PENGEMBANGAN STANDARISASI KATEGORI KELUHAN APLIKASI MYPELINDO BERBASIS LATENT DIRICHLET ALLOCATION DAN KLASIFIKASI

Muwirotul Hasanah[1)], Wiyli Yustanti[2)]

[1)] Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya, Surabaya, Indonesia
Email: muwirotulhasanah.22016@mhs.unesa.ac.id
[2)] Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya, Surabaya, Indonesia
Email: wiyliyustanti@unesa.ac.id

Abstract
The number of complaints received through the MyPelindo app has increased in line with its high usage for operational and administrative activities by employees. The large volume of complaint data received makes manual identification and grouping of complaints less effective and time-consuming. Therefore, the aim of this study is to extract and standardise complaint topics using the Latent Dirichlet Allocation (LDA) method, build a classification model based on the obtained topics, and implement this model in a website-based system to automatically predict the category of user complaints. The study was conducted using a text mining approach with a Knowledge Discovery in Databases (KDD) methodology, which includes the preprocessing stage, topic modelling using LDA and topic standardisation based on cosine similarity. The results of the standardisation were then used as category labels in the classification process. The best model was then implemented in a web-based system using the Flask framework. The results of the study showed that the optimal number of topics obtained was eight, with the highest coherence score of 0.353. The standardisation process successfully simplified the eight topics into three main categories: Ganti Perangkat & Error Umum, Masalah Login & Akun, and Absensi & Koreksi Data. Based on the classification evaluation results, the best model was Logistic Regression, achieving an accuracy rate of 96.05% and an F1-score of 0.96. The implementation of the web-based system also successfully predicted the category of complaints based on user input.

Keywords: User Complaints, MyPelindo, Latent Dirichlet Allocation, Standardisation, Cosine Similarity.

Abstrak
Keluhan pengguna pada aplikasi MyPelindo terus meningkat seiring dengan tingginya penggunaan aplikasi dalam aktivitas operasional dan administrasi karyawan. Banyaknya data keluhan yang masuk menyebabkan proses identifikasi dan pengelompokan keluhan secara manual menjadi kurang efektif dan membutuhkan waktu yang cukup lama. Oleh karena itu, penelitian ini bertujuan untuk melakukan ekstraksi dan standarisasi topik keluhan menggunakan metode Latent Dirichlet Allocation (LDA), membangun model klasifikasi berdasarkan hasil topik yang diperoleh, serta mengimplementasikan model tersebut ke dalam sistem berbasis website untuk memprediksi kategori keluhan pengguna secara otomatis. Penelitian dilakukan menggunakan pendekatan text mining dengan metodologi Knowledge Discovery in Databases (KDD) yang meliputi tahap preprocessing, topic modelling menggunakan LDA, standarisasi topik berdasarkan cosine similarity, Selanjutnya, hasil standarisasi digunakan sebagai label kategori dalam proses klasifikasi. Model terbaik kemudian diimplementasikan ke dalam sistem berbasis website menggunakan framework Flask. Hasil penelitian menunjukkan bahwa jumlah topik optimal yang diperoleh adalah sebanyak 8 topik dengan nilai coherence score tertinggi sebesar 0,353. Proses standarisasi berhasil menyederhanakan delapan topik tersebut menjadi tiga kategori utama, yaitu Ganti Perangkat & Error Umum, Masalah Login & Akun, serta Absensi & Koreksi Data. Berdasarkan hasil evaluasi klasifikasi, model terbaik yakni Logistic Regression memperoleh nilai akurasi sebesar 96,05% dengan nilai f1-score sebesar 0,96. Implementasi sistem berbasis website juga berhasil melakukan prediksi kategori keluhan secara otomatis berdasarkan input teks pengguna.

Kata Kunci: Keluhan Pengguna, MyPelindo, Latent Dirichlet Allocation, Standarisasi, Cosine Similarity.

