Terbit *online* pada laman : http://teknosi.fti.unand.ac.id/

Jurnal Nasional Teknologi dan Sistem Informasi

| ISSN (Print) 2460-3465 | ISSN (Online) 2476-8812 |

Artikel Penelitian

Analisis Sentimen Berbasis SHAP untuk Mendukung Perencanaan Strategis Sistem Layanan M-Paspor

Hagi Semara Putera [a], I Made Lanang Putra Pringgadhana [b], Moh. Zulfikar Zauzi [c], Gede Rasben Dantes [d], Gede Indrawan [e], I Made Agus Oka Gunawan [f]

[abcde] *Program Studi Magister Ilmu Komputer, Universitas Pendidikan Ganesha, Jalan Udayana, Singaraja 81116, Indonesia*
[f] *Program Studi Manajemen Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Bali, Badung 80361, Indonesia*

INFORMASI ARTIKEL

Sejarah Artikel:
Diterima Redaksi: 15 Juni 2026
Revisi Akhir: 19 Agustus 2026
Diterbitkan *Online*: 29 Agustus 2026

KATA KUNCI

Analisis Sentimen,
M-Paspor,
XGBoost,
SHAP,
Pelayanan Publik

KORESPONDENSI

E-mail: hagi@student.undiksha.ac.id

A B S T R A C T

Transformasi digital layanan publik melalui aplikasi M-Paspor menghadapi tantangan signifikan terkait stabilitas sistem dan kepuasan pengguna. Analisis sentimen berbasis *machine learning* selama ini mampu mengklasifikasikan ulasan pengguna, namun bersifat *black box* sehingga sulit diketahui faktor spesifik penyebab sentimen negatif, sehingga upaya perbaikan layanan menjadi kurang terarah. Melalui pendekatan *Explainable AI*, metode SHAP diimplementasikan untuk membongkar mekanisme internal algoritma XGBoost, untuk mengidentifikasi kontributor utama sentimen negatif secara visual dan terukur. Penelitian ini mengintegrasikan model XGBoost dengan metode SHAP melalui tahapan pengumpulan data ulasan pengguna, praproses teks, ekstraksi fitur, klasifikasi sentimen, dan interpretasi nilai SHAP, sehingga menghasilkan identifikasi kuantitatif fitur dominan. Hasil algoritma XGBoost menunjukkan akurasi 92%. Kata positif yaitu "lancar" (1,80), "mantap" (1,22), "membantu" (1,02) serta kata negatif "lambat" (-0,77), "mempersulit" (-0,73), "buruk" (-0,43), mengindikasikan adanya kendala pada sistem verifikasi dan alur antarmuka. Temuan ini diformulasikan ke dalam kerangka kerja strategis yang merekomendasikan prioritas perbaikan pada aspek kecepatan respon sistem, penyederhanaan prosedur, dan stabilitas teknis.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi digital dalam berbagai sektor, termasuk layanan publik berbasis mobile maupun website. Pemanfaatan aplikasi mobile memungkinkan masyarakat untuk mengakses layanan secara lebih cepat, efisien dan fleksibel tanpa dibatasi oleh ruang dan waktu. Transformasi digital dalam sektor publik di Indonesia, khususnya pada sistem keimigrasian melalui aplikasi M-Paspor, telah menjadi pilar utama efisiensi birokrasi. Akumulasi ulasan pengguna di *platform* digital menyediakan data yang masif namun tidak terstruktur, sehingga sulit bagi pengambil kebijakan untuk mengidentifikasi prioritas perbaikan secara cepat. Fenomena akumulasi ulasan pengguna di *platform* digital seperti *Google play store* menyediakan sumber data yang sangat kaya akan wawasan, namun memiliki kompleksitas tinggi karena sifat data teks yang tidak terstruktur, ambigu, dan sering kali tidak seimbang secara statistik [1]. Dimana disisi lain hasil dari *review* tersebut mencerminkan pengalaman, persepsi, serta tingkat kepuasan pengguna, sehingga menjadi sumber data yang penting dalam mengevaluasi kualitas layanan digital berbasis pengguna. Seiring dengan meningkatnya jumlah, diperlukan pendekatan yang mampu mengolah data teks secara otomatis dan sistematis. Dalam ranah Perencanaan Strategis Sistem Informasi (PSSI), kegagalan organisasi dalam mengekstraksi nilai strategis dari umpan balik digital ini dapat berujung pada terjadinya misalkulasi dalam penetapan prioritas pengembangan sistem dan alokasi sumber daya TI [2]. Oleh karena itu, diperlukan metodologi *knowledge discovery* yang mampu mentransformasi ulasan mentah menjadi rekomendasi kebijakan yang konkret dan akurat [3].

https://doi.org/10.25077/TEKNOSI.v12i2.2026.340-349

Meskipun penelitian terdahulu telah banyak mengadopsi algoritma *Machine Learning* konvensional seperti *Support Vector Machine* (SVM) untuk klasifikasi sentimen dengan capaian akurasi yang tinggi [1], [2], algoritma tersebut masih menyimpan tantangan fundamental terkait aspek transparansi. Algoritma klasifikasi teks sering kali beroperasi sebagai *black box*, untuk menerima luaran berupa label sentimen tanpa memahami alasan logis atau fitur dominan yang mendasari keputusan algoritma tersebut. Dalam konteks tata kelola layanan publik, keterbatasan ini menjadi hambatan serius karena keputusan strategis seperti peningkatan kapasitas server atau perbaikan alur antarmuka memerlukan dasar argumentasi yang transparan dan dapat dipertanggungjawabkan. Penggunaan algoritma *Extreme Gradient Boosting* (XGBoost) memang telah terbukti unggul dalam menangani data tabular dan teks yang kompleks dengan efisiensi tinggi [4], namun kekuatan prediktif ini tetap memerlukan instrumen tambahan untuk menjembatani celah antara hasil komputasi dan pemahaman manajerial.

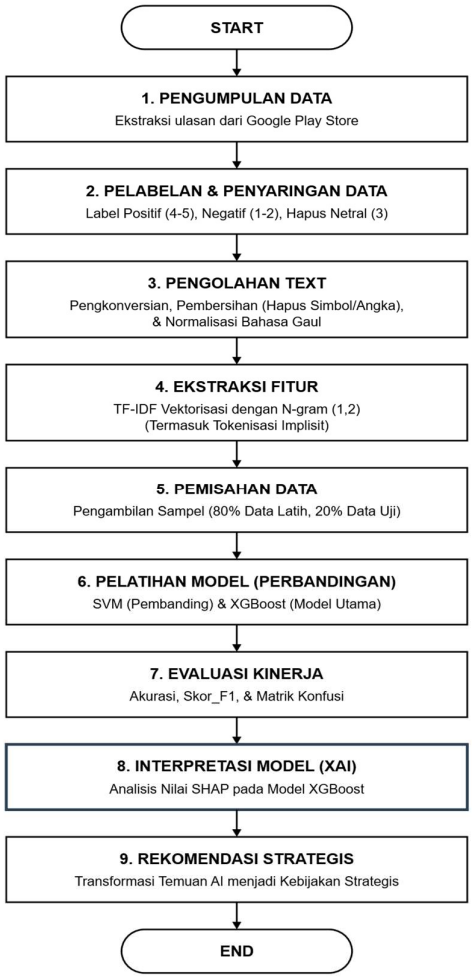
Sebagian besar penelitian dalam analisis sentimen masih berfokus pada peningkatan performa model seperti akurasi, presisi dan *recall* [5], [6]. Pendekatan ini mengandung asumsi bahwa model dengan tingkat akurasi yang tinggi sudah cukup untuk menghasilkan pengaruh yang berkualitas. Padahal, model dengan performa tinggi belum tentu mampu memberikan pemahaman yang mendalam mengenai faktor yang mempengaruhi hasil klasifikasi, sehingga cenderung bersifat *black box.* Keterbatasan tersebut menjadi penting terutama dalam konteks layanan publik yang menuntut transparansi dan akuntabilitas. Kebaruan yang diusulkan dalam penelitian ini terletak pada integrasi metode SHAP untuk membongkar mekanisme internal algoritma XGBoost secara mendalam dan terukur. Berbeda dengan analisis sentimen tradisional yang hanya berfokus pada hasil akhir akurasi, pendekatan *Explainable AI* (XAI) melalui *SHapley Additive exPlanations* (SHAP) memungkinkan visualisasi kontribusi setiap kata kunci secara kuantitatif terhadap pembentukan sentimen pengguna [7].

