data testing yang digunakan. Namun agar menghasilkan analisis yang lebih terfokus, penelitian ini hanya mengambil 30 topik teratas yang memiliki nilai representasi yang lebih tinggi. Berikut ini merupakan gambaran distribusi sentimen berdasarkan topiknya.



Gambar 6. Distribusi Sentimen Per Topik (BERTopic)

Analisa topik juga dilakukan di setiap sentimennya untuk mengetahui topik apa yang menjadi pembahasan pengguna untuk analisa yang lebih mendalam. Berikut merupakan hasil ekstraksi topik yang menjadi representasi pengalaman pengguna di sentimen negatif.

Tabel 8. BERTopic *Extraction* (Negatif)

Topik	Representasi	Representasi Ulasan	Interpretasi
di_bisa_tidak_ini	['di', 'bisa', 'tidak', 'ini', 'apk', 'aplikasi', 'mau', 'ada', 'susah','daftar','otp', 'nik', 'ga']	"registrasi sulit banget, kode otp ga masuk masuk ke sms padahal nomor tlp udah bener, tiba aplikasinya log out, terus registrasi lagi notifikasi nomor tlp sudah digunakan. lah kan belom sempet masukin kode otp. terus pas mau coba login ternyata belom terdaftar. gimana sih"	Banyak pengguna pertama tidak dapat mendaftarkan akun karena terbilang NIK/ nomor telepon sudah terdaftar dan banyak juga pengguna yang kesulitan dalam menerima OTP.

tiket_jam_harus_batas	['tiket', 'jam', 'harus', 'antrian', '10', 'beli', 'sebelum', 'menit', 'keberangkatan', 'gabisa']	"kenapa update terbaru pemesanan maximal hanya 4 tiket pemesanan, tidak bisa 10 tiket lagi, jadi kembali ke aplikasi yang dulu kala, yang sekali pesan 10 tiket saja susah dapat tiketnya, apalagi ini harus bolak balik pesan karena maximal hanya 4. Ditambah mau bayar susah, pas isi saldo kaspro harus nunggu 40 menit an, pas masuk saldo gabisa dipake buat bayar tiket, yang harusnya pesen dapat kursi... nunggu saldo masuk sampe kursi pada abis"	Pengguna mengeluhkan maksimal pembelian tiket yang dibatasi ke 4 tidak yang sebelumnya bisa pesan hingga 10 tiket. Pengguna lain juga mengeluhkan masalah pembayaran yang lama hingga menyebabkan keterlambatan keberangkatan akibat menunggu pembayaran berhasil.
jadwal_server_waktu	['00', 'jadwal', 'server', 'waktu', 'pelanggan', 'menunggu', 'time', 'jam', 'habis', 'pesan']	"server lawak, kebanyakan loading, pake biznet 150 mbps aja ga ngangkat, apk suka bug, detik 00.00 langsung klik, pas di dalem daftar penumpang ga nongol, harus ngulang masuk lagi, sudah habis lah tiketnya, sudah 2 hari kejadian seperti ini, penumpang ga muncul, padahal saya akurat pake time.is, lawak lawak server mudah penuh, tiap ganti page selalu lama dan rawan timeout"	Pengguna merasakan menunggu *load* konten yang lama bahkan aplikasi tidak menampilkan informasi apapun *(blank screen)* meskipun sudah memiliki jaringan yang kuat. Banyak juga pengguna yang menunggu hingga pukul 00.00 agar terhindar dari masalah server penuh tetapi masih mengalami masalah yang sama.
aplikasi_usulan_gunakan_bikin	['aplikasi', 'usulan', 'gunakan', 'bikin', 'pemerintah', 'digunakan', 'tapi', 'loading', 'buka', 'selalu']	"bagaimana ini aplikasi pemerintah tapi gabisa digunakan , masa settiap mau mengusulkan muncul pemberitahuan koneksi buruk, padahal saya gunakan untuk youtube an full hd lancar bagaimana bisa di apk ini terjadi seperti itu, masih banyak bugnya tolong diperbaiki supaya segera bisa digunakan masyarakat untuk mengakses"	Pengguna tidak dapat mengakses layanan dari aplikasi dan mendapatkan pesan "jaringan lemah" yang sebenarnya menurut pengguna jaringan yang mereka miliki sudah cepat dan tidak mengalami masalah yang sama di aplikasi lain.
jelek_banget_gabisa_nanya	['jelek', 'banget', 'gabisa', 'nanya', 'webnya', 'bgt',	"dulu bisa bayar lewat indomaret dan linkaja. sekarang udah gabisa. udah topup ovo shopeepay ternyata error mulu. kalo emang gabisa dan harus pake kaipay	Pengguna merasa masalah terbesar berada pada proses pemesanan hingga pembayaran karena prosesnya yang sulit dan

