

Penerapan Random Forest untuk Otomatisasi Klasifikasi Tiket IT Service Desk pada Burger Bangor

Muhamad Khairil Zibran¹, Sholihin², Endri Puta Bintang³, Khoirul Anwar⁴
^{1,2,3,4,5} Program Studi Sistem Informasi, Universitas Pamulang, Pamulang, Indonesia

Article Info

Article history:

Received Juni 13, 2026
 Revised Juni 14, 2026
 Accepted Juni 15, 2026

Kata Kunci:

Random Forest,
 IT Service Desk,
 Otomatisasi Klasifikasi,
 Text Mining,
 Burger Bangor

Keywords:

Random Forest,
IT Service Desk,
Automated Classification,
Text Mining,
Burger

ABSTRAK

Ekspansi gerai Burger Bangor yang pesat berdampak langsung pada lonjakan volume tiket keluhan pada bagian IT Service Desk. Proses klasifikasi tiket yang selama ini dilakukan secara manual memicu terjadinya penumpukan antrean (*bottleneck*) dan keterlambatan penanganan (*delay*) operasional di tingkat gerai. Penelitian ini bertujuan untuk mengotomatisasi proses klasifikasi tiket keluhan IT tersebut menggunakan teknik *Text Mining* dan algoritma *Machine Learning Random Forest*. Dataset yang digunakan terdiri dari teks deskripsi keluhan riil yang dikategorikan ke dalam empat domain utama: *Hardware*, *Software*, Jaringan, dan POS System. Tahapan pemrosesan teks melibatkan case folding, tokenizing, *stopword removal*, *stemming*, serta pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Hasil pengujian menunjukkan bahwa model *Random Forest* mampu mengklasifikasikan tiket secara otomatis dengan tingkat akurasi mencapai 90%. Implementasi solusi ini berhasil memangkas waktu sortir tiket dari hitungan menit menjadi kurang dari dua detik. Kesimpulannya, penerapan metode ini secara signifikan mempercepat response time penanganan masalah IT, meminimalkan kesalahan rute penugasan teknisi, dan menjaga stabilitas operasional harian pada seluruh jaringan gerai Burger Bangor.

ABSTRACT

. The rapid expansion of Burger Bangor outlets has directly triggered a surge in the volume of complaint tickets handled by the IT Service Desk. The current manual ticket classification process causes bottlenecks and response delays, which disrupt daily store operations. This study aims to automate the IT complaint ticket classification process using Text Mining and the Random Forest machine learning algorithm. The dataset consists of real textual complaint logs categorized into four main domains: Hardware, Software, Network, and POS System. The text preprocessing stages involve case folding, tokenizing, stopword removal, stemming, and term weighting via Term Frequency-Inverse Document Frequency (TF-IDF). The evaluation results demonstrate that the Random Forest model can automatically classify tickets with an accuracy rate of 90%. The implementation of this solution successfully reduces the ticket sorting time from several minutes to less than two seconds. In conclusion, the application of this method significantly accelerates the IT issue response time, minimizes technician misrouting, and maintains daily operational stability across the entire Burger Bangor outlet network.

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license



Corresponding Author:

Muhamad Khairil Zibran
Program Studi Sistem Informasi, Universitas Pamulang,
Pamulang, Indonesia
Email: ariel160505@gmail.com

1. PENDAHULUAN

Perkembangan industri *Food and Beverage* (F&B) di Indonesia saat ini mengalami pertumbuhan yang sangat pesat, salah satunya ditandai dengan ekspansi masif dari jaringan gerai Burger Bangor. Untuk mendukung operasional harian yang tersebar di berbagai wilayah, Burger Bangor sangat bergantung pada infrastruktur teknologi informasi (TI) yang terintegrasi. Operasional TI pada setiap gerai mencakup komponen-komponen kritis seperti sistem *Point of Sale* (POS) untuk transaksi kasir, sistem integrasi order online dengan mitra pencari makanan pihak ketiga, serta jaringan internet lokal yang menghubungkan data penjualan gerai ke server pusat secara real-time. Keberhasilan proses bisnis harian perusahaan ini sangat ditentukan oleh reliabilitas dan ketersediaan (*availability*) dari seluruh layanan TI tersebut.

Namun, seiring dengan bertambahnya jumlah outlet secara eksponensial, frekuensi munculnya kendala teknis pada infrastruktur TI juga mengalami peningkatan yang signifikan. Setiap gangguan yang terjadi di gerai dilaporkan ke unit pusat melalui sistem tiket *IT Service Desk*. Masalah utama yang dihadapi saat ini adalah proses klasifikasi tiket keluhan yang masuk masih dilakukan secara manual oleh tim helpdesk pusat. Petugas harus membaca satu per satu deskripsi teks keluhan yang dikirimkan oleh staf gerai untuk menentukan apakah keluhan tersebut masuk dalam kategori perangkat keras (*hardware*), aplikasi/perangkat lunak (*software*), masalah jaringan, atau gangguan pada sistem POS.