Tujuan utama dari riset ini adalah untuk membuktikan bahwa transparansi model AI dapat secara signifikan memperkuat proses perencanaan strategis sistem informasi. Dengan memanfaatkan dataset sebanyak 3.500 ulasan yang telah dikurasi, penelitian ini melakukan analisis mendalam terhadap korelasi antara fitur teknis yang muncul dengan penurunan kepuasan publik. Analisis ini sangat krusial bagi pengelola aplikasi M-Paspor untuk bertransformasi dari pola pemeliharaan sistem yang bersifat reaktif menuju perencanaan strategi pengembangan yang proaktif dan presisi. Hasil dari penelitian ini diharapkan dapat menjadi pedoman bagi instansi pemerintah dalam menyusun *roadmap* teknologi yang tidak hanya canggih secara teknis, tetapi juga selaras dengan kebutuhan masyarakat, guna memastikan keberlanjutan kualitas layanan publik digital di Indonesia. Dengan pendekatan ini, penelitian tidak hanya menghasilkan klasifikasi sentimen, tetapi juga mampu mengidentifikasi faktor utama yang mempengaruhi persepsi pengguna.

2. METODE

Penelitian ini menggunakan dataset yang terdiri dari 3.500 ulasan bersih aplikasi M-Paspor yang telah dikurasi sebelumnya. Tahap awal melibatkan ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan rentang n-gram (1,2) untuk menangkap konteks frasa. Dataset dibagi menjadi 2.800 data latih (80%) dan 700 data uji (20%). SVM dan XGBoost, diimplementasikan sebagai pembanding performa. Evaluasi model dilakukan melalui *Confusion Matrix* dan *Classification Report*. Tahap akhir adalah implementasi *TreeExplainer* dari pustaka SHAP pada algoritma XGBoost untuk memvisualisasikan fitur dominan yang memengaruhi sentimen negatif, yang kemudian disintesis menjadi rekomendasi perencanaan strategis.

Penelitian ini mengusulkan penggunaan XAI melalui metode SHAP untuk membongkar *black box* model klasifikasi XGBoost. Kebaruan penelitian ini terletak pada transformasi hasil analisis sentimen menjadi rekomendasi strategis yang transparan, yang memungkinkan instansi terkait melakukan intervensi infrastruktur secara tepat sasaran. Penelitian ini bertumpu pada tiga pilar utama dalam tinjauan pustaka, yaitu: (1) analisis sentimen sebagai metode untuk mengekstrak opini pengguna dari teks ulasan; (2) algoritma XGBoost dan SVM yang terbukti unggul dalam klasifikasi teks; serta (3) SHAP sebagai kerangka kerja XAI untuk menginterpretasi prediksi model. Berikut akan diuraikan secara sistematis konsep dari penelitian ini berdasarkan teori-teori fundamental dan temuan-temuan empiris dari penelitian sebelumnya.



Gambar 1. Alur Penelitian

2.1. Alur Penelitian

Alur penelitian ini mencakup siklus pengolahan data mulai dari tahap *text preprocessing* hingga evaluasi model menggunakan

normalisasi *confusion matrix* untuk mendapatkan hasil yang digunakan pada XAI melalui integrasi metode SHAP.

Gambar 1 merupakan alur penelitian yang dirancang secara sistematis untuk menghasilkan kerangka analisis komprehensif dalam mengkaji sentimen pengguna aplikasi M-Paspor. Tahapan penelitian ini dimulai dari pengumpulan data berupa ulasan pengguna M-Paspor dari *Google play store* melalui proses *scraping,* kemudian dilanjutkan dengan pelabelan dan penyaringan data dengan memberi label sentimen biner, yaitu 0 untuk negatif dan 1 untuk positif, serta menghapus ulasan dengan rating netral agar klasifikasi lebih terfokus. Selanjutnya, data teks diproses pada tahap *text preprocessing* melalui *case folding*, pembersihan simbol dan angka, serta normalisasi kata agar lebih konsisten. Dataset kemudian dibagi pada tahap pemisahan data dengan metode pengambilan sampel menjadi 80% data latih dan 20% data uji agar proporsi kelas tetap seimbang. Pada model training, digunakan dua algoritma yaitu SVM sebagai model dasar dan XGBoost sebagai model utama, yang selanjutnya dievaluasi menggunakan akurasi*,* skor F1 dan *confusion matrix* untuk menilai performa klasifikasi. Model XGBoost kemudian dianalisis pada tahap interpretasi model menggunakan metode SHAP untuk mengidentifikasi fitur atau kata yang paling berpengaruh terhadap hasil prediksi sentimen. Kemudian seluruh hasil analisis diterjemahkan kedalam rekomendasi strategis yang dapat digunakan sebagai dasar perencanaan kebijakan dan pengembangan layanan M-Paspor secara lebih tepat sasaran.

2.2. Data Scraping

Penelitian ini menggunakan pengumpulan data melalui ulasan di *Google play store* pada aplikasi M-Paspor. Pengambilan data menggunakan *Google play scraper* pada server aplikasi M-Paspor (id.go.imigrasi.paspor_online) dengan 3.500 data ulasan bersih [3]. Data diambil pada tanggal 11 April 2026 dengan pembaruan sistem pada versi 7.4.16 pada aplikasi M-Paspor.

2.3. Data Labeling

Pelabelan data pada penelitian ini menggunakan ulasan 1,2 adalah sentimen negatif dan ulasan 4,5 adalah sentimen positif, sementara pada ulasan 3 merupakan netral. Hasil penelitian sebelumnya menunjukkan adanya hubungan antara polaritas sentimen dan keadaan emosional, di mana sentimen positif umumnya merepresentasikan perasaan bahagia dan puas, sementara sentimen negatif mencerminkan emosi seperti marah, sedih, kecewa, dan takut [3].

2.4. Data Preprocessing

Proses *preprocessing* data merupakan tahap krusial karena kualitas data yang buruk akan menghasilkan model *machine learning* yang tidak optimal. Keputusan yang baik selalu didasarkan pada data yang berkualitas [8]. Data yang telah tervektorisasi kemudian dibagi menjadi data latih dan data uji untuk memastikan model memiliki kemampuan generalisasi yang baik dan terhindar dari masalah *overfitting* [9].