	'aja', 'topup', 'pakek', 'anehhh']	bilang dari awal. terlanjur topup gini terus gimana? balikin aplikasi kai yang lama"	sering gagal dalam proses pembayaran.
update_se makin_ka lah	['update', 'semakin', 'bukannya', 'bulan', 'harapan', 'kalah', 'aplikasi', 'tambah']	"menyesal sekali update aplikasi ini, harapan saat di update bisa lebih baik lagi, namun itu hanya sekedar harapan, semakin di update semakin ga bisa melakukan pembayaran padahal koneksi bagus pame wifi pula. enak pake aplikasi dulu sebelum update please kai perbaiki lagi" dan "tidak bisa muncul aplikasinya.. sdh instal uninstal tetap sama"	Setelah adanya *update*, aplikasi menjadi semakin buruk bahkan tidak bisa diakses. Banyak pengguna yang melakukan uninstall aplikasi dan mengunduhnya kembali dengan harapan dapat menerima layanan dari aplikasi namun hasilnya tetap tidak dapat mengakses layanan dari aplikasi.
pemesana n_tiket_k aipay_sal do	['pemesanan', 'tiket', 'kadaluarsa', 'kecewa', 'nya', 'loket', 'aplikasi', 'hilang', 'saldo', 'kaipay']	"sudah ke 2x nya saya pesan melalui pembayaran qris dan keterangan nya berhasil tapi tiket tidak muncul sama sekali di notifikasi pun juga tidak ada pemberitahuan sama sekali. waktu pengajuan di mintain kode pemesanan dll, secara di notif tidak muncul gimna lihatnya. sekelas bumn harusnya aplikasi nya di perbaharui lagi sistemnya" dan "mengecewakan, topup kaipay masuknya setelah nyaris 3hari, saldo kadang muncul kadang hilang, diminta untuk aktivasi lagi"	Pengguna merasakan kekecewaan karena pembelian tiket kereta api yang dilakukan di aplikasi tidak bisa digunakan karena appnya error sehingga data tidak terbuka, Selain itu juga banyak pengguna yang kecewa dengan KAIPAY dikarenakan sudah melakukan *top up* tetapi tidak masuk di aplikasi sehingga saldonya hilang begitu saja.
kereta_te mpat_sya _kampun g	['kereta', 'tempat', 'sya', 'kampung', 'tujuan', 'lokal', 'api', 'jadwal', 'kerja', 'malem']	"biasanya lancar2 saja. tapi hari ini mau pesan tiket untuk besok, sy pesan pulang pergi.yang jadwal kereta lokal jam 8 malem. itu tidak ada. sedangkan jika pesan nya hanya satu tiket itu kereta yang jam 8 ada. biasanya gak begini.mohon di perbaiki, karena kereta lokal ini sangat penting untuk kita yang kerja cikampek Jakarta"	Banyak pengguna serupa yang memiliki masalah dalam memesan kereta lokal.

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

bayar_samsat_uda_pajak	['bayar', 'samsat', 'uda', 'pajak', 'denda', 'udan', 'validasi', 'gede', 'ga', 'aplikasi']	"pendaftaran pengesahan stnk tidak dapat diproses, perlu validasi nik dan data lainnya ke samsat hadeh sama aj donk klo harus ke samsat langsung ane bayar pak gak perlu daftar online segala mana posisi diluar propinsi mau setor kas negara ni susah amat pak online pak gak perlu ke samsat napa mana aplikasi samsat satunya propinsi nilai 1,9 lagi daftar gak bisa2 hade gak diurus po ma petugasnya sama aja boonk donk"	Pengguna mengeluhkan proses validasi NIK dan KTP yang harus datang langsung ke SAMSAT sehingga merasa aplikasi tidak berguna karena tetap harus pergi ke SAMSAT.
pajak_biaya_rb_bayar	['pajak', 'biaya', 'rb', 'bayar', 'dki', 'lg', 'semahal', 'taat', 'sanksi', 'signal']	"kita kan masyarakat sudah dikenakan berbagai macam pajak salah satu nya pkb. kenapa aplikasi ini dikenakan biaya juga dan mahal lagi 10 rb dan biaya bank nya 10 rb. disaat pemda menggenjot pad dari pajak dan menghapus2in sanksi penunggak pajak. sementara yang taat bayar pajak malah disuruh bayar lebih. ini kan layanan publik sudah seharusnya gratis karena biaya aplikasi dari pajak dan biaya admin bank ga semahal itu. toko online ja biaya layanan 1 rb dan ongkir cuma 7 rb ini pos signal 12.5 rb"	Banyak pengguna yang mengeluhkan tingginya biaya lainnya dalam proses pembayaran pajak melalui aplikasi SIGNAL

Melalui hasil ekstraksi BERTopic pada sentimen negatif, dapat terlihat hasil interpretasinya yang beragam. Pada Tabel 8, terdapat kolom representasi ulasan yang merupakan teks ulasan mentah pengguna. Kelas sentimen ini didominasi dengan keluhan pengguna mengenai masalah pendaftaran akun yang sulit dengan permasalahan yang beragam mulai dari tidak menerima OTP hingga permasalahan NIK ataupun nomor telepon yang telah terdaftar di sistem meskipun pengguna belum pernah mendaftarkan sebelumnya. Kemudian pengguna juga merasa kecewa dengan performa aplikasi yang sering kali mengalami *timeout*, *blank screen*, dan juga respon aplikasi yang lambat terhadap permintaan pengguna. Permasalahan lain seperti kegagalan pembayaran juga menjadi permasalahan penting karena menyangkut kerugian finansial dan waktu pengguna. Namun terdapat juga ulasan pengguna dengan sentimen netral seperti yang terdapat pada Tabel 9.

