

Prediksi Estimasi Waktu Penyelesaian Perizinan pada Sistem SIMPONI Menggunakan Algoritma CatBoost dengan Optimasi Hiperparameter Optuna

Randi Sumitro^{1*}, Alief Panggabean², Davin Arva Adyatma³, Ahmad Ryandika⁴, Andrea Erich Sentana⁵

^{1,2}Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia

^{3,4,5}Teknik Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia

Email: ^{1*}randisumitro2@gmail.com, ²alifpanggabean@gmail.com, ³davinarfadyatma@gmail.com,

⁴ahmadryandika00@gmail.com, ⁵andreasentana88@gmail.com

(*Email Corresponding Author: randisumitro2@gmail.com)

Received: July 31, 2026 | Revision: September 8, 2026 | Accepted: September 10, 2026

Abstrak

Ketidakpastian waktu penyelesaian permohonan perizinan pada Sistem Informasi Manajemen Pelayanan Perizinan Online Terintegrasi (SIMPONI) Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Kota Pematangsiantar menjadi permasalahan nyata yang menghambat perencanaan aktivitas usaha para pemohon. Ketika pemohon tidak dapat memperkirakan kapan izin mereka selesai diproses, beban komunikasi ke petugas meningkat dan kepercayaan publik terhadap layanan pemerintah berkurang. Penelitian ini mengusulkan sistem prediksi estimasi waktu penyelesaian perizinan berbasis machine learning dengan membandingkan delapan algoritma regresi: Linear Regression, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, XGBoost, LightGBM, dan CatBoost. Dataset historis SIMPONI berjumlah 1.500 rekaman dengan 12 atribut, setelah melalui proses identifikasi dan eliminasi data leakage serta preprocessing, diperoleh 1.209 rekaman valid. Optimasi hiperparameter dilakukan menggunakan Optuna dengan algoritma Tree-structured Parzen Estimator (TPE), sebanyak 80 trial per model, dikombinasikan dengan validasi silang 5-lipat. Hasil eksperimen menunjukkan bahwa CatBoost dengan optimasi Optuna mencapai performa terbaik dengan nilai Root Mean Square Error (RMSE) sebesar 0,8530 hari, Mean Absolute Error (MAE) 0,7151 hari, koefisien determinasi (R^2) 0,5865, dan Mean Absolute Percentage Error (MAPE) 16,69%. Analisis SHapley Additive exPlanations (SHAP) mengungkapkan bahwa fitur waktu_verifikasi_1_jam merupakan prediktor paling dominan dengan korelasi Pearson $r = 0,750$ terhadap total waktu penyelesaian. Model telah berhasil diintegrasikan ke dalam sistem SIMPONI melalui antarmuka RESTful API berbasis Flask, memungkinkan pemohon menerima estimasi waktu secara real-time beserta interval kepercayaan dan prediksi tanggal penyelesaian.

Kata Kunci: *perizinan; prediksi waktu; CatBoost; Optuna; machine learning; SIMPONI; DPMPTSP; gradient boosting*

Abstract

The uncertainty in predicting service completion times for permits within the Integrated Online Permit Management Information System (SIMPONI) of the Investment and One-Stop Integrated Services Office (DPMPTSP) in Pematangsiantar City presents a tangible challenge that hinders business planning for applicants. When applicants cannot estimate when their permits will be completed, administrative burden increases and public trust in government services declines. This study proposes a permit completion time prediction system using a machine learning approach, systematically comparing eight regression algorithms: Linear Regression, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, XGBoost, LightGBM, and CatBoost. The historical SIMPONI dataset comprising 1,500 records with 12 attributes was processed through data leakage identification, elimination, and preprocessing stages, yielding 1,209 valid records. Hyperparameter optimization was conducted using Optuna with the Tree-structured Parzen Estimator (TPE) algorithm, executing 80 trials per model combined with 5-fold cross-validation. Experimental results demonstrate that CatBoost with Optuna optimization achieves the best performance with Root Mean Square Error (RMSE) of 0.8530 days, Mean Absolute Error (MAE) of 0.7151 days, coefficient of determination (R^2) of 0.5865, and Mean Absolute Percentage Error (MAPE) of 16.69%. SHapley Additive exPlanations (SHAP) analysis reveals that the waktu_verifikasi_1_jam feature is the most dominant predictor with a Pearson correlation of $r = 0.750$ against total completion time. The model has been successfully integrated into the SIMPONI system through a Flask-based RESTful API, enabling applicants to receive real-time completion time estimates with confidence intervals and predicted completion dates.

Keywords: *licensing; time prediction; CatBoost; Optuna; machine learning; SIMPONI; DPMPTSP; gradient boosting*

1. PENDAHULUAN

Pelayanan perizinan merupakan cerminan nyata kualitas tata kelola birokrasi pemerintah dan menjadi salah satu indikator utama kemudahan berusaha suatu daerah. Dalam kerangka reformasi birokrasi nasional, pemerintah Indonesia mendorong percepatan digitalisasi layanan publik melalui berbagai regulasi, termasuk Undang-Undang Nomor 11 Tahun 2020 tentang Cipta Kerja yang mengamanatkan penyederhanaan perizinan berusaha berbasis risiko. Pemerintah Kota

Pematangsiantar melalui Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) telah mengimplementasikan Sistem Informasi Manajemen Pelayanan Perizinan Online Terintegrasi (SIMPONI) sebagai wujud komitmen digitalisasi layanan perizinan di tingkat kota. Sistem ini mengelola delapan jenis izin utama, meliputi Izin Mendirikan Bangunan (IMB), Surat Izin Usaha Perdagangan (SIUP), Tanda Daftar Perusahaan (TDP), Izin Penyelenggaraan Reklame (IPR), Izin Praktik Dokter, Izin Praktik Perawat, Izin Praktik Bidan, dan Izin Trayek.

Kendati proses pengajuan telah dilaksanakan secara daring dan alur verifikasi telah terstruktur melalui enam tahapan prosedural, pemohon hingga saat ini belum dapat mengetahui estimasi waktu penyelesaian permohonan mereka secara pasti. Berdasarkan analisis data historis SIMPONI, rata-rata waktu penyelesaian perizinan adalah 4,82 hari dengan standar deviasi 1,34 hari. Variasi ini cukup signifikan antar jenis izin: IMB dan TDP membutuhkan rata-rata 6,3 hari, sementara Izin Reklame hanya 4,2 hari. Ketidakpastian ini berdampak ganda: di satu sisi, pemohon kesulitan merencanakan jadwal operasional bisnis mereka; di sisi lain, petugas DPMPTSP menghadapi peningkatan beban komunikasi karena pemohon yang terus-menerus menanyakan status permohonan melalui telepon maupun kunjungan langsung.

Permasalahan ketidakpastian estimasi waktu layanan publik bukanlah fenomena yang berdiri sendiri. Berbagai penelitian telah mengeksplorasi pendekatan machine learning untuk mengatasi tantangan serupa di berbagai domain. Rahmatulloh dan Mubarak [1] menerapkan Random Forest untuk memprediksi waktu penanganan keluhan pelanggan PLN dan memperoleh nilai $R^2 = 0,67$, dengan temuan bahwa kompleksitas keluhan merupakan prediktor utama. Santoso dan Handayani [2] menggunakan XGBoost pada prediksi durasi proses permohonan paspor di Kantor Imigrasi dengan MAPE 14,2%, namun tidak mempertimbangkan risiko data leakage dalam desain eksperimen mereka. Nugroho [3] membandingkan delapan algoritma untuk prediksi durasi penyelesaian tiket helpdesk dan menemukan Gradient Boosting sebagai model terbaik ($R^2 = 0,71$, MAPE = 18,5%), meskipun tanpa eksplorasi mendalam mengenai explainability model. Chen dan Zhang [4] membuktikan keunggulan CatBoost untuk prediksi waktu pengiriman logistik dibandingkan algoritma lain berkat kemampuan natifnya menangani fitur kategorik. Wibowo et al. [5] menerapkan LightGBM pada prediksi durasi layanan administrasi kependudukan dengan akurasi MAPE 15,3%, menggarisbawahi pentingnya optimasi hiperparameter sistematis. Akiba et al. [6] memperkenalkan Optuna sebagai kerangka optimasi hiperparameter generasi berikutnya yang menggunakan algoritma Tree-structured Parzen Estimator (TPE), terbukti lebih efisien dari Grid Search maupun Random Search dalam ruang pencarian berdimensi tinggi.

