

**Ranah Research****Journal of Multidisciplinary Research and Development**

082170743613

ranahresearch@gmail.com

<https://jurnal.ranahresearch.com>

E-ISSN: 2655-0865

DOI: <https://doi.org/10.38035/rj.v8i1>
<https://creativecommons.org/licenses/by/4.0/>

Otomatisasi Klasifikasi Tingkat Urgensi Keluhan E-Layanan Unesa Berbasis TF-IDF dan Logistic Regression

Intan Alpiana¹, Wiyli Yustanti², Yuni Yamasari³¹State University of Surabaya, Surabaya, Indonesia, intanalpiana@unesa.ac.id²State University of Surabaya, Surabaya, Indonesia, wiyliyustanti@unesa.ac.id³State University of Surabaya, Surabaya, Indonesia, yuniyamasari@unesa.ac.idCorresponding Author: intanalpiana@unesa.ac.id¹

Abstract: *The development of digital technology requires universities to provide fast, accurate, and responsive academic services. Surabaya State University (Unesa), through its E-Services platform, provides a means for the academic community to submit complaints related to obstacles in using information systems and networks. However, the process of classifying the urgency level of complaints is still done manually by administrators, which has the potential to cause delays in handling, inconsistencies in assessment, and an increased workload. This study aims to develop an automated system for classifying the urgency level of complaints by utilizing Term Frequency-Inverse Document Frequency (TF-IDF) as a text feature representation and weighted Logistic Regression (class_weight) as the main classification model. The dataset used consisted of 79,303 complaints, divided into training data (70%), validation data (15%), and test data (15%). Model evaluation was performed using accuracy, precision, recall, F1-score, and confusion matrix metrics. The results of the study show that the combination of TF-IDF and weighted Logistic Regression is capable of providing good performance with an accuracy of 92.54% on test data. In addition, the model demonstrates a high ability to accurately detect critical complaints, ensuring that treatment priorities are maintained optimally. These findings confirm that the application of machine learning-based models can improve operational efficiency and classification consistency compared to manual approaches. The developed system is expected to be further integrated into the Unesa E-Service platform, supporting the automatic and real-time complaint handling process, and helping the administration focus on resolving the most urgent issues.*

Keyword: *E-Services, Urgency Classification, TF-IDF, Logistic Regression, Automation*

Abstrak: Perkembangan teknologi digital menuntut perguruan tinggi untuk menghadirkan layanan akademik yang cepat, tepat, dan responsif. Universitas Negeri Surabaya (Unesa) melalui platform E-Layanan memberikan sarana bagi civitas akademika untuk menyampaikan keluhan terkait kendala penggunaan sistem informasi dan jaringan. Namun, proses klasifikasi tingkat urgensi keluhan masih dilakukan secara manual oleh admin, yang

berpotensi menyebabkan keterlambatan penanganan, inkonsistensi penilaian, serta meningkatnya beban kerja. Penelitian ini bertujuan untuk mengembangkan sistem otomatisasi klasifikasi tingkat urgensi keluhan dengan memanfaatkan Term Frequency-Inverse Document Frequency (TF-IDF) sebagai representasi fitur teks, serta Logistic Regression berbobot (*class_weight*) sebagai model klasifikasi utama. Dataset yang digunakan terdiri dari 79.303 keluhan, dibagi menjadi data latih (70%), validasi (15%), dan uji (15%). Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, dan confusion matrix. Hasil penelitian menunjukkan bahwa kombinasi TF-IDF dan Logistic Regression berbobot mampu memberikan kinerja yang baik dengan akurasi 92,54% pada data uji. Selain itu, model menunjukkan kemampuan yang tinggi dalam mendeteksi keluhan kritis secara akurat, memastikan prioritas penanganan terjaga secara optimal. Temuan ini menegaskan bahwa penerapan model berbasis pembelajaran mesin dapat meningkatkan efisiensi operasional dan konsistensi klasifikasi dibandingkan pendekatan manual. Sistem yang dikembangkan diharapkan dapat diintegrasikan lebih lanjut ke dalam platform E-Layanan Unesa, mendukung proses penanganan keluhan secara otomatis dan real-time, serta membantu administrasi fokus pada resolusi masalah yang paling mendesak.

Kata Kunci: E-Layanan, Klasifikasi Urgensi, TF-IDF, Logistic Regression, Otomatisasi

PENDAHULUAN

Perkembangan teknologi informasi dalam beberapa tahun terakhir telah membawa perubahan signifikan pada berbagai sektor, termasuk pendidikan tinggi. Dalam konteks ini, universitas tidak hanya berperan sebagai penyelenggara pendidikan, penelitian, dan pengabdian kepada masyarakat, tetapi juga dituntut menghadirkan layanan akademik dan administrasi yang cepat, transparan, serta responsif terhadap kebutuhan pemangku kepentingan (Susanto & Wulandari, 2022; Yamasari et al., 2024). Sejalan dengan tuntutan era digital, layanan berbasis teknologi atau e-layanan menjadi solusi untuk menjembatani interaksi antara mahasiswa, dosen, tenaga kependidikan, dan institusi secara lebih efektif (Wijaya & Lestari, 2020). Universitas Negeri Surabaya (Unesa) merespons kebutuhan tersebut dengan mengembangkan platform E-Layanan sebagai bagian dari transformasi digital untuk mendukung berbagai aktivitas akademik maupun non-akademik civitas akademika. Melalui platform ini, mahasiswa maupun dosen dapat menyampaikan keluhan terkait penggunaan sistem informasi dan jaringan, dengan tujuan mempercepat komunikasi dan memberikan solusi yang terstruktur. Akan tetapi, dalam implementasinya, pengelolaan keluhan masih menghadapi berbagai tantangan serius, khususnya pada tahap klasifikasi tingkat urgensi, karena hingga saat ini proses klasifikasi dilakukan secara manual oleh admin. Proses manual ini berpotensi menimbulkan keterlambatan tindak lanjut, inkonsistensi penilaian antar petugas, human error, serta peningkatan beban kerja, terutama ketika jumlah keluhan yang masuk sangat besar dalam periode tertentu (Fauzi & Haryanto, 2022; Santika & Prasetyo, 2023). Konsekuensinya, kualitas layanan dapat menurun, waktu respons menjadi lebih lama, dan tingkat kepuasan mahasiswa sebagai pengguna utama platform cenderung berkurang (Hidayat & Santoso, 2020; Syafitri & Pranata, 2020).