Tabel 9. BERTopic *Extraction* (Netral)

Topik	Representasi	Representasi Ulasan	Interpretasi
selain_pilihan_ganti_terus	['selain', 'pilihan', 'ganti', 'terus', 'diberikan', 'pertahankan',	"sekedar masukan, mohon ditambahkan pilihan untuk diambil sendiri di kantor samsat terdekat	Pengguna memberikan saran agar metode pengambilan STNK dapat lebih fleksibel.

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

	'indonesia', 'ekspedisi', 'diambil', 'pt']	juga diberikan beberapa pilihan jasa ekspedisi selain pt pos Indonesia"	
tapi_aplikasi_nya_di	['tapi', 'aplikasi', 'nya', 'di', 'tidak', 'tiket', 'bisa', 'verifikasi', 'saya', 'bagus']	"ini aplikasi sudah lumayan bagus tapi kenapa tiba tiba ada login atau register padahal diawal saya sudah login tapi tetap ada kayak gitu entah berapa kali terus sayakan lupa kata sandi terus saya pencet tapi tulisannya proses lupa kata sandi gagal terus saya mau pesan tiket gimana"	Pengguna merasa aplikasinya sudah bagus tetapi terdapat area yang harus diperbaiki seperti proses verifikasi dan daftar yang membingungkan untuk sebagian besar pengguna.
dan_stnk_lebih_bisa	['dan', 'stnk', 'lebih', 'bisa', 'utk', 'biaya', 'mungkin', 'bayar', 'lama', 'hari']	"untuk aplikasinya sih okeee, tapi trnyata seminggu baru nyampe stnk. klo gk sempetnya sehari mending bayar langsung aja, krna online sama aja malah lebih lama, memang sih dapet soft copy nya, tpi better bayar lngsung kalo lama mah"	Banyak pengguna yang merasa pengiriman stnk yang lama tetapi tetap merasa aplikasinya sudah cukup baik dalam merespon permintaan pengguna.
gempa_setelah_notifikasi_gempanya	['gempa', 'setelah', 'notifikasi', 'gempanya', 'dibuka', 'notif', 'lain', 'gk', 'tolong', 'bbrp']	"tolong buatkan suara notifikasi apknya, supaya orang langsung tahu notif gempa secara real time"	Pengguna memberikan masukan agar notifikasi gempa dapat lebih *real-time*.
cuaca_dini_peringatan_widget	['cuaca', 'dini', 'peringatan', 'widget', 'bogor', 'wilayah', 'ke', 'beberapa', 'bisa', 'alamat']	"saran min, di cuaca harusnya bisa masukin alamat lebih dari 1, jadi bisa mantau cuaca beberapa daerah tanpa ganti ganti alamat" dan "tolong untuk aplikasi bmkg tambahkan fitur widget cuaca dalam sehari terdiri dari beberapa jam di jamin oke lah ... kalau sudah saya kasih rating deh ... saya berharap ada yah kak"	Pengguna memberikan masukan agar dapat ditambahkan *widget* cuaca agar tidak perlu membuka aplikasi terus menerus dan ditambahkan untuk melihat cuaca di berbagai alamat.

Pada kelas sentimen netral, banyak pengguna yang merasa aplikasi yang digunakan sudah cukup baik tetapi masih membutuhkan adanya perbaikan atau pembaharuan agar lebih nyaman untuk digunakan. Pengguna dengan ulasan netral didominasi dengan pemberian saran atas fitur yang mereka inginkan seperti metode pengiriman yang lebih fleksibel, notifikasi bencana alam yang *real-time*, penambahan *widget* cuaca, dan juga masukan lainnya.

Representasi ulasan pada Tabel 9, menunjukkan kutipan ulasan dalam kategori sentimen netral dengan pola ulasan

menyampaikan apresiasi yang disertai dengan harapan akan pengembangan fitur. Pola ulasan seperti ini menyebabkan sentimen netral cenderung bersifat ambigu karena terdapat unsur positif dan juga negatifnya yaitu kebutuhan akan perbaikan/pengembangan aplikasi secara bersamaan. Hal ini menjadi tantangan bagi ketiga model klasifikasi sentimen (SVM, Naïve Bayes, dan IndoBERT) dalam menentukan kelas sentimen secara tepat.

Tabel 10. BERTopic *Extraction* (Positif)