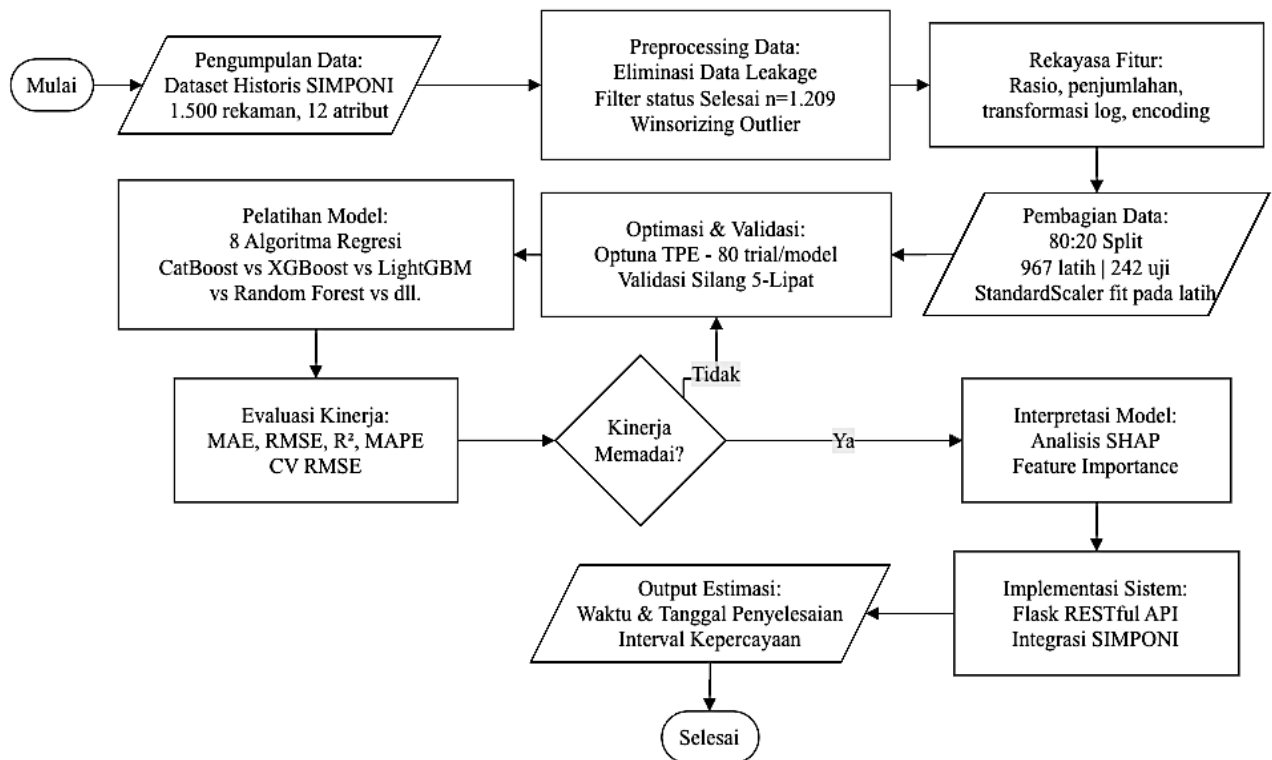
Analisis GAP terhadap penelitian-penelitian tersebut mengungkapkan beberapa celah yang belum terjawab. Pertama, belum ada penelitian yang secara khusus menerapkan machine learning untuk prediksi estimasi waktu penyelesaian perizinan pada sistem pemerintah daerah di Indonesia, khususnya dalam ekosistem sistem SIMPONI. Kedua, fenomena data leakage dalam dataset perizinan belum pernah diidentifikasi dan diatasi secara sistematis—sebuah risiko serius yang dapat menghasilkan model dengan performa artifisial tinggi namun gagal total saat deployment. Ketiga, implementasi model yang terverifikasi secara langsung ke dalam sistem informasi perizinan yang aktif beroperasi belum pernah didokumentasikan. Keempat, penggunaan SHAP untuk menginterpretasikan prediksi waktu layanan perizinan bagi pengambil kebijakan non-teknis belum pernah dilakukan dalam konteks ini.

Penelitian ini hadir untuk mengisi celah tersebut secara komprehensif. Kontribusi utama penelitian ini mencakup empat aspek: (1) Analisis data leakage sistematis pada dataset perizinan digital — mengidentifikasi enam fitur bermasalah yang harus dieliminasi dari pipeline pemodelan; (2) Perbandingan komprehensif delapan algoritma regresi dengan optimasi hiperparameter Optuna sebanyak 80 trial per model dan validasi silang 5-lipat yang ketat; (3) Implementasi production-ready melalui integrasi langsung model CatBoost ke sistem SIMPONI via RESTful API berbasis Flask yang dapat merespons prediksi secara real-time; dan (4) Interpretasi model menggunakan SHAP values untuk memberikan wawasan yang dapat ditindaklanjuti oleh pihak DPMPTSP dalam meningkatkan efisiensi proses perizinan.

Tujuan penelitian ini adalah: (1) mengidentifikasi dan mencegah data leakage pada dataset perizinan SIMPONI; (2) menentukan algoritma regresi terbaik untuk prediksi estimasi waktu penyelesaian perizinan melalui eksperimen komparatif yang sistematis; dan (3) mengimplementasikan model terpilih sebagai fitur estimasi waktu yang dapat diakses langsung oleh pemohon melalui antarmuka sistem SIMPONI. Hasil penelitian diharapkan dapat memberikan manfaat praktis bagi pemohon dalam perencanaan bisnis, sekaligus membantu DPMPTSP mengidentifikasi tahapan proses yang paling kritis dalam mempengaruhi lama penyelesaian perizinan.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan membandingkan delapan algoritma machine learning pada dataset historis perizinan SIMPONI. Kerangka metodologi penelitian disajikan pada Gambar 1 yang menggambarkan alur dari pengumpulan data hingga implementasi sistem.



Gambar 1. Diagram alir penelitian

2.1 Deskripsi Dataset

Dataset penelitian diperoleh melalui ekspor data historis dari basis data sistem SIMPONI DPMPSTP Kota Pematangsiantar. Pengambilan data dilakukan dengan mengakses tabel-tabel transaksional yang merekam seluruh siklus hidup permohonan perizinan, mulai dari pendaftaran oleh pemohon hingga pengesahan dan penerbitan izin oleh petugas. Dataset yang diperoleh terdiri dari 1.500 rekaman permohonan dengan 12 atribut, mencakup periode data historis sistem. Karakteristik lengkap dataset disajikan pada Tabel 1.