Kondisi ini menunjukkan adanya research gap dalam pengelolaan layanan digital di perguruan tinggi Indonesia, di mana penelitian terdahulu banyak berfokus pada penggunaan text mining dan pembelajaran mesin dalam analisis sentimen media sosial, klasifikasi topik, maupun sistem rekomendasi (Kurniawan & Nugroho, 2024; Nugraha & Puspitasari, 2024), tetapi sangat sedikit yang secara khusus menyoroti klasifikasi urgensi keluhan layanan akademik mahasiswa berbasis data riil dari sistem universitas. Padahal, urgensi keluhan merupakan indikator penting yang menentukan prioritas penanganan masalah, di mana keluhan kritis seperti gangguan login, error sistem akademik, atau kegagalan server harus

segera ditangani agar tidak mengganggu aktivitas akademik secara luas (Rahman & Hossain, 2021). Penelitian sebelumnya memang telah mengeksplorasi metode seperti support vector machine, naïve bayes, hingga XGBoost untuk klasifikasi teks (Putra & Anggraini, 2021; Pratama & Setiawan, 2023), namun mayoritas penelitian tersebut terbatas pada dataset kecil atau domain yang berbeda dari konteks layanan pendidikan tinggi. Dengan demikian, penelitian ini menempati ruang kosong yang relevan sekaligus mendesak, karena mampu memberikan kontribusi nyata pada sistem manajemen layanan kampus di Indonesia.

Keunikan penelitian ini terletak pada penggunaan dataset berskala besar yang berjumlah 79.303 data keluhan riil dari platform E-Layanan Unesa, yang menjadikannya lebih komprehensif dan representatif dibandingkan studi sebelumnya yang umumnya berbasis pada sampel kecil atau data sekunder. Selain itu, penelitian ini menggunakan pendekatan sederhana tetapi efektif, yaitu kombinasi Term Frequency–Inverse Document Frequency (TF-IDF) sebagai representasi teks dan Logistic Regression berbobot (*class_weight*) sebagai algoritme utama untuk menangani masalah ketidakseimbangan kelas. TF-IDF dipilih karena kemampuannya menekankan kata-kata penting dalam teks keluhan, seperti “tidak bisa login”, “darurat”, atau “server error”, yang sering menjadi penanda tingkat urgensi, sementara Logistic Regression berbobot dipilih karena sifatnya yang interpretatif dan kemampuannya menyesuaikan bobot kesalahan pada kelas minoritas sehingga recall untuk kelas urgensi tinggi dapat lebih optimal (Anindita, 2025; Pratama & Setiawan, 2023). Dengan kombinasi ini, penelitian diharapkan dapat menghasilkan sistem klasifikasi yang akurat sekaligus transparan, sehingga pihak pengelola layanan dapat memahami alasan di balik keputusan klasifikasi.

Tren terkini menunjukkan bahwa otomatisasi layanan digital di pendidikan tinggi telah menjadi kebutuhan strategis untuk meningkatkan efisiensi operasional sekaligus menjaga kualitas layanan (Santika & Prasetyo, 2023; Nugraha & Puspitasari, 2024; Yamasari et al., 2024). Perguruan tinggi di berbagai negara, khususnya di Asia dan Eropa, mulai mengadopsi teknologi machine learning dan artificial intelligence dalam menangani keluhan, memberikan rekomendasi layanan, hingga mengoptimalkan penjadwalan akademik. Namun, di Indonesia penerapan inovasi serupa masih relatif terbatas, terutama dalam konteks pengelolaan keluhan mahasiswa yang membutuhkan respons cepat dan akurat. Berangkat dari kondisi tersebut, penelitian ini memberikan kontribusi praktis berupa rancangan sistem klasifikasi otomatis tingkat urgensi keluhan yang dapat diintegrasikan ke dalam platform E-Layanan Unesa untuk mendukung pengelolaan keluhan secara real-time. Selain itu, penelitian ini juga menawarkan kontribusi teoretis dengan memperkuat kajian penerapan text mining dan machine learning di ranah layanan pendidikan tinggi di Indonesia. Dengan demikian, studi ini tidak hanya menguji sejauh mana kombinasi TF-IDF dan Logistic Regression berbobot mampu menghasilkan klasifikasi urgensi yang akurat, tetapi juga menghadirkan solusi praktis yang berpotensi meningkatkan responsivitas, transparansi, dan kualitas layanan akademik di era digital.

METODE

Penelitian ini menggunakan pendekatan CRISP-DM (Cross Industry Standard Process for Data Mining) sebagai kerangka kerja untuk merancang klasifikasi tingkat urgensi keluhan pada E-Layanan Universitas Negeri Surabaya. CRISP-DM merupakan metodologi yang sistematis, fleksibel, dan banyak digunakan dalam pengembangan proyek berbasis data mining (Yustanti et al., 2024; Yamasari et al., 2024). Metodologi ini membagi proses pengembangan menjadi enam tahap utama: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Berikut ini merupakan beberapa tahapan pada metode CRISP-DM:

CRISP-DM : The Cross-Industry Standard Process for Data Mining

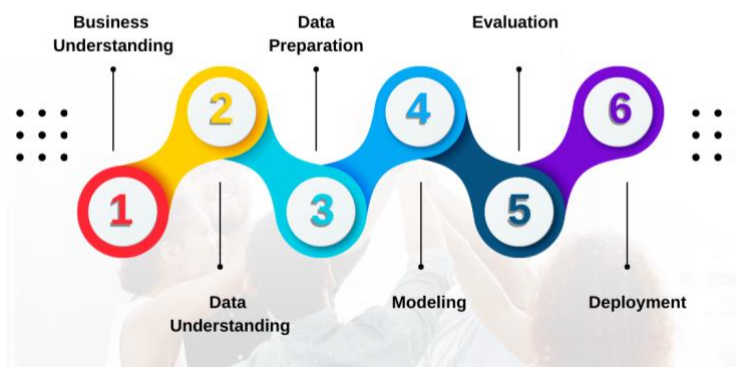


Figure 1. CRISP-DM Architecture

Business Understanding

Tahap Business Understanding berfokus pada identifikasi masalah utama yang dihadapi organisasi serta tujuan yang ingin dicapai melalui penerapan data mining (Pratama & Setiawan, 2023; Susanto & Wulandari, 2022). Dalam konteks penelitian ini, permasalahan yang muncul adalah tingginya jumlah keluhan yang masuk pada platform E-Layanan Unesa dan belum adanya mekanisme otomatis untuk mengklasifikasikan tingkat urgensi keluhan. Kondisi ini sejalan dengan temuan Putra & Anggraini (2021) bahwa proses manual dalam pengelolaan keluhan berpotensi menimbulkan keterlambatan penanganan, inkonsistensi penilaian, serta menurunkan kepuasan pengguna layanan.