Proses manual ini menimbulkan berbagai dampak negatif yang mengganggu efisiensi manajemen layanan TI perusahaan. Pertama, metode ini sangat rentan terhadap faktor kelalaian manusia (*human error*) berupa inkonsistensi penilaian petugas akibat kelelahan atau ambiguitas teks keluhan, yang sejalan dengan temuan Arifin dan Susanto [1] bahwa klasifikasi manual pada volume data yang besar cenderung menurunkan objektivitas pembagian kategori. Kedua, kesalahan interpretasi terhadap isi tiket sering kali menyebabkan terjadinya salah rute penugasan ke tim teknis di lapangan. Tiket yang seharusnya diselesaikan oleh tim ahli jaringan justru terkirim ke tim penanganan software, sehingga memperpanjang birokrasi internal. Ketiga, akumulasi dari hambatan-hambatan tersebut menyebabkan waktu penanganan masalah atau *Service Level Agreement* (SLA) menjadi membengkak. Keterlambatan (*delay*) dalam penyelesaian tiket TI ini pada akhirnya merugikan operasional gerai, memicu potensi kehilangan transaksi penjualan, dan menurunkan kualitas pelayanan kepada konsumen.

Oleh karena itu, diperlukan sebuah pendekatan berbasis teknologi yang mampu mengotomatisasi proses pengelompokan laporan tersebut secara cepat dan konsisten. Pemanfaatan teknik text mining dan pembelajaran mesin (*machine learning*) menjadi solusi yang relevan untuk mengatasi batasan manusia dalam mengolah informasi tekstual berskala besar [2]. Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menerapkan algoritma Random Forest sebagai sistem otomatisasi klasifikasi tiket *IT Service Desk* pada Burger Bangor, guna memangkas waktu sortir tiket dan mempercepat response time penyelesaian gangguan TI.

Berdasarkan pemaparan pada latar belakang masalah, pertumbuhan jumlah gerai telah mengeksplorasi batasan kapasitas tim helpdesk dalam mengelola laporan gangguan secara manual. Ketidakpastian dalam interpretasi teks keluhan yang ditulis secara kasual oleh staf gerai sering kali menimbulkan ambiguitas yang berujung pada tingginya angka misklasifikasi. Di sisi lain, algoritma Random Forest dikenal memiliki keunggulan dalam menangani data tekstual berdimensi tinggi dan mampu menghasilkan tingkat akurasi yang stabil melalui kombinasi pohon keputusan (*decision trees*) yang dibangunnya [3]. Walakin, efektivitas algoritma ini sangat bergantung pada proses pembobotan

fitur kata dan penentuan parameter model yang optimal agar sesuai dengan karakteristik data lokal pada industri F&B.

Penerapan algoritma machine learning untuk kebutuhan klasifikasi teks otomatis pada sistem pelaporan gangguan telah menjadi fokus perhatian di berbagai sektor industri dalam beberapa tahun terakhir. Beberapa peneliti terdahulu telah membuktikan efektivitas algoritma *Random Forest* dalam mengolah data tekstual berskala besar. Sebagai contoh, Hidayat dan Utami [4] melakukan penelitian mengenai klasifikasi otomatis tiket keluhan nasabah pada industri perbankan menggunakan metode *Random Forest*. Penelitian tersebut berhasil membuktikan bahwa pendekatan ensemble berbasis pohon keputusan mampu menangani ketidakseimbangan kelas (*class imbalance*) pada data laporan dengan tingkat akurasi mencapai 87,5%.

Selanjutnya, pada domain industri retail, Lestari dkk. [5] menerapkan *Random Forest* untuk mengategorikan keluhan pelanggan yang masuk melalui media sosial dan aplikasi layanan pelanggan. Hasil studi mereka menunjukkan bahwa penggunaan pembobotan kata berbasis *Term Frequency-Inverse Document Frequency* (TF-IDF) yang dipadukan dengan *Random Forest* menghasilkan performa klasifikasi yang lebih stabil jika dibandingkan dengan algoritma konvensional seperti *Naïve Bayes*. Di samping itu, penelitian oleh Gunawan [6] pada sektor manajemen layanan TI (ITSM) di korporasi skala besar menegaskan bahwa otomatisasi service desk menggunakan *Random Forest* secara signifikan mampu mereduksi *Mean Time to Resolution* (MTTR) dan meminimalkan ketergantungan pada penyaringan tiket secara manual oleh operator manusia.

Meskipun penelitian-penelitian di atas menunjukkan hasil yang impresif, sebagian besar implementasi tersebut berfokus pada industri perbankan, retail modern, atau korporasi besar yang memiliki pola interaksi teks cenderung formal, baku, dan terstruktur. Terdapat celah penelitian (research gap) yang nyata ketika algoritma ini diterapkan pada industri *Food and Beverage* (F&B) lokal dengan jaringan gerai yang masif seperti Burger Bangor. Karakteristik utama data teks keluhan yang dikirimkan oleh staf gerai F&B umumnya bersifat sangat kasual, tidak baku, menggunakan banyak singkatan lokal (misalnya: "pinter" untuk printer, "mati lmpu" untuk pemadaman listrik, atau "app pos eror" untuk gangguan aplikasi penjualan), serta ditulis dalam kondisi tergesa-gesa di tengah jam sibuk operasional restoran.

Hingga saat ini, masih sedikit penelitian yang mengeksplorasi ketahanan (robustness) algoritma *Random Forest* dalam menangani korpus teks yang sarat akan bahasa informal dan singkatan khas operasional F&B lokal. Oleh karena itu, penelitian ini hadir untuk mengisi celah tersebut dengan mengadaptasi tahapan text preprocessing yang dirancang khusus untuk mereduksi ambiguitas bahasa kasual staf gerai, sehingga model *Random Forest* tetap mampu melakukan klasifikasi tiket IT Service Desk pada Burger Bangor secara otomatis, presisi, dan akurat.