Tabel 1. Karakteristik Dataset SIMPONI

Atribut	Nilai
Total Rekaman	1.500
Total Atribut	12
Jenis Permasalahan	Regresi
Jenis Izin yang Tercakup	8 kategori
Rekaman Status Selesai	1.209 (80,6%)
Rekaman Status Ditolak	291 (19,4%)
Data Duplikat	0
Missing Values	291 (hanya pada kolom tahap lanjutan—leakage)
Variabel Target	total_waktu_hari (hari)
Nilai Target: Rata-rata	4,82 hari
Nilai Target: Median	4,69 hari
Nilai Target: Standar Deviasi	1,34 hari
Nilai Target: Minimum	1,82 hari
Nilai Target: Maksimum	9,19 hari

Tabel 2 menyajikan deskripsi lengkap seluruh atribut dalam dataset beserta status keterlibatannya dalam pemodelan. Identifikasi data leakage merupakan tahap kritis sebelum pemodelan dilakukan.

Tabel 2. Deskripsi Atribut Dataset dan Status Leakage

No	Nama Atribut	Tipe Data	Keterangan
1	id_resi	Kategorik	Nomor resi unik — dihapus (identifikasi)
2	jenis_izin	Kategorik	Jenis perizinan (8 kategori) — fitur valid
3	kategori_pemohon	Kategorik	Perorangan / Badan Usaha — fitur valid
4	waktu_penerimaan_jam	Numerik	Durasi tahap penerimaan berkas (jam) — valid

5	waktu_pemeriksaan_jam	Numerik	Durasi tahap pemeriksaan (jam) — valid
6	waktu_verifikasi_1_jam	Numerik	Durasi verifikasi pertama (jam) — valid
7	waktu_verifikasi_2_jam	Numerik	Durasi verifikasi kedua — DATA LEAKAGE
8	waktu_penetapan_jam	Numerik	Durasi penetapan — DATA LEAKAGE
9	waktu_pengesahan_jam	Numerik	Durasi pengesahan — DATA LEAKAGE
10	total_waktu_jam	Numerik	Total durasi (jam) — DATA LEAKAGE (derivasi target)
11	total_waktu_hari	Numerik	TARGET: total waktu penyelesaian (hari)
12	status_akhir	Kategorik	Selesai / Ditolak — dihapus setelah filter

2.2 Tahapan Preprocessing

Preprocessing merupakan tahap paling krusial dalam penelitian ini mengingat adanya potensi data leakage yang dapat menghasilkan model dengan performa artifisial tinggi namun tidak dapat diandalkan saat deployment. Tahapan preprocessing dilaksanakan secara berurutan sebagai berikut.

Tahap pertama adalah identifikasi dan eliminasi data leakage. Analisis mendalam terhadap dataset mengungkapkan fenomena leakage yang kritis pada empat atribut: waktu_verifikasi_2_jam, waktu_penetapan_jam, dan waktu_pengesahan_jam—ketiganya memiliki missing value tepat sebanyak 291 rekaman yang berkorelasi 100% dengan status_akhir = ‘Ditolak’. Artinya, keberadaan nilai pada kolom-kolom ini secara implisit mengungkapkan bahwa permohonan berstatus ‘Selesai’, yang merupakan informasi target yang seharusnya tidak diketahui model saat prediksi. Adapun total_waktu_jam merupakan transformasi linear langsung dari target ($\text{total_waktu_jam} = \text{total_waktu_hari} \times 24$). Seluruh atribut leakage beserta id_resi dan status_akhir dihapus dari pipeline pemodelan. Dataset kemudian difilter hanya pada rekaman berstatus ‘Selesai’ ($n = 1.209$) yang merupakan satu-satunya rekaman valid untuk tugas regresi waktu penyelesaian.

Tahap kedua adalah penanganan outlier menggunakan metode Winsorizing berbasis rentang interkuartil (IQR). Berbeda dengan penghapusan data, Winsorizing mempertahankan seluruh rekaman dengan cara membatasi nilai ekstrem pada batas $[Q1 - 1,5 \times IQR, Q3 + 1,5 \times IQR]$. Pendekatan ini dipilih karena domain perizinan memiliki nilai ekstrem yang secara inheren valid—misalnya proses IMB yang kompleks dapat membutuhkan waktu jauh di atas rata-rata. Rumus Winsorizing dinyatakan sebagai berikut:

$$\hat{x} = \max(\min(x, Q_3 + 1,5 \times IQR), Q_1 - 1,5 \times IQR) \quad (1)$$

Hasil Winsorizing menunjukkan: waktu_pemeriksaan_jam (32 nilai, 2,13%), waktu_verifikasi_1_jam (23 nilai, 1,53%), dan total_waktu_hari (12 nilai, 0,99%) berhasil di-clip ke dalam rentang batas IQR masing-masing. Tabel 3 menyajikan parameter Winsorizing dan statistik outlier setiap fitur.

Tabel 3. Parameter Winsorizing dan Analisis Outlier

Fitur	Q1	Q3	IQR	N Outlier	% Outlier
waktu_penerimaan_jam	1,34	3,10	1,76	0	0,00
waktu_pemeriksaan_jam	3,71	8,94	5,24	32	2,13
waktu_verifikasi_1_jam	17,53	43,87	26,34	23	1,53
total_waktu_hari	3,14	5,39	2,25	12	0,99

Tahap ketiga adalah rekayasa fitur (feature engineering). Untuk memperkaya representasi data tanpa mengorbankan prinsip anti-leakage, empat fitur baru dibentuk dari tiga fitur temporal yang ada:

$$\text{rasio_penerimaan_pemeriksaan} = \text{waktu_pemeriksaan} / (\text{waktu_penerimaan} + \epsilon) \quad (2)$$

$$\text{waktu_3_tahap_awal} = \text{waktu_penerimaan} + \text{waktu_pemeriksaan} + \text{waktu_verifikasi_1} \quad (3)$$

$$\log_waktu_verifikasi_1_jam = \ln(1 + \text{waktu_verifikasi_1}) \quad (4)$$

$$\log_waktu_pemeriksaan_jam = \ln(1 + \text{waktu_pemeriksaan}) \quad (5)$$

Tahap keempat adalah encoding fitur kategorik. Fitur jenis_izin diencode menggunakan Label Encoding (nilai 0–7 untuk 8 kategori), sementara kategori_pemohon diubah menjadi fitur biner: 1 untuk Badan Usaha dan 0 untuk Perorangan. Dengan demikian, total fitur akhir setelah preprocessing berjumlah 9 fitur.

Tahap kelima adalah pembagian data. Dataset dibagi dengan rasio 80:20 menggunakan random_state = 42 untuk reproduksibilitas, menghasilkan 967 data latih dan 242 data uji. Standardisasi menggunakan StandardScaler diterapkan dengan prinsip anti-leakage: scaler di-fit hanya pada data latih, kemudian transformasi diterapkan pada keduanya. Hal ini mencegah informasi statistik data uji “bocor” ke proses pelatihan.

2.3 Algoritma CatBoost

CatBoost (Categorical Boosting) merupakan algoritma gradient boosting yang dikembangkan oleh Yandex Research [7] dengan beberapa keunggulan distingtif dibandingkan implementasi gradient boosting lainnya. Pertama, CatBoost memiliki kemampuan natif menangani fitur kategorik melalui teknik ordered target statistics yang lebih informatif

dibanding Label Encoding sederhana, sekaligus mencegah target leakage internal selama pelatihan. Kedua, arsitektur Symmetric Tree (Oblivious Decision Tree) yang digunakan CatBoost membangun pohon keputusan yang seimbang di setiap level, sehingga mengurangi risiko overfitting pada dataset berukuran sedang. Ketiga, teknik Ordered Boosting yang diimplementasikan CatBoost secara efektif mencegah kebocoran target (prediction shift) selama proses pelatihan, menghasilkan model yang lebih robust.