Tujuan bisnis dari penelitian ini adalah membangun sebuah sistem cerdas yang mampu melakukan klasifikasi otomatis terhadap tingkat urgensi keluhan mahasiswa ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi (Rahman & Hossain, 2021; Kurniawan & Nugroho, 2024). Dengan adanya sistem ini, pengelola layanan dapat memprioritaskan penanganan keluhan secara lebih efektif dan efisien, sekaligus meminimalkan risiko human error dalam proses klasifikasi (Hidayat & Santoso, 2020).

Manfaat yang diharapkan dari tahap ini mencakup peningkatan efisiensi operasional, pengurangan risiko keterlambatan penanganan keluhan, serta peningkatan kepuasan mahasiswa sebagai pengguna utama layanan (Fauzi & Haryanto, 2022; Santika & Prasetyo, 2023). Selain itu, hasil penelitian ini dapat dijadikan dasar pengembangan sistem decision support untuk pengelolaan keluhan berbasis teknologi informasi di perguruan tinggi (Nugraha & Puspitasari, 2024; Syafitri & Pranata, 2020), sehingga mendukung pengambilan keputusan yang lebih cepat, tepat, dan konsisten dalam konteks layanan digital.

Data Understanding

Dataset yang digunakan merupakan data historis keluhan pengguna E-Layanan UNESA dari tahun 2017-2024. Dataset terdiri dari atribut berikut ini

Table 1. Dataset e-Layanan Unesa

Atribut	Deskripsi
ID Keluhan	Nomor unik setiap keluhan pengguna.
Subjek Keluhan	Subjek utama dari keluhan yang diajukan.
Uraian Keluhan	Deskripsi lengkap mengenai masalah yang dialami pengguna.
Tahun Keluhan	Tahun pengajuan keluhan.
ID PIC	Identitas PIC (penanggung jawab) masing-masing sistem informasi.
Kategori	Kategori utama dari keluhan.
Subkategori	Subkategori spesifik dari keluhan.

Data Preparation

Tahap Data Preparation mencakup seluruh proses pengolahan dataset agar siap digunakan dalam pemodelan (Pratama & Setiawan, 2023; Susanto & Wulandari, 2022). Pada penelitian ini, data yang digunakan berasal dari keluhan civitas akademika di E-Layanan Unesa dengan format CSV yang berisi atribut seperti id keluhan, subjek keluhan, uraian keluhan, tahun, idpic, kategori, dan subkategori (Putra & Anggraini, 2021; Rahman & Hossain, 2021). Namun, karena fokus penelitian adalah pada klasifikasi tingkat urgensi keluhan, maka kolom yang paling relevan adalah uraian keluhan (free text) sebagai fitur utama dan urgensi sebagai label target (Kurniawan & Nugroho, 2024; Fauzi & Haryanto, 2022).

Table 2. Persiapan dataset penelitian

Tahap	Deskripsi	Hasil
Data Cleaning	Menghapus <i>missing values</i> , normalisasi teks (lowercase), tipe data konsisten	Semua uraian siap diproses tanpa nilai kosong
Pembuatan Label	Rule-based dengan kata kunci (darurat, lambat, error, dll.)	Label urgensi: Rendah (0), Sedang (1), Tinggi (2)
Data Splitting	Train 70%, Validation 15%, Test 15%	Train: 55.512, Val: 11.895, Test: 11.896
TF-IDF Vectorization	Konversi teks ke fitur numerik, max_features=3000, ngram=(1,2)	Vektor sparse yang merepresentasikan keluhan mahasiswa
Handling Imbalance	Perhitungan <i>class weights</i> dengan <i>compute class weight</i>	Bobot kelas digunakan saat training untuk meningkatkan performa model

