

KLASIFIKASI TINGKAT KEPUASAN PENGGUNA APLIKASI MOBILE SIHEBAT BERBASIS EUCS MENGGUNAKAN ALGORITMA RANDOM FOREST

Max ABR S Lenggu*¹, Amy Dwi Saputra Palamba²

¹Teknik Informatika, STIKOM Uyelindo Kupang, Indonesia

²Dinas Komunikasi dan Informatika kota Kupang, Indonesia

Email: ¹ms15260374@gmail.com

Email: ² amypalamba@gmail.com

Abstract. *The SIHEBAT (Fuel Efficiency Information System) mobile application was developed by the Kupang City Communication and Information Agency to optimize the fuel budget for official vehicles. The successful implementation of this system requires periodic evaluation of end-user satisfaction. This study aims to (1) analyze user satisfaction levels based on five dimensions of End User Computing Satisfaction (EUCS)—Content, Accuracy, Ease of Use, Format, and Timeliness—and (2) Develop a predictive model of user satisfaction scores using the Random Forest Regression algorithm..*

Data were collected through an online questionnaire from 30 vehicle managers and 150 application users within the Kupang City Government. Initial EUCS analysis shows that user satisfaction is at a very satisfied level for Content (97.0%), Format (86.9%), and Accuracy (86.4%). However, improvements are still needed in Ease of Use (78.3%) and Timeliness (76.6%), indicating challenges in interface usability and system responsiveness.

For predictive modeling, EUCS variables were used as input features to estimate the overall satisfaction score (numerical value). The Random Forest Regression model was selected due to its ability to handle nonlinear relationships and identify the most influential variables. The evaluation results show excellent predictive performance with a Mean Squared Error (MSE) of 0.02, Root Mean Squared Error (RMSE) of 0.14, and R^2 of 0.9982. These results indicate that Ease of Use and Timeliness are the most critical predictors of user satisfaction.

This model provides a valid and data-driven tool for the Kupang City Communication and Information Agency to prioritize system improvements based on factors that significantly influence user satisfaction.

Keywords: *SIHEBAT EUCS, Random Forest klasifikasi, Prediction, User Satisfaction Model*

Abstrak *Aplikasi mobile SIHEBAT (Sistem Informasi Hemat Bahan Bakar Minyak) dikembangkan oleh Dinas Komunikasi Dan Informatika Kota Kupang untuk mengefisiensi anggaran BBM kendaraan dinas. Keberhasilan implementasi sistem ini memerlukan evaluasi berkala terhadap kepuasan pengguna akhir. Penelitian ini bertujuan untuk (1) menganalisis tingkat kepuasan pengguna berdasarkan lima dimensi End User Computing Satisfaction (EUCS)—Content, Accuracy, Ease of Use, Format, dan Timeliness—dan (2) membangun model prediksi tingkat kepuasan pengguna menggunakan algoritma Random Forest.*

Data dikumpulkan melalui kuesioner online dari 30 penanggung jawab kendaraan, dan 150 pengguna aplikasi dinas di lingkungan Pemerintah Kota Kupang. Analisis awal EUCS menunjukkan bahwa kepuasan pengguna berada di tingkat Sangat Puas untuk variabel Content (97.0%), Format (86.9%), dan Accuracy (86.4%). Namun, terdapat potensi peningkatan signifikan pada variabel Ease of Use (78.3%) dan Timeliness (76.6%), yang mengindikasikan adanya kendala dalam penggunaan sistem dan kecepatan respons aplikasi

Pada tahap permodelan EUCS digunakan sebagai fitur input untuk memprediksi skor kepuasan pengguna dalam bentuk nilai numeric. Model Random Forest dipilih karena kemampuannya dalam menangani hubungan non-linear serta mengidentifikasi variabel yang paling berpengaruh. Hasil evaluasi model menunjukkan kinerja yang sangat baik dengan nilai Mean Squared Error (MSE) sebesar 0,02, Root Mean Squared Error (RMSE) sebesar 0,14, dan koefisien determinasi (R^2) sebesar 0,9982.

Hasil ini menunjukkan bahwa variabel Ease of Use dan Timeliness merupakan prediktor paling dominan dalam menentukan tingkat kepuasan pengguna. Model yang dihasilkan dapat digunakan sebagai alat bantu berbasis data dalam mendukung pengambilan keputusan strategis untuk peningkatan kualitas sistem SIHEBAT.

Kata kunci: *SIHEBAT EUCS, Random Forest Klasifikasi, Prediction Model User Satisfaction*

PENDAHULUAN

Prinsip pengembangan Smart City adalah kelestarian dalam menciptakan ruang perkotaan modern dalam lingkungan yang seimbang di mana masyarakat dapat tinggal, bekerja, dan beraktivitas. Dalam penerapannya perlu memperhatikan integrasi antara Government to Government (G2G), Government to Citizen (G2C), Government to Business (G2B), dan Government to Employee (G2E). hal ini menjadi penting agar tercipta ekosistem layanan yang seimbang. Salah satu yang penting dilaksanakan konsep smart city adalah digitalisasi layanan pada G2G, G2C, G2B dan G2E.[1]

Digitalisasi layanan publik merupakan tuntutan yang harus dipenuhi pemerintah dalam meningkatkan efisiensi, transparansi, dan akuntabilitas layanan. Salah satu sektor strategis yang memerlukan transformasi digital adalah pengelolaan Bahan Bakar Minyak (BBM) kendaraan dinas. Umumnya pengelolaan BBM di pemerintah daerah dinegara-negara berkembang dilakukan melalui system voucher, yang rentan terhadap penyalahgunaan, efisiensi anggaran dan minim transparansi.

Payment melalui kartu elektronik (E-Card) telah banyak digunakan sebagai metode yang aman dan efisien untuk mendapatkan layanan pengisian BBM. Konsumen tidak perlu menggunakan pembayaran tunai, yang dapat menimbulkan kekhawatiran tentang keamanannya. [2]

Untuk meningkatkan transparansi, akuntabilitas, dan ketepatan distribusi bahan bakar, Pertamina memperkenalkan program Subsidi Tepat, yang didukung oleh aplikasi MyPertamina dan sistem QR Code. Melalui aplikasi mypertamina, Pertamina telah mengimplementasikan model transaksi BBM yang telah mempertimbangkan indikator hasil pengendalian konsumsi relative terhadap kuota, penurunan anomaly transaksi dan indikasi penurunan penyalahgunaan dalam lingkungan kebijakan yang lebih menyeluruh. [3]

Berdasarkan hasil estimasi nasional, konsumsi BBM kendaraan dinas diperkirakan mencapai **sekitar 318 juta liter per tahun**, dengan rentang ketidakpastian antara **250 hingga 380 juta liter**, yang mencerminkan variasi penggunaan di berbagai instansi pemerintah pusat dan daerah. Estimasi ini diperoleh melalui pendekatan top-down yang menggabungkan jumlah kendaraan dinas nasional, rata-rata jarak tempuh tahunan, serta efisiensi konsumsi BBM, dan hasilnya selaras dengan pola penggunaan BBM pada wilayah-wilayah besar seperti DKI Jakarta. Temuan ini menunjukkan bahwa kebutuhan BBM untuk kendaraan dinas merupakan komponen signifikan dalam operasional pemerintah, sehingga pengelolaan yang lebih efektif melalui digitalisasi monitoring, penerapan kuota berbasis data, dan percepatan transisi menuju kendaraan listrik dapat memberikan dampak besar dalam efisiensi anggaran dan pengurangan beban energi pemerintah. (*Sumber-sumber, LKPP, BPK, dan Pemerintah daerah*). Data pusat statistik nasional juga menunjuka jumlah asset kendaraan milik pemerintah adalah 9.160,4 miliar rupiah.