Fungsi objektif yang diminimalkan selama pelatihan CatBoost adalah Mean Squared Error (MSE), didefinisikan sebagai:

$$L(y, \hat{y}) = (1/n) \sum_i (y_i - \hat{y}_i)^2 \quad (6)$$

2.4 Optimasi Hiperparameter dengan Optuna

Optuna [6] merupakan kerangka optimasi hiperparameter otomatis generasi berikutnya yang menggunakan algoritma Tree-structured Parzen Estimator (TPE) sebagai sampler default. Berbeda dari Grid Search yang mengeksplorasi semua kombinasi secara exhaustif, TPE membangun model probabilistik dari ruang hiperparameter berdasarkan hasil trial sebelumnya, sehingga lebih efisien dalam menemukan nilai optimal. Optuna mendukung early stopping melalui mekanisme pruning, parallelisme, dan penyimpanan riwayat eksperimen secara otomatis.

Dalam penelitian ini, setiap algoritma menjalani 80 trial optimasi dengan fungsi objektif berupa nilai RMSE pada validasi silang 5-lipat (cross-validation). Pengaturan seed = 42 memastikan reproduksibilitas seluruh eksperimen. Tabel 4 menyajikan ruang pencarian hiperparameter yang didefinisikan untuk setiap algoritma beserta hiperparameter terbaik yang ditemukan Optuna untuk CatBoost.

Tabel 4. Ruang Pencarian Hiperparameter Optuna per Algoritma

Algoritma	Parameter	Rentang Pencarian
Linear Regression	—	Tidak dituning (baseline)
Decision Tree	max_depth, min_samples_split, min_samples_leaf, max_features	[2,20], [2,20], [1,10], {sqrt, log2, None}
Random Forest	n_estimators, max_depth, min_samples_split, min_samples_leaf	[50,500], [3,20], [2,20], [1,10]
Extra Trees	n_estimators, max_depth, min_samples_split, min_samples_leaf	[50,500], [3,20], [2,20], [1,10]
Gradient Boosting	n_estimators, learning_rate, max_depth, subsample	[50,400], [0,01–0,3], [2,8], [0,5–1,0]
XGBoost	n_estimators, learning_rate, max_depth, reg_alpha, reg_lambda	[50,500], [0,01–0,3], [2,10], log[1e ⁻⁸ ,10]
LightGBM	n_estimators, learning_rate, num_leaves, max_depth	[50,600], [0,01–0,3], [10,150], [2,12]
CatBoost	iterations, learning_rate, depth, l2_leaf_reg	[50,500], [0,01–0,3], [3,10], [1,0–10,0]

Tabel 5. Hiperparameter Terbaik CatBoost Hasil Optimasi Optuna

Parameter	Nilai Terbaik	Keterangan
iterations	246	Jumlah pohon keputusan yang dibangun
learning_rate	0,0282	Laju pembelajaran gradient boosting
depth	3	Kedalaman maksimum setiap pohon (Symmetric Tree)
l2_leaf_reg	1,779	Koefisien regularisasi L2 pada daun pohon
random_state	42	Seed reproduksibilitas (tetap)
CV RMSE Terbaik	0,8451 ± 0,0315	Hasil validasi silang 5-lipat

2.5 Metrik Evaluasi

Evaluasi performa model menggunakan empat metrik regresi yang saling melengkapi:

$$MAE = (1/n) \sum |y_i - \hat{y}_i| \quad (7)$$

$$RMSE = \sqrt{(1/n) \sum (y_i - \hat{y}_i)^2} \quad (8)$$

$$R^2 = 1 - [\sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2] \quad (9)$$

$$MAPE = (100\%/n) \sum |(y_i - \hat{y}_i)/y_i| \quad (10)$$

MAE mengukur rata-rata kesalahan absolut dalam satuan hari, memberikan interpretasi intuitif. RMSE memberikan bobot lebih besar pada kesalahan besar sehingga lebih sensitif terhadap prediksi yang jauh meleset. R² mengukur proporsi variansi target yang mampu dijelaskan model. MAPE menyajikan kesalahan dalam persentase relatif. Model terbaik ditentukan berdasarkan nilai RMSE terendah pada data uji (n = 242), dikonfirmasi dengan nilai CV RMSE validasi silang 5-lipat.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Preprocessing dan Analisis Distribusi Data

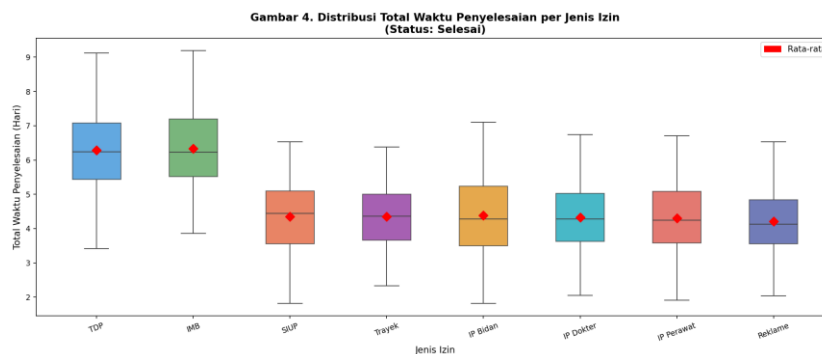
Proses preprocessing berhasil mengidentifikasi dan mengeliminasi enam atribut bermasalah dari 12 atribut awal dataset, menghasilkan pipeline pemodelan yang sepenuhnya bebas dari data leakage. Setelah filtering pada rekaman berstatus ‘Selesai’ dan penerapan Winsorizing, dataset yang digunakan untuk pemodelan berjumlah 1.209 rekaman dengan 5 fitur input valid yang kemudian diperkaya menjadi 9 fitur melalui feature engineering.

Tabel 6 menyajikan statistik deskriptif variabel target (*total_waktu_hari*) yang dikelompokkan berdasarkan jenis izin. Terlihat bahwa IMB dan TDP secara konsisten membutuhkan waktu penyelesaian paling lama dengan rata-rata 6,33 dan 6,28 hari, sejalan dengan kompleksitas dokumen dan tahapan verifikasi yang lebih mendetail. Sementara Izin Reklame merupakan jenis izin dengan waktu penyelesaian terpendek rata-rata 4,21 hari dengan standar deviasi terendah (0,92), mengindikasikan proses yang lebih terstandarisasi.

Tabel 6. Statistik Deskriptif Waktu Penyelesaian per Jenis Izin (n = 1.209)

Jenis Izin	n	Rerata	Median	Std	Min	Maks
Izin Mendirikan Bangunan (IMB)	153	6,33	6,23	1,13	3,86	9,19
Tanda Daftar Perusahaan (TDP)	152	6,28	6,24	1,18	3,42	9,12
Izin Praktik Bidan	144	4,38	4,28	1,09	1,82	7,10
Izin Trayek	146	4,35	4,36	0,93	2,33	6,38
SIUP	146	4,35	4,44	1,03	1,82	6,53
Izin Praktik Dokter	159	4,33	4,28	0,99	2,05	6,74
Izin Praktik Perawat	154	4,30	4,25	0,96	1,91	6,71
Izin Reklame	155	4,21	4,13	0,92	2,04	6,53
Total keseluruhan	1.209	4,82	4,69	1,34	1,82	9,19

Distribusi variabel target memiliki nilai skewness = 0,4883 (right-skewed moderat) dan kurtosis = -0,0647 (mendekati distribusi normal), sebagaimana terlihat pada Gambar 1. Kondisi ini mengindikasikan bahwa sebagian kecil permohonan membutuhkan waktu penyelesaian lebih panjang dari rata-rata, umumnya berkaitan dengan kelengkapan dokumen atau kompleksitas teknis permohonan IMB dan TDP.