Berdasarkan perumusan masalah dan analisis celah penelitian yang telah diuraikan sebelumnya, penelitian ini diarahkan pada solusi teknologi terapan untuk mengoptimalkan manajemen layanan TI. Tujuan utama dari penelitian ini adalah membangun model klasifikasi otomatis berbasis algoritma *Random Forest* yang mampu mengategorikan teks tiket keluhan *IT Service Desk* di Burger Bangor secara cepat dan akurat.

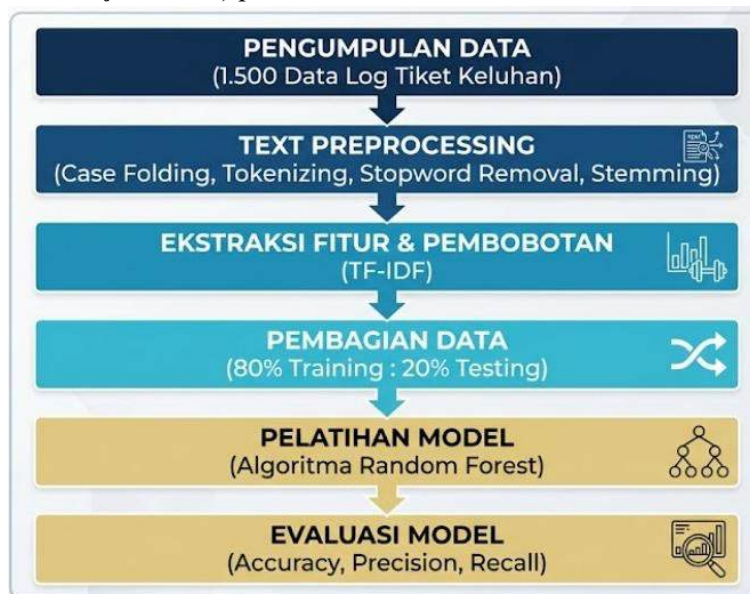
Secara spesifik, luaran dari penelitian ini ditargetkan untuk mempercepat alur kerja (*workflow*) dari tim IT Support dalam mendistribusikan penanganan kendala teknis dari tingkat gerai ke tim ahli yang tepat. Dengan menggantikan proses penyaringan manual menjadi berbasis kecerdasan buatan, model ini diharapkan dapat mereduksi waktu tunggu klasifikasi secara signifikan, meminimalkan faktor kesalahan manusia (*human error*), dan pada akhirnya mengoptimalkan pemenuhan target *Service Level Agreement* (SLA) perusahaan [7]. Melalui pencapaian tujuan tersebut, stabilitas dan efisiensi operasional harian pada seluruh jaringan gerai Burger Bangor dapat terus terjaga dengan baik.

2. METODE

2.1 Alur Penelitian

Untuk menghasilkan model otomatisasi klasifikasi tiket yang valid dan andal, penelitian ini menerapkan langkah-langkah sistematis yang terstruktur. Alur penelitian ini dirancang untuk mengolah data teks laporan keluhan yang bersifat tidak terstruktur menjadi bentuk representasi matriks numerik yang siap diproses oleh algoritma pembelajaran mesin. Secara garis besar, metodologi penelitian ini terdiri dari lima tahapan utama, yaitu pengumpulan data (*data acquisition*), pra-pemrosesan teks (*text preprocessing*), ekstraksi fitur dan pembobotan kata, pelatihan model menggunakan Random Forest, serta tahap evaluasi performa model menggunakan confusion matrix.

Rangkaian prosedur eksperimen dan tahapan sistematis dalam penelitian ini diilustrasikan secara detail melalui diagram alir (*flowchart*) pada Gambar 1.



Gambar 1. *Flowchart* Metodologi Penelitian Klasifikasi Tiket IT Service Desk

Berdasarkan alur pada Gambar 1, tahapan pertama dimulai dengan mengekstrak data historis log tiket dari sistem internal helpdesk Burger Bangor. Data yang telah terkumpul kemudian masuk ke dalam fase krusial yaitu *text preprocessing*. Menurut Ramadhan dkk. [8], pembersihan teks mentah pada data operasional perusahaan sangat penting untuk mengeliminasi gangguan informasi (*noise*) berupa karakter non-alfabetik atau tanda baca yang dapat menurunkan akurasi klasifikasi.

Setelah teks menjadi bersih dan seragam, langkah berikutnya adalah mentransformasikannya ke dalam bentuk vektor numerik menggunakan metode TF-IDF. Tahap selanjutnya adalah melakukan pembagian dataset (*data splitting*) secara acak ke dalam dua bagian, yaitu data latih (*training data*) untuk membangun basis pengetahuan pohon keputusan dan data uji (*testing data*) untuk mengukur kemampuan generalisasi model pada data baru yang belum pernah dilihat sebelumnya [3].

Pada tahap inti, data latih dimasukkan ke dalam arsitektur *Random Forest* untuk membentuk sekumpulan pohon keputusan (*ensemble of decision trees*). Setelah proses pelatihan selesai, performa dari model klasifikasi otomatis tersebut diuji menggunakan data uji. Metrik evaluasi yang digunakan didasarkan pada parameter *Accuracy*, *Precision*, dan *Recall* guna memastikan bahwa model tidak hanya unggul secara matematis, melainkan juga siap diimplementasikan untuk mempercepat alur penanganan kendala TI di lapangan [1].