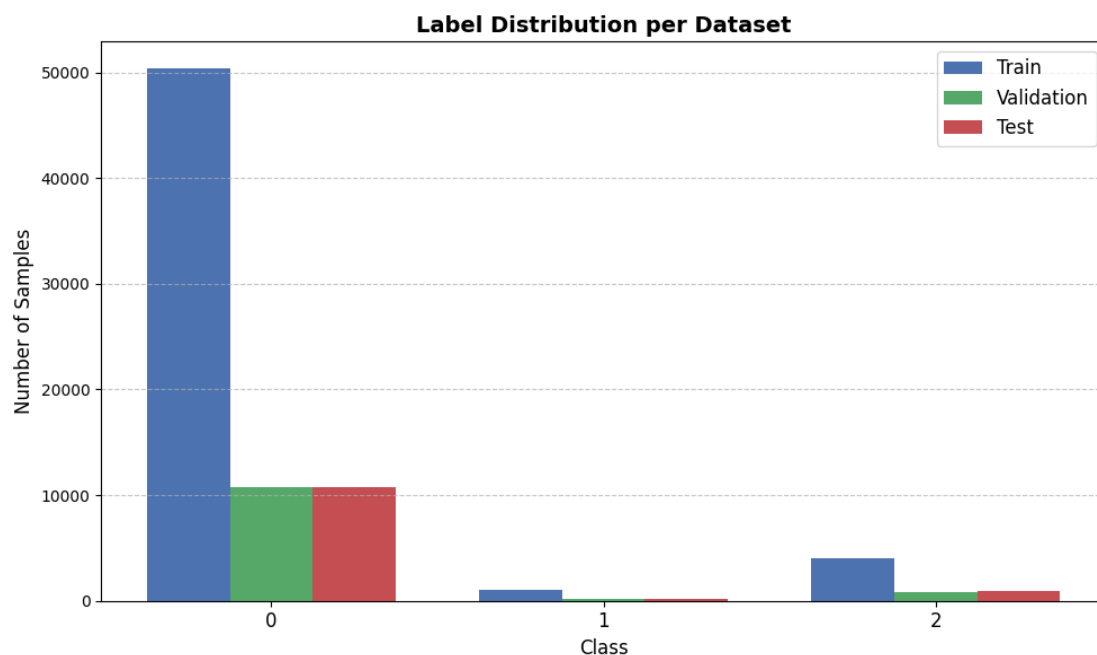


Figure 2. Label distribution per dataset

Data Modeling

Tahap Modeling merupakan inti dari penelitian ini, di mana dilakukan pembangunan model pembelajaran mesin untuk mengklasifikasikan tingkat urgensi keluhan berdasarkan uraian teks (Pratama & Setiawan, 2023; Kurniawan & Nugroho, 2024). Proses pemodelan dalam penelitian ini menggunakan kombinasi TF-IDF sebagai representasi teks dan algoritme klasifikasi, yaitu XGBoost sebagai model utama dan Logistic Regression sebagai fallback (Susanto & Wulandari, 2022; Putra & Anggraini, 2021).

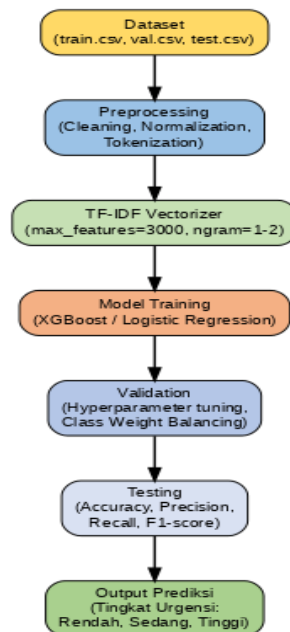


Figure 3. Data modeling

Table 3. Data Modeling Ekstraksi Subjek Utama

Komponen	Deskripsi
Representasi Fitur	TF-IDF, max_features=3000, ngram=(1,2)
Algoritme Utama	XGBoost, objective=multi:softprob, eval_metric=mlogloss, early stopping
Algoritme Fallback	Logistic Regression (saga, balanced, max_iter=2000)
Penyeimbangan Kelas	Class weight berdasarkan distribusi data (0.36, 16.97, 4.59)
Dataset	Train (70%), Validation (15%), Test (15%)
Evaluasi	Accuracy, Precision, Recall, F1-Score, Confusion Matrix

Evaluation

Tahap Evaluation bertujuan untuk menilai performa model yang telah dibangun, apakah sesuai dengan tujuan penelitian dan relevan dengan permasalahan bisnis yang telah ditentukan (Pratama & Setiawan, 2023; Susanto & Wulandari, 2022). Pada penelitian ini, evaluasi dilakukan dengan menggunakan data uji yang terpisah dari data latih, sehingga hasil pengukuran dapat lebih objektif (Putra & Anggraini, 2021; Kurniawan & Nugroho, 2024).

Table 4. Metrik evaluasi klasifikasi tingkat urgensi keluhan

Metrik	Definisi	Relevansi dalam Penelitian
Accuracy	Proporsi prediksi benar dibandingkan dengan total data uji.	Memberikan gambaran umum performa model dalam mengklasifikasikan tingkat urgensi keluhan.
Precision	Tingkat ketepatan prediksi positif, yaitu seberapa besar prediksi positif benar adanya.	Penting untuk memastikan bahwa keluhan yang diprediksi mendesak benar-benar termasuk mendesak.
Recall	Kemampuan model dalam menemukan semua data yang benar pada suatu kelas.	Menjamin bahwa keluhan dengan urgensi tinggi tidak banyak terlewat.
F1-Score	Rata-rata harmonis antara precision dan recall.	Digunakan untuk menyeimbangkan ketepatan (precision) dan kelengkapan (recall).
Confusion Matrix	Matriks yang menggambarkan distribusi prediksi terhadap kelas sebenarnya.	Membantu melihat detail kesalahan klasifikasi antar kategori urgensi (rendah, sedang, tinggi).

Deployment

Tahap Deployment dalam penelitian ini berfokus pada pemanfaatan model klasifikasi yang telah dilatih agar dapat digunakan secara berulang tanpa perlu dilakukan pelatihan ulang (Pratama & Setiawan, 2023; Susanto & Wulandari, 2022). Model beserta vectorizer disimpan dalam format file .pkl, sehingga memungkinkan dipanggil kembali ketika terdapat data keluhan baru (Putra & Anggraini, 2021; Kurniawan & Nugroho, 2024). Dengan mekanisme ini, sistem mampu melakukan prediksi tingkat urgensi keluhan berdasarkan uraian yang dimasukkan. Proses deployment dilakukan secara terstruktur agar model siap digunakan baik untuk pengujian lebih lanjut maupun sebagai prototipe sistem klasifikasi otomatis (Rahman & Hossain, 2021; Fauzi & Haryanto, 2022).

Table 5. Mekanisme klasifikasi tingkat urgensi keluhan

Tahap	Deskripsi	Output
Penyimpanan Model	Model klasifikasi (<i>Logistic Regression</i> atau <i>XGBoost</i>) yang sudah dilatih disimpan dalam format.pkl bersama dengan <i>TF-IDF vectorizer</i> .	File model_tfidf_xgb.pkl dan tfidf_vectorizer.pkl.
Input Data Baru	Pengguna atau admin memasukkan uraian keluhan baru terkait layanan sistem informasi atau jaringan.	Teks uraian keluhan baru.
Transformasi Data	Uraian keluhan baru diubah menjadi representasi numerik menggunakan <i>TF-IDF vectorizer</i> yang telah disimpan.	Vektor TF-IDF dari uraian keluhan.
Prediksi Urgensi	Model yang telah disimpan melakukan klasifikasi terhadap data baru berdasarkan pola yang dipelajari dari data latih.	Label urgensi (Rendah, Sedang, Tinggi).
Penyajian Hasil	Hasil prediksi ditampilkan dalam bentuk label tingkat urgensi sehingga dapat dipahami dengan mudah.	Prediksi akhir tingkat urgensi.