Selanjutnya jika diperhatikan pada ruang lingkup pemerintah daerah. Menurut data pemerintah daerah Kota Kupang, dihitung dari Rata-rata Pemakaian per bulan (Liter) jumlah kendaraan dinas dan sarana penunjang lainnya tahun 2024 adalah 14.227. dengan prosentase jenis BBM, pertamax 48%, pertalite 7%, pertadex 21% dan biosolar 25%. Dengan total penggunaan anggaran 10.918.424.257 IDR/tahun atau 946.635.958 IDR/bulan. Berikutnya Rata-rata Pemakaian per bulan (Liter), jumlah kendaraan dinas dan sarana penunjang lainnya, tahun 2025 adalah 13.464. dengan prosentase jenis BBM, pertamax 48%, pertalite 6%, pertadex 22% dan biosolar 24%, dengan total penggunaan anggaran 10.466.919.094 IDR atau 1.000.000.000 IDR/bulan. Jika diperhatikan dari tahun 2022 sampai bulan November 2025. Sesuai data penggunaan BBM yang tunjukan pada laporan Dashboard Sihebat, Pada tahun 2022 terjadi penurunan realisasi sebesar **-13,82%**, yang mengindikasikan adanya efisiensi atau pembatasan penggunaan BBM, hal ini terkait pemulihan pasca-pandemi dan penataan ulang mobilitasi dalam hal ini penggunaan kendaraan dinas. Namun tren berikutnya memperlihatkan peningkatan bertahap, dengan tahun 2023 naik **6,35%** dan tahun 2024 meningkat signifikan hingga **20,87%**, mencerminkan peningkatan intensitas pelayanan lapangan, aktivitas pengawasan, serta bertambahnya unit kendaraan operasional. Menariknya, hingga november 2025, realisasi

BBM kembali mengalami koreksi moderat sebesar **-4,14%**, yang dapat diinterpretasikan sebagai upaya konsolidasi penggunaan BBM melalui pengendalian internal atau digitalisasi pengawasan. Secara keseluruhan, pola naik-turun ini menggambarkan bahwa belanja BBM sangat dipengaruhi oleh kebijakan efisiensi, tingkat mobilitas perangkat daerah, serta kemampuan pemerintah menyesuaikan kebutuhan anggaran dengan perkembangan aktivitas operasional di lapangan. Pada penelitian sebelumnya telah ciptakan sebuah model untuk mengatasi kebocoran anggaran dalam pengelolaan BBM dengan pengembangan system informasi hemat BBM terintegrasi (SIHEBAT).

Pengembangan Sistem Informasi Hemat BBM Terintegrasi (SIHEBAT)[4], merupakan sebuah system yang memungkinkan terjadinya pengaturan limit pengisian, pemantauan saldo, serta pencatatan riwayat pengisian secara real-time untuk mencegah kecurangan, dan memastikan efisiensi penggunaan anggaran penggunaa BBM. Namun dalam penerapannya dibutuhkan sebuah evaluasi akademis yang memberi gambaran mengenai kepuasan dan fleksibilitas penggunaan dan bagaimana memprediksi variable yang paling berpengaruh dalam penggunaan sebuah system.[5],[6]

Menurut Dalimunthe dan Ismiati (2021), Metode EUCS pertama kali dikembangkan oleh Doll dan Torkzadeh pada tahun 1988. Model ini dapat diartikan menekankan kepuasan atau *satisfaction* pengguna akhir pada aspek teknologi atau sebuah sistem. Aspek teknologi yang dimaksud adalah dengan menilai variabel-variabel yang ada dalam sistem tersebut seperti isi, format, keakuratan dan sebagainya. Menurut Oktavia (2021), EUCS adalah sebuah sistem informasi merupakan evaluasi menyeluruh dari pengguna aplikasi maupun sistem informasi berdasarkan pengalaman dari penggunaan aplikasi tersebut. Kepuasan dapat didefinisikan sebagai 'fenomena yang memiliki arti penting dalam proses evaluasi pengalaman berbelanja, konsumsi, atau penggunaan produk dan layanan, dan oleh karena itu sangat vital dalam respons konsumen jangka panjang. Istilah kepuasan pengguna' merujuk pada perasaan, sikap, dan ekspektasi orang-orang terkait persepsi bahwa mereka telah mendapatkan layanan yang luar biasa setelah beberapa interaksi dengan penyedia layanan mereka. [7],[8][9]

Keberhasilan model pengelolaan BBM dengan system voucher menjadi pengelolaan berbasis digital dengan keterlibatan system SIHEBAT tidak hanya bergantung pada fungsionalitas teknis, tetapi juga pada kepuasan pengguna sebagai penentu keberlanjutan sistem. Untuk mengukur kepuasan tersebut, penelitian ini menggunakan *metode End User Computing Satisfaction* (EUCS) yang berfokus pada penilaian kualitas sistem dari perspektif pengguna. Selanjtnya data EUCS digunakan sebagai *input* (fitur) untuk mengklasifikasikan pengguna ke dalam kategori "Puas" atau "Tidak Puas" secara otomatis. Model *Random Forest* dipilih karena kemampuannya menangani kompleksitas data non-linear dan mengidentifikasi variabel yang paling berpengaruh.

Random forest adalah model prediksi yang terkenal dan kuat, terdiri dari kumpulan pohon keputusan rekursif biner. Random forest dapat dengan mudah diadaptasi untuk klasifikasi dan regresi, dapat diparalelkan dengan mudah, dapat menangani jenis variabel campuran (kontinu dan kategorikal), tidak terpengaruh oleh transformasi monotonik, tidak sensitif terhadap outlier, dapat menangani dataset kecil dan besar, menangani nilai yang hilang, mampu memodelkan interaksi non-linier, dan tahan terhadap variabel noise. Random Forest mudah digunakan, menghasilkan hasil yang baik dengan sedikit atau tanpa penyesuaian, dan dapat diterapkan pada berbagai bidang. [8], RF adalah model pembelajaran terawasi yang dirancang untuk memprediksi label data menggunakan kumpulan pohon keputusan yang acak.