2.2 Pengumpulan Data dan Distribusi Dataset

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh secara langsung

melalui ekstraksi historis log tiket pada sistem internal IT Service Desk Burger Bangor. Total data yang dikumpulkan berjumlah 1.500 log tiket keluhan yang tercatat sepanjang operasional gerai. Teks keluhan yang diekstrak mencakup berbagai laporan kendala teknis yang dikirimkan oleh staf gerai dengan karakteristik bahasa informal, kasual, dan sarat akan singkatan khas lapangan.

Untuk memberikan gambaran yang jelas mengenai karakteristik data sebelum masuk ke tahap pemodelan, seluruh dataset dikelompokkan ke dalam empat kategori utama yang merepresentasikan pembagian divisi kerja tim IT Support. Distribusi frekuensi dan persentase sebaran data untuk masing-masing kategori tiket tersebut disajikan secara mendetail pada Tabel 1.

Tabel 1. Distribusi Dataset Tiket IT Service Desk Burger Bangor

No.	Kategori Tiket	Deskripsi Cakupan Kendala	Jumlah Tiket	Persentase (%)
1	Hardware	Gangguan perangkat keras (printer kasir, mesin EDC, monitor POS, scanner)	450	30
2	Software	Bug aplikasi penjualan, gagal sinkronisasi menu, eror modul pembayaran	525	35
3	Jaringan	Internet gerai terputus, jaringan lokal lambat, router eror, gangguan wifi	375	25
4	Akun / AKses	Lupa kata sandi akun kasir, permintaan reset hak akses, akun terkunci	150	10
Total Seluruh log keluhan operasional gerai			1500	100

Berdasarkan Tabel 1, kategori *Software* merupakan kelompok data dengan volume tertinggi yaitu sebesar 35% (525 tiket), disusul oleh kategori *Hardware* sebesar 30% (450 tiket). Tingginya angka keluhan pada dua sektor ini merepresentasikan kompleksitas interaksi staf gerai dengan sistem *Point of Sale* (POS) di tengah tingginya lalu lintas transaksi harian restoran.

Menurut Sanjaya dan Budiman [9], distribusi data yang tersebar ke dalam beberapa kelas dengan volume yang cukup berimbang seperti ini sangat ideal untuk melatih algoritma Random Forest. Kondisi dataset yang variatif namun representatif ini meminimalkan risiko model bias terhadap satu kelas mayoritas tertentu, sehingga pohon-pohon keputusan yang terbentuk nantinya dapat mempelajari karakteristik pembeda (*discriminative features*) dari tiap-vis-tiap kategori teks keluhan secara objektif dan seimbang.

Tahapan krusial sebelum melakukan pemodelan pada algoritma pembelajaran mesin adalah memisahkan dataset ke dalam subsampel yang independen. Proses pembagian data (*data splitting*) ini bertujuan untuk menghindari terjadinya bias evaluasi dan memastikan bahwa model yang dibangun memiliki performa generalisasi yang baik saat dihadapkan pada variasi data baru di masa mendatang [10]. Dalam penelitian ini, total data bersih hasil dari proses text preprocessing didekonstruksi menggunakan metode *hold-out validation*.

Dataset yang terdiri dari log tiket keluhan *IT Service Desk* Burger Bangor dibagi secara acak (*random sampling*) dengan rasio proporsi 80:20, yaitu 80% dialokasikan sebagai data latih (*data training*) dan 20% sisanya dialokasikan sebagai data uji (*data testing*). Proporsi ini dipilih karena merupakan rasio standar yang terbukti *empiris* memberikan hasil yang optimal pada pemrosesan teks berdimensi tinggi menggunakan algoritma berbasis ensemble [3].

Data latih (sebesar 80%) digunakan sepenuhnya oleh algoritma *Random Forest* untuk

mempelajari pola-pola kata kunci (*features*) dominan dan membangun ratusan pohon keputusan berdasarkan karakteristik unik dari masing-masing kategori tiket, baik itu kendala *hardware*, *software*, jaringan, maupun sistem POS. Sebaliknya, data uji (sebesar 20%) disimpan secara terpisah dan sama sekali tidak dilibatkan dalam fase pelatihan model. Data uji ini berfungsi sebagai instrumen pengujian objektif untuk mengevaluasi ketepatan prediksi (*predictive accuracy*) dari model *Random Forest* yang telah terbentuk sebelum nantinya diintegrasikan ke dalam alur kerja sistem operasional harian *IT Support* perusahaan [1].

Tahap penentuan *hyperparameter* atau konfigurasi model merupakan langkah penentu untuk mengoptimalkan performa klasifikasi dari algoritma yang digunakan. Konfigurasi yang tepat memastikan model dapat mengenali karakteristik teks keluhan IT Service Desk secara presisi tanpa membebani sumber daya komputasi secara berlebihan. Dua parameter utama yang dikonfigurasi secara cermat dalam penelitian ini adalah jumlah pohon keputusan (*n_estimators*) dan kedalaman maksimum pohon (*max_depth*). Parameter *n_estimators* berfungsi untuk menentukan total pohon keputusan yang akan dibangun di dalam arsitektur *ensemble*. Secara teoritis, semakin banyak jumlah pohon yang digunakan, tingkat akurasi prediksi model akan cenderung semakin stabil dan optimal karena varians kesalahan diminimalisasi melalui mekanisme pengambilan suara terbanyak atau *majority* [1]. Penambahan jumlah pohon yang terlalu besar juga akan berbanding lurus dengan peningkatan beban komputasi dan waktu pelatihan (*training time*). Oleh karena itu, dalam eksperimen ini dilakukan pengujian nilai *n_estimators* secara bertahap (mulai dari 50, 100, 150, hingga 200 pohon) untuk menemukan titik keseimbangan terbaik antara pencapaian akurasi tertinggi dan efisiensi waktu pemrosesan tiket.