2.5. Data Split

Tahap *data split*, dilakukan pembobotan istilah menggunakan metode TF-IDF untuk mengukur tingkat kepentingan setiap kata dalam suatu dokumen dengan mempertimbangkan seberapa sering kata tersebut muncul dalam dokumen tersebut dibandingkan dengan frekuensinya di seluruh dataset. Secara teknis, TF-IDF bekerja melalui dua komponen: *term frequency* (TF) yang menghitung frekuensi kemunculan suatu kata dalam satu ulasan, dan *inverse document frequency* (IDF) yang memberikan bobot lebih rendah pada kata yang muncul di hampir seluruh dokumen karena kurang diskriminatif, serta bobot lebih tinggi pada kata yang jarang muncul namun mencirikan sentimen tertentu [10] Data dikumpulkan sebanyak 3500 ulasan (positif 511, negatif 2989). Selanjutnya, pada data latih dilakukan pembobotan istilah menggunakan metode TF-IDF untuk mengukur tingkat kepentingan setiap kata dengan rumus:

$$W_{i,j} = tf_{i,j} \text{ x } \log\left(\frac{N}{df_i}\right) \quad (1)$$

di mana $tf_{i,j}$ menyatakan frekuensi kemunculan kata i dalam dokumen j, N adalah jumlah total dokumen dalam korpus, dan df_i adalah jumlah dokumen yang mengandung kata i. Setelah proses pembobotan selesai, data kemudian dibagi menjadi data latih (2800 ulasan, 80%) dan data uji (700 ulasan, 20%) yang digunakan untuk menguji kinerja model [11]. TF-IDF relevan digunakan untuk klasifikasi sentimen pada data layanan publik dalam konteks analisis layanan publik di Indonesia [10].

2.6. SHapley Additive exPlanations (SHAP)

Penelitian ini mengintegrasikan pendekatan XAI melalui metode SHAP untuk menginterpretasikan hasil prediksi. SHAP mendasarkan perhitungannya pada konsep Shapley *value* dari *cooperative game theory*, yang memberikan pendekatan interpretatif untuk mengatribusikan tingkat kepentingan setiap fitur terhadap *output* model [9], [12]. Kontribusi setiap kata terhadap prediksi dihitung sebagai selisih antara hasil prediksi model ketika kata tersebut disertakan dan ketika tidak disertakan, sehingga diperoleh nilai kontribusi yang dapat diinterpretasikan secara konsisten [12]. Penggunaan SHAP memungkinkan identifikasi fitur-fitur teks dominan yang memengaruhi sentimen secara transparan, sehingga hasil teknis model tidak lagi bersifat *black box* melainkan dapat dipertanggungjawabkan secara ilmiah dengan *trustworthiness* suatu model AI erat kaitannya dengan interpretabilitasnya, dan SHAP menjadi salah satu pendekatan untuk menjawab kebutuhan tersebut [12]. Pada tahap implementasi, digunakan *TreeExplainer* dari pustaka SHAP, yaitu varian algoritma SHAP yang dioptimalkan khusus untuk model berbasis *decision tree* seperti XGBoost [13]. Metodologi interpretasi ini memungkinkan transformasi temuan data menjadi rekomendasi strategis sistem informasi yang aplikatif bagi pengelola aplikasi M-Paspor [14], [15]. SHAP digunakan untuk menghasilkan dua bentuk interpretasi yaitu, lokal dan global [13]

2.7. Extreme Gradient Boosting (XGBoost)

Algoritma XGBoost dikenal memiliki performa tinggi dalam pengolahan data tabular dan teks, namun bersifat kompleks sehingga sulit diinterpretasikan secara langsung. Secara konseptual, XGBoost bekerja dengan menggabungkan beberapa pohon keputusan yang lemah secara berurutan, di mana setiap pohon baru dibuat untuk memperbaiki kesalahan residual dari pohon-pohon sebelumnya melalui proses iteratif yang meminimalkan fungsi *loss*, sehingga akurasi model meningkat secara bertahap [4], [16], [17]. XGBoost dipilih karena keunggulannya dalam melakukan *ensemble learning* secara iteratif yang terbukti andal dalam memprediksi perilaku

https://doi.org/10.25077/TEKNOSI.v12i2.2026.340-349

pengguna [16], [18]. Alur ini juga mencakup optimasi parameter model melalui teknik *tunning*, penyesuaian parameter yang tepat pada XGBoost dapat meningkatkan nilai evaluasi model secara signifikan dibandingkan dengan penggunaan parameter standar [19]. Berikut adalah fungsi dari algoritma XGBoost.

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (2)$$

Merujuk pada Chen dan Guestrin [9], [10], [16], suku pertama $\sum i\ l(\hat{y}i,yi)$ merupakan fungsi *loss* yang mengukur selisih antara nilai prediksi ($\hat{y}i$) dan label aktual (yi), sedangkan suku kedua $\sum k\ \Omega(fk)$ merupakan komponen regularisasi yang mengendalikan kompleksitas struktur tiap *tree fk* agar model tidak mengalami *overfitting* [9], [16]. Dalam praktiknya, parameter seperti gamma berfungsi mengontrol proses *pruning*, sementara *reg_alpha* memberikan penalti tambahan untuk menekan kompleksitas model [20]. Skor akhir klasifikasi diperoleh dari akumulasi *output* seluruh *tree* yang telah dilatih secara berurutan, sehingga prediksi akhir merupakan hasil kombinasi dari banyak model sederhana yang saling melengkapi [4], [16].

2.8. Support Vector Machine (SVM)

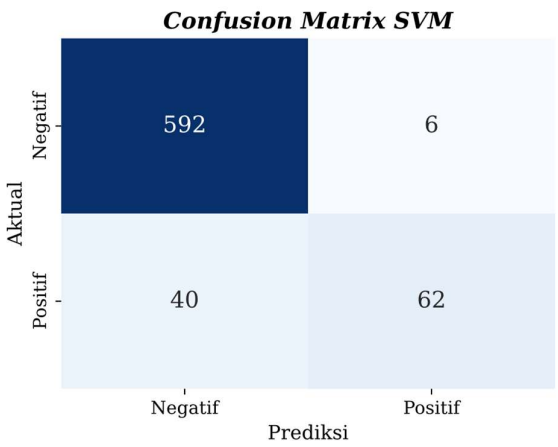
Metode SVM merupakan algoritma klasifikasi yang merepresentasikan setiap contoh teks sebagai titik dalam ruang vektor, dengan memisahkan kategori dengan *hyperplane* yang memiliki margin seluas mungkin [21]. SVM melakukan klasifikasi dengan cara mencari *hyperplane* terbaik untuk memisahkan data ke dalam beberapa kelas sentimen [22]. Pada data yang bersifat tidak *linear*, SVM menggunakan fungsi kernel untuk memetakan data ke ruang berdimensi lebih tinggi agar dapat memisahkan kelas secara optimal, dengan kernel *Radial Basis Function* (RBF) sebagai salah satu yang umum digunakan [22]. Metode SVM digunakan sebagai dasar pembanding. Penggunaan SVM sebagai algoritma dasar didasarkan pada stabilitasnya dalam menangani data dimensi tinggi untuk mengklasifikasikan kepuasan layanan publik [23], [24].

3. HASIL

Berdasarkan metode yang digunakan, bab ini menyajikan hasil pengujian komparatif menggunakan algoritma SVM dan XGBoost dalam mengklasifikasi sentimen pada ulasan M-Paspor. Model XGBoost dipilih untuk menghasilkan nilai SHAP sebagai acuan strategis dalam merumuskan prioritas perbaikan dan pengembangan sistem informasi layanan M-Paspor.