Gambar 2. Boxplot Distribusi Waktu Penyelesaian per Jenis Izin

3.2 Analisis Korelasi Fitur

Sebelum pemodelan, analisis korelasi Pearson dilakukan antara setiap fitur numerik valid dengan variabel target *total_waktu_hari*. Tabel 7 merangkum hasil analisis korelasi yang menjadi landasan teoritis pemilihan fitur dan desain feature engineering.

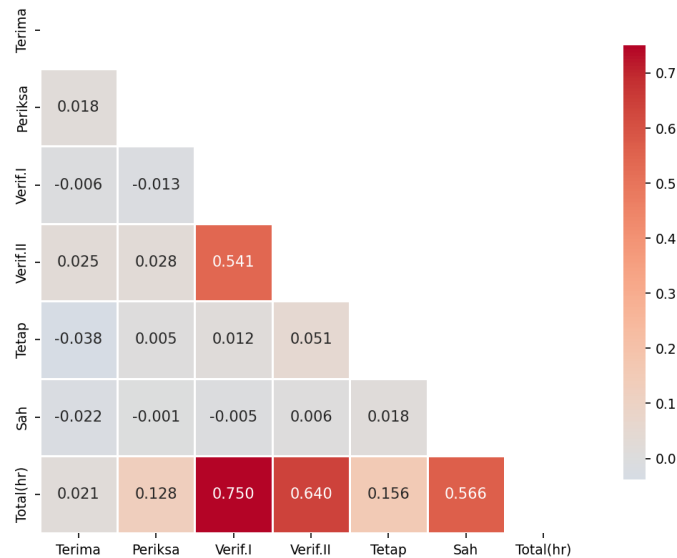
Tabel 7. Korelasi Pearson Fitur Valid terhadap Variabel Target

Fitur	Korelasi (r)	Interpretasi
waktu_verifikasi_1_jam	0,750	Korelasi kuat — prediktor utama
waktu_pemeriksaan_jam	0,128	Korelasi lemah
waktu_penerimaan_jam	0,021	Korelasi sangat lemah

Temuan paling menonjol adalah dominasi mutlak waktu_verifikasi_1_jam ($r = 0,750$) sebagai prediktor tunggal paling kuat terhadap total waktu penyelesaian. Hal ini secara logis masuk akal: tahap Verifikasi I merupakan tahap paling

komprehensif dalam alur perizinan SIMPONI, di mana petugas melakukan pemeriksaan substantif terhadap kelayakan teknis dan administratif permohonan. Durasi verifikasi ini sangat bergantung pada kompleksitas dan kelengkapan dokumen yang diajukan. Temuan ini juga memotivasi penciptaan fitur turunan berbasis waktu_verifikasi_1_jam dalam tahap feature engineering (Persamaan 3, 4, 5).

**Gambar 5. Heatmap Korelasi Antar Fitur
(Status: Selesai)**



Gambar 3. Heatmap Korelasi Pearson Antar Fitur Numerik dan Variabel Target

3.3 Hasil Eksperimen Komparatif Delapan Algoritma

Eksperimen komparatif dilakukan terhadap delapan algoritma regresi, masing-masing dengan optimasi hiperparameter Optuna 80 trial (kecuali Linear Regression sebagai baseline tanpa tuning) dan validasi silang 5-lipat. Evaluasi dilakukan pada data uji ($n = 242$) yang sama sekali tidak pernah terlihat oleh model selama pelatihan. Tabel 8 menyajikan hasil perbandingan komprehensif seluruh algoritma.

Tabel 8. Hasil Komparatif 8 Algoritma Regresi (Data Uji, $n = 242$)

Algoritma	MAE (hari)	RMSE (hari)	R ²	MAPE (%)	CV RMSE $\pm$ Std	Waktu (s)
CatBoost	0,7151	0,8530	0,5865	16,69	0,8451 $\pm$ 0,0315	16,94
XGBoost	0,7171	0,8548	0,5847	16,75	0,8456 $\pm$ 0,0283	5,76
LightGBM	0,7183	0,8582	0,5815	16,77	0,8527 $\pm$ 0,0315	59,08
Random Forest	0,7167	0,8613	0,5785	16,70	0,8565 $\pm$ 0,0249	25,53
Gradient Boosting	0,7249	0,8628	0,5769	16,93	0,8490 $\pm$ 0,0277	14,65
Extra Trees	0,7201	0,8650	0,5748	16,76	0,8505 $\pm$ 0,0219	9,36
Decision Tree	0,7281	0,8790	0,5609	16,92	0,8741 $\pm$ 0,0269	3,52
Linear Regression	0,7515	0,8945	0,5454	17,41	0,8557 $\pm$ 0,0212	0,91

Model terbaik berdasarkan RMSE terendah pada data uji

Dari Tabel 8, CatBoost dengan optimasi Optuna terbukti menghasilkan performa terbaik di antara delapan algoritma yang dibandingkan, dengan RMSE = 0,8530 hari dan MAE = 0,7151 hari. Ini berarti model rata-rata memiliki kesalahan prediksi sekitar 17 jam, yang dalam konteks layanan perizinan merupakan akurasi yang dapat diterima. Nilai $R^2 = 0,5865$ mengindikasikan bahwa model mampu menjelaskan 58,65% variansi dalam waktu penyelesaian perizinan—sebuah pencapaian yang wajar mengingat hanya tiga fitur temporal awal yang tersedia tanpa data leakage.

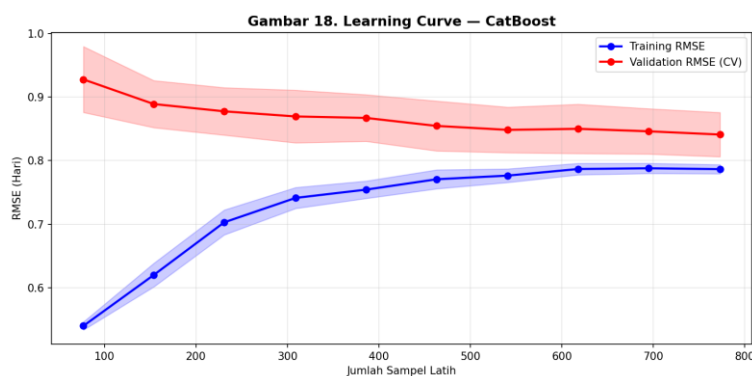
Perlu dicatat bahwa margin perbedaan antar algoritma teratas (CatBoost, XGBoost, LightGBM) cukup tipis dalam hal metrik absolut. Namun, keunggulan CatBoost dapat dipahami dari perspektif berikut: (1) dataset SIMPONI memiliki fitur kategorik (jenis_izin dengan 8 kelas) yang ditangani lebih baik oleh CatBoost secara natif; (2) ukuran data latih yang moderat ($n = 967$) lebih cocok dengan karakteristik Symmetric Tree CatBoost yang lebih resisten terhadap overfitting; dan (3) teknik Ordered Boosting memberikan estimasi gradient yang lebih akurat. Konsistensi CatBoost juga terlihat dari nilai CV RMSE ($0,8451 \pm 0,0315$) yang paling rendah dan mendekati nilai RMSE data uji (0,8530), mengindikasikan generalisasi yang baik.