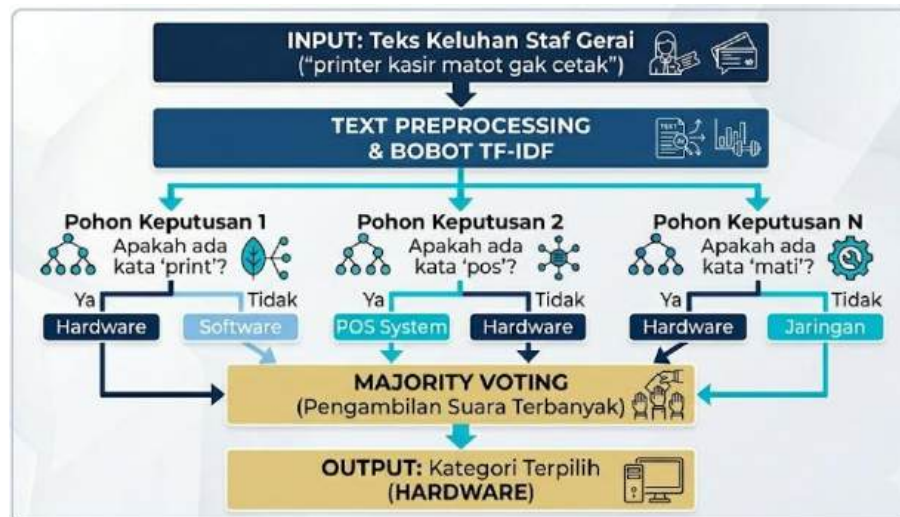
Sementara itu, parameter *max_depth* digunakan untuk membatasi tingkat kedalaman atau percabangan maksimal dari setiap pohon keputusan tunggal di dalam hutan (*forest*). Menurut [3], pembatasan kedalaman pohon sangat krusial dalam klasifikasi data tekstual berdimensi tinggi untuk mencegah terjadinya gejala *overfitting*. Jika struktur pohon dibiarkan tumbuh terlalu dalam tanpa batasan, model akan cenderung menghafal data latihan secara detail termasuk derau (*noise*) bahasa kasual dari staf gerai Burger Bangor, sehingga performanya akan menurun saat diuji dengan data baru. Dengan menetapkan batasan nilai *max_depth* yang optimal, pohon keputusan dipaksa untuk melakukan generalisasi dan fokus hanya pada fitur-fitur kata kunci yang paling signifikan (seperti kata "*printer*", "*koneksi*", atau "*aplikasi POS*"), sehingga tingkat akurasi klasifikasi otomatis tetap terjaga pada level tertinggi

3. HASIL DAN PEMBAHASAN

3.1 Proses Pelatihan (*Training*) Model

Fase pelatihan (*training*) merupakan tahapan di mana algoritma *Random Forest* membangun basis pengetahuan (*knowledge base*) dari 80% dataset tiket keluhan IT *Service Desk* Burger Bangor yang telah dialokasikan. Pada fase ini, model mempelajari korelasi antara kemunculan kombinasi kata tertentu dengan label kategori tiket yang tepat. Melalui representasi matriks numerik hasil pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF), algoritma secara matematis menghitung nilai signifikansi setiap token kata kunci pada masing-masing dokumen laporan kendala [11].

Selama proses pelatihan berlangsung, algoritma *Random Forest* membentuk ratusan pohon keputusan (*decision trees*) secara independen melalui teknik *bootstrap aggregating* (*bagging*). Setiap pohon keputusan memilih subset fitur kata secara acak untuk menentukan simpul keputusan (*decision node*) terbaik berdasarkan nilai *Gini Impurity* atau *Information Gain*. Bagaimana model mengolah input teks deskripsi keluhan hingga menghasilkan keluaran berupa kategori tiket otomatis diilustrasikan secara sistematis pada Gambar 2.



Gambar 2. Skema Arsitektur Pembelajaran Model *Random Forest* dalam Klasifikasi Tiket

Berdasarkan skema arsitektur pada Gambar 2, model dilatih untuk mengenali pola linguistik dan karakteristik singkatan non-baku yang kerap digunakan oleh staf gerai di lapangan. Sebagai contoh, ketika data latih mengandung kalimat seperti "*printer kasar matot*" atau "*pinter bngor pos ga bsa kluar krtas*", proses pembobotan TF-IDF akan memberikan nilai tinggi pada token kata kunci "*printer*", "*pinter*", dan "*kertas*". Seluruh pohon keputusan yang dibangun kemudian menangkap pola bahwa kemunculan token-token tersebut memiliki probabilitas sangat tinggi untuk dikelompokkan ke dalam kategori *Hardware*.