PENDAHULUAN

Aplikasi digital telah menjadi salah satu infrastruktur utama dalam penyediaan layanan publik dan swasta. aplikasi berbasis seluler dan web semakin banyak digunakan oleh masyarakat untuk memenuhi berbagai kebutuhan, seperti komunikasi, transportasi, perbankan, dan layanan logistik. Berdasarkan laporan *We Are Social* dan *Hootsuite* (2024), tercatat lebih dari 221 juta penduduk Indonesia, atau sekitar 79,5% dari total populasi, telah menggunakan internet. Dari jumlah tersebut, 98,3% mengakses layanan digital melalui perangkat seluler. Selaras dengan data tersebut, Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (Susenas) juga mencatat peningkatan penggunaan internet, dari 62,10% pada tahun 2021 menjadi 66,48% pada tahun 2022. Fakta ini menegaskan tingginya adopsi teknologi serta mencerminkan keterbukaan masyarakat Indonesia terhadap arus informasi dan pergeseran menuju era masyarakat berbasis informasi.

Di Indonesia, jumlah keluhan telah meningkat sejalan dengan tren penggunaan aplikasi digital yang semakin luas. Menurut studi Kementerian Komunikasi dan Informatika (Kominfo, 2023), pemerintah telah menerima sekitar 40% lebih banyak keluhan terkait layanan digital dalam tiga tahun terakhir. Sementara itu, Kementerian Perdagangan (2022) melaporkan bahwa platform dan aplikasi digital mencatat lebih dari 92% dari total keluhan konsumen di industri perdagangan elektronik (e-commerce). Berdasarkan data ini, keluhan terkait aplikasi telah menjadi masalah serius yang perlu ditangani secara sistematis.

Hal ini juga terjadi dalam konteks bisnis yang dijalankan oleh Badan Usaha Milik Negara (BUMN) di sektor logistik dan manufaktur, seperti PT Pelabuhan Indonesia (Pelindo), yang mengembangkan aplikasi sebagai bagian dari proses transformasi digital. Aplikasi ini dirancang untuk mendukung inisiatif Indonesia Digital Economy 2025, bertujuan untuk membantu pengguna mengakses layanan pelabuhan secara terintegrasi, mulai dari penempatan kontainer secara real-time dan pengolahan jasa hingga pembayaran digital melalui integrasi dengan sistem perbankan. Menurut laporan BUMN (2023), aplikasi seperti MyPelindo merupakan contoh implementasi Gerakan Nasional 1000 Startup Digital, yang bertujuan meningkatkan produktivitas tenaga kerja nasional sebesar 30% melalui teknologi mobile-first.

Laporan Transformasi Digital BUMN (Kementerian BUMN, 2024) mengungkapkan bahwa sekitar 70% BUMN di sektor logistik dan transportasi telah mengadopsi aplikasi mobile untuk pelayanan pelanggan, dengan Pelindo menjadi salah satu pionir melalui aplikasi MyPelindo yang sudah diunduh lebih dari 50.000 kali sejak diluncurkan pada 2022 (Google Play Store, 2024).Menurut deskripsinya, MyPelindo dikembangkan sebagai aplikasi terintegrasi untuk komunikasi antara perusahaan dan karyawannya. Aplikasi ini memudahkan akses ke data karyawan, termasuk gaji, kompetensi, dan informasi biografi; memberikan umpan balik; mendistribusikan informasi internal; serta mengirimkan kuesioner dan surat edaran (Google Play Store, 2024).Meskipun banyak orang telah mengunduh aplikasi ini, data empiris dari ulasan pengguna mengindikasi adanya kesenjangan (gap) pada aspek pengalaman pengguna. Distribusi penilaian yang tertahan pada angka 3,6/5 di App Store merepresentasikan adanya beberapa aspek aplikasi yang belum optimal.

Sistem pengelolaan keluhan pada aplikasi sering menghadapi masalah klasifikasi yang tidak konsisten, sehingga menurunkan kepuasan pengguna dan memperlambat proses penanganan (Zhang, Chen, & Wang, 2019; Sun & Wang, 2020). Penelitian di industri transportasi digital Indonesia juga menemukan hal serupa, di mana pencatatan manual menyebabkan ketidakkonsistenan dan keterlambatan penyelesaian keluhan (Rahayu, Ishartani, & Setyowati, 2020). Idealnya, manajemen keluhan mampu melakukan kategorisasi yang jelas dan konsisten agar setiap aduan terarah ke unit teknis yang relevan. Namun, praktik di lapangan masih mengandalkan proses manual oleh petugas customer service yang rentan bias, tidak terstandarisasi, serta kategori yang terlalu umum. Hal ini menimbulkan kesenjangan antara kebutuhan akan sistem klasifikasi baku dengan kondisi nyata yang masih belum terotomasi.