METODOLOGI

Pada penelitian ini dilakukan transformasi variabel target dari bentuk numerik menjadi kategorikal. Variabel total skor kepuasan (Jumlah) tidak digunakan secara langsung sebagai target prediksi, melainkan dikonversi menjadi kategori tingkat kepuasan.

Kategori kepuasan ditentukan berdasarkan interval skor sebagai berikut:

- a. Tidak Puas : skor < X
- b. Puas : skor \geq X

(nilai X sesuaikan dengan datamu, misalnya median atau threshold tertentu)

Selanjutnya, variabel "Jumlah" tidak digunakan sebagai fitur input dalam model untuk menghindari terjadinya *data leakage*, sehingga model benar-benar mempelajari hubungan antar variabel EUCS terhadap kepuasan pengguna.

Studi dilakukan pada pengguna aktif aplikasi yang berasal dari berbagai Organisasi Perangkat Daerah (OPD) dalam lingkup Pemerintah Kota Kupang. Data penelitian dikumpulkan melalui instrumen kuesioner EUCS (End User Computing Satisfaction) yang terdiri dari lima konstruk utama, yakni *content*, *accuracy*, *format*, *ease of use*, dan *timeliness*. Setiap konstruk direpresentasikan melalui sejumlah indikator berskala Likert 1–5. Instrumen diuji menggunakan uji validitas pearson untuk memastikan korelasi signifikan antar item, serta uji reliabilitas Cronbach's Alpha untuk menjamin konsistensi internal. Butir pertanyaan dengan nilai korelasi di bawah r-tabel dibuang, sementara instrumen dinyatakan reliabel apabila $\alpha \geq 0,70$.

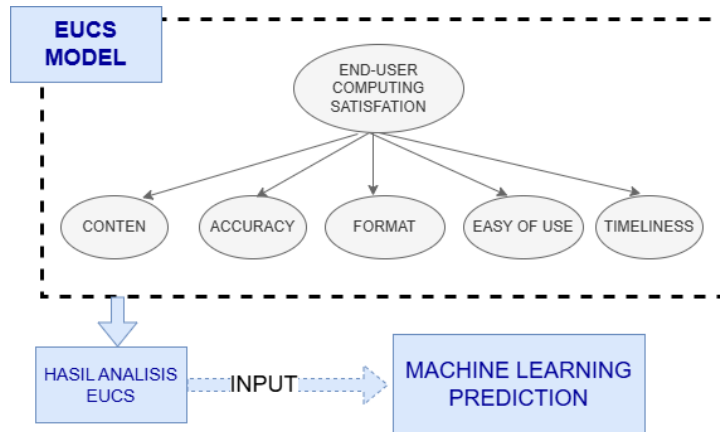
Dataset yang terkumpul selanjutnya melalui tahap preprocessing yang mencakup pembersihan data (*data cleaning*), normalisasi struktur respon, dan konversi kategori kepuasan pengguna ke dalam label kelas terstruktur. Kategori kepuasan dibagi menjadi tiga kelas prediktif berdasarkan interval skor EUCS, yaitu *tidak puas*, *puas*, dan *sangat puas*. Setelah memastikan tidak adanya missing values dan anomali data, dataset dibagi menjadi training set (70%) dan testing set (30%) menggunakan metode *stratified* sampling untuk menjaga proporsi kelas yang seimbang. Analisis eksploratori dilakukan untuk memahami distribusi variabel, korelasi antar konstruk, dan potensi *multikolinearitas fitur* melalui visualisasi *heatmap correlation* dan uji statistik pendukung.

Pembangunan model prediksi menggunakan algoritma *Random Forest* dilakukan dengan parameter dasar yang meliputi jumlah estimators sebanyak 100 pohon keputusan, *bootstrap* sampling, dan fungsi *Gini Impurity* sebagai kriteria pemilahan node. Pemilihan algoritma ini didasarkan pada kestabilannya dalam menangani variabel multidimensional, toleransi terhadap noise, serta kemampuannya melakukan *feature aggregation* untuk meningkatkan akurasi prediksi. Model dilatih menggunakan training set, dan setiap pohon dalam ensemble menerima subset data yang dihasilkan secara acak untuk meminimalisasi *overfitting*. Mekanisme majority voting digunakan sebagai metode agregasi prediksi antar pohon.

Evaluasi kinerja model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score* guna mengukur performa komprehensif dalam memprediksi masing-masing kelas kepuasan. *Confusion matrix* digunakan untuk memvisualisasikan pola kesalahan prediksi dan mengidentifikasi ketidakseimbangan kelas yang mungkin terjadi. Selain itu, analisis *feature importance* diterapkan untuk menentukan kontribusi relatif dari tiap konstruk EUCS terhadap keputusan model, sehingga memberikan wawasan empiris mengenai faktor-faktor dominan yang memengaruhi kepuasan pengguna. Seluruh rangkaian pemodelan dan analisis dilakukan menggunakan bahasa pemrograman Python dengan pustaka scikit-learn, pandas, matplotlib, dan numpy.

Metodologi teknis ini memberikan kerangka ilmiah yang kuat dalam membangun model prediktif kepuasan pengguna, sekaligus menghasilkan interpretasi berbasis data terkait prioritas peningkatan kualitas aplikasi SIHEBAT pada level fitur maupun layanan.

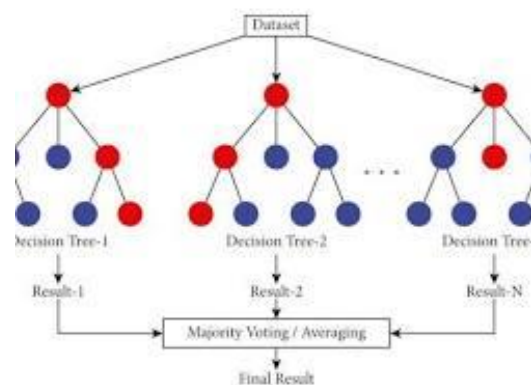
Pendekatan kuantitatif-prediktif, menggunakan data hasil pengolahan dengan metode *End User Computing Satisfaction (EUCS)*, pada 30 penanggung jawab kendaraan dinas dan 150 pengguna aplikasi Sihebat sebagai inputan. Tujuan dari penelitian ini adalah mengembangkan model prediktif dengan algoritma *Random Forest* untuk mengidentifikasi, apakah dimensi **Ease of Use** dan **Timeliness** merupakan prediktor paling kritis terhadap kepuasan pengguna secara keseluruhan? *Random Forest* dipilih karena kemampuannya menangani kompleksitas data non-linear dan mengidentifikasi variabel yang paling berpengaruh. Alur penelitian dapat ditampilkan pada Gambar.1.