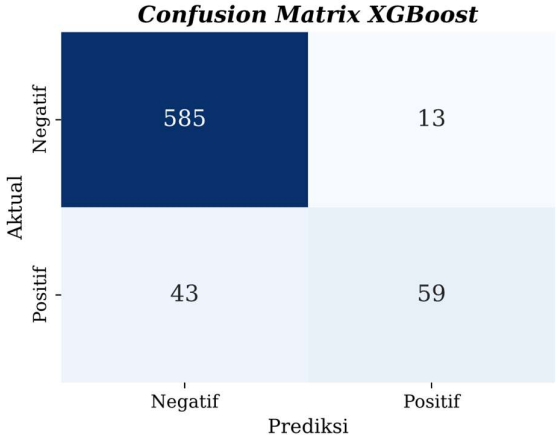
3.1. Pengujian Algoritma SVM dan XGBoost

Pengujian dari kedua algoritma dapat dijadikan pendukung pengambilan keputusan strategis berbasis layanan publik. Berbeda dengan penelitian [25] yang menerapkan SVM pada survei kepuasan layanan publik secara umum tanpa integrasi model *explainable*, penelitian ini menggabungkan hasil klasifikasi SVM/XGBoost dengan interpretasi SHAP untuk menghasilkan peta prioritas perbaikan berjenjang yang secara langsung dapat diadopsi dalam perencanaan sistem informasi instansi. Berikut hasil perbandingan performa dari SVM dan XGBoost yang menjadi landasan dalam perumusan rekomendasi strategi sistem informasi.



Gambar 2. Hasil Normalisasi *Confusion Matrix* SVM

Gambar 2 merupakan hasil dari algoritma SVM dalam mengklasifikasikan sentimen ulasan M-Paspor menunjukkan kemampuan yang stabil dalam memetakan data teks ke dalam ruang dimensi tinggi. Berdasarkan hasil normalisasi *Confusion Matrix*, model ini mampu mengisolasi fitur-fitur linguistik yang kompleks untuk menentukan batas keputusan yang optimal antara sentimen positif dan negatif. Penggunaan SVM pada data berbahasa Indonesia terbukti relevan karena fleksibilitasnya dalam menangani masalah non-linear melalui penggunaan kernel mengenai optimasi parameter untuk sentimen media sosial [22]. Selain itu, reliabilitas SVM dalam survei kepuasan layanan publik telah divalidasi, dimana model ini efektif untuk memberikan hasil klasifikasi yang konsisten pada data opini masyarakat [24].

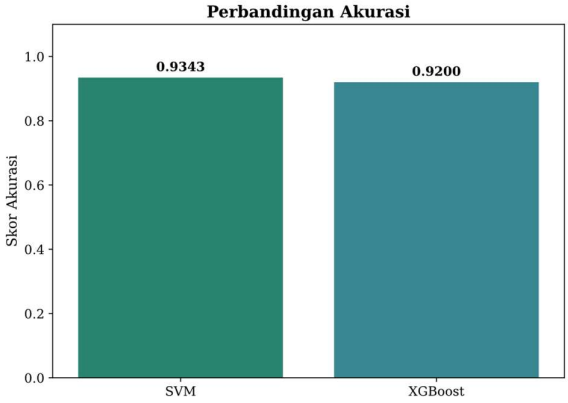


Gambar 3. Hasil Normalisasi *Confusion Matrix* XGBoost

Gambar 3 merupakan hasil dari algoritma XGBoost menunjukkan bahwa XGBoost memiliki tingkat presisi yang tinggi, di mana proses peningkatan secara iteratif mampu memperbaiki kesalahan klasifikasi dari tahapan sebelumnya. Keunggulan XGBoost dalam menghasilkan prediksi yang lebih tajam pada dataset skala besar yang menonjolkan efisiensi algoritma ini dalam mengelola pola data yang kompleks [17]. Stabilitas XGBoost dalam berbagai skenario pengujian menjadikannya instrumen yang kuat untuk mendukung pengambilan keputusan strategis, sejalan dengan metodologi optimasi [20].

3.2. *Perbandingan Akurasi dan Pembobotan Sentiman*

Proses pada penelitian ini melibatkan kinerja matrik yang terdiri dari *precision, recall, f1-score* untuk mendapatkan nilai akurasi dan pembobotan dari algoritma SVM dan XGBoost.



Gambar 4. Perbandingan Akurasi

Gambar 4 merupakan perbandingan akurasi antara SVM dengan nilai 0.9343 (93.43%) dan XGBoost dengan nilai 0.92 (92%) dalam mengklasifikasikan sentimen ulasan M-Paspor. Penggunaan SVM pada data teks ulasan terbukti mampu menghasilkan batas keputusan yang stabil meskipun menghadapi dimensi fitur yang tinggi yang menekankan bahwa optimasi parameter pada SVM sangat efektif untuk meningkatkan akurasi analisis sentimen pada *platform* media sosial [22]. Selain itu, reliabilitas SVM dalam memberikan tingkat presisi yang konsisten pada data survei layanan publik di Indonesia menunjukkan bahwa algoritma ini memiliki ketahanan yang baik terhadap variansi bahasa nonformal yang sering ditemukan pada ulasan pengguna [24].

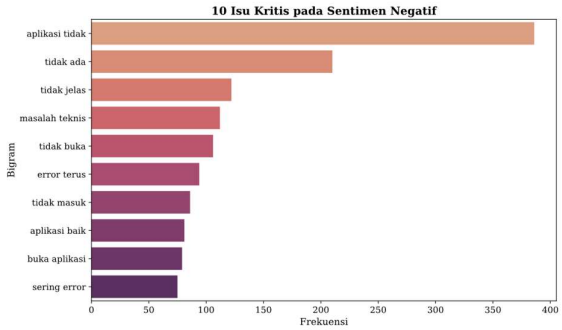
Meskipun XGBoost menunjukkan nilai lebih rendah, efektivitas XGBoost dalam menghasilkan akurasi yang lebih tajam melalui proses peningkatan iteratif yang berhasil mengoptimalkan klasifikasi sentimen dengan menggabungkan XGBoost dan teknik penanganan data [25]. Keunggulan komparatif ini juga diperkuat oleh penelitian lain yang menonjolkan kemampuan XGBoost dalam menangkap pola tren kompleks secara lebih akurat dibandingkan model konvensional [17], [20]. Penerapan ini digunakan untuk pendefinisian kelas target yang menentukan keberhasilan deteksi kredibilitas informasi atau sentimen [26]. Berikut adalah pengukuran sentimen negatif dan positifnya.

Gambar 5. Sentimen Negatif

Gambar 5 merupakan sentimen negatif yang dianalisis sistem. Terdapat kata "buruk", "mempersulit", "lambat" yang muncul dan mempengaruhi nilai dari sentimen negatif. Dimana dapat disimpulkan keluhan utama pengguna bukan pada layanan kantor imigrasinya, melainkan pada stabilitas sistem digital dari aplikasi yang tidak bisa diakses dan durasi waktu tunggu.

Gambar 6. Sentimen Positif

Gambar 6 merupakan sentimen positif yang dianalisis sistem. Terdapat kata "mantap", "membantu", "lancar" yang muncul pada sentimen positif. Ini menunjukkan bahwa jika aplikasi berjalan lancar, pengguna sebenarnya merasa prosedur imigrasi menjadi jauh lebih efisien.