Perkembangan teknologi analisis text atau text mining memberikan alternatif baru dalam melakukan kategorisasi dan mengklasifikasikan jenis kendala yang masuk.salah satu metode yang digunakan yakni Latent Dirichlet Allocation (LDA) adalah teknik yang banyak digunakan untuk mengidentifikasi topik tersembunyi dalam kumpulan data teks, seperti keluhan pelanggan terkait program. dengan LDA, frekuensi kata-kata tertentu dapat digunakan untuk mengidentifikasi pola keluhan secara lebih objektif. Di sisi lain, telah dibuktikan bahwa pendekatan model machine learning yakni algoritma klasifikasi berhasil dalam mengidentifikasi teks secara akurat (Cortes & Vapnik, 1995). Dengan menggabungkan kedua pendekatan ini, model hibrida yang dapat secara otomatis mengelompokkan keluhan dari pengguna aplikasi ke dalam kategori yang ditentukan dapat dibuat.

Sejumlah penelitian terdahulu menunjukkan bahwa pendekatan machine learning, khususnya kombinasi topik modeling (LDA) dan algoritma klasifikasi, mampu meningkatkan keakuratan kategorisasi teks (Zhang et al., 2019; Sun & Wang, 2020). Namun, penerapan model ini dalam konteks pengelolaan keluhan aplikasi BUMN di Indonesia masih sangat terbatas. Dengan kata lain, terdapat ruang penelitian yang luas untuk menguji efektivitas metode LDA dan Machine learning dalam kasus riil seperti aplikasi MyPelindo.

Urgensi penelitian ini semakin kuat apabila dikaitkan dengan dampak keluhan aplikasi terhadap kepuasan pengguna dan kinerja organisasi. Jika keluhan tidak ditangani secara cepat dan tepat, perusahaan berisiko kehilangan kepercayaan pengguna, mengalami penurunan reputasi, bahkan potensi kerugian finansial akibat migrasi pelanggan ke layanan alternatif. Dari sisi akademik, penelitian ini berkontribusi pada pengayaan literatur tentang penerapan metode machine learning dalam bidang manajemen keluhan digital di Indonesia, khususnya dengan fokus pada model standarisasi kategori keluhan.

Berdasarkan uraian tersebut, jelas terlihat adanya kebutuhan untuk mengembangkan standar kategori keluhan

 ISSN 3110-8245 (E)

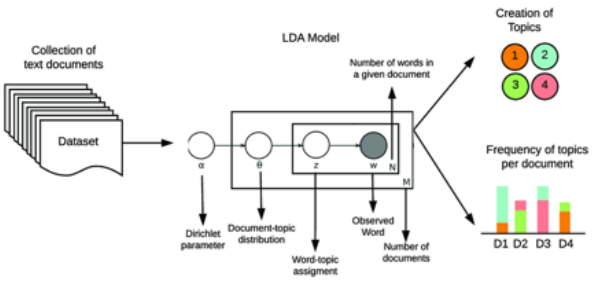
aplikasi yang terintegrasi dengan metode analisis teks modern. Penelitian ini diarahkan untuk menjawab bagaimana LDA dan algoritma klasifikasi dapat digunakan dalam membangun sistem klasifikasi keluhan yang konsisten, efisien, dan akurat pada aplikasi MyPelindo. Dari sinilah kemudian muncul pertanyaan penelitian yang menjadi dasar pelaksanaan kajian ini.