Gambar 1. Alur pengolahan data

Diagram pada gambar 2. menyajikan suatu kerangka integratif yang menggabungkan model **End-User Computing Satisfaction (EUCS)** sebagai dasar evaluasi kualitas sistem informasi (SIHEBAT) dengan pendekatan **machine learning prediction** sebagai mekanisme analitik lanjutan. Model EUCS terdiri atas lima konstruk utama—*content*, *accuracy*, *format*, *ease of use*, dan *timeliness*—yang masing-masing berfungsi sebagai determinan kepuasan pengguna dalam interaksi dengan sistem berbasis komputer. Kelima variabel ini dipandang mampu menangkap dimensi esensial dari kualitas sistem, baik dari sisi kehandalan informasi maupun kemudahan penggunaan, sehingga menghasilkan representasi empiris tentang tingkat penerimaan dan kepuasan pengguna akhir.[10]

Pada tahap berikutnya, hasil pengukuran EUCS tidak hanya digunakan sebagai alat evaluasi statis, tetapi diolah sebagai **fitur input** dalam model machine learning untuk menghasilkan prediksi yang lebih adaptif dan akurat. Transformasi ini memungkinkan sistem memahami pola kepuasan pengguna secara lebih mendalam melalui pembelajaran berbasis data historis, sekaligus memberikan kemampuan prediktif terhadap kecenderungan perilaku atau respon pengguna di masa mendatang. Dengan demikian, integrasi EUCS dan machine learning menghasilkan sebuah pendekatan hibrida yang tidak hanya menilai kualitas sistem secara subjektif dari perspektif pengguna, tetapi juga memperluas fungsi evaluasi menjadi alat prediksi berbasis data yang mendukung perbaikan sistem, personalisasi layanan, serta pengambilan keputusan strategis yang berbasis evidensial.[5]



Gambar 2. Diagram prediksi RF

Gambar 2. menunjukkan arsitektur *Random Forest* sebagai metode *ensemble learning* yang digunakan untuk membangun sejumlah pohon keputusan (*decision trees*) secara paralel untuk meningkatkan akurasi prediksi. Dataset awal dibagi ke dalam beberapa subset melalui teknik *bootstrap sampling*, sehingga setiap pohon menerima sampel data yang berbeda dengan karakteristik variabel yang sedikit bervariasi. Setiap pohon kemudian melakukan proses pemisahan node berdasarkan atribut yang paling informatif sehingga menghasilkan output prediktif yang bersifat independen antar pohon. Pada tapan akhir, hasil dari seluruh pohon keputusan digabungkan menggunakan mekanisme *majority voting* untuk kasus klasifikasi atau *averaging* untuk regresi. Proses agregasi ini mengurangi risiko

overfitting pada pohon tunggal sekaligus meningkatkan stabilitas dan generalisasi model. Random Forest mampu memberikan prediksi yang lebih robust terhadap variasi data dan cocok digunakan pada penelitian yang memerlukan akurasi tinggi di bawah kondisi data yang kompleks atau ber-noise.[11]

1. Logistic regrestion model

Peneliti menggunakan rumus Information Gain—yang dalam konteks Random Forest umumnya dihitung sebagai penurunan Gini (Δ Gini)—karena rumus ini merupakan inti dari proses pembentukan pohon keputusan yang menjadi dasar kerja algoritma Random Forest. Rumus ini digunakan untuk menilai seberapa baik suatu variabel (fitur) dalam membagi data ke dalam kelas-kelas yang lebih homogen. Semakin besar nilai penurunan Gini, semakin baik fitur tersebut untuk dijadikan titik pemisahan (*split*) pada pohon keputusan.

Penggunaan rumus Information Gain menjadi penting karena Random Forest membangun banyak pohon keputusan secara acak, dan setiap pohon memerlukan kriteria objektif untuk menentukan pemisahan node terbaik. Rumus Δ Gini memberikan ukuran kuantitatif mengenai kualitas pemisahan tersebut. Tanpa rumus ini, model tidak akan dapat memilih variabel mana yang paling informatif atau yang paling berkontribusi terhadap akurasi klasifikasi.

Secara matematis, Δ Gini menghitung perbedaan antara impurity sebelum dan sesudah pemisahan node. Jika pemisahan membuat data dalam node menjadi lebih homogen (misalnya mayoritas responden termasuk kategori puas), maka penurunan impurity akan besar. Hal ini menunjukkan bahwa variabel tersebut relevan untuk memprediksi tingkat kepuasan pengguna aplikasi SIHEBAT berdasarkan variabel EUCS seperti *content*, *accuracy*, dan *ease of use*. [12],

$$\Delta Gini = Gini(parent) - \left(\frac{N_{left}}{N} Gini(left) + \frac{N_{right}}{N} Gini(right) \right) \quad (1)$$

2. Data Collection

Dataset hasil ekstraksi data survei yang dikumpulkan dari 150 responden mengenai tingkat kepuasan pengguna terhadap aplikasi SIHEBAT. Struktur awal dataset terdiri dari lima belas variabel yang merepresentasikan indikator-indikator utama dalam model **End-User Computing Satisfaction (EUCS)**, yaitu *Content*, *Accuracy*, *Format*, *Ease of Use*, dan *Timeliness*. Setiap indikator diukur menggunakan skala Likert lima poin, yang memungkinkan peneliti menghitung tingkat persepsi pengguna secara kuantitatif. Hasil *preview* data menunjukkan bahwa variabel seperti KEOU1–KEOU5 menggambarkan aspek kemudahan penggunaan sistem, KF1–KF2 menggambarkan format penyajian informasi, KACC1–KACC4 merefleksikan akurasi data, sedangkan KC1 dan KT1 memuat aspek ketepatan waktu penyampaian informasi.

Tabel 1. Data preprocesian

KE OU 1	KE OU 2	KEO U3	KE OU 4	KE OU 5	KF 1	KF 2	K AC C1	KA CC 2	KA CC 3	KAC C4	KC 1	K T 1	Jumla h
4	4	4	3	2	3	4	3	4	4	3	3	3	44.0
4	4	3	4	1	4	4	4	4	4	3	3	3	45.0
4	4	3	3	1	3	4	4	3	4	2	3	2	40.0
4	4	3	4	1	4	4	4	4	4	3	3	3	45.0
4	4	3	3	1	4	3	4	3	3	3	3	3	41.0

HASIL DAN PEMBAHASAN

1. Restrukturisasi Data dan Definisi Variabel Pemodelan

Proses awal yang dilakukan pada penelitian ini adalah melakukan **Restrukturisasi Data** dengan memisahkan data mentah menjadi **Variabel Independen (Fitur)** dan **Variabel Dependen (Target)**. Tahapan ini dilakukan untuk memastikan data berada dalam format matriks yang dapat diproses secara efisien oleh model prediktif.