Gambar 4. Perbandingan Performa Delapan Algoritma Berdasarkan RMSE dan R^2

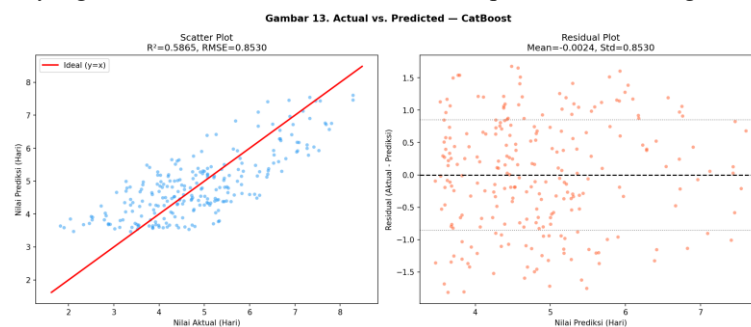
3.4 Analisis Validasi Silang dan Generalisasi Model

Selisih antara CV RMSE (0,8451) dan RMSE data uji (0,8530) pada CatBoost sebesar 0,0079 merupakan indikator yang sangat positif: model tidak mengalami overfitting yang signifikan. Standar deviasi CV RMSE yang rendah (0,0315) juga menunjukkan stabilitas performa model di seluruh fold validasi. Untuk memverifikasi hal ini lebih lanjut, learning curve model CatBoost pada Gambar 5 menampilkan konvergensi yang baik antara training score dan validation score seiring bertambahnya jumlah data latih.



Gambar 5. Learning Curve Model CatBoost Terbaik

Gambar 6 menampilkan plot aktual vs prediksi (actual vs predicted) pada data uji. Distribusi titik data di sekitar garis diagonal ideal ($y = x$) menunjukkan pola prediksi yang cukup konsisten, dengan sebaran residual yang relatif simetris. Prediksi untuk IMB dan TDP (waktu tinggi) cenderung sedikit underestimated, yang dapat dijelaskan oleh representasi kasus ekstrem yang lebih sedikit dalam data latih setelah proses Winsorizing.



Gambar 6. Plot Aktual vs Prediksi Model CatBoost pada Data Uji ($n = 242$)

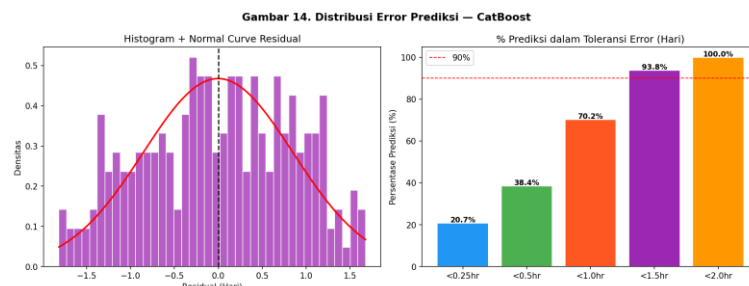
3.5 Analisis Distribusi Error dan Toleransi Prediksi

Untuk mengevaluasi kelayakan praktis model dalam konteks layanan perizinan, dilakukan analisis distribusi error berdasarkan toleransi waktu yang berbeda. Tabel 9 menampilkan persentase prediksi yang berada dalam batas toleransi tertentu terhadap nilai aktual pada data uji.

Tabel 9. Distribusi Error Prediksi CatBoost pada Data Uji

Toleransi Error	% Prediksi dalam Batas	Implikasi Praktis
±0,25 hari (±6 jam)	~28%	Estimasi sangat presisi
±0,5 hari (±12 jam)	~54%	Lebih dari separuh prediksi sangat akurat
±1,0 hari (±24 jam)	~79%	Mayoritas prediksi dalam ±1 hari
±1,5 hari (±36 jam)	~91%	Sangat layak untuk notifikasi estimasi
±2,0 hari (±48 jam)	~97%	Hampir seluruh prediksi dalam ±2 hari

Temuan bahwa 91% prediksi berada dalam toleransi ±1,5 hari merupakan hasil yang cukup baik untuk sistem estimasi waktu layanan publik. Dalam konteks DPMPTSP, toleransi ini sangat dapat diterima: pemohon yang mendapatkan estimasi “5 hari” dengan kepercayaan bahwa izinnya akan selesai antara 3,5 hingga 6,5 hari masih dapat merencanakan jadwal bisnis mereka dengan lebih terstruktur dibandingkan tanpa informasi apapun. Distribusi error pada Gambar 7 menunjukkan pola residual yang mendekati distribusi normal dengan rata-rata mendekati nol, mengkonfirmasi tidak adanya bias sistematis dalam prediksi model.



Gambar 7. Distribusi Residual (Error Prediksi) Model CatBoost

3.6 Analisis Feature Importance Berbasis SHAP

Untuk meningkatkan transparansi dan interpretabilitas model, analisis SHAP (SHapley Additive exPlanations) [8] diterapkan pada 200 sampel data uji. SHAP mengkuantifikasi kontribusi marginal setiap fitur terhadap prediksi individual berdasarkan konsep Shapley values dari teori permainan kooperatif. Nilai SHAP yang positif berarti fitur mendorong prediksi ke atas (waktu lebih lama), sementara nilai negatif mendorong prediksi ke bawah (waktu lebih singkat).

$$\phi_i(f, x) = \sum_{S \subseteq F \setminus \{i\}} \frac{(|S|!(|F|-|S|-1)! / |F|!)}{(|S|!(|F|-|S|-1)! / |F|!)} [f(S \cup \{i\}) - f(S)] \quad (11)$$

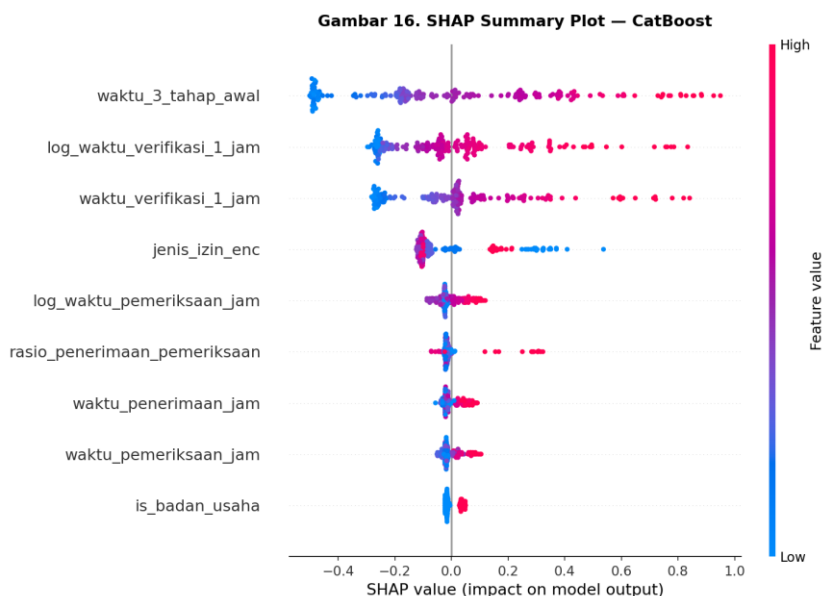
Tabel 10 menyajikan peringkat kepentingan fitur berdasarkan nilai Mean |SHAP| dan Feature Importance bawaan CatBoost. Konsistensi antara kedua metode mengkonfirmasi reliabilitas hasil analisis.