Gambar 7. Bigram

Gambar 7 merupakan hasil bigram dari sistem dimana kata "tidak ada" dan "aplikasi tidak" adalah pasangan kata yang paling sering muncul yang menunjukkan bahwa kendala utama pengguna M-Paspor berpusat pada kegagalan operasional sistem. Selain itu, munculnya kata "aplikasi baik" dan "buka aplikasi" mempertegas bahwa masalah aksesibilitas bukan sekadar insiden tunggal, melainkan pola yang masif.

Integrasi visualisasi ini sangat penting, ekstraksi informasi melalui analisis *N-gram* terbukti efektif dalam mendeteksi kredibilitas masalah pada *platform* media sosial atau ulasan daring [26]. Pada hasil visualisasi ini dapat memberikan landasan kuat bagi perumusan rekomendasi strategi perbaikan infrastruktur teknologi pada aplikasi M-Paspor.

Temuan ini memperlihatkan pola yang berbeda dari studi analisis sentimen aplikasi komersial pada umumnya, di mana keluhan pengguna biasanya berpusat pada fitur produk atau proses transaksi [1], [2]. Pada konteks aplikasi layanan publik seperti M-Paspor, hasil kombinasi frekuensi kata dan bigram justru menunjukkan bahwa sumber sentimen negatif dominan berasal dari kegagalan infrastruktur teknis (server tidak responsif, aplikasi tidak dapat diakses), bukan pada kualitas prosedur birokrasi itu sendiri. Temuan ini memberikan gambaran baru bahwa evaluasi kepuasan pengguna pada layanan digital

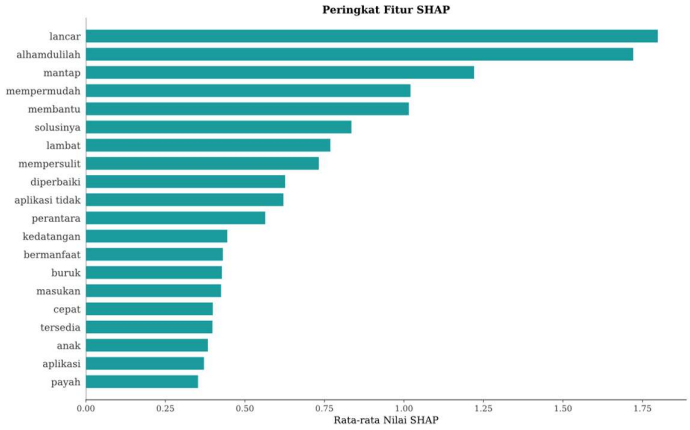
https://doi.org/10.25077/TEKNOSI.v12i2.2026.340-349

pemerintah perlu memisahkan secara eksplisit antara kualitas layanan administratif dan keandalan sistem digital, dua dimensi yang selama ini sering digabung dalam studi sentimen layanan publik konvensional [25].

3.3. Interpretasi SHAP

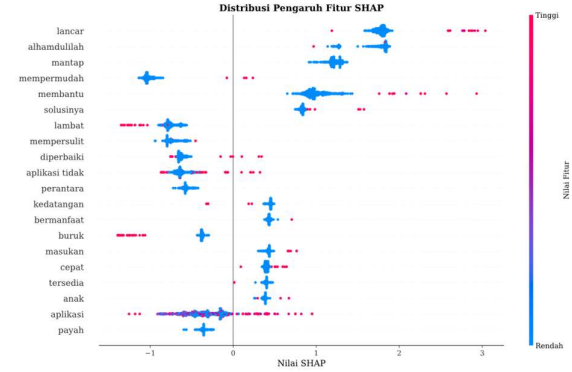
Kontribusi penelitian ini tidak berhenti pada perbandingan angka akurasi antar algoritma, melainkan melanjutkan interpretasi model XGBoost melalui pendekatan *Explainable AI* (XAI) menggunakan SHAP. Pendekatan ini menjawab kesenjangan pada studi-studi analisis sentimen ulasan aplikasi sebelumnya [1], [2], [25] yang umumnya berhenti pada pelaporan metrik akurasi, presisi, dan recall tanpa menjelaskan fitur linguistik spesifik mana yang secara kuantitatif mendorong prediksi ke arah sentimen tertentu, sehingga rekomendasi perbaikan yang dihasilkan cenderung bersifat umum dan tidak berbasis bukti kontribusi fitur. Dalam penelitian ini, peringkat fitur SHAP digunakan untuk melihat fitur secara global yang paling memengaruhi model secara keseluruhan, serta pengaruh fitur SHAP digunakan untuk melihat distribusi dampak fitur yang secara spesifik mendorong prediksi ke arah negatif atau positif. Berikut adalah rumus SHAP yang merupakan dasar untuk menentukan peringkat dan pengaruh SHAP.

$$\phi_i(f,x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)] \quad (3)$$



Gambar 8. Peringkat Fitur SHAP

Gambar 8 merupakan hasil terhadap fitur-fitur yang memengaruhi model klasifikasi dilakukan menggunakan metode SHAP untuk memberikan transparansi pada luaran algoritma XGBoost. Berdasarkan hasil peringkat SHAP, menjelaskan kata kunci yang berpengaruh secara keseluruhan terhadap fitur-fitur teks yang berkaitan dengan aspek layanan diidentifikasi memiliki bobot kontribusi global yang paling signifikan dalam menentukan arah sentimen. Penggunaan visualisasi bar plot ini sangat krusial untuk memetakan urgensi setiap fitur secara objektif [14]. SHAP mampu memberikan penjelasan yang komprehensif pada tingkat fitur dalam model klasifikasi teks. Identifikasi fitur dominan ini menjadi instrument penting bagi manajemen dalam memahami variabel kunci yang memengaruhi persepsi pengguna aplikasi secara sistematis [15].



Gambar 9. Pengaruh Hasil Fitur SHAP

Gambar 9 merupakan hasil fitur pengaruh dari analisis SHAP untuk mengevaluasi pengaruh distribusi nilai fitur terhadap arah prediksi sentimen, baik positif maupun negatif. Visualisasi diatas menunjukkan kata "lambat" merupakan salah satu kata yang mendorong ulasan negatif pada nilai SHAP. Warna merah menunjukkan tingginya pengaruh kata tersebut. Nilai SHAP yang positif atau negatif pada gambar diatas merepresentasikan bagaimana keberadaan suatu variabel mampu meningkatkan atau menurunkan probabilitas klasifikasi pada kelas tertentu [27]. Pendekatan ini memberikan kontribusi pada setiap ulasan yang efektif dalam menangani analisis data berskala besar untuk menghasilkan interpretasi yang dapat dipertanggungjawabkan secara ilmiah [12]. Dengan demikian, integrasi visualisasi peringkat dan pengaruh SHAP ini menyediakan bukti empiris yang kuat bagi PSSI dalam mengoptimalkan layanan M-Paspor berdasarkan ulasan pengguna.

Secara keseluruhan, integrasi antara perbandingan performa algoritma klasifikasi, analisis distribusi kata melalui frekuensi kata dan bigram, serta interpretasi SHAP menghasilkan temuan bahwa kegagalan sistem digital, bukan kualitas prosedur administratif, menjadi determinan utama sentimen negatif pada layanan publik berbasis aplikasi. Temuan ini memperkuat argumen bahwa evaluasi sentimen pada aplikasi layanan pemerintah memerlukan kerangka interpretasi yang berbeda dari aplikasi komersial, sekaligus menunjukkan bahwa pendekatan berbasis SHAP dapat menjembatani kesenjangan antara hasil klasifikasi statistik dan rekomendasi strategis yang dapat ditindaklanjuti secara langsung oleh pengembang sistem dan kontribusi yang belum banyak dieksplorasi pada studi analisis sentimen aplikasi layanan publik di Indonesia.