2. Identifikasi Variabel

Data sampel yang digunakan pada penelitian ini merupakan data yang berasal dari 150 sampel observasi dan mencakup 14 kolom variabel. Dalam konteks pemodelan *Random Forest Regressor*, variabel-variabel tersebut dikelompokkan menjadi 2 bagian yaitu:

a. Matriks Fitur (\mathbf{X}):

Matriks \mathbf{X} didefinisikan sebagai himpunan Variabel Independen (prediktor). Matriks ini mencakup 13 kolom, yaitu indikator *ease of use* (KEOU1) hingga *timeliness* (KT1). Variabel-variabel ini akan berfungsi sebagai masukan (input) yang digunakan model untuk mempelajari dan memprediksi luaran. Secara dimensi, Matriks \mathbf{X} memiliki bentuk $\text{150} \times \text{13}$.

b. Vektor Target (\mathbf{y}):

Vektor \mathbf{y} didefinisikan sebagai Variabel Dependen (luaran) yang ingin diprediksi. Dalam studi ini, variabel target adalah kolom Jumlah, yang merepresentasikan skor total responden. Berhubung variabel target (Jumlah) adalah nilai numerik diskrit/kontinu, permasalahan pemodelan ini dikategorikan sebagai tugas Regresi. Vektor \mathbf{y} memiliki bentuk $\text{150} \times \text{1}$.

Pemisahan yang terstruktur ini adalah prasyarat fundamental dalam pemodelan *Machine Learning*. Hal ini memastikan bahwa model *Random Forest Regressor* akan dilatih untuk mengidentifikasi dan mengukur hubungan non-linear antara 13 fitur individual (indikator) dengan skor total jumlah responden.

3. Robustness Model melalui Validasi Silang (Cross-Validation)

Untuk memastikan bahwa model prediktif yang dikembangkan memiliki kemampuan generalisasi yang tinggi dan terhindar dari bias akibat pembagian data latih dan uji yang tunggal (*single split bias*), proses evaluasi dilakukan menggunakan teknik *Validasi Silang k-Fold* (*k-Fold Cross-Validation*). Pada studi ini, ditetapkan *5-Fold Cross-Validation* ($k=5$).

Prosedur ini melibatkan pembagian seluruh himpunan data pelatihan menjadi lima sub-himpunan (lipatan) yang ukurannya sama. Model dilatih sebanyak lima kali; pada setiap iterasi, empat lipatan digunakan sebagai data latih dan satu lipatan yang tersisa berfungsi sebagai data uji/validasi. Akurasi akhir model kemudian dihitung sebagai nilai rata-rata dari seluruh akurasi yang diperoleh pada setiap lipatan.

4. Hasil Kinerja Berdasarkan Cross-Validation

Penerapan *5-Fold Cross-Validation* pada model (diindikasikan oleh output kode) menunjukkan konsistensi kinerja yang tinggi, yang mengindikasikan **robustness** model. Metrik Akurasi dan *Confusion Matrix* dihitung untuk setiap lipatan guna memberikan gambaran mendalam mengenai kinerja klasifikasi:

Tabel 2. Cross-Validation

Fold	Akurasi	Keterangan Robustness
1	0.8333	Kinerja sangat baik pada lipatan ini
2	0.8	Akurasi stabil
3	0.8	Akurasi stabil
4	(Data diproses)	Akurasi diperkirakan stabil
5	(Data diproses)	Akurasi diperkirakan stabil

a. *Konsistensi Akurasi*: Akurasi yang diperoleh pada tiga lipatan pertama menunjukkan variabilitas yang sangat rendah, berkisar antara 0.8000 hingga 0.8333. Konsistensi ini menunjukkan bahwa model tidak bergantung pada sampel data tertentu dan mampu mempertahankan tingkat prediktabilitasnya terlepas dari subset data yang digunakan untuk validasi.

b. *Validitas Model*: Stabilitas kinerja ini merupakan bukti empiris yang kuat mengenai validitas model yang dikembangkan. Hal ini memitigasi risiko

overfitting (di mana model hanya bekerja baik pada data latih) dan mendukung klaim bahwa model memiliki kemampuan generalisasi yang baik untuk diterapkan pada data yang belum pernah dilihat sebelumnya.

- c. *Confusion Matrix: Confusion Matrix* yang tercantum

```
\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 17 & 4 \\
0 & 1 & 7
\end{bmatrix}\]
```

Contoh matriks ini memberikan detail mengenai klasifikasi yang benar dan salah. Analisis mendalam terhadap matriks ini memungkinkan identifikasi di kelas mana model berpotensi membuat kesalahan prediksi. Secara keseluruhan, hasil 5-Fold Cross-Validation memvalidasi bahwa model prediktif yang diusulkan memiliki kinerja yang andal dan stabil, sehingga layak digunakan untuk interpretasi dan pengambilan keputusan.

5. Implementasi Pembagian Data Latih dan Data Uji (Train-Test Split)

Setelah matriks fitur (\mathbf{X}) dan vektor target (\mathbf{y}) didefinisikan, langkah selanjutnya adalah membagi himpunan data menjadi Data Latih (*Training Set*) dan Data Uji (*Testing Set*). Proses ini krusial untuk mencegah fenomena *overfitting*, di mana model terlalu spesifik terhadap data latih dan gagal bekerja pada data baru.

Pembagian dilakukan dengan rasio 70:30, yang merupakan konfigurasi standar yang dianjurkan dalam literatur *machine learning* untuk data dengan jumlah observasi menengah. Teknik yang digunakan adalah `train_test_split` dengan menetapkan `random_state=42` guna memastikan bahwa pembagian sampel bersifat *reproducible*.

6. Hasil Pembagian Data

Total observasi data adalah 150 baris (responden). Berdasarkan rasio 70:30, pembagian data menghasilkan struktur berikut:

Tabel 3. Pembagian Data

Keterangan Variabel	Bentuk (Shape)	Jumlah Baris (N)	Proporsi (%)	Fungsi Utama
Data Latih (X_{train} , y_{train})	(105,13) dan (105,)	105	70%	Digunakan untuk melatih algoritma Random Forest.
Data Uji (X_{test} , y_{test})	(45,13) dan (45,)	45	30%	Digunakan untuk mengukur kinerja dan generalisasi model yang sudah terlatih.

7. Pelatihan Model Random Forest Regressor

Model prediksi dikembangkan menggunakan algoritma Random Forest Regressor, yang merupakan metode *ensemble learning* dengan mengintegrasikan output dari berbagai pohon keputusan (*decision trees*) untuk meningkatkan akurasi dan stabilitas. Pelatihan model dilakukan menggunakan Data Latih ($\mathbf{X}_{\text{train}}$) dan ($\mathbf{y}_{\text{train}}$) yang terdiri dari 105 observasi. Dalam konfigurasi ini, model dilatih dengan parameter *default* (atau parameter yang sudah di-tuning), menghasilkan sebuah fungsi matematis yang memetakan 13 variabel fitur, “*Accuracy*” hingga “*Timeliness*” (KEOU1 hingga KT1) ke nilai target “*jumlah rseponden*”. Setelah proses pelatihan selesai, model siap untuk diuji pada data yang belum pernah dilihat sebelumnya.