Topik	Representasi	Representasi Ulasan	Interpretasi
dan_untuk_aplikasi_mempermudah	['dan', 'untuk', 'aplikasi', 'mempermudah', 'bagus', 'sangat', 'ini', 'cepat']	"terimakasih atas pelayanan sebagus ini admin signal dan tim yang bertugas, sudah 2 kali ini ngurus perpanjangan stnk dan alhamdulillah sampai dengan cepat di antar oleh kurir yang bertugas, sekali lagi terima kasih admin dan seluruh tim yang bertugas"	Pengguna merasa aplikasinya sudah bagus dan cepat dalam merespon pengguna sehingga mempermudah tidak harus langsung datang ke kantor.
tiket_kereta_pesan_api	['tiket', 'kereta', 'pesan', 'api', 'jadwal', 'pemesanan', 'kai', 'hal', 'perjalanan', 'memudahkan']	"aplikasi kai access sangat memudahkan beli tiket kereta api dan sangat mudah untuk melakukan pengecekan jadwal perjalanan kereta api di seluruh indonesia"	Pengguna merasakan aplikasi KAI Access memudahkan dalam pemesanan tiket kereta api dan juga untuk melihat jadwal kereta.
tinggal_lewat_kurir_kalo	['tinggal', 'lewat', 'kurir', 'kalo', 'repot', 'hari', 'rumah', 'tapi', 'sampai', 'dikirim']	"aslinya saya baru tau ada aplikasi sehebat ini coba klo dari dulu dulu mungkin gk akan lagi repot repot ngantri desak desakan buat bayar pajak di samsat,jadi makin tenang bayar pajak tanpa harus repot,cukup duduk santai di rumah dan nunggu cukup 2 3 hari stnk sudah d tangan"	Pengguna merasa terbantu karena dapat dikirim langsung ke rumah.
pajak_bayar_nama_kedepannya	['pajak', 'bayar', 'nama', 'kedepannya', 'bisa', 'gak', 'lembar', 'taunan', 'polisi', 'semoga']	"makin mudah untuk bayar pajak semoga kedepannya di lembar pajak udah langsung ada qrcodnya jadi gak perlu unduh sendiri lagian kan dibuang juga lembar lamanya"	Pengguna merasa dimudahkan dalam proses pembayaran pajak tetapi tetap harus ada pembaharuan agar dapat lebih baik lagi.
semua_sangat_update_aplikasi	['semua', 'sangat', 'update', 'aplikasi', 'lebih', 'aplikasinya',	"semakin di update semakin luar biasa pakai versi lama lebih asyik buat hp jadul saya"	Merasa sangat terbantu khususnya karena aplikasi dapat berjalan

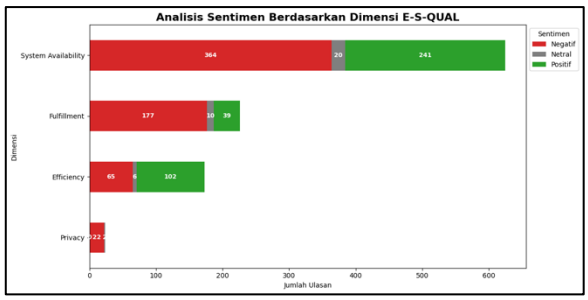
	'membantu', 'bumn', 'asyik']		baik di perangkat lama pengguna.
mudah_simple_simpel_penggunaan	['mudah', 'simple', 'simpel', 'penggunaan', 'dimengerti']	"'simple dan mudah, sangat membantu dalam penggunaan nya"	Aplikasinya *simple*, mudah untuk dioperasikan dan membantu.
hari_lama_baru_kemaren	['hari', 'lama', 'baru', 'kemaren', 'minggu', 'nunggu', 'sih', 'cepet']	"alhamdulillah setelah kejadian kemaren lama yang bikin panik takutnya ga dikirim surat pajak motor nya ternyta cepet sih hari ini baru nyampek aplikasi bantu banget buat orang yang mager kluar kek aku"	Pengguna merasa *service delivery*nya yang cepat dan permintaan bisa diselesaikan dari aplikasi tanpa harus datang yang cocok untuk pengguna yang *introvert* atau sedang malas keluar dan sedang di luar kota.
perjalanan_kereta_nyaman_api	['perjalanan', 'kereta', 'nyaman', 'api', 'lengkap', 'pemesanan', 'fitur', 'khawatir', 'disaat', 'perjalanan']	"perjalanan naik kereta menjadi lebih mudah, lancar, aman dan menyenangkan karena gak khawatir tiket habis kalo beli langsung. utk kedepannya lebih baik lg dlm segala hal maju terus kai"	Mempermudah perjalanan dengan kereta karena tidak khawatir tidak kebagian tiket kereta jika beli langsung di loket.
bagus_aplikasi_informatif_interface	['bagus', 'aplikasi', 'informatif', 'interface', 'gelap', 'aplikasi', 'pajek', 'pembayran', 'pemberitahuan', 'request']	"aplikasi udah bagus, tapi saya ingin request ada mode gelap di aplikasi ini terimakasih"	Aplikasinya sudah bagus, lebih bagus lagi jika ditambahkan adanya *dark mode.*
online_ganti_melalui_jika	['online', 'ganti', 'melalui', 'jika', 'banyak', 'plat', 'juga', 'dana', 'metode', 'usulan']	"sangat mmbantu lah lebih bagus nya lgi sih fitur nya di tambah lagi dong. saya berharap sih fitur untuk pajak 5 tahunan bisa juga melalui online. karna ada juga mau ganti plat tpi kendaraan nya di luar propinsi"	Aplikasi SIGNAL sudah bagus, tetapi lebih bagus lagi jika bisa melakukan pembayaran pajak 5 tahunan dan ganti plat kendaraan.

Topik dengan sentimen positif banyak diberikan oleh pengguna yang merasa terbantu dengan aplikasi layanan digital yang memungkinkan untuk mendapat layanan masyarakat tanpa harus datang ke kantor pemerintahan. Pada Tabel 10, dapat terlihat masyarakat yang puas dan merasa terbantu seperti dalam hal pembayaran pajak dan juga pemesanan tiket kereta api. Selain merasa puas, pengguna juga memberikan masukkan yang membangun seperti fitur-fitur yang dapat

meningkatkan pengalaman pengguna dari aplikasi yang digunakan.