TINJAUAN PUSTAKA

Topik Modelling *Latent Dirichlet Allocation* (LDA)

Latent Dirichlet Allocation (LDA) merupakan salah satu metode *unsupervised machine learning* yang digunakan untuk menemukan struktur topik tersembunyi (*latent topics*) dalam sekumpulan dokumen teks. Metode ini pertama kali diperkenalkan oleh David M. Blei, Andrew Y. Ng, dan Michael I. Jordan (2003) sebagai model probabilistik generatif yang mengasumsikan bahwa setiap dokumen merupakan kombinasi dari beberapa topik, sedangkan setiap topik tersusun atas distribusi probabilitas kata-kata tertentu. Dengan pendekatan tersebut, LDA mampu mengidentifikasi pola hubungan antar kata sehingga menghasilkan kelompok topik yang mewakili isi dokumen secara otomatis.

Dalam proses pembentukan topik, LDA bekerja melalui distribusi probabilitas berbasis distribusi Dirichlet. Setiap dokumen memiliki proporsi topik tertentu, sedangkan setiap topik memiliki distribusi kata yang berbeda. Algoritma kemudian melakukan proses inferensi menggunakan metode seperti *Gibbs Sampling* atau *Variational Bayesian* untuk memperkirakan distribusi topik yang paling sesuai terhadap data teks. Hasil akhirnya berupa daftar kata-kata dominan pada setiap topik beserta bobot probabilitasnya sehingga dapat digunakan sebagai dasar interpretasi kategori dokumen.



Gambar 1. Model LDA

Pada penelitian yang memanfaatkan data ulasan maupun keluhan pelanggan, LDA telah banyak digunakan karena mampu mengurangi subjektivitas dalam proses identifikasi tema keluhan. Melalui pemodelan topik, kumpulan teks yang awalnya tidak terstruktur dapat dikelompokkan menjadi beberapa tema utama seperti masalah login, kegagalan transaksi, gangguan sistem, maupun kualitas layanan. Pendekatan ini membantu organisasi memahami pola keluhan pelanggan secara lebih sistematis dibandingkan proses klasifikasi manual.

Dalam konteks penelitian ini, LDA digunakan sebagai tahap awal untuk membentuk kategori keluhan berdasarkan pola kemunculan kata pada ulasan aplikasi MyPelindo. Topik-topik yang dihasilkan kemudian dianalisis untuk menentukan label kategori keluhan yang akan digunakan sebagai dasar pembentukan dataset klasifikasi. Dengan demikian, proses kategorisasi menjadi lebih objektif karena didasarkan pada struktur data yang terbentuk secara otomatis, bukan semata-mata berdasarkan interpretasi peneliti.

Klasifikasi (*Text Classification*)

Klasifikasi merupakan salah satu teknik dalam *supervised machine learning* yang bertujuan mengelompokkan suatu data ke dalam kelas atau kategori tertentu berdasarkan pola yang dipelajari dari data berlabel. Pada klasifikasi teks, algoritma mempelajari karakteristik setiap kategori berdasarkan representasi fitur dokumen sehingga mampu memprediksi kategori dokumen baru secara otomatis. Proses ini banyak diterapkan dalam analisis sentimen, deteksi spam, klasifikasi berita, hingga sistem pengelolaan keluhan pelanggan.

Tahapan klasifikasi teks umumnya dimulai dari proses *text preprocessing*, seperti *case folding*, *tokenizing*, penghapusan *stopword*, dan *stemming*. Selanjutnya dokumen diubah menjadi representasi numerik menggunakan metode seperti *Term Frequency-Inverse Document Frequency* (TF-IDF). Representasi numerik tersebut kemudian digunakan sebagai masukan bagi algoritma klasifikasi untuk membangun model prediksi.

Berbagai algoritma dapat digunakan dalam klasifikasi teks, seperti *Naïve Bayes*, *Support Vector Machine* (SVM), *Decision Tree*, *Random Forest*, maupun *Logistic Regression*. Pemilihan algoritma bergantung pada karakteristik data serta tingkat kompleksitas permasalahan. Dalam penelitian teks berukuran sedang hingga besar, algoritma SVM sering dipilih karena memiliki kemampuan generalisasi yang baik terhadap data berdimensi tinggi serta mampu menghasilkan tingkat akurasi yang tinggi.