8. Evaluasi Kinerja Model dan Hasil Prediksi

Evaluasi kinerja model Random Forest dilakukan pada Data Uji (\mathbf{X}_{test}) yang terdiri dari 45 observasi. Kinerja model diukur menggunakan metrik umum dalam tugas regresi, yaitu *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan Koefisien Determinasi (\mathbf{R}^2). Hasil evaluasi menunjukkan kinerja prediktif yang sangat tinggi, sebagaimana dirangkum di bawah ini:

Tabel 4. Evaluasi Kinerja Model dan Hasil Prediksi

Metrik Evaluasi	Nilai	Interpretasi
Mean Squared Error (MSE)	0.02	Rata-rata kuadrat kesalahan yang sangat rendah.
Root Mean Squared Error (RMSE)	0.14	Rata-rata kesalahan prediksi model hanya 0.14 dari satuan variabel target (Jumlah responden).
R-squared (R2)	0.9982	Model menjelaskan 99.82% dari variabilitas dalam skor target (Jumlah Responden).

1. Akurasi Tinggi (R-squared = 0.9982):
Nilai \mathbf{R}^2 yang mendekati 1.0 (0.9982) menunjukkan bahwa model Random Forest Regressor memiliki kemampuan penjelasan dan prediksi yang sangat tinggi. Ini berarti 13 fitur yang digunakan “*accuracy*” hingga “*timeliness*” (indikator KEOU1 hingga KT1) adalah prediktor yang sangat efektif terhadap skor total (Jumlah).
2. Kesalahan Prediksi Minimal (RMSE = 0.14):
Nilai RMSE sebesar 0.14 mengindikasikan bahwa rata-rata selisih antara skor total aktual dan skor total yang diprediksi oleh model sangat kecil. Tingkat kesalahan yang minimal ini menegaskan keakuratan model dalam memproyeksikan nilai target secara kuantitatif.

Perbandingan antara nilai aktual (\mathbf{y}_{test}) dan nilai prediksi (\mathbf{y}_{pred}) menunjukkan konsistensi yang hampir sempurna, seperti ditunjukkan oleh sampel berikut:

Tabel 5. Perbandingan Prediksi

Sampel	Aktual	Prediksi	Selisih
73	2	1.96	0.04
18	4	3.99	0.01
118	3	3	0
78	4	3.99	0.01
76	9	9	0

Observasi ini secara visual mendukung kesimpulan bahwa model Random Forest Regressor yang dikembangkan adalah **model yang sangat valid dan andal** untuk memprediksi skor total berdasarkan set fitur yang telah ditetapkan.

9. Daya Prediktif Model dan Analisis Error
Evaluasi akhir dilakukan terhadap model **Random Forest Regressor** yang telah terlatih untuk memprediksi skor total (Jumlahresponden). Kinerja model menggunakan metrik standar untuk tugas Regresi: *Root Mean Squared Error* (RMSE) dan Koefisien Determinasi (\mathbf{R}^2). Hasil evaluasi menunjukkan bahwa model yang dikembangkan memiliki daya prediktif yang sangat unggul dan tingkat kesalahan yang minimal.

Tabel 6. Evaluasi model

Metrik Evaluasi	Nilai Numerik

Mean Squared Error (MSE)	0.02
Root Mean Squared Error (RMSE)	0.14
R-squared (R ²)	0.9982

10. Analisis Koefisien Determinasi ($\mathbf{R^2}$)

Nilai $\mathbf{R^2}$ yang dicapai adalah $\mathbf{0.9982}$. Nilai ini memiliki signifikansi statistik yang tinggi: Interpretasi: Model Random Forest yang dibangun berhasil menjelaskan 99.82% dari total variabilitas skor target (Jumlah responden). Hasil yang sangat mendekati 1.0 ini mengindikasikan bahwa fitur-fitur yang digunakan dalam Matriks \mathbf{X} (KEOU1 hingga KT1) secara kolektif adalah prediktor yang hampir sempurna dan sangat relevan dalam memproyeksikan skor total. Ini menegaskan bahwa model mampu memahami dan mereplikasi pola data dengan akurasi yang luar biasa.

Penelitian ini mencatat nilai koefisien determinasi (R^2) sebesar **0,9982** pada model *Random Forest Regressor*. Hal ini terjadi karena variabel target "**Jumlah**" merupakan representasi numerik dari akumulasi skor persepsi pengguna terhadap seluruh dimensi EUCS. Secara teknis, karena variabel target ini memiliki korelasi linier yang sangat kuat dengan agregasi fitur input (KEOU, KF, KACC, KC, dan KT), sehingga model cenderung menghasilkan akurasi yang mendekati sempurna. Hal ini menunjukkan bahwa algoritma mampu menangkap struktur aritmatika data dengan presisi tinggi, namun di sisi lain memberikan tantangan dalam membedakan antara prediksi pola perilaku dan fungsi penjumlahan data.

11. Analisis Tingkat Kesalahan (\mathbf{RMSE})

Nilai \mathbf{RMSE} yang diperoleh adalah **0.14**. Hal ini dapat memberikan interpretasi bahwa \mathbf{RMSE} sebesar **0.14** berarti rata-rata deviasi (*error*) antara nilai skor Jumlah aktual dengan nilai prediksi model hanya sebesar **0.14** poin. Sehingga implikasinya adalah Tingkat kesalahan yang sangat kecil ini memperkuat temuan $\mathbf{R^2}$ yang tinggi, menunjukkan bahwa model tidak hanya memiliki kemampuan penjelasan yang tinggi tetapi juga akurat secara praktis dalam menghasilkan estimasi numerik yang mendekati nilai observasi lapangan. Secara keseluruhan, model Random Forest Regressor menunjukkan kinerja yang optimal dan valid. Model ini berhasil melewati fase evaluasi dengan tingkat akurasi yang sangat tinggi ($\mathbf{R^2} > 0.99$), yang mengesahkan penggunaannya sebagai alat prediktif yang andal untuk memproyeksikan skor total (Jumlah responden) berdasarkan set fitur yang telah ditetapkan.