Aspect Mapping E-S-QUAL

Pemetaan aspek dilakukan pada menggunakan 1.400 *data training* untuk menjaga konsistensi sentimen penelitian dengan fokus pada 30 topik teratas menjadi representasi topik yang paling sering dibahas pengguna. Hal ini dilakukan agar hasil pemetaan dapat terarahkan pada topik yang lebih relevan untuk menghasilkan hasil evaluasi yang lebih akurat. Setiap topik dalam urutan 30 teratas dipetakan ke dalam 4 dimensi E-S-QUAL, yaitu *efficiency*, *system availability, fulfillment*, dan *privacy* berdasarkan item keterangan penilaian dimensi E-S-QUAL yang terdapat pada Tabel 4. Setelah 30 topik teratas sudah diklasifikasikan, kemudian *keyword* representasi dari 30 topik teratas itu, diterapkan juga ke seluruh topik untuk hasil penilaian yang lebih menyeluruh. Dengan pendekatan ini, topik yang terdapat *keyword* seperti 'cepat' dan 'mudah' akan masuk ke dalam dimensi *efficiency*, begitu juga dengan *keyword* lainnya di dimensi lain.



Gambar 7. ABSA Hasil *Aspect Mapping* E-S-QUAL

Hasil penilaian E-S-QUAL dengan ABSA melalui *aspect mapping* menunjukkan bahwa *system availability* menjadi aspek utama yang menjadi faktor berkurangnya kualitas *e-service* dari sampel aplikasi yang diteliti terlihat dari jumlah sentimen negatifnya yang tertinggi di angka 58,24% (364 ulasan negatif). Hal-hal yang berkaitan dengan ketersediaan layanan seperti *crash, freeze, service timeout, blank screen* dan hal serupa menjadi *key factors* dalam berkurangnya nilai pada aspek *system availability.* Namun dalam aspek system availability terdapat juga jumlah sentimen positif yang cukup banyak di angka 38,56% (241 ulasan positif) namun keseluruahan sentimen dalam aspek ini masih didominasi oleh sentimen yang negatif. Aspek *fulfillment* juga memberikan kontribusi besar terhadap kurangnya kualitas *e-service* dikarenakan sentimen negatifnya yang lebih dominan di angka 78.32% (177 ulasan negatif). Dapat diinterpretasikan bahwa mayoritas dari aplikasi yang diteliti belum dapat dengan baik memenuhi kebutuhan pengguna terhadap layanan masyarakat digital yang diselenggarakan. Berdasarkan pengujian yang dilakukan, aspek *efficiency* menjadi aspek yang mayoritas sentimen penggunanya condong positif di angka 58.96% (102 ulasan positif). Efisiensi yang positif turut dirasakan pengguna dalam mengakses layanan seperti pembayaran pajak yang cepat tanpa proses yang sulit untuk dimengerti. Namun, aspek *privacy* masih menjadi tantangan bagi pemerintah Indonesia untuk menjaga keamanan data pengguna. Hal ini dikarenakan banyak kasus serupa yang terjadi bahwa baik nomor telepon maupun Nomor Identitas Kependudukan (NIK) pengguna sudah terdaftar pada sistem walaupun pengguna belum pernah mendaftar sebelumnya sehingga muncul kekhawatiran dari pengguna akan keamanan data mereka. Meskipun aspek *privacy* memiliki persentase nilai negatif hingga 91.67%, tetapi aspek *fulfillment* dan *system availability* tetap menjadi pengaruh utama dilihat dari jumlah representasi datanya yang lebih banyak.

Terdapat juga hubungan sebab-akibat terutama pada aspek system availability terhadap aspek fulfilment yang juga terlihat pada representasi ulasan pada Tabel 8. Aspek

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

fulfilment gagal terlihat dari pengguna yang berusaha untuk membeli tiket kereta dengan KAIPAY gagal akibat kurangnya ketersediaan layanan sehingga proses topup saldo memakan waktu 3 hari untuk bisa masuk ke akun pengguna. Kasus serupa juga banyak dialami pengguna dimana aplikasi tidak dapat merespon permintaan OTP untuk verifikasi sehingga pengguna tidak dapat menggunakan layanan digital pemerintah.

Visualization

Untuk mendapatkan gambaran lebih mengenai kata kunci yang sering menjadi pembahasan pengguna, visualisasi menggunakan *word cloud* dilakukan agar dapat merepresentasikan frekuensi kata yang sering muncul di setiap sentimennya. Kata-kata yang tampil dengan ukuran lebih besar menandakan frekuensi munculnya kata pada ulasan semakin sering disebutkan pengguna. Berikut merupakan *word cloud* pada setiap sentimennya.