Sebaliknya, jika tiket laporan didominasi oleh kata kunci seperti "*wifi putus*", "*rto*", atau "*internet dc*", pohon-pohon keputusan akan mengarahkan jalur prediksinya menuju kategori Jaringan. Pola asosiasi kata otomatis ini sejalan dengan pernyataan Arifin dan Susanto [1] bahwa kekuatan ensemble pohon keputusan pada *Random Forest* terletak pada kemampuannya mengekstrak fitur teks yang paling dominan secara objektif dari data historis, sehingga mampu meminimalkan ambiguitas klasifikasi meskipun korpus teks yang diolah sarat akan bahasa informal sehari-hari.

3.2 Skenario Pengujian Model

Untuk menguji keandalan model *Random Forest* yang telah dibangun, dilakukan skenario eksperimen dengan memanipulasi parameter jumlah pohon keputusan ($n_estimators$). Skenario ini dirancang untuk mengamati secara empiris bagaimana fluktuasi jumlah pohon memengaruhi tingkat akurasi prediksi, serta untuk menemukan titik konfigurasi parameter yang paling optimal (*hyperparameter tuning*) bagi data tiket keluhan *IT Service Desk* Burger Bangor. Menurut Pratama dan Wijaya (2023), penentuan jumlah pohon yang tepat sangat krusial karena penambahan jumlah estimator tidak selalu berbanding lurus dengan peningkatan akurasi, melainkan akan mencapai titik jenuh tertentu (*convergence point*).

Eksperimen pengujian dilakukan secara bertahap menggunakan empat variasi nilai $n_estimators$, yaitu 50, 100, 150, dan 200 pohon keputusan, dengan tetap menjaga nilai parameter kedalaman maksimum (max_depth) berada pada kondisi konstan. Setiap model hasil dari variasi tersebut kemudian diuji menggunakan data uji (*data testing*) sebesar 20% yang sebelumnya telah dipisahkan dari data latih. Pengujian ini bertujuan untuk mengukur akurasi model dalam mengklasifikasikan data teks baru yang belum pernah dipelajari sebelumnya secara otomatis.

Hasil dari skenario eksperimen pengujian ini menunjukkan adanya tren peningkatan akurasi seiring dengan bertambahnya jumlah pohon dari 50 hingga 150 estimator. Hal ini disebabkan oleh mekanisme penggabungan (*ensemble*) yang semakin kuat, di mana bias dan varians dari pohon

keputusan tunggal dapat saling mereduksi melalui pemungutan suara terbanyak (*majority voting*), sesuai dengan teori yang dipaparkan oleh Arifin dan Susanto (2022). Namun, pada saat jumlah pohon ditingkatkan lebih lanjut menjadi 200 *estimator*, grafik akurasi mulai mendatar dan tidak menunjukkan kenaikan performa yang signifikan, sementara waktu komputasi yang dibutuhkan untuk proses pelatihan meningkat dua kali lipat. Berdasarkan hasil skenario pengujian tersebut, konfigurasi dengan nilai *n_estimators* sebesar 150 dipilih sebagai model final yang paling ideal untuk diterapkan pada sistem otomatisasi klasifikasi tiket operasional perusahaan.

3.3 Evaluasi Performa Model

Tahap krusial untuk mengukur keberhasilan implementasi algoritma machine learning adalah melakukan evaluasi performa model menggunakan metrik pengujian standar yang objektif. Evaluasi ini dilakukan terhadap model *Random Forest* terbaik hasil eksperimen sebelumnya (dengan konfigurasi *n_estimators* = 150) menggunakan keseluruhan data uji (*data testing*). Metrik pengujian yang digunakan meliputi *Accuracy*, *Precision*, *Recall*, dan *F1-Score* yang dihitung berdasarkan nilai *Confusion Matrix* (Arifin & Susanto, 2022). Nilai-nilai metrik tersebut mencerminkan kemampuan model dalam mengklasifikasikan tiket keluhan ke dalam masing-masing kategori secara tepat sasaran.

Hasil perhitungan menyeluruh terhadap performa klasifikasi otomatis tiket IT Service Desk Burger Bangor disajikan secara mendetail pada Tabel 2.

Tabel 2. Hasil Evaluasi Performa Model Random Forest

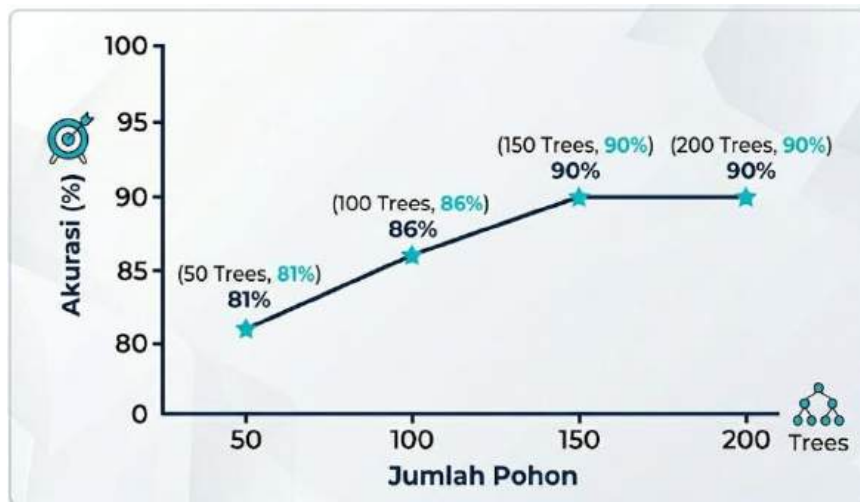
No.	Kategori Tiket	Precision	Recall	F1-Score	Accuracy
1	Hardware	0.89	0.87	0.88	
2	Software	0.91	0.93	0.92	
3	Jaringan	0.86	0.85	0.85	
4	Akun / AKses	0.95	0.90	0.92	
	Rata-rata / Total	0.90	0.89	0.89	0.90