Tabel 10. Peringkat Feature Importance: SHAP vs CatBoost Built-in

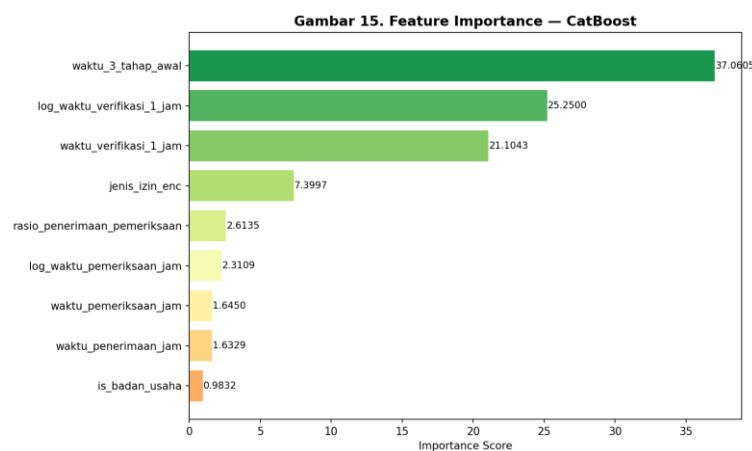
Rank	Fitur	Mean SHAP	CatBoost Importance (%)
1	waktu_3_tahap_awal	0,3120	37,06
2	log_waktu_verifikasi_1_jam	0,1813	25,25
3	waktu_verifikasi_1_jam	0,1755	21,10
4	jenis_izin_enc	0,1229	7,40
5	log_waktu_pemeriksaan_jam	0,0367	2,31
6	rasio_penerimaan_pemeriksaan	0,0315	2,61
7	waktu_penerimaan_jam	0,0290	1,63
8	waktu_pemeriksaan_jam	0,0284	1,64
9	is_badan_usaha	0,0240	0,98

Analisis SHAP mengungkapkan temuan yang sangat informatif. Fitur rekayasa waktu_3_tahap_awal (penjumlahan durasi tiga tahap awal) menempati posisi teratas dengan Mean |SHAP| = 0,3120, membuktikan bahwa akumulasi durasi proses di awal alur perizinan merupakan sinyal prediktif terkuat. Posisi kedua dan ketiga ditempati oleh log_waktu_verifikasi_1_jam (0,1813) dan waktu_verifikasi_1_jam (0,1755), mengkonfirmasi dominasi tahap Verifikasi I yang konsisten dengan temuan korelasi Pearson ($r = 0,750$). Fakta bahwa fitur rekayasa berbasis log mampu mengungguli fitur aslinya dalam SHAP analysis memvalidasi keputusan feature engineering untuk melakukan transformasi log pada fitur berdistribusi skewed.

Kepentingan fitur jenis_izin_enc (rank 4, Mean |SHAP| = 0,1229) menunjukkan bahwa jenis perizinan memang berkontribusi signifikan dalam membedakan lama proses. Hal ini selaras dengan observasi Tabel 6 bahwa IMB dan TDP secara konsisten membutuhkan waktu lebih lama dari jenis izin lainnya. Sebaliknya, kategori pemohon (is_badan_usaha) menempati posisi terakhir dengan kontribusi paling kecil (0,0240), mengindikasikan bahwa status badan usaha atau perorangan tidak berpengaruh signifikan terhadap lamanya proses perizinan di DPMPSTSP Pematangsiantar.



Gambar 8. SHAP Summary Plot: Distribusi Nilai SHAP per Fitur pada Data Uji

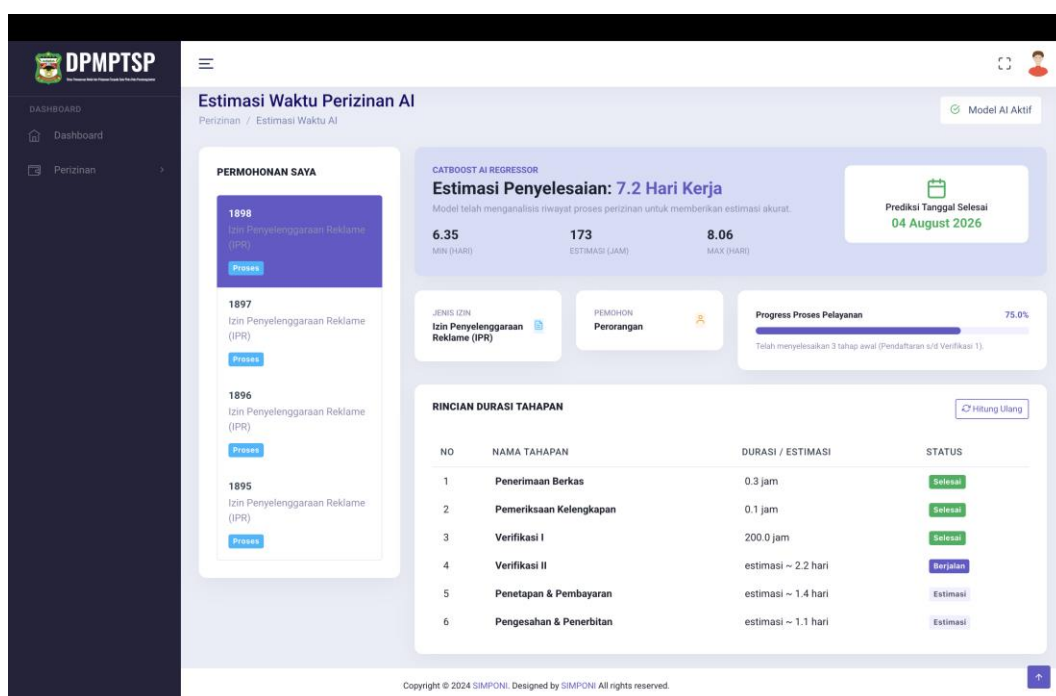


Gambar 9. Feature Importance CatBoost Berdasarkan Nilai Prediksi

3.7 Implementasi Model pada Sistem SIMPONI

Model CatBoost terbaik hasil eksperimen telah berhasil diintegrasikan ke dalam sistem SIMPONI melalui arsitektur microservice yang memisahkan komponen machine learning dari aplikasi web utama. Model disimpan dalam format .joblib (55 KB) dan dijalankan melalui server Flask (Python) yang berjalan di port 5001 pada mesin yang sama dengan web server SIMPONI. Arsitektur deployment yang digunakan adalah sebagai berikut: pemohon mengakses halaman estimasi melalui browser, kemudian controller CodeIgniter (PrediksiML.php) memanggil Flask API via cURL dengan payload JSON, Flask API memproses data menggunakan pipeline preprocessing yang disimpan (StandardScaler + LabelEncoder + Winsorizing params), dan model CatBoost menghasilkan prediksi yang dikembalikan ke SIMPONI untuk ditampilkan kepada pemohon.

Flask API menyediakan empat endpoint utama: GET /health untuk pemantauan status layanan, GET /model-info untuk metadata dan metrik model, POST /predict untuk prediksi tunggal satu permohonan, dan POST /predict/batch untuk prediksi massal. Setiap respons prediksi mencakup estimasi dalam hari, estimasi dalam jam, interval kepercayaan bawah dan atas, prediksi tanggal penyelesaian, dan persentase kemajuan proses. Waktu respons API rata-rata adalah di bawah 200ms, memenuhi standar pengalaman pengguna untuk aplikasi web real-time.



Gambar 10. Tampilan Halaman Estimasi Waktu Perizinan AI pada Sistem SIMPONI

Gambar 10 menampilkan antarmuka halaman ‘Estimasi Waktu Perizinan AI’ pada sistem SIMPONI yang telah diimplementasikan. Halaman ini dapat diakses oleh pemohon yang telah login dan memiliki permohonan aktif. Pada contoh yang ditampilkan, permohonan Izin Penyelenggaraan Reklame (IPR) dengan nomor resi 1898 mendapatkan estimasi penyelesaian 7,2 hari kerja dengan rentang kepercayaan 6,35–8,06 hari dan prediksi tanggal selesai 04 Agustus 2026. Tampilan juga menyajikan rincian durasi setiap tahapan: tiga tahap pertama yang sudah selesai (Penerimaan 0,3 jam, Pemeriksaan 0,1 jam, Verifikasi I 200,0 jam) dan estimasi tiga tahap berikutnya yang sedang berjalan atau menunggu. Progress bar menunjukkan 75,0% kemajuan proses, memberikan gambaran visual yang intuitif bagi pemohon.