Gambar 8. *Word Cloud* Sentimen Negatif

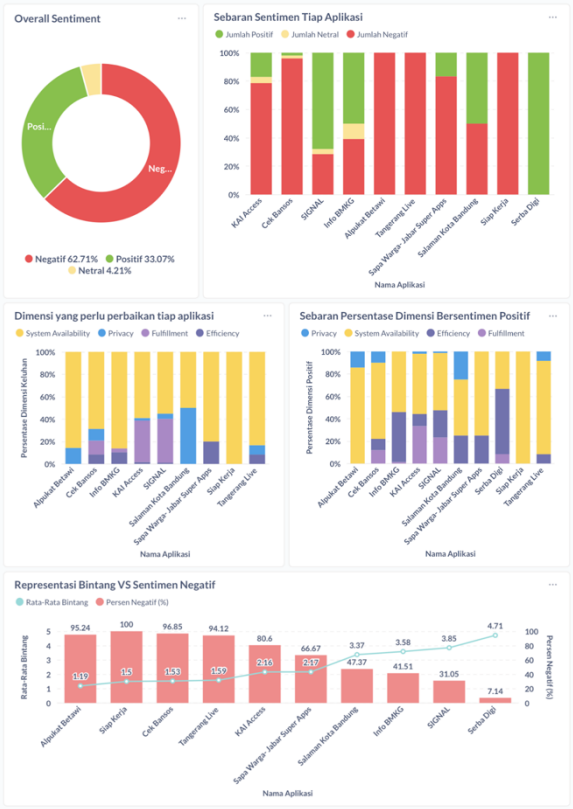


Gambar 9. *Word Cloud* Sentimen Netral



Gambar 10. *Word Cloud* Sentimen Positif

Selain penggambaran dengan *word cloud, data visualization* juga dilakukan dengan menggunakan *dashboard* Metabase. Pada Gambar 11, terdapat *dashboard* yang berisikan beberapa informasi yang lebih mendalam dan komprehensif.



Gambar 11. Dashboard Visualisasi Keseluruhan

Berdasarkan *dashboard* pada Gambar 11, didapatkan sentimen keseluruhan pengguna yang 62.71% negatif. Dengan mayoritas ulasan masyarakat yang negatif, hal ini menunjukkan adanya permasalahan yang membutuhkan perhatian dari pemerintah Indonesia untuk segera diperbaiki. Terutama aplikasi yang representasi datanya memiliki sentimen negatif

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

100% seperti Tangerang Live, Alpukat Betawi dan Siap Kerja. Meskipun begitu, aplikasi lain juga tetap membutuhkan perhatian dari pemerintah untuk dapat memberikan pelayanan digital yang berkualitas tinggi. Sedangkan, dalam melihat sentimen positif pada tiap aplikasinya, Serba Digi berhasil menjadi satu-satunya aplikasi yang memperoleh 100% sentimen positif dalam pengujian ini. Namun, hasil sebaran sentimen tiap aplikasi ini juga dipengaruhi oleh jumlah data representasi proposional tiap aplikasi seperti pada Tabel 1.

Melihat sentimen pengguna berdasarkan bintang yang diberikan tergolong cukup representatif, hal ini dikarenakan semakin banyak bintang yang diberikan, jumlah sentimen negatifnya juga menurun. Secara garis besar, ulasan terhadap *system availability* bersifat negatif (70%), tetapi banyak juga pengguna yang menilai positif tentang *system availability*-nya. Hal ini menunjukkan adanya masalah inkonsistensi dan kompatibilitas antar perangkat dan pengguna yang memerlukan inspeksi lebih dalam agar setiap pengguna memiliki pengalaman pengguna yang baik dalam mengakses layanan digital pemerintah Indonesia.

Persentase aspek *system availability* menjadi masalah utama dan dominan pada 9 aplikasi yang diteliti. Sedangkan pada aspek *privacy*, aplikasi Sapa Warga Jabar Super Apps menjadi aplikasi dengan persentase tinggi yang memerlukan perbaikan jika dibandingkan dengan aplikasi lainnya. Sedangkan, terlihat pada aplikasi Serba Digi, dengan terpenuhinya aspek *fulfilment* dan *system availability*, maka hasilnya juga akan berdampak pada aspek *efficiency* yang meraih sentimen positif

KESIMPULAN DAN SARAN

Hasil dari penelitian yang dilakukan, ditemukan bahwa sentimen pengguna aplikasi layanan publik pemerintah Indonesia yang cenderung negatif dengan persentase sebesar 62.71%. Proses klasifikasi sentimen dilakukan dengan menggunakan 3 model (IndoBERT, SVM, dan NB) dimana ditemukan hasil pengujian dari *Fine-Tuned* IndoBERT memberikan hasil yang lebih unggul mencapai akurasi sebesar 91% jika dibandingkan dengan SVM (84%) dan NB(80%). *Aspect-Based Sentiment Analysis* (ABSA) dilakukan dengan memanfaatkan akurasi IndoBERT yang tinggi dalam menentukan sentimennya, kemudian dilanjutkan dengan analisis pemetaan BERTopic ke dimensi E-S-QUAL topik negatif yang mendominasi meliputi kesulitan pendaftaran akun (masalah NIK/OTP), masalah waktu/kecepatan *(server timeout, blank screen)*, dan kegagalan dalam proses pemesanan/pembayaran (tiket, pajak). Sebaliknya, topik positif difokuskan pada rasa "mudah," "cepat," dan "membantu" dalam mempermudah layanan tanpa harus datang ke kantor, terutama terkait pemesanan tiket kereta api dan pembayaran pajak. Hasil dari pemetaan aspek E-S-QUAL, ditemukan hasil bahwa dimensi *system availability* dan *fulfillment* menjadi area masalah yang paling krusial. Kedua dimensi ini memiliki persentase negatif tinggi (58.24% dan 75.6%), didukung jumlah ulasan yang tergolong cukup banyak yang menunjukkan adanya masalah operasional pada layanan. Meskipun demikian, dimensi *efficiency* menjadi kekuatan utama layanan dengan respons positif dominan (58.96%). Di sisi lain, topik netral didominasi oleh saran perbaikan fitur (misalnya, metode pengiriman yang fleksibel, notifikasi *real-time*). Sementara itu, isu *privacy* adalah isu kualitas yang paling sensitif, mencatat persentase negatif yang sangat tinggi (91.67%) meski memiliki ulasan representatif yang sedikit.