4. PEMBAHASAN

Berdasarkan hasil penelitian terhadap 3.500 data ulasan yang telah melalui tahap *preprocessing* dan dibagi dengan skema 80:20. Melalui analisis yang dilakukan pada penelitian ini untuk mengolah data, diperoleh performa yang baik untuk mendukung penerapan algoritma XGBoost pada Metode SHAP pada perencanaan strategi sistem informasi di aplikasi M-Paspor.

4.1. Normalisasi Kata

Tabel 1 merupakan tahapan normalisasi ulasan *Google play store* dari 3.500 data, agar nantinya kata yang digunakan dapat memperbaiki analisis sentimen pada ejaan [21]. Berikut tabel normalisasi kata yang tidak baku.

Tabel 1. Sampel Normalisasi Kata Tidak Baku

Kategori	Kata Mentah (*Row Data*)	Kata Normal (*Output*)
Negasi	ngk, ngg, gk, gak, ga, tdk, g	tidak
Aplikasi/Teknis	eror, apk, apl, app, lemot, stuck	error, aplikasi, lambat, berhenti
Keterangan	sgt, bgt, banget, lbh, udh, sdh, mulu	sangat, sekali, lebih, sudah
Kata Ganti	sya, sy, aq, gw, km, lu, lo	saya, kamu
Kata Hubung	yg, utk, dr, krn, kalo, kl, tp	yang, untuk, dari, karena, kalau, tapi

Normalisasi istilah tersebut membantu algoritma XGBoost dalam mengidentifikasi fitur-fitur keluhan utama pengguna secara lebih tajam [25]. Dengan menormalisasi *custom slang* ini, interpretasi fitur melalui nilai SHAP menjadi lebih akurat karena fitur-fitur yang dianggap penting oleh model merujuk pada kosakata yang konsisten dan memiliki makna yang jelas dalam konteks pelayanan publik [14].

4.2. Text Processing

Tabel 2 menjelaskan bagaimana tahapan dari *text processing* dengan menggunakan 4 tahapan. antara lain adalah *Raw Data*, *Case Folding*, *Cleansing* dan *Custom Slang*. Berikut merupakan tabel *text processing* stepnya.

Tabel 2. Step *Text Processing*

Peninjauan	Isi Text	Penerapan
1. *Raw Data*	"Request Time Out mulu, udh coba berbagai solusi. Please improve and fix the bug"	Input dari *Google play store*
2. *Case Folding*	"request time out mulu, udh coba berbagai solusi. please improve and fix the bug"	Konversi kehuruf kecil.
3. *Cleansing*	"request time out mulu udh coba berbagai solusi please improve and fix the bug"	Hilangkan tanda baca (titik, koma)
4. *Custom Slang*	"request time out “terus sudah” coba berbagai solusi please improve and fix the bug"	Normalisasi : “mulu” menjadi “terus”, “udh” menjadi “sudah”.

Tabel 2 diatas menjelaskan bagaimana tahap *text preprocessing* dilakukan secara sistematis untuk mentransformasi data ulasan mentah menjadi format yang optimal bagi model klasifikasi. Proses ini diawali dengan *case folding* untuk menyelaraskan seluruh karakter menjadi huruf kecil, diikuti dengan tahap *cleansing* guna membersihkan *noise* seperti angka, tanda baca, dan karakter non-alfabet lainnya yang tidak merepresentasikan bobot sentimen. Untuk mengatasi hambatan linguistik pada ulasan aplikasi seluler, dilakukan penerapan kamus *custom slang* untuk menormalisasi istilah tidak baku dan singkatan ke dalam bentuk formal Bahasa Indonesia.

4.3. Klasifikasi Hasil SVM dan XGBoost

Klasifikasi hasil ini merupakan hasil dari perhitungan Algoritma SVM dan XGBoost yang memberikan hasil sebagai berikut. Hasil ini digunakan untuk perhitungan menggunakan Metode SHAP sebagai rekomendasi strategis.

Tabel 3. Klasifikasi SVM dan XGBoost

SVM				
	Presisi	*Recall*	Skor F1	*Support*
Negatif	0.94	0.99	0.96	598
Positif	0.91	0.61	0.73	102
Akurasi			0.93	700
Rata-rata makro	0.92	0.80	0.85	700
Rata-rata berat	0.93	0.93	0.93	700
XGBoost				
	Presisi	*Recall*	Skor F1	*Support*
Negatif	0.93	0.98	0.95	598
Positif	0.82	0.58	0.68	102
Akurasi			0.92	700
Rata-rata makro	0.88	0.78	0.82	700
Rata-rata berat	0.92	0.92	0.91	700

Berdasarkan Tabel 3, algoritma SVM mencapai akurasi keseluruhan sebesar 93%, sementara XGBoost mencapai akurasi 92% pada dataset uji yang terdiri dari 700 ulasan (598 negatif dan 102 positif). Meskipun selisih akurasi keduanya tipis, terdapat perbedaan kinerja dalam mengidentifikasi kelas positif. Pada sentimen negatif, kedua algoritma bekerja dengan baik. Algoritma SVM dan XGBoost dengan presisi (0,94 dan 0,93), *recall* (0,99 dan 0,98), dan skor F1 (0,96 dan 0,95). Ini menunjukkan keduanya sangat andal dalam mendeteksi ulasan negatif, yang didominasi oleh kata-kata seperti "buruk," "susah," dan "error,". Pada sentimen positif, SVM sedikit mengungguli XGBoost: presisi (0,91 vs. 0,82), *recall* (0,61 vs. 0,58), dan skor F1 (0,73 vs. 0,68).

Secara keseluruhan, kedua algoritma memberikan kinerja yang sebanding dan sangat baik untuk tugas klasifikasi sentimen pada data yang tidak seimbang. SVM direkomendasikan jika fokus penelitian adalah pada peningkatan presisi dan *recall* untuk kelas positif, sementara XGBoost tetap menjadi alternatif yang kompetitif dengan waktu pelatihan yang lebih cepat.

4.4. Analisis Hasil

Berdasarkan interpretasi algoritma XGBoost menggunakan nilai SHAP pada gambar 7, maka disusunlah rekomendasi strategis yang mengacu pada konsep pengembangan sistem informasi dan prioritas kebutuhan perangkat lunak yang dapat disusun dalam bentuk skala prioritas, yang digunakan untuk menentukan urutan

https://doi.org/10.25077/TEKNOSI.v12i2.2026.340-349

implementasi berdasarkan tingkat kepentingan dan rilis pengembangan [28]. Dibawah ini rekomendasi pengembangan aplikasi M-Paspor adalah sebagai berikut.