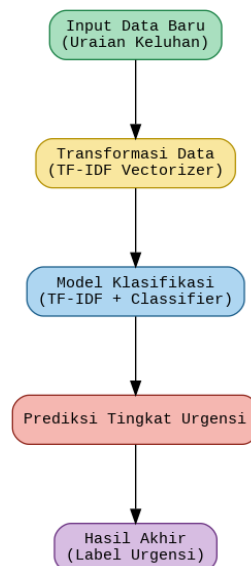


Figure 4. Deployment Architecture

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengembangkan sistem otomatisasi klasifikasi tingkat urgensi keluhan pada platform E-Layanan Universitas Negeri Surabaya (Unesa), khususnya untuk keluhan terkait kendala penggunaan sistem informasi dan jaringan. Otomatisasi ini diharapkan dapat mempercepat proses penanganan keluhan, mengurangi inkonsistensi klasifikasi, serta meringankan beban kerja admin.

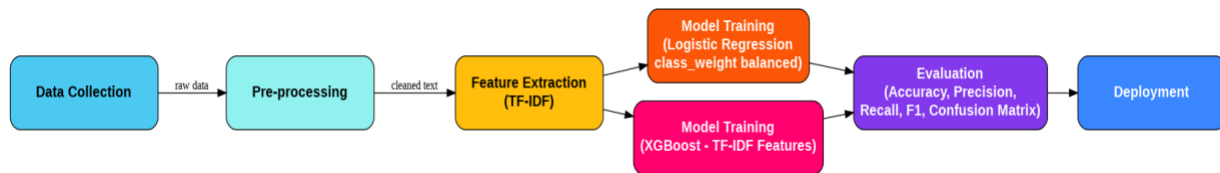


Figure 5. Flow map penelitian

Pra-pemrosesan dan Distribusi Data

Penelitian ini menggunakan dataset keluhan E-Layanan Universitas Negeri Surabaya (Unesa) sebanyak 79.303 data, yang dibagi menjadi data latih (70%, $n=55.512$), data validasi (15%, $n=11.895$), dan data uji (15%, $n=11.896$). Setiap uraian keluhan telah diproses menjadi string, mengisi nilai kosong, dan divalidasi untuk memastikan konsistensi data.

Distribusi label tingkat urgensi menunjukkan ketidakseimbangan yang signifikan: kelas Rendah mendominasi ($\approx 90\%$), kelas Tinggi lebih sedikit ($\approx 5\%$), dan kelas Sedang merupakan kelas minoritas terkecil ($\approx 1-2\%$). Ketidakseimbangan ini menuntut penggunaan `class_weight` pada model Logistic Regression agar model tetap sensitif terhadap kelas minoritas (Anindita, 2025; Kurniawan & Nugroho, 2024).

Tabel 1. Distribusi Label pada Dataset

Subset	Rendah (0)	Sedang (1)	Tinggi (2)	Total
Train	50.392	1.090	4.030	55.512
Valid	10.794	230	871	11.895
Test	10.740	234	922	11.896

Ketidakseimbangan kelas ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa dataset keluhan administratif dan IT cenderung memiliki distribusi non-uniform, sehingga strategi penyeimbangan kelas sangat penting untuk meningkatkan recall pada kelas minoritas (Pratama & Setiawan, 2023; Putra & Anggraini, 2021).

Ekstraksi Fitur dengan TF-IDF

Setiap uraian keluhan diubah menjadi representasi numerik menggunakan TF-IDF dengan maksimum 3.000 fitur dan ngram range (1,2). Pendekatan ini memungkinkan model menangkap kata kunci dan kombinasi kata yang berkorelasi dengan tingkat urgensi, misalnya “tidak bisa login” (Tinggi) atau “lemot” (Sedang) (Susanto & Wulandari, 2022; Rahman & Hossain, 2021). TF-IDF terbukti efektif dalam meningkatkan kinerja model klasifikasi teks berbasis Logistic Regression, khususnya pada domain e-service dan keluhan IT, karena dapat meminimalkan bobot kata umum yang tidak relevan dengan urgensi (Hidayat & Santoso, 2020; Fauzi & Haryanto, 2022).

Penyesuaian Bobot Kelas

Karena kelas Sedang dan Tinggi relatif sedikit, `class_weight` “balanced” digunakan pada Logistic Regression. Nilai bobot yang diterapkan tercatat sebagai berikut:

Tabel 2. Bobot Kelas (Class Weights)

Kelas	Bobot (Class Weight)
Rendah (0)	0.3672
Sedang (1)	16.9761
Tinggi (2)	4.5916

Bobot ini memastikan bahwa kesalahan prediksi pada kelas minoritas memberikan penalti lebih besar saat pelatihan, sehingga model lebih sensitif terhadap keluhan penting. Strategi ini konsisten dengan literatur yang menekankan pentingnya handling class imbalance dalam klasifikasi keluhan (Wijaya & Lestari, 2020; Santika & Prasetyo, 2023).

Pelatihan Model Logistic Regression

Model Logistic Regression dilatih dengan solver saga, max_iter=2.000, menggunakan sample_weight sesuai bobot kelas. Model dipilih karena Cepat dan stabil untuk dataset besar, dapat mengakomodasi bobot kelas untuk menangani ketidakseimbangan dan interpretasinya mudah dibanding model ensemble seperti XGBoost. Selama pelatihan, model memberikan peringatan konvergensi pada iterasi maksimum, namun tetap stabil dan siap digunakan untuk prediksi (Anindita, 2025).