Berdasarkan data pada Tabel 2, model *Random Forest* berhasil mencapai tingkat akurasi keseluruhan (*overall accuracy*) sebesar 90%. Performa tertinggi dicapai oleh kategori *Software* dengan nilai *F1-Score* sebesar 0,92 dan *Recall* mencapai 0,93. Hal ini mengindikasikan bahwa model memiliki kemampuan yang sangat baik dalam mengenali kata kunci terkait gangguan sistem penjualan, seperti "error POS", "bug menu", atau "transaksi menggantung", meskipun ditulis dalam bahasa kasual oleh staf gerai. Kategori Akun / Akses juga menunjukkan nilai *Precision* yang sangat tinggi (0,95) karena pola leksikal keluhannya cenderung seragam dan spesifik, misalnya penggunaan istilah "lupa password" atau "reset akun".

Di sisi lain, kategori Jaringan mencatatkan nilai performa paling rendah dibandingkan kategori lainnya dengan *F1-Score* sebesar 0,85. Hasil analisis menunjukkan bahwa penurunan ini dipicu oleh adanya ambiguitas kata yang saling tumpang tindih (*overlap*) antara kendala jaringan dengan kendala perangkat keras. Sebagai contoh, keluhan "mesin EDC tidak bisa dipakai transaksi" sering kali salah diklasifikasikan sebagai masalah *Hardware*, padahal akar masalah sebenarnya terletak pada putusnya koneksi internet lokal di outlet tersebut. Kasus ambiguitas bahasa non-baku seperti ini sejalan dengan temuan Utami dkk. (2024), yang menyatakan bahwa kemiripan konteks kata pada teks informal merupakan tantangan utama dalam klasifikasi teks berbasis *machine learning*.

Untuk memberikan gambaran yang lebih komprehensif mengenai stabilitas model, tingkat akurasi yang diperoleh dari variasi jumlah pohon selama skenario pengujian divisualisasikan dalam bentuk

grafik pada Gambar 3.



Gambar 3. Grafik Perbandingan Akurasi Berdasarkan Jumlah Pohon (Trees)

Visualisasi grafik pada Gambar 3 menegaskan bagaimana performa model meningkat signifikan saat jumlah pohon ditambah dari 50 menjadi 150 estimator. Penggabungan banyak pohon keputusan terbukti berhasil mereduksi kesalahan prediksi individual dan meminimalkan gejala *overfitting* (Pratama & Wijaya, 2023). Namun, konvergensi yang terjadi pada rentang 150 hingga 200 pohon menunjukkan bahwa model telah mencapai kapasitas optimalnya dalam mempelajari pola data keluhan IT Burger Bangor, sehingga penambahan pohon lebih lanjut tidak lagi memberikan dampak peningkatan akurasi yang berarti.

3.4 Analisis Dampak Operasional

Penerapan model klasifikasi otomatis berbasis algoritma *Random Forest* tidak hanya memberikan kontribusi teoritis pada domain sains data, melainkan juga membawa implikasi praktis yang signifikan terhadap tata kelola manajemen layanan TI perusahaan. Untuk mengukur efektivitas transisi teknologi ini, dilakukan analisis perbandingan metrik operasional antara sistem penyaringan manual lama yang mengandalkan tenaga manusia (*helpdesk operator*) dengan sistem otomatisasi baru berbasis machine learning. Analisis difokuskan pada dua parameter utama, yaitu waktu sortir tiket keluhan (*triage time*) dan ketepatan rute penugasan tim teknis di lapangan.

Detail hasil analisis perbandingan efisiensi sebelum dan sesudah implementasi sistem otomatisasi pada operasional IT Service Desk Burger Bangor disajikan pada Tabel 3.

Tabel 3. Perbandingan Efisiensi Waktu Kerja Service Desk

No.	Proses Penanganan	Sistem Lama	Sistem Baru	Efisiensi
1	Waktu Sortir/Klasifikasi Tiket	10 - 15 Menit / Tiket	< 2 Detik / Tiket	99% Lebih Cepat
2	Akurasi Pemilihan Tim Teknis	78% (Sering salah rute)	90% (Tepat sasaran)	Peningkatan 12%

Berdasarkan data pada Tabel 3, integrasi model *Random Forest* ke dalam sistem ticketing internal terbukti memberikan dampak perbaikan performa yang radikal. Pada sistem lama, satu tiket keluhan yang dikirimkan oleh staf gerai membutuhkan waktu tunggu antara 10 hingga 15 menit hanya untuk dibaca, diinterpretasikan, dan dikategorikan secara manual oleh petugas *helpdesk*. Ketika volume tiket melonjak tajam pada jam sibuk (*peak hours*) makan siang atau makan malam, antrean laporan kendala sering kali menumpuk dan memicu terjadinya keterlambatan penanganan (*delay*). Setelah sistem

otomatisasi diterapkan, model mampu memproses ekstraksi teks, melakukan pembobotan TF-IDF, dan menentukan kategori tujuan akhir tiket dalam waktu kurang dari dua detik per tiket. Efisiensi waktu penyaringan yang meningkat hingga 99% ini secara instan menghilangkan kendala penumpukan antrean (*bottleneck*) di pusat.