Tabel 4. Tabel Rekomendasi Strategis Berbasis Nilai SHAP

Kategori	Prioritas	Target Perbaikan
Jangka Pendek (0-6 Bulan)	Tertinggi	Eliminasi kegagalan fungsi aplikasi (*aplikasi tidak, tidak ada, error terus*)
Jangka Pendek (0-6 Bulan)	Tinggi	Peningkatan kejelasan informasi & navigasi (*tidak jelas, tidak baik, tidak masuk, sering error*)
Jangka Menengah (6-12 Bulan)	Sedang	Reduksi kompleksitas verifikasi yang terlalu mendetail (*mendalam telitas*)
Jangka Panjang (>12 Bulan)	Rendah	Pemeliharaan dan peningkatan aspek positif (*aplikasi baik, baik aplikasi*)

Pada tabel 4 merupakan hasil penyusunan rencana untuk transformasi temuan teknis menjadi langkah manajerial yang terukur. Prioritas perbaikan difokuskan pada penanganan kegagalan fungsi dan ketidakjelasan informasi dalam jangka pendek, diikuti oleh penyederhanaan verifikasi di jangka menengah, serta pemeliharaan aspek positif di jangka panjang. Penerapan strategi jangka pendek hingga panjang ini bertujuan untuk meningkatkan kepercayaan pengguna terhadap ekosistem digital pemerintahan [25]. Pemahaman mendalam terhadap sentimen ulasan di *Google Play Store* memungkinkan pengembang melakukan perbaikan yang tepat sasaran.

Hasil akurasi yang diperoleh perlu dibandingkan secara eksplisit dengan hasil penelitian terdahulu yang menerapkan algoritma serupa pada domain analisis sentimen ulasan aplikasi. Ramadani et al. melaporkan bahwa SVM secara konsisten memberikan performa terbaik dibandingkan *Decision Tree* dan *Logistic Regression* dalam klasifikasi sentimen ulasan aplikasi Netflix, dengan akurasi rata-rata 88,18% dan nilai tertinggi 92,69% pada pengujian *K-Fold Cross Validation* [1]. Kurniawan et al. dalam penelitian terhadap ulasan aplikasi Ajaib melaporkan SVM mencapai akurasi tertinggi sebesar 91%, mengungguli *Random Forest* (85%) dan *Naive Bayes* (83%) [2]. Dibandingkan dengan kedua penelitian tersebut, akurasi SVM pada penelitian ini (93%) sedikit lebih tinggi, yang mengindikasikan bahwa kombinasi normalisasi *custom slang* dan pembobotan fitur yang diterapkan pada tahap *preprocessing* cukup efektif dalam meningkatkan kualitas representasi teks.

Sementara itu, untuk algoritma XGBoost, Setiyono et al. melaporkan bahwa penerapan XGBoost tanpa teknik *balancing* pada ulasan aplikasi Threads hanya menghasilkan akurasi 87,60% dengan *recall* kelas negatif yang rendah (0,69), namun setelah diterapkan teknik SMOTE untuk mengatasi ketidakseimbangan kelas, akurasi meningkat signifikan menjadi 97,49% dengan *recall* kelas negatif mencapai 0,99 [25]. Dibandingkan dengan temuan tersebut, akurasi XGBoost pada penelitian ini (92%) berada di antara kedua kondisi tersebut, lebih tinggi dibandingkan XGBoost tanpa SMOTE pada penelitian Setiyono et al., namun masih di bawah performa XGBoost yang telah diperkuat SMOTE. Hal ini konsisten dengan kondisi pada penelitian ini, di mana ketidakseimbangan kelas yang signifikan belum ditangani dengan teknik *balancing*, sehingga berdampak pada rendahnya *recall* kelas positif (0,58 untuk XGBoost dan 0,61 untuk SVM). Dengan demikian, dapat disimpulkan bahwa performa SVM dan XGBoost pada penelitian ini berada pada level yang kompetitif dibandingkan studi-studi sejenis, namun penerapan teknik penyeimbangan kelas seperti SMOTE berpotensi memberikan peningkatan performa lebih lanjut, khususnya pada kemampuan model mengenali kelas positif.

Dengan mengintegrasikan nilai SHAP ke dalam perencanaan sistem informasi, instansi dapat beralih dari penanganan masalah yang bersifat reaktif menuju pengembangan sistem yang lebih proaktif dan berorientasi pada pengguna [14], [15].

Penelitian ini memiliki keterbatasan data ulasan yang hanya bersumber dari *Google play store*, sehingga tidak mewakili opini dari *platform* lain, serta ketidakseimbangan kelas yang signifikan berpotensi menyebabkan bias terhadap kelas mayoritas [18]. Berbeda dengan penelitian sebelumnya yang hanya berfokus pada interpretasi fitur SHAP tanpa menghasilkan rekomendasi strategis terstruktur [14], [15], penelitian ini telah menerapkan Metode SHAP pada domain aplikasi M-Paspor untuk membuat rekomendasi strategis sebagai prioritas perbaikan berjangka. Namun, keterbatasan linguistik terkait bahasa daerah dan bahasa gaul tertentu dapat memengaruhi kontribusi fitur dalam menafsirkan nilai SHAP, adanya perubahan kebijakan migrasi atau pembaruan aplikasi setelah periode pengumpulan data juga dapat mengubah pola sentimen pengguna. Oleh karena itu, rekomendasi strategis yang dihasilkan dari nilai SHAP perlu divalidasi lebih lanjut oleh pakar domain, serta cakupan data dan periode waktu harus diperluas pada penelitian mendatang untuk meningkatkan generalisasi temuan.

5. KESIMPULAN

Penelitian ini berhasil mengimplementasikan analisis sentimen pada 3.500 ulasan M-Paspor menggunakan SVM dan XGBoost, di mana SVM mencapai akurasi tertinggi (93,43%) dibandingkan XGBoost (92%). Penerapan SHAP pada XGBoost mengungkap fitur paling berpengaruh terhadap sentimen negatif, yaitu "aplikasi tidak", "tidak ada", dan "tidak jelas", yang menjadi dasar rekomendasi strategis prioritas penanganan kegagalan fungsi serta peningkatan kejelasan informasi. Temuan ini memberikan wawasan berharga bagi pengembang M-Paspor untuk memahami persepsi pengguna secara lebih mendalam, sehingga integrasi SHAP dalam perencanaan sistem informasi memungkinkan instansi beralih dari pendekatan reaktif menuju pengembangan sistem yang proaktif dan berorientasi pada pengguna. Penelitian selanjutnya diharapkan dapat memperluas cakupan data, menerapkan teknik penyeimbang kelas seperti SMOTE, membandingkan metode berbasis *deep learning* seperti IndoBERT, serta mengintegrasikan analisis temporal untuk memantau perubahan sentimen pengguna M-Paspor secara lebih dinamis dari waktu ke waktu.

DAFTAR PUSTAKA

[1] N. C. Ramadani, I. Tahyudin, and A. Shouni Barkah, "Perbandingan Algoritma Support Vector Machine, Decision Tree, dan Logistic Regresion Pada Analisis

Sentimen Ulasan Aplikasi Netflix," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 10, no. 2, pp. 110–117, Aug. 2024, doi: 10.25077/teknosi.v10i2.2024.110-117.

[2] N. D. Kurniawan, P. R. Ferdian, and N. Hidayati, "Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 11, no. 1, pp. 87–97, May 2025, doi: 10.25077/teknosi.v11i1.2025.87-97.

[3] E. T. Pardede, K. D. Tania, and M. Afrina, "Knowledge Discovery: Analisis Sentimen dan Emosi WhatsApp Business dengan Machine learning dan Deep Learning," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 11, no. 3, pp. 310–318, Jan. 2026, doi: 10.25077/teknosi.v11i3.2025.310-318.

[4] E. Mustika Sari, C. Sabila, R. Fakhrizal Adam, and R. Kurniawan, "Analisis dan Prediksi Indeks Kualitas Udara Jakarta: Penerapan Algoritma XGBoost," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 11, no. 2, pp. 161–169, Sep. 2025, doi: 10.25077/teknosi.v11i2.2025.161-169.