Pada penelitian ini, proses klasifikasi dilakukan setelah kategori keluhan diperoleh melalui pendekatan LDA. Dataset yang telah memiliki label kategori kemudian digunakan sebagai data latih (*training data*) sehingga model mampu mengklasifikasikan keluhan baru ke dalam kategori yang sesuai. Pendekatan tersebut menghasilkan sistem klasifikasi otomatis yang dapat mempercepat proses identifikasi jenis keluhan pengguna aplikasi MyPelindo.

Coherence Score

Coherence Score merupakan ukuran evaluasi yang digunakan untuk menilai kualitas topik yang dihasilkan oleh model *topic modelling*. Nilai ini mengukur tingkat keterkaitan semantik antar kata dalam suatu topik sehingga dapat diketahui apakah topik tersebut mudah dipahami dan memiliki makna yang jelas. Semakin tinggi nilai *coherence*, semakin baik kualitas topik yang dihasilkan oleh model LDA.

Evaluasi menggunakan *Coherence Score* menjadi penting karena jumlah topik (*number of topics*) yang dipilih akan memengaruhi kualitas hasil pemodelan. Jika jumlah topik terlalu sedikit, beberapa tema berbeda dapat tergabung menjadi satu. Sebaliknya, jika jumlah topik terlalu banyak, topik yang dihasilkan menjadi terlalu spesifik dan sulit

diinterpretasikan. Oleh karena itu, pemilihan jumlah topik dilakukan dengan membandingkan nilai *coherence* pada beberapa variasi jumlah topik.

Beberapa jenis *Coherence Score* yang umum digunakan adalah C_V, UMass, C_UCI, dan C_NPMI. Di antara berbagai metode tersebut, C_V merupakan ukuran yang paling banyak digunakan karena memiliki korelasi tinggi terhadap interpretasi manusia terhadap kualitas topik. Nilai C_V dihitung berdasarkan kombinasi *Normalized Pointwise Mutual Information* (NPMI) dan *cosine similarity* antar vektor kata sehingga menghasilkan evaluasi yang lebih stabil.

Dalam penelitian ini, *Coherence Score* digunakan untuk menentukan jumlah topik terbaik pada model LDA. Model dengan nilai *coherence* tertinggi dipilih sebagai model yang paling representatif dalam menggambarkan struktur keluhan pengguna aplikasi MyPelindo.

Cosine Similarity

Cosine Similarity merupakan metode pengukuran tingkat kemiripan dua dokumen berdasarkan sudut yang terbentuk antara dua buah vektor. Berbeda dengan pengukuran berbasis jarak Euclidean, *Cosine Similarity* lebih menitikberatkan pada arah vektor dibandingkan panjangnya sehingga sangat sesuai digunakan pada data teks yang memiliki dimensi tinggi.

Nilai *Cosine Similarity* berada pada rentang 0 hingga 1. Nilai mendekati 1 menunjukkan bahwa dua dokumen memiliki tingkat kemiripan yang tinggi, sedangkan nilai mendekati 0 menunjukkan bahwa kedua dokumen memiliki sedikit kesamaan. Sebelum dilakukan perhitungan *Cosine Similarity*, dokumen biasanya direpresentasikan terlebih dahulu ke dalam bentuk TF-IDF sehingga setiap kata memiliki bobot sesuai tingkat kepentingannya.

Dalam penelitian klasifikasi teks, *Cosine Similarity* sering digunakan untuk mengukur kemiripan antara dokumen baru dengan dokumen referensi maupun antar topik hasil LDA. Pengukuran ini membantu memastikan bahwa dokumen ditempatkan pada kategori yang paling sesuai berdasarkan karakteristik isi teks.

Pada penelitian ini, *Cosine Similarity* dimanfaatkan sebagai metode pendukung untuk mengukur kedekatan antar dokumen maupun antara dokumen dengan kategori hasil pemodelan topik. Penggunaan metode ini diharapkan mampu meningkatkan konsistensi dalam proses klasifikasi keluhan pengguna aplikasi MyPelindo.

Manajemen Data Keluhan

Manajemen data keluhan merupakan proses sistematis dalam mengumpulkan, menyimpan, mengelompokkan, menganalisis, serta menindaklanjuti keluhan pelanggan sebagai dasar peningkatan kualitas layanan. Dalam era transformasi digital, data keluhan tidak hanya berfungsi sebagai media komunikasi antara pelanggan dan perusahaan, tetapi juga menjadi sumber informasi strategis dalam pengambilan keputusan.