Secara keseluruhan, tantangan utama bagi pemerintah adalah mengubah sentimen negatif pengguna ini menjadi positif. Hal ini

DOI : https://doi.org/10.36341/rabit.v11i1.7143 *Corresponding Author* : Nathanael Abel Adrielvino

memerlukan prosedur yang tepat dan terarah pada akar masalah *system availability* dan *fulfillment* serta peningkatan keamanan data, yang pada akhirnya akan memperkuat kepercayaan pengguna dan mendorong adopsi layanan publik digital yang lebih luas dan berkelanjutan. Pemerintah perlu untuk memastikan setiap pengembangan sistem yang dilakukan dapat menjawab kebutuhan pengguna dan juga memanfaatkan *feedback* pengguna dalam pengambilan keputusan untuk kebutuhan pengembangan yang berkelanjutan. Dengan demikian, proses pengembangan layanan digital pemerintah tidak hanya bersifat *top-down*, tetapi juga melibatkan perspektif Masyarakat sebagai pengguna layanan.

Untuk penelitian selanjutnya, analisis sentimen dapat diperluas dengan menggunakan sumber data dari App Store sebagai representasi pengguna dari pengguna iOS, sehingga cakupan persepsi pengguna terhadap kualitas layanan aplikasi digital pemerintah dapat lebih luas. Selain itu, penelitian mendatang juga dapat mengeksplorasi metode yang lebih efektif dalam menangani distribusi data terutama untuk kelas sentimen netral agar menghasilkan analisis yang lebih akurat. Dengan pengembangan tersebut, penelitian selanjutnya diharapkan dapat memberikan gambaran yang lebih komprehensif terhadap persepsi pengguna serta memperkuat hasil evaluasi terkait kualitas layanan digital pemerintah.

DAFTAR PUSTAKA

[1] M. Adnan, M. Ghazali, and N. Z. S. Othman, "E- participation within the context of e-government initiatives: A comprehensive systematic review," *Telematics. Informatics. Reports.*, vol. 8, p. 100015, Dec. 2022, doi: 10.1016/j.teler.2022.100015.

[2] D. Afrizal and Mu. Wallang, "Attitude on intention to use e-government in Indonesia," *Indonesian. Journal. Electrical. Engineering. Computer. Science.*, vol. 22, no. 1, p. 435, Apr. 2021, doi: 10.11591/ijeecs.v22.i1.pp435-441.

[3] N. Aleisa, "Key factors influencing the e-government adoption: a systematic literature review," *Journal. Innovative. Digital. Transformation.*, vol. 1, no. 1, pp. 14–31, Aug. 2024, doi: 10.1108/JIDT-09-2023-0016.

[4] N. M. Doran *et al.*, "E-Government Development—Effectiveness in the European Union," *Electronics*, vol. 12, no. 3, p. 641, Jan. 2023, doi: 10.3390/electronics12030641.

[5] S. Nookhao and S. Kiattisin, "Achieving a successful e-government: Determinants of behavioral intention from Thai citizens' perspective," *Heliyon*, vol. 9, no. 8, p. e18944, Aug. 2023, doi: 10.1016/j.heliyon.2023.e18944.

[6] A. Kanaan, A. AL-Hawamleh, A. Abulfaraj, H. M. Al-Kaseasbeh, and A. Alorfi, "The effect of quality, security and privacy factors on trust and intention to use e-government services," *International. Journal. Data Network. Science.*, vol. 7, no. 1, pp. 185–198, 2023, doi: 10.5267/j.ijdns.2022.11.004.

[7] Research and Development Agency, Ministry of Home Affairs, Indonesia, S. Aminah, H. Saksono, and Research and Development Agency, Ministry of Home Affairs, Indonesia, "Digital Transformation of the Government: A Case Study in Indonesia," *Jurnal. Komunikasi. Malaysian. Journal. Communication.*, vol. 37, no. 2, pp. 272–288, June 2021, doi: 10.17576/JKMJC-2021-3702-17.

[8] A. Alkraiji and N. Ameen, "The impact of service quality, trust and satisfaction on young citizen loyalty towards government e-services," *Information. Technology. People*, vol. 35, no. 4, pp. 1239–1270, June 2022, doi: 10.1108/ITP-04-2020-0229.

[9] D. P. Faeni, "SERVQUAL measures: Indonesian government healthcare (BPJS) from a human resource perspective," *Journal. Infrastructure. Policy*

Development., vol. 8, no. 2, Dec. 2023, doi: 10.24294/jipd.v8i2.2271.

[10] R. I. Perwira, V. A. Permadi, D. I. Purnamasari, and R. P. Agusdin, "Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content," *Journal. Information. System. Engineering. Business. Intelligence.*, vol. 11, no. 1, pp. 30–40, Mar. 2025, doi: 10.20473/jisebi.11.1.30-40.