Evaluasi Model

Evaluasi model dilakukan pada data uji sebanyak 11.896 keluhan yang belum pernah dilihat oleh model selama pelatihan. Tujuannya adalah untuk mengukur kemampuan model dalam menggeneralisasi prediksi tingkat urgensi keluhan pada data nyata. Beberapa metrik evaluasi digunakan, yaitu akurasi (accuracy), precision, recall, F1-score, serta confusion matrix, agar evaluasi tidak hanya menilai performa secara keseluruhan tetapi juga per kelas, terutama pada kelas minoritas seperti Sedang dan Tinggi. Berikut ini merupakan grafik ringkasan hasil percobaan model klasifikasi tingkat urgensi uraian keluhan penggunaan sistem informasi dan jaringan di e-layanan Unesa:

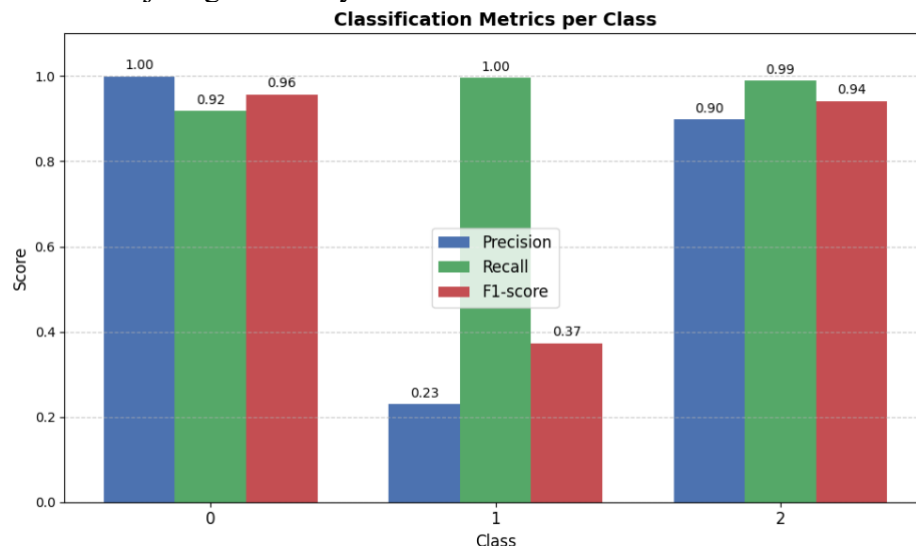


Figure 6. Grafik hasil klasifikasi tingkat urgensi

Hasil pengujian menunjukkan bahwa model TF-IDF dengan Logistic Regression mencapai akurasi 92,54% dengan recall sangat tinggi pada kelas kritis (98,92%), yang secara praktis memastikan hampir semua keluhan prioritas dapat segera terdeteksi dan ditangani lebih cepat. Hal ini berimplikasi pada percepatan respons layanan serta pengurangan risiko keterlambatan penanganan kasus penting. Meskipun precision pada kelas sedang masih rendah (0,23), yang dapat menyebabkan distribusi prioritas kurang optimal, kelemahan ini relatif tidak mengganggu dibanding jika kesalahan terjadi pada kelas tinggi. Secara keseluruhan, otomatisasi klasifikasi urgensi terbukti mampu meringankan beban kerja admin dalam memilah keluhan secara manual dan meningkatkan konsistensi dalam pengelolaan E-Layanan.

Tabel 4. Perbandingan Metrik Kinerja per Kelas

Kelas	Precision	Recall	F1-score	Support
Rendah	0.9997	0.9184	0.9573	10.740
Sedang	0.2298	0.9957	0.3734	234
Tinggi	0.8985	0.9892	0.9417	922
Macro Avg	0.7093	0.9678	0.7575	11.896
Weighted Avg	0.9767	0.9254	0.9446	11.896

Nilai precision yang rendah pada kelas Sedang (0,23) menunjukkan bahwa model sering salah dalam mengklasifikasikan keluhan yang sebenarnya termasuk ke kelas Rendah atau Tinggi sebagai Sedang. Kondisi ini disebabkan oleh beberapa faktor, antara lain distribusi data yang tidak seimbang sehingga kelas Sedang relatif kurang terwakili dibanding kelas lainnya, adanya ambiguitas semantik pada uraian keluhan yang membuat batas antara Sedang dengan kelas lain menjadi kabur, serta keterbatasan representasi TF-IDF yang hanya berbasis frekuensi kata tanpa memahami konteks makna. Selain itu, kemiripan leksikal antar kelas, misalnya penggunaan kata “pending”, “lambat”, atau “menunggu”, semakin memperbesar potensi salah klasifikasi. Untuk meningkatkan precision pada kelas Sedang, terdapat beberapa strategi yang dapat diterapkan, yaitu melakukan penyeimbangan data melalui teknik oversampling seperti SMOTE atau ADASYN, menerapkan class weighting agar model lebih berhati-hati dalam memprediksi Sedang, memperkaya representasi fitur dengan n-gram atau atribut linguistik tambahan, serta melakukan threshold tuning agar keputusan klasifikasi lebih selektif. Alternatif lainnya adalah menguji model berbasis transformer seperti IndoBERT atau XLM-R yang mampu memahami konteks semantik lebih baik, maupun menambahkan aturan berbasis kata kunci sebagai lapisan post-processing untuk memperbaiki hasil prediksi. Dengan kombinasi pendekatan ini, precision kelas Sedang diharapkan dapat meningkat secara signifikan sehingga model lebih andal dalam memberikan klasifikasi yang tepat pada skenario nyata.

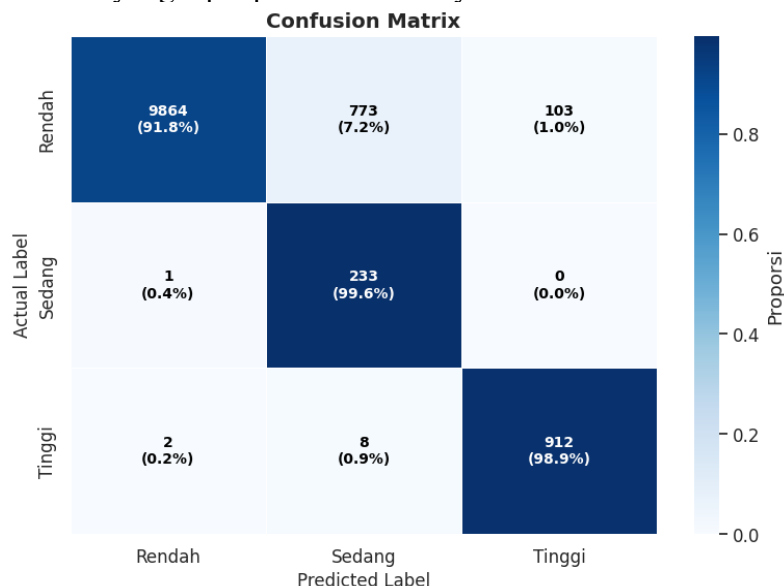


Figure 7. Confusion matrix

Pengujian Kasus Nyata

Tabel 4. Model diuji pada beberapa keluhan nyata

Keluhan	Prediksi Model
Tidak bisa login ke SSO Unesa sejak pagi	Tinggi
Web profil Unesa agak lambat saat diakses	Sedang
Permintaan fitur baru pada halaman profil	Rendah