Data keluhan umumnya berasal dari berbagai sumber, seperti aplikasi mobile, media sosial, surat elektronik, pusat layanan pelanggan (*call center*), maupun formulir pengaduan daring. Sebagian besar data tersebut berbentuk teks tidak terstruktur sehingga memerlukan proses *text mining* sebelum dapat dianalisis secara efektif. Tanpa adanya sistem pengelolaan yang baik, informasi penting dalam keluhan pelanggan sulit dimanfaatkan sebagai dasar evaluasi layanan.

Manajemen keluhan yang efektif melibatkan beberapa tahapan, yaitu penerimaan keluhan, validasi data, klasifikasi, penugasan kepada unit terkait, penyelesaian masalah, hingga evaluasi tingkat kepuasan pelanggan. Standarisasi kategori keluhan menjadi salah satu aspek penting karena menentukan kecepatan proses distribusi aduan kepada unit teknis yang bertanggung jawab.

Dalam penelitian ini, data keluhan diperoleh dari ulasan pengguna aplikasi MyPelindo pada platform distribusi aplikasi. Data tersebut kemudian melalui proses pembersihan data, pembentukan topik menggunakan LDA, pelabelan kategori, serta klasifikasi otomatis sehingga dihasilkan sistem pengelolaan keluhan yang lebih konsisten, efisien, dan berbasis data.

Rapid Application Development (RAD)

Rapid Application Development (RAD) merupakan model pengembangan perangkat lunak yang menekankan proses pembangunan sistem secara cepat melalui pendekatan prototyping dan keterlibatan aktif pengguna. Model ini diperkenalkan oleh James Martin pada awal 1990-an sebagai alternatif dari model pengembangan perangkat lunak tradisional yang membutuhkan waktu relatif lama.

RAD terdiri atas beberapa tahapan utama, yaitu *Requirements Planning*, *User Design*, *Construction*, dan *Cutover*. Pada tahap *Requirements Planning*, kebutuhan sistem diidentifikasi bersama pengguna. Tahap *User Design* berfokus pada pembuatan prototipe antarmuka dan alur sistem. Selanjutnya, tahap *Construction* dilakukan pengembangan aplikasi berdasarkan prototipe yang telah disepakati. Tahap terakhir adalah *Cutover*, yaitu implementasi sistem, pengujian akhir, pelatihan pengguna, serta proses penerapan ke lingkungan operasional.

Keunggulan RAD terletak pada kemampuannya menghasilkan aplikasi dalam waktu relatif singkat karena proses pengembangan dilakukan secara iteratif dengan umpan balik pengguna yang berkelanjutan. Pendekatan ini sangat sesuai diterapkan pada sistem informasi yang membutuhkan penyempurnaan antarmuka maupun fitur secara bertahap berdasarkan kebutuhan pengguna.

Dalam penelitian ini, metode RAD digunakan sebagai kerangka pengembangan sistem klasifikasi keluhan berbasis web. Pendekatan tersebut memungkinkan proses pembangunan aplikasi dilakukan secara cepat melalui evaluasi prototipe secara berulang sehingga sistem yang dihasilkan mampu memenuhi kebutuhan pengguna dalam melakukan pengelolaan, kategorisasi, dan analisis data keluhan aplikasi MyPelindo secara efektif.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *Knowledge Discovery in Databases* (KDD) sebagai kerangka kerja utama dalam proses pengolahan data. Tahapan KDD meliputi pemilihan data (*selection*),

prapemrosesan (*preprocessing*), transformasi (*transformation*), *data mining*, serta evaluasi hasil. Pendekatan ini dipilih karena mampu mengolah data teks dalam jumlah besar secara sistematis untuk menghasilkan informasi yang relevan dari kumpulan keluhan pengguna aplikasi MyPelindo. Seluruh proses penelitian dilakukan secara bertahap mulai dari pengumpulan data, pengolahan teks, pembentukan topik, klasifikasi, hingga evaluasi performa model.