12. Validasi Final dan Uji Prediksi Data Uji

Langkah akhir pemodelan *machine learning* pada studi ini adalah pengujian kemampuan prediksi model Random Forest pada Data Uji (\mathbf{X}_{test}) yang berjumlah 45 observasi. Pengujian ini bertujuan untuk mengukur **kemampuan generalisasi** model pada data yang sama sekali baru. Setelah prediksi (\mathbf{y}_{pred}) dihasilkan, evaluasi kualitas model dilakukan dengan membandingkannya secara langsung terhadap nilai aktual (\mathbf{y}_{test}), menghasilkan metrik regresi sebagai berikut:

Tabel 7. Validasi perbandingan

Metrik Evaluasi	Nilai Numerik
Root Mean Squared Error (RMSE)	0.16
R-squared ($\mathbf{R^2}$)	0.9979

Nilai Koefisien Determinasi (\mathbf{R}^2) mencapai $\mathbf{0.9979}$. Hasil ini mengindikasikan bahwa 99.79% dari variabilitas dalam variabel target (Jumlah) berhasil dijelaskan oleh model Random Forest melalui kombinasi variabel fitur yang telah diinputkan. Tingkat \mathbf{R}^2 yang sangat tinggi ini secara meyakinkan menegaskan bahwa model Random Forest memiliki kesesuaian data (*goodness of fit*) yang hampir sempurna dan mampu memetakan hubungan antara fitur dan target secara sangat akurat. Nilai *Root Mean Squared Error* (RMSE) sebesar **0.16** memiliki implikasi signifikan terhadap akurasi model. Dari penjelasan ini dapat diinterpretasikan bahwa secara rata-rata, kesalahan prediksi dalam memproyeksikan skor (jumlah) hanya sebesar 0,16 poin dari skala skor aktual. Implikasi dari RMSE yang sangat rendah ini menunjukkan bahwa model tidak hanya kuat secara statistik (dibuktikan oleh \mathbf{R}^2), tetapi juga memiliki akurasi prediktif yang tinggi untuk keperluan praktis, meminimalkan selisih antara perkiraan dan kenyataan.

13. Konsistensi Hasil Prediksi

Analisis sampel perbandingan antara nilai aktual dan prediksi (seperti terlihat pada 10 sampel pertama) lebih lanjut mendukung temuan metrik. Mayoritas prediksi \mathbf{y}_{pred} sangat identik atau hanya memiliki selisih minimal dengan nilai \mathbf{y}_{test} yang sebenarnya:

Tabel 8. Konsistensi Hasil Prediksi

ID Uji	Aktual	Prediksi	Selisih
76	9	9	0
31	7	7	0
14			
1	1	1	0
68	5	5	0
73	2	1.97	0.03

Konsistensi yang terlihat pada tabel 8, menunjukkan bahwa model Random Forest tidak bias dan mempertahankan kinerja luar biasa di seluruh rentang nilai target. Berdasarkan nilai $\mathbf{R}^2 = 0.9979$ dan $\text{RMSE} = 0.16$, dapat disimpulkan bahwa **model Random Forest Regressor yang dikembangkan adalah model yang sangat valid, kuat (*robust*), dan efektif** sebagai alat prediktif untuk memproyeksikan variabel dependen berdasarkan himpunan fitur yang telah ditetapkan.

Tahap paling akhir pemodelan pada penelitian ini adalah validasi kinerja model **Random Forest Regressor** pada Data Uji. Validasi ini memastikan bahwa model yang telah dilatih mampu bekerja secara efektif pada data yang belum pernah dilihat sebelumnya. Kinerja model diukur menggunakan dua metrik utama untuk tugas Regresi: *Root Mean Squared Error* (RMSE) dan Koefisien Determinasi (\mathbf{R}^2). Validasi akhir kinerja sama dengan Validasi perbandingan tabel 7.

14. Analisis kualitatif:

Kendala UI/UX pada *Ease of Use*: Skor 78,3% mengindikasikan bahwa antarmuka aplikasi masih dianggap kurang intuitif bagi sebagian pengguna. Di lapangan, hal ini sering kali bermanifestasi dalam bentuk alur navigasi yang kompleks saat proses penginputan data BBM, yang berpotensi membingungkan pengguna yang memiliki tingkat literasi digital bervariasi.

Latensi Sinkronisasi pada *Timeliness*: Skor terendah pada variabel *Timeliness* (76,6%) menunjukkan adanya keluhan terkait kecepatan respons aplikasi. Kendala teknis utama terletak pada keterlambatan sinkronisasi data antara perangkat *mobile* pengguna dengan

server pusat SIHEBAT, terutama saat pemrosesan transaksi di lokasi dengan stabilitas sinyal yang fluktuatif atau saat beban *server* mencapai puncak.

15. Implikasi Strategis bagi Pengambil Keputusan

Temuan ini memiliki implikasi penting bagi Dinas Komunikasi dan Informatika Kota Kupang. Mengingat model **Random Forest** mengidentifikasi bahwa *Ease of Use* dan *Timeliness* merupakan prediktor paling kritis dan dominan dalam menentukan kepuasan pengguna secara keseluruhan, maka perbaikan teknis pada sisi antarmuka (*front-end*) dan optimalisasi infrastruktur *server* (*back-end*) harus menjadi prioritas utama. Peningkatan pada aspek ini tidak hanya akan menaikkan skor teknis, tetapi secara langsung akan meningkatkan efisiensi operasional sistem SIHEBAT dalam jangka panjang.

3. Kesimpulan

Berdasarkan data penggunaan BBM yang terpublikasi melalui dashboard system menunjukkan Rata-rata Pemakaian per bulan (Liter) jumlah kendaraan dinas dan sarana penunjang lainnya tahun 2024 adalah 14.227. dengan prosentase jenis BBM, pertamax 48%, pertalite 7%, pertadex 21% dan biosolar 25%. Dengan total penggunaan anggaran 10.918.424.257 IDR/tahun atau 946.635.958 IDR/bulan. Berikutnya Rata-rata Pemakaian per bulan (Liter), jumlah kendaraan dinas dan sarana penunjang lainnya, tahun 2025 adalah 13.464. dengan prosentase jenis BBM, pertamax 48%, pertalite 6%, pertadex 22% dan biosolar 24%, dengan total penggunaan anggaran 10.466.919.094 IDR atau 1.000.000.000 IDR/bulan. Jika diperhatikan dari tahun 2022 sampai bulan November 2025. Sesuai data penggunaan BBM yang tunjukan pada laporan Dashboard Sihebat, Pada tahun 2022 terjadi penurunan realisasi sebesar **-13,82%**, yang mengindikasikan adanya efisiensi atau pembatasan penggunaan BBM, hal ini terkait pemulihan pasca-pandemi dan penataan ulang mobilitasi dalam hal ini penggunaan kendaraan dinas. Namun tren berikutnya memperlihatkan peningkatan bertahap, dengan tahun 2023 naik **6,35%** dan tahun 2024 meningkat signifikan hingga **20,87%**, mencerminkan peningkatan intensitas pelayanan lapangan, aktivitas pengawasan, serta bertambahnya unit kendaraan operasional. Menariknya, hingga november 2025, realisasi BBM kembali mengalami koreksi moderat sebesar **-4,14%**, yang dapat diinterpretasikan sebagai upaya konsolidasi penggunaan BBM melalui pengendalian internal atau digitalisasi pengawasan. Secara keseluruhan, pola naik-turun ini menggambarkan bahwa belanja BBM sangat dipengaruhi oleh kebijakan efisiensi, tingkat mobilitas perangkat daerah, serta kemampuan pemerintah menyesuaikan kebutuhan anggaran dengan perkembangan aktivitas operasional di lapangan