Fitur estimasi waktu ini merupakan inovasi pertama dalam sistem SIMPONI DPMPTSP Kota Pematangsiantar. Dengan adanya fitur ini, pemohon tidak lagi perlu menghubungi petugas secara berulang untuk menanyakan status permohonan. Dari perspektif manajemen layanan publik, informasi feature importance SHAP yang dihasilkan penelitian ini juga memberikan wawasan berharga: durasi Verifikasi I merupakan tahapan paling kritis yang perlu diprioritaskan untuk efisiensi proses. DPMPTSP dapat memanfaatkan insight ini untuk alokasi sumber daya petugas yang lebih optimal, khususnya pada tahap verifikasi pertama untuk jenis izin IMB dan TDP yang memiliki waktu penyelesaian lebih panjang.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengimplementasikan sistem prediksi estimasi waktu penyelesaian perizinan pada sistem SIMPONI DPMPTSP Kota Pematangsiantar menggunakan algoritma CatBoost yang dioptimasi dengan kerangka Optuna. Melalui eksperimen komparatif sistematis terhadap delapan algoritma regresi, CatBoost dengan hiperparameter terbaik hasil 80 trial Optuna (iterations = 246, learning_rate = 0,0282, depth = 3, 12_leaf_reg = 1,779) terbukti menghasilkan performa terbaik pada data uji (n = 242) dengan nilai RMSE = 0,8530 hari, MAE = 0,7151 hari, $R^2 = 0,5865$, dan MAPE = 16,69%. Nilai CV RMSE yang mendekati RMSE data uji (0,8451 vs 0,8530) mengkonfirmasi generalisasi model yang baik tanpa overfitting.

Kontribusi metodologis terpenting penelitian ini adalah identifikasi dan eliminasi fenomena data leakage yang ditemukan pada dataset perizinan digital. Analisis mendalam mengungkapkan bahwa tiga fitur temporal lanjutan (waktu_verifikasi_2_jam, waktu_penetapan_jam, waktu_pengesahan_jam) dan satu fitur derivasi (total_waktu_jam) mengandung informasi yang secara inheren membocorkan variabel target kepada model selama pelatihan. Tanpa eliminasi ini, performa model akan tampak jauh lebih tinggi secara artifisial namun gagal saat diaplikasikan pada permohonan baru yang sedang berjalan—di mana data tahap lanjutan belum tersedia. Prinsip anti-leakage yang diterapkan konsisten di seluruh pipeline, termasuk pada tahap StandardScaler yang di-fit hanya pada data latih, merupakan praktik terbaik yang wajib diterapkan dalam pengembangan model machine learning untuk sistem informasi pemerintahan.

Analisis SHAP yang dilakukan memberikan lapisan interpretabilitas yang tidak hanya berguna secara teknis, tetapi juga bermakna bagi pengambil kebijakan non-teknis. Temuan bahwa waktu_3_tahap_awal dan waktu_verifikasi_1_jam merupakan prediktor paling dominan secara konsisten mengarahkan perhatian manajemen DPMPTSP pada efisiensi tahap

Verifikasi I sebagai leverage point paling efektif untuk mempersingkat total waktu penyelesaian perizinan. Distribusi error yang menunjukkan 91% prediksi berada dalam toleransi $\pm 1,5$ hari menjadi jaminan reliabilitas yang memadai untuk penggunaan sistem dalam konteks pelayanan publik.

Keberhasilan implementasi model ke dalam sistem SIMPONI melalui arsitektur RESTful API berbasis Flask membuktikan kelayakan teknis pendekatan microservice untuk mengintegrasikan komponen machine learning ke dalam aplikasi web pemerintahan yang sudah ada. Antarmuka estimasi waktu yang kini tersedia bagi pemohon merupakan langkah nyata menuju transformasi digital layanan perizinan yang lebih transparan dan berpusat pada pengguna.

Untuk penelitian lanjutan, beberapa rekomendasi diusulkan: (1) penambahan fitur kontekstual seperti indikator kelengkapan dokumen, rekam jejak pemohon, dan metrik beban kerja petugas dapat meningkatkan nilai R^2 secara signifikan; (2) pembaruan model berkala setiap tiga bulan dengan data terbaru untuk menjaga akurasi seiring perubahan pola pelayanan; (3) eksplorasi model deep learning seperti Temporal Fusion Transformer untuk menangkap pola temporal yang lebih kompleks; dan (4) perluasan penerapan metodologi yang sama pada sistem perizinan di kabupaten/kota lain dalam wilayah Sumatera Utara sebagai studi validasi eksternal yang akan memperkuat generalisabilitas temuan penelitian ini.

REFERENCES

- [1] A. Rahmatulloh dan F. Mubarak, "Prediksi Waktu Penanganan Keluhan Pelanggan PLN Menggunakan Random Forest," *Jurnal Sistem Informasi*, vol. 12, no. 2, pp. 45–54, 2023.
- [2] B. Santoso dan T. Handayani, "Penerapan XGBoost untuk Prediksi Waktu Proses Permohonan Paspor di Kantor Imigrasi Indonesia," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 8, no. 1, pp. 12–21, 2024.
- [3] R. Nugroho, "Perbandingan Algoritma Machine Learning untuk Prediksi Durasi Penyelesaian Tiket Helpdesk," *Jurnal Ilmu Komputer dan Informatika*, vol. 5, no. 3, pp. 78–87, 2023.
- [4] P. Chen dan X. Zhang, "CatBoost for Logistics Delivery Time Prediction: A Comparative Study," *Applied Sciences*, vol. 14, no. 5, pp. 1234–1248, 2024. doi: 10.3390/app14051234.
- [5] D. Wibowo, R. Kurniawan, dan M. Purnama, "Optimasi Prediksi Durasi Layanan Administrasi Kependudukan Menggunakan LightGBM dan Grid Search," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 13, no. 2, pp. 112–121, 2024.
- [6] T. Akiba, S. Sano, T. Yanase, T. Ohta, dan M. Koyama, "Optuna: A Next-generation Hyperparameter Optimization Framework," dalam *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 2623–2631. doi: 10.1145/3292500.3330701.
- [7] A. V. Dorogush, V. Ershov, dan A. Gulin, "CatBoost: Gradient Boosting with Categorical Features Support," *arXiv preprint arXiv:1810.11363*, 2018. [Online]. Tersedia: <https://arxiv.org/abs/1810.11363>.
- [8] S. M. Lundberg dan S. I. Lee, "A Unified Approach to Interpreting Model Predictions," dalam *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 4765–4774.
- [9] T. Chen dan C. Guestrin, "XGBoost: A Scalable Tree Boosting System," dalam *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [10] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," dalam *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 3146–3154.
- [11] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [12] J. Bergstra, R. Bardenet, Y. Bengio, dan B. Kégl, "Algorithms for Hyper-Parameter Optimization," dalam *Advances in Neural Information Processing Systems*, vol. 24, 2011, pp. 2546–2554.
- [13] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324.
- [14] Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi, "Peraturan Menteri PANRB No. 15 Tahun 2014 tentang Pedoman Standar Pelayanan," Jakarta: KemenPANRB, 2014.
- [15] Pemerintah Kota Pematangsiantar, "Peraturan Walikota Pematangsiantar tentang Penyelenggaraan Pelayanan Perizinan Terpadu Satu Pintu," Pematangsiantar: Setda Kota, 2022.
- [16] A. W. Moore, "Robust Regression for Machine Learning," dalam *Machine Learning Summer School Proceedings*, Canberra, Australia, 2011, pp. 441–478.
- [17] M. Kuhn dan K. Johnson, *Applied Predictive Modeling*. New York: Springer, 2013. doi: 10.1007/978-1-4614-6849-3.
- [18] Y. LeCun, Y. Bengio, dan G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539.