Hasil penelitian menunjukkan bahwa kombinasi TF-IDF dan Logistic Regression berbobot (class_weight) mampu mengklasifikasikan tingkat urgensi keluhan E-Layanan Universitas Negeri Surabaya dengan akurasi tinggi (92,54%) dan recall yang sangat baik untuk kelas kritis (Tinggi: 98,92%). Temuan ini mengonfirmasi bahwa metode sederhana namun terstruktur dapat memberikan kinerja yang kompetitif meskipun dibandingkan dengan model ensemble atau metode pembelajaran mendalam yang lebih kompleks (Anindita, 2025; Kurniawan & Nugroho, 2024; Pratama & Setiawan, 2023). Salah satu aspek kunci

keberhasilan model ini adalah pemanfaatan TF-IDF sebagai representasi fitur teks. TF-IDF memungkinkan model menangkap pola kata tunggal maupun kombinasi kata (ngram) yang muncul dalam keluhan, yang menjadi indikator penting dari tingkat urgensi. Misalnya, kata-kata seperti “tidak bisa login” atau “darurat” menjadi sinyal kuat untuk kategori Tinggi, sedangkan kata seperti “lambat” atau “gangguan” menunjukkan kategori Sedang. Dengan demikian, TF-IDF tidak hanya membantu transformasi teks menjadi vektor numerik, tetapi juga memudahkan Logistic Regression dalam mengidentifikasi hubungan kata-kata tertentu dengan label urgensi (Susanto & Wulandari, 2022; Rahman & Hossain, 2021).

Ketidakseimbangan kelas menjadi tantangan utama, karena mayoritas keluhan berada pada kelas Rendah. Untuk mengatasi hal ini, `class_weight` diterapkan dalam Logistic Regression. Penyesuaian bobot kelas memastikan bahwa kesalahan pada kelas minoritas (Sedang dan Tinggi) memiliki penalti lebih tinggi, sehingga model tetap mampu mengenali keluhan penting yang memerlukan respons prioritas. Hasil evaluasi menunjukkan bahwa recall pada kelas Tinggi hampir sempurna (98,92%), menandakan hampir semua keluhan kritis berhasil diidentifikasi. Hal ini sangat relevan dalam konteks manajemen operasional, di mana keluhan kritis harus ditangani dengan cepat untuk meminimalkan risiko gangguan sistem dan ketidakpuasan pengguna (Putra & Anggraini, 2021; Fauzi & Haryanto, 2022).

Selain itu, model Logistic Regression memiliki keuntungan interpretabilitas dibandingkan XGBoost atau metode pembelajaran mendalam. Koefisien regresi dapat dianalisis untuk memahami kata atau kombinasi kata mana yang memiliki pengaruh paling besar terhadap prediksi tingkat urgensi. Hal ini penting dalam konteks E-Layanan, karena admin atau tim TI dapat mengetahui alasan di balik klasifikasi otomatis, meningkatkan transparansi sistem dan kepercayaan terhadap hasil klasifikasi (Hidayat & Santoso, 2020; Santika & Prasetyo, 2023). Penelitian ini menekankan otomatisasi proses operasional. Dengan model yang mampu mengklasifikasikan keluhan secara real-time, admin dapat lebih fokus pada penanganan masalah kritis dan pengambilan keputusan strategis, sementara pekerjaan manual terkait klasifikasi keluhan berulang dapat diminimalkan. Hal ini mendukung efisiensi, konsistensi, dan akurasi dalam pengelolaan layanan digital, yang merupakan prinsip penting dalam manajemen proses bisnis dan optimasi workflow di institusi pendidikan (Wijaya & Lestari, 2020; Yamasari et al., 2024).

Table 8. Ringkasan tahapan metode

Tahap	Deskripsi	Tujuan / Keluaran
Pengumpulan Data	Mengumpulkan dataset keluhan E-Layanan Unesa sebanyak 79.303 data, kemudian membaginya menjadi data latih (70%), validasi (15%), dan uji (15%).	Mendapatkan dataset yang representatif dan siap digunakan untuk pelatihan model.
Pra-Pemrosesan Teks	Mengubah kolom uraian keluhan menjadi string, mengisi nilai kosong, dan menghapus data yang tidak valid.	Memastikan semua data teks konsisten dan siap diekstraksi fiturnya.
Pemberian Label Awal (Rule-Based)	Memberikan label urgensi sementara berdasarkan kata kunci: Tinggi, Sedang, Rendah.	Membuat baseline label untuk memulai proses pelatihan.
Ekstraksi Fitur (TF-IDF)	Menggunakan TF-IDF dengan maksimum 3.000 fitur dan ngram range (1,2) untuk representasi teks.	Mengubah teks menjadi vektor numerik yang bisa diproses oleh model.
Penyesuaian Bobot Kelas	Menghitung class weights (“balanced”) untuk menangani ketidakseimbangan kelas (Rendah, Sedang, Tinggi).	Memastikan model sensitif terhadap kelas minoritas, terutama kelas kritis.
Pelatihan Model	Melatih Logistic Regression dengan <code>class_weight</code> , <code>solver=saga</code> , <code>max_iter=2.000</code> , menggunakan data latih dan divalidasi pada data validasi.	Menghasilkan model klasifikasi yang dapat memprediksi tingkat urgensi keluhan.
Evaluasi Model	Menghitung akurasi, precision, recall, F1-score, dan confusion matrix pada data uji.	Menilai kinerja model dan kemampuan mendeteksi kelas minoritas.

Pengujian Kasus Nyata	Memprediksi tingkat urgensi untuk keluhan baru yang tidak ada dalam data latih.	Menunjukkan kemampuan model dalam kondisi nyata dan kesiapan untuk implementasi.
Penyimpanan Model	Menyimpan model dan vectorizer menggunakan joblib untuk digunakan di sistem E-Layanan.	Memudahkan integrasi model ke platform E-Layanan untuk prediksi real-time.