[11] C. Fiarni and C. Cellose, "Sentiment Analysis Of Indonesian Video Streaming Application Services Reviews Using Fine-Tuning Indobert And Aspect Modeling," *Journal. Computer. Science. Informatics.*, no. 0, p. 1, 2024, doi: 10.5455/JCSI.20240915104407.

[12] M. Jefri, R. Fauzi, and R. Y. Fa'rifah, "Sentiment-Aware Feature Recommendations for Maternal Mental-Health Apps via IndoBERT and BERTopic on Indonesian TikTok Data," in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, Yogyakarta, Indonesia: IEEE, June 2025, pp. 575–580. doi: 10.1109/ICERA66156.2025.11087326.

[13] A. R. G. Herdany and T. Oktavia, "Analyzing User Experience in Indonesian Investment Platforms: a Hybrid Approach Using Sentiment Analysis and BERTopic Modeling," in *2025 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, Bali, Indonesia: IEEE, July 2025, pp. 144–149. doi: 10.1109/IAICT65714.2025.11100416.

[14] A. Alslaity and R. Orji, "Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions," *Behaviour. Information. Technology.*, vol. 43, no. 1, pp. 139–164, Jan. 2024, doi: 10.1080/0144929X.2022.2156387.

[15] Y. N. Khoiril Umat, D. Rusyda Nafsyi, D. Kusumaningsih, and L. Hakim, "Analisis Faktor Yang Mempengaruhi Pemilihan Gubernur Daerah Khusus Jakarta Menggunakan Algoritma Naive Bayes Dan Regresi Logistik," *Rabit Jurnal. Teknologi. Dan Sistem. Informasi. Univrab*, vol. 9, no. 2, pp. 211–224, July 2024, doi: 10.36341/rabit.v9i2.4778.

[16] D. Mustafa Abdullah and A. Mohsin Abdulazeez, "Machine Learning Applications based on SVM Classification A Review," *Qubahan Acad. Joural.*, vol. 1, no. 2, pp. 81–90, Apr. 2021, doi: 10.48161/qaj.v1n2a50.

[17] C. Hasya Yutika, W. Maharani, and A. Romadhony, "Aspect Extraction on Indonesia Beauty Product Review Using Pre-Trained Language Model," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*, Kuta, Bali, Indonesia: IEEE, July 2024, pp. 219–224. doi: 10.1109/ICoDSA62899.2024.10652031.

[18] R. A. Pranata, I. Hidayah, and S. A. Ishaq Alfarozi, "Aspect Based Sentiment Analysis of PLN Customer Complaints Data Using BERT to Improve Services," in *2024 16th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Bali, Indonesia: IEEE, Oct. 2024, pp. 258–263. doi: 10.1109/ICITEE62483.2024.10808802.

[19] N. Rihaadatul'Aisy and E. W. Pamungkas, "Sentiment Analysis of Indonesian News Texts Using IndoBERT and IndoRoBERTa," in *2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)*, Surakarta, Indonesia: IEEE, June 2025, pp. 1–6. doi: 10.1109/SIML65326.2025.11080939.

[20] Abd. C. Fauzan and K. Hikmah, "Implementasi Algoritma Naive Bayes Dalam Analisis Polarisasi Opini Masyarakat Terkait Vaksin COVID-19," *Rabit Jurnal. Teknologi. Dan Sistem. Informasi. Univrab*, vol. 7, no. 2, pp. 122–128, July 2022, doi: 10.36341/rabit.v7i2.2403.

[21] F. D. Pratama, I. Zufria, and T. Triase, "Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar," *Rabit*

Jurnal. Teknologi. Dan Sistem. Informasi. Univrab, vol. 7, no. 1, pp. 77–84, Jan. 2022, doi: 10.36341/rabit.v7i1.2217.

[22] H. M. Ramdhan, M. Dwifebri Purbolaksono, and B. Bunyamin, "Sentiment Analysis of Beauty Product Reviews Using the IndoBERT Method and Naive Bayes Classification," in *2024 12th International Conference on Information and Communication Technology (ICoICT)*, Bandung, Indonesia: IEEE, Aug. 2024, pp. 397–404. doi: 10.1109/ICoICT61617.2024.10698198.

[23] H. Lee, S. Hong Lee, K. Re Lee, and J. Hyun Kim, "ESG Discourse Analysis Through BERTopic: Comparing News Articles and Academic Papers," *Comput. Mater. Contin.*, vol. 75, no. 3, pp. 6023–6037, 2023, doi: 10.32604/cmc.2023.039104.

[24] R. R. Ahmed, D. Streimikiene, Z. A. Channar, R. H. Soomro, and J. Streimikis, "E-banking Customer Satisfaction and Loyalty: Evidence from Serial Mediation through Modified E-S-QUAL Model and Second-Order PLS-SEM," *Eng. Econ.*, vol. 32, no. 5, pp. 407–421, Dec. 2021, doi: 10.5755/j01.ee.32.5.28997.

[25] M. V. D. Rada *et al.*, "Integrating E-S-QUAL Model and Analytical Hierarchy Process to Assess Consumers' Preferences in Selected Online Shopping Platforms in Manila, Philippines," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Rome, Europe: IEOM Society International, July 2022, pp. 94–106. doi: 10.46254/EU05.20220025.