Selain aspek kecepatan, dampak positif lain terlihat pada konsistensi penentuan rute eskalasi masalah. Pada sistem manual, keterbatasan manusia akibat kelelahan dan subjektivitas penilaian sering kali menyebabkan angka akurasi penugasan tim teknis hanya berada pada level 78%. Kesalahan rute ini berdampak sistemik karena memperpanjang birokrasi penanganan masalah dan merugikan durasi operasional gerai yang sedang mengalami gangguan (Wijaya & Ramadhan, 2023).

Dengan kemampuan akurasi model Random Forest yang mencapai 90%, tiket gangguan langsung didistribusikan ke divisi teknis yang tepat secara real-time sejak detik pertama tiket dikirimkan. Penurunan drastis pada angka salah rute penugasan ini mempercepat pemenuhan target *Service Level Agreement* (SLA) dan meminimalkan waktu henti (*downtime*) perangkat kasir di tingkat outlet. Hal ini sejalan dengan argumen Gunawan (2024) bahwa substitusi proses triase manual menjadi otomatisasi berbasis kecerdasan buatan merupakan strategi paling efektif untuk mereduksi *Mean Time to Resolution* (MTTR) sekaligus menjaga reliabilitas rantai operasional bisnis ritel modern yang memiliki ratusan jaringan cabang.

4. KESIMPULAN

Algoritma pembelajaran mesin (*machine learning*) berbasis *Random Forest* telah berhasil diimplementasikan dengan sangat baik untuk mengotomatisasi proses klasifikasi teks tiket keluhan *IT Service Desk* pada jaringan gerai Burger Bangor. Konfigurasi model optimal yang dicapai melalui skenario pengujian dengan parameter jumlah pohon (*n_estimators*) sebesar 150 mampu menghasilkan tingkat akurasi prediksi keseluruhan (*overall accuracy*) yang tinggi dan stabil, yaitu mencapai 90%.

Implementasi sistem otomatisasi ini terbukti memberikan dampak perbaikan yang radikal terhadap efisiensi tata kelola layanan TI perusahaan. Model yang dibangun mampu memotong waktu sortir dan triase tiket secara drastis, dari yang semula membutuhkan waktu 10 hingga 15 menit per tiket pada sistem penyaringan manual, menjadi kurang dari dua detik per tiket. Kecepatan pemrosesan ini secara instan mengeliminasi penumpukan antrean laporan (*bottleneck*) di tingkat pusat, mengurangi faktor kesalahan manusia (*human error*) dalam penugasan tim teknis, serta mempercepat pemenuhan target *Service Level Agreement* (SLA) demi menjaga stabilitas operasional harian di seluruh outlet Burger Bangor.

REFERENSI

- [1] Arifin, M., & Susanto, A. (2022). Klasifikasi Teks Laporan Menggunakan Random Forest. *Jurnal Sains Komputer*, 5(2), 112-120.
- [2] Siregar, R. (2024). Otomatisasi IT Service Desk Pada Industri Retail. *Jurnal Teknologi Informasi Terapan*, 8(1), 45-52
- [3] Pratama, A. R., & Wijaya, K. (2023). Analisis Sentimen dan Klasifikasi Teks Menggunakan Ensembel Random Forest. *Jurnal Teknologi dan Sistem Komputer*, 11(1), 34-42.
- [4] Hidayat, T., & Utami, W. (2021). Klasifikasi Otomatis Tiket Keluhan Nasabah Perbankan Menggunakan Metode Random Forest. *Jurnal Rekayasa Sistem dan Teknologi Informasi*, 5(3), 412-419.
- [5] Lestari, D. P., Wahyuni, S., & Kusuma, A. (2023). Penerapan TF-IDF dan Random Forest untuk Kategorisasi Keluhan Pelanggan pada Industri Retail. *Jurnal Ilmiah Teknologi Informasi*, 21(1), 55-64.
- [6] Gunawan, I. (2024). Otomatisasi IT Service Desk Menggunakan Pembelajaran Mesin untuk Reduksi Mean Time to Resolution. *Jurnal Teknologi Sistem Informasi*, 10(2), 88-97
- [7] Wijaya, A., & Ramadhan, F. (2023). Optimasi Workflow IT Support Melalui Penerapan Klasifikasi

- Tiket Otomatis. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(4), 1821-1829.
- [8] Ramadhan, F., Sanjaya, R., & Budiman, A. (2023). Pengaruh Teknik Preprocessing Terhadap Akurasi Klasifikasi Keluhan Publik Berbahasa Indonesia. *Jurnal Ilmu Komputer dan Informatika*, 9(2), 145-153.
- [9] Sanjaya, R., & Budiman, A. (2022). Teknik Pengolahan Data Tidak Terstruktur Menggunakan Text Mining. *Jurnal Sains Komputer dan Teknologi*, 4(1), 45-53.
- [10] Kusuma, H., & Hartono, B. (2025). Evaluasi Ketahanan Algoritma Ensemble pada Data Tekstual Berdimensi Tinggi. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 12(1), 15-24.
- [11] Utami, W., Lestari, D. P., & Wahyuni, S. (2024). Optimasi Pembobotan TF-IDF untuk Klasifikasi Bahasa Non-Baku pada Media Sosial. *Jurnal Ilmiah Teknologi Informasi Terapan*, 10(3), 210-218.