[5] L. Ashbaugh and Y. Zhang, "A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning," *Computers*, vol. 13, no. 12, Dec. 2024, doi: 10.3390/computers13120340.

[6] M. R. Alifi, D. C. U. Lieharyani, B. P. Sudimulya, and M. R. Maulidhan, "Implementation of IndoNLU Pre-Trained Model for Aspect-Based Sentiment Analysis of Indonesian Stock News," *JURNAL TEKNIK INFORMATIKA*, vol. 16, no. 2, pp. 151–160, Dec. 2023, doi: 10.15408/jti.v16i2.33791.

[7] A. Koushik, M. Manoj, and N. Nezamuddin, "SHapley Additive exPlanations for Explaining Artificial Neural Network Based Mode Choice Models," *Transportation in Developing Economies*, vol. 10, Mar. 2024, doi: 10.1007/s40890-024-00200-6.

[8] M. T. Syamkalla, S. Khomsah, and Y. S. R. Nur, "Implementasi Algoritma Catboost Dan Shapley Additive Explanations (SHAP) Dalam Memprediksi Popularitas Game Indie Pada Platform Steam," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 777–786, Aug. 2024, doi: 10.25126/jtiik.1148503.

[9] B. Widoyono, M. F. Nadhif, and R. A. Eryadi, "Evaluating Single and Hybrid Feature Selection for Rainfall Prediction Using XGBoost," *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, vol. 9, no. 1, pp. 50–61, 2026, doi: 10.24014/ijaidm.v9i1.39110.

[10] M. Donita Maharani, I. Gusti Ayu Agung Diatri Indradewi, and I. Nyoman Saputra Wahyu Wijaya, "Perbandingan Performa Tf-Idf dan Bow pada Analisis Sentimen BPJS Kesehatan Menggunakan XGBoost," *Jurnal Pendidikan Teknologi dan Kejuruan*, vol. 23, no. 1, 2026, doi: 10.23887/jptk-undiksha.v23i1.109604.

[11] F. S. Shantika and Z. Abidin, "Sentiment Analysis of Jobstreet Application Reviews on Google Play Store Using Support Vector Machine Algorithm with Adaptive Synthetic," *Recursive Journal of Informatics*, vol. 3, no. 1, pp. 99–107, Oct. 2025, doi: 10.15294/rji.v3i2.11891.

[12] J. H. Choi, Y. Choi, K. S. Lee, K. H. Ahn, and W. Y. Jang, "Explainable Model Using Shapley Additive Explanations Approach on Wound Infection after Wide Soft Tissue Sarcoma Resection: 'Big Data' Analysis Based on Health Insurance Review and Assessment Service Hub," *Medicina (Lithuania)*, vol. 60, no. 2, Feb. 2024, doi: 10.3390/medicina60020327.

[13] S. M. Lundberg *et al.*, "From local explanations to global understanding with explainable AI for trees," *Nat. Mach. Intell.*, vol. 2, no. 1, pp. 56–67, 2020, doi: 10.1038/s42256-019-0138-9.

[14] C. Dewi, B. J. Tsai, and R. C. Chen, "Shapley Additive Explanations for Text Classification and Sentiment Analysis of Internet Movie Database," in *Communications in Computer and Information Science*, Springer Science and Business Media Deutschland GmbH, 2022, pp. 69–80. doi: 10.1007/978-981-19-8234-7_6.

[15] A. S. Iffadah, Trimono, and Dwi Arman Prasetya, "Shapley Additive Explanations Interpretation of the XGBoost Model in Predicting Air Quality in Jakarta," *Jurnal Riset Informatika*, vol. 7, no. 3, pp. 119–127, Jun. 2025, doi: 10.34288/jri.v7i3.366.

[16] S. Hakkal and A. A. Lahcen, "XGBoost To Enhance Learner Performance Prediction," *Computers and Education: Artificial Intelligence*, vol. 7, Dec. 2024, doi: 10.1016/j.caeai.2024.100254.

[17] K. Arunika and L. J. E. Dewi, "Perbandingan Model Sarima, Exponential Smoothing, dan Xgboost untuk Prediksi Penjualan Super Store," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.8173.

[18] W. Wang *et al.*, "A User Purchase Behavior Prediction Method Based on XGBoost," *Electronics (Switzerland)*, vol. 12, no. 9, May 2023, doi: 10.3390/electronics12092047.

[19] H. Wijaya, D. P. Hostiadi, and E. Triandini, "Optimization XGBoost Algorithm Using Parameter Tunning in Retail Sales Prediction," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 13, no. 3, Dec. 2024, doi: 10.23887/janapati.v13i3.82214.

[20] K. A. Wirakusuma, P. Novita, P. Dewi, and Y. E. Aryanto, "Optimasi Model Prediksi Extreme Gradient Boosting dengan Genetic Algorithm Untuk Produksi dan Produktivitas Padi," *Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 4, no. 4, pp. 378–392, 2025, doi: 10.55123.

[21] I. Bagus Nyoman Wijana Manuaba, G. Rasben Dantes, and G. Indrawan, "Analisis Sentimen Data Provider Layanan Internet Pada Twitter Menggunakan Support Vector Machine (SVM) Dengan Penambahan Algoritma Levenshtein Distance," Mar. 2022. doi: 10.47970/siskom-kb.v5i2.261.

[22] I. P. Dedy, W. Darmawan, G. Aditra Pradnyana, I. Bagus, and N. Pascima, "Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram," *SINTECH JOURNAL*, Apr. 2023, doi: 10.31598.

[23] N. Putu, A. Rainita, I. Md, D. Maysanjaya, and G. S. Mahendra, "Perbandingan Performa Algoritma Naive Bayes dan Support Vector Machine pada Analisis Sentimen Implementasi Program KIP-Kuliah," *Print) Jurnal TEKNOSIA*, vol. 19, no. 2, pp. 76–87, doi: 10.33369/teknosia.v19i02.43479.

[24] G. Indrawan, H. Setiawan, and A. Gunadi, "Multi-class SVM Classification Comparison for Health Service Satisfaction Survey Data in Bahasa," *HighTech and Innovation Journal*, vol. 3, no. 4, pp. 425–442, Dec. 2022, doi: 10.28991/HIJ-2022-03-04-05.

[25] Bayu Nur Setiyono, Nadia Annisa Maori, and Teguh Tamrin, "Analisis Sentimen Ulasan Pengguna Aplikasi Threads di Google Play Menggunakan Algoritma XGBoost Dengan Penguatan SMOTE," *Jurnal Informatika Dan Tekonologi Komputer (JITEK)*, vol. 5, no. 2, pp. 119–130, Jul. 2025, doi: 10.55606/jitek.v5i2.6832.

[26] N. Y. Hassan, W. H. Gomaa, G. A. Khoriba, and M. H. Haggag, "Credibility detection in Twitter using word N-gram analysis and supervised machine learning techniques," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 1, pp. 291–300, Feb. 2020, doi: 10.22266/ijies2020.0229.27.

[27] G. Feretzakis *et al.*, "Integrating Shapley Values into Machine Learning Techniques for Enhanced Predictions of Hospital Admissions," *Applied Sciences (Switzerland)*, vol. 14, no. 13, Jul. 2024, doi: 10.3390/app14135925.

[28] Y. Bugayenko *et al.*, "Prioritizing tasks in software development: A systematic literature review," *PLoS One*, vol. 18, no. 4 April, Apr. 2023, doi: 10.1371/journal.pone.0283838.