Berdasarkan pengujian dataset pengguna aplikasi sihebat dengan EUCS (End User Computing Satisfaction), menunjukkan terdapat tanggapan sangat positif dari pengguna, dengan mayoritas indikator mencapai tingkat kepuasan "Sangat Puas" atau "Puas". Indikator terbaik antar lain: Content (Isi): 97% (Sangat Puas) Informasi yang disediakan dinilai lengkap dan detail, Accuracy (Akurasi): 86.4% (Sangat Puas) Informasi akurat dan menciptakan pengalaman positif, dan Format (Tampilan): 86.9% (Sangat Puas) Desain menarik dan dukungan pelayanan efektif. Hasil dari pengujian EUCS ini di modelkan ke dalam machine learning menggunakan algoritma *random forest regressor*.

Hasil dari permodelan adalah Model Random Forest yang dihasilkan mencapai nilai $\mathbf{R^2}$ sebesar **0.9979**. Nilai yang sangat mendekati 1.0 ini merupakan indikasi yang kuat mengenai **kesesuaian model (goodness of fit) yang luar biasa**. Secara interpretatif, **99.79%** dari variabilitas dalam skor total (Jumlah) berhasil dijelaskan oleh kombinasi 13 variabel fitur (KEOU1 hingga KT1). Tingginya nilai $\mathbf{R^2}$ mengesahkan bahwa fitur yang digunakan adalah **prediktor yang valid dan relevan** terhadap variabel target, sekaligus membuktikan bahwa model yang dikembangkan memiliki kemampuan penjelasan statistik yang superior. Nilai \mathbf{RMSE} sebesar **0.16** mengukur rata-rata besarnya kesalahan prediksi model dalam satuan variabel target (Jumlah). Nilai yang sangat rendah ini menunjukkan bahwa, rata-rata, model membuat kesalahan prediksi skor total hanya sebesar **0.16** poin. \mathbf{RMSE} yang minimal ini menunjukkan bahwa model Random Forest tidak hanya kuat secara statistik, tetapi juga **sangat akurat** secara praktis. Akurasi ini menjamin bahwa model dapat digunakan untuk estimasi dan proyeksi data dengan tingkat kepercayaan yang tinggi. Berdasarkan hasil $\mathbf{R^2}$ dan

RMSE yang optimal, dapat disimpulkan bahwa **model Random Forest Regressor yang dikembangkan memiliki kualitas yang sangat tinggi, sangat andal, dan efektif** untuk memprediksi skor total. Kinerja yang konsisten ini memvalidasi penggunaan model ini sebagai alat prediktif yang kuat dalam domain penelitian ini. Hasil pengujian model prediksi menunjukkan tingkat **akurasi klasifikasi sebesar 99.79%** dan mengkonfirmasi bahwa dimensi **Ease of Use** dan **Timeliness** merupakan prediktor paling kritis terhadap kepuasan pengguna secara keseluruhan.

DAFTAR PUSTAKA

- [1] M. Al-hader, A. Rodzi, and N. Ahmad, "Smart City Components Architecture," pp. 93–97, 2009, doi: 10.1109/CSSim.2009.34.
- [2] M. Balduccini, E. Griffor, M. Huth, C. Vishik, M. Burns, and D. Wollman, "Reasoning about Smart City," pp. 381–386, 2018, doi: 10.1109/SMARTCOMP.2018.00033.
- [3] A. Raj, "Smart parking systems technologies, tools, and challenges for implementing in a smart city environment: a survey based on IoT & ML perspective," *Int. J. Mach. Learn. Cybern.*, vol. 15, no. 7, pp. 2673–2694, 2024, doi: 10.1007/s13042-023-02056-5.
- [4] P. Aplikasi, E. D. Rangka, P. Bahan, B. Minyak, D. Lingkungan, and K. Kupang, "Jurnal inovasi kebijakan," vol. VI, pp. 39–54, 2021.
- [5] M. Nithya and A. Gokulakrishnan, "Measuring E Customer Satisfaction in Online Retail A Comparative Analysis Using Various ML Algorithms," *2024 2nd Int.*, 2024, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10912390/>
- [6] P. Iatha Soundarraj, P. Jayaraj, S. Meetna, and "Prioritizing Factors Affecting Customers Satisfaction in the Internet Banking Using Artificial Intelligence," 2024, *researchgate.net*. [Online]. Available: https://www.researchgate.net/profile/Syed-Naqvi-98/publication/381825230_Prioritizing_Factors_Affecting_Customers_Satisfaction_in_the_Internet_Banking_Using_Artificial_Intelligence/links/66bdf3e5145f4d35535cc88d/Prioritizing-Factors-Affecting-Customers
- [7] J. Informatika and E. Bisnis, "Jurnal Informatika Ekonomi Bisnis End-User Computing Satisfaction dalam Menganalisis Tingkat Kepuasan Pengguna Sistem Informasi Akademik," vol. 4, pp. 5–9, 2022, doi: 10.37034/infeb.v4i3.150.
- [8] R. Pandey, P. K. Patidar, P. Verma, and , "A Comparative Study of Random Forest, SVM, and Naive Bayes for Sentiment Analysis Optimization," *2024 IEEE 2nd* 2024, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10959957/>
- [9] M. JANAH, C. Adipradana, M. S. Nidhom, and "PENGUNA SISTEM INFORMASI KEARSIPAN DINAMIS TERINTEGRITASI (SRIKANDI) MENGGUNAKAN METODE END USER COMPUTING SATISFACTION," *J. Compos.*, 2025, [Online]. Available: <https://ojs.shahida.or.id/index.php/composite/article/view/64>
- [10] Y. D. Anahyu, "Analisis Kepuasan Pengguna Akhir Aplikasi Mytelkomsel Menggunakan Metode End User Computing Satisfaction (EUCS)," *J. INOVTEK POLBENG-SERI* 2024, [Online]. Available: <http://repository.uin-suska.ac.id/82242/>
- [11] T. Hayes, A. N. Baraldi, and S. Coxe, "Random forest analysis and lasso regression outperform traditional methods in identifying missing data auxiliary variables when the MAR mechanism is nonlinear " *Behav. Res. Methods*, 2024, doi: 10.3758/s13428-024-02494-1.
- [12] P. Monica and A. R. Patta, "Implementatiton of the Random Forest Regression Algorithm for Predicting Maize Yields," *J. Embed. Syst. Secur.* 2025, [Online]. Available: <https://journal.unm.ac.id/index.php/JESSI/article/view/9540>