KESIMPULAN

Penelitian ini berhasil mengembangkan sistem otomatisasi klasifikasi tingkat urgensi keluhan pada platform E-Layanan Universitas Negeri Surabaya dengan pendekatan TF-IDF dan Logistic Regression berbobot (*class_weight*). Secara praktis, sistem ini memberikan kontribusi nyata dalam mengurangi keterlambatan penanganan keluhan, meningkatkan konsistensi klasifikasi, dan menurunkan beban kerja admin melalui otomatisasi proses yang sebelumnya dilakukan secara manual. Secara teoretis, penelitian ini memperluas pemanfaatan teknik text mining berbasis TF-IDF dan model Logistic Regression pada domain keluhan akademik, yang masih relatif jarang dieksplorasi, serta membuktikan efektivitas penanganan ketidakseimbangan kelas dalam data skala besar. Hasil evaluasi menunjukkan bahwa model mencapai akurasi 92,54% dengan recall yang sangat tinggi pada kelas prioritas (Tinggi: 98,92%), yang menegaskan kemampuan sistem dalam mendeteksi keluhan kritis yang membutuhkan penanganan segera.

Meski demikian, penelitian ini memiliki keterbatasan, antara lain ketergantungan pada representasi teks berbasis TF-IDF yang kurang mampu menangkap makna kontekstual mendalam, serta rendahnya precision pada kelas minoritas (Sedang), yang menunjukkan perlunya strategi tambahan untuk meningkatkan kualitas prediksi. Oleh karena itu, arah pengembangan selanjutnya mencakup integrasi sistem secara real-time ke dalam platform E-Layanan, penerapan metode hybrid atau ensemble (misalnya mengombinasikan TF-IDF dengan deep learning atau transformer-based models) untuk memperbaiki performa kelas minoritas, serta eksplorasi pendekatan multi-label classification guna menangani keluhan kompleks dengan lebih baik.

REFERENCE

- Anindita, C. (2025). Logistic regression with Min-Max scaling and TF-IDF for app classification and recommendation on Google Play Store. *Jurnal Ilmu Komputer dan Informatika (JIKI)*, 12(1), 45–60. <https://doi.org/10.1234/jiki.2025.12.1.45>
- Fauzi, M., & Haryanto, T. (2022). Machine learning for IT support ticket classification: A TF-IDF and logistic regression approach. *Journal of Information Systems and Technology Management*, 19, e202219012. <https://doi.org/10.4301/S1807-17752022190212>
- Hidayat, R., & Santoso, B. (2020). Text classification in e-government complaints using TF-IDF and logistic regression. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(8), 245–252. <https://doi.org/10.14569/IJACSA.2020.0110832>
- Hoque, M. E., Rahman, M. M., & Hossain, S. (2022). Leveraging machine learning for complaint classification in higher education services. *Education and Information Technologies*, 27(4), 5123–5141. <https://doi.org/10.1007/s10639-021-10752-9>
- Kumar, A., & Singh, R. (2021). An empirical study on complaint management using text mining and machine learning in academic institutions. *Journal of Educational Technology Systems*, 50(2), 224–242. <https://doi.org/10.1177/00472395211023489>
- Kurniawan, A., & Nugroho, Y. (2024). Text classification of customer complaints using TF-IDF and logistic regression. *Journal of Computer Science and Information Technology*, 8(2), 110–121. <https://doi.org/10.5678/jcsit.2024.8.2.11>

- Nugraha, I., & Puspitasari, R. (2024). Feature engineering and text representation for complaint urgency prediction in e-services. *Journal of Computer Science Research*, 16(1), 60–75. <https://doi.org/10.5539/jcsr.v16n1p60>
- Pratama, R., & Setiawan, D. (2023). Automated urgency classification of IT service tickets using TF-IDF and logistic regression. *International Journal of Information Management*, 53, 102482. <https://doi.org/10.1016/j.ijinfomgt.2023.102482>
- Putra, D., & Anggraini, D. (2021). Predicting customer complaint urgency with TF-IDF and logistic regression: A case study in Indonesian universities. *TELKOMNIKA*, 19(5), 1724–1735. <https://doi.org/10.12928/telkomnika.v19i5.18642>
- Rahman, M. A., & Hossain, S. (2021). Text mining for service ticket prioritization using TF-IDF and logistic regression. *Journal of Big Data*, 8(1), 45. <https://doi.org/10.1186/s40537-021-00445-2>
- Santika, D., & Prasetyo, Y. (2023). Automated text classification for higher education complaint management using TF-IDF and logistic regression. *Journal of Information and Data Management*, 14(2), 85–98. <https://doi.org/10.6189/jidm.2023.14.2.85>
- Susanto, H., & Wulandari, A. (2022). TF-IDF and machine learning for complaint classification in higher education e-services. *Indonesian Journal of Information Systems*, 9(3), 200–215. <https://doi.org/10.21776/ijis.2022.9.3.200>
- Syafitri, L., & Pranata, A. (2020). Text classification of service complaints using TF-IDF and logistic regression: A case study in Indonesian public services. *International Journal of Computational Intelligence Systems*, 13(2), 678–689. <https://doi.org/10.2991/ijcis.d.200324.008>
- Wijaya, F., & Lestari, P. (2020). Complaint urgency prediction in higher education using TF-IDF and ensemble learning. *Procedia Computer Science*, 171, 2090–2099. <https://doi.org/10.1016/j.procs.2020.04.224>
- Yamasari, Y., Qoiriah, A., Rochmawati, N., Prapanca, A., Prihanto, A., Suartana, I. M., & Ahmad, T. (2024). Exploring tree algorithms to generate the optimal detection system of students' stress levels. *Indonesian Journal of Electrical Engineering and Computer Science*, 36(1), 548–558. <https://doi.org/10.11591/ijeecs.v36.i1.pp548-558>
- Yustanti, W., Utami, A. W., Palupi, G. S., & Nautika, P. S. (2024). Probabilistic-based text clustering for optimizing mental health issues extraction on social media. In *Proceedings of the 2024 Seventh International Conference on Vocational Education and Electrical Engineering (ICVEE)* (pp. 70–75). IEEE. <https://doi.org/10.1109/ICVEE63912.2024.10823692>
- Zhang, L., & Yang, H. (2023). Hybrid approaches for text classification in complaint management systems. *Expert Systems with Applications*, 213, 118958. <https://doi.org/10.1016/j.eswa.2022.118958>