Data yang digunakan merupakan data sekunder berupa 8.603 keluhan pengguna aplikasi MyPelindo yang berasal dari sistem internal pengaduan pelanggan selama periode 3 Januari 2025 hingga 13 April 2026. Dataset terdiri atas beberapa atribut, antara lain nomor tiket, waktu pelaporan, status tiket, wilayah pelapor, serta kolom **Keterangan** yang berisi deskripsi keluhan pengguna dalam bentuk teks tidak terstruktur. Pada penelitian ini, analisis difokuskan pada kolom **Keterangan** karena memuat informasi utama mengenai permasalahan yang dialami pengguna. Sebelum diproses lebih lanjut, data diperiksa untuk memastikan tidak terdapat data duplikat maupun data yang tidak lengkap sehingga kualitas dataset tetap terjaga.

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data teks agar siap dianalisis menggunakan metode *machine learning*. Tahapan ini meliputi *cleaning*, yaitu menghapus karakter khusus, angka, tanda baca, URL, emoji, serta data duplikat; *case folding* untuk mengubah seluruh huruf menjadi huruf kecil; *tokenizing* untuk memecah kalimat menjadi token kata; *stemming* menggunakan algoritma stemming bahasa Indonesia guna mengubah kata berimbuhan menjadi kata dasar; serta *stopword removal* untuk menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap proses analisis. Hasil dari tahapan ini berupa kumpulan token bersih yang merepresentasikan isi keluhan secara lebih efektif.

Selanjutnya dilakukan tahap transformasi data sesuai kebutuhan masing-masing metode analisis. Untuk proses *topic modelling*, data diubah menjadi representasi *Bag-of-Words* (BoW) melalui pembentukan *dictionary* dan *corpus* sebagai masukan bagi algoritma *Latent Dirichlet Allocation* (LDA). Jumlah topik terbaik ditentukan menggunakan nilai *Coherence Score* sehingga topik yang dihasilkan memiliki tingkat koherensi semantik yang tinggi. Setelah topik diperoleh, peneliti melakukan interpretasi terhadap kata-kata dominan pada setiap topik untuk membentuk standar kategori keluhan yang digunakan sebagai label dalam proses klasifikasi.

Tahap berikutnya adalah klasifikasi keluhan menggunakan pendekatan *supervised machine learning*. Data yang telah diberi label berdasarkan hasil pemodelan LDA kemudian ditransformasikan menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) sehingga setiap dokumen direpresentasikan dalam bentuk vektor numerik. Vektor tersebut digunakan sebagai masukan bagi algoritma klasifikasi untuk membangun model yang mampu mengategorikan keluhan baru secara otomatis. Selain itu, pengukuran *Cosine Similarity* dimanfaatkan untuk mengevaluasi tingkat kemiripan antar dokumen maupun kesesuaian dokumen terhadap kategori yang terbentuk sehingga konsistensi hasil klasifikasi dapat ditingkatkan.

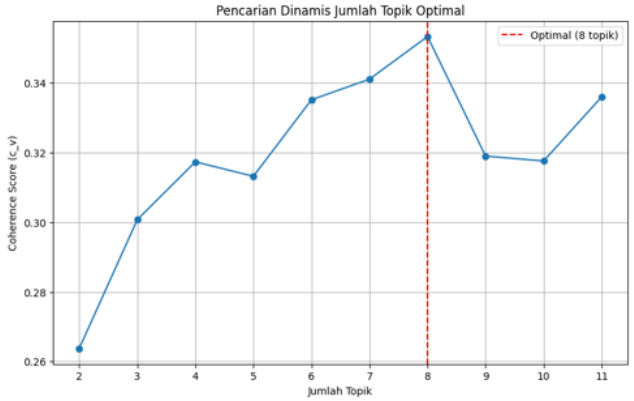
Gambar 2. Kerangka Penelitian

Tahap terakhir adalah evaluasi model dan pengembangan sistem berbasis metode *Rapid Application Development* (RAD). Kinerja model klasifikasi dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* berdasarkan hasil pengujian terhadap data uji. Model dengan performa terbaik kemudian diimplementasikan ke dalam aplikasi berbasis web yang dikembangkan menggunakan pendekatan RAD melalui tahapan *requirements planning*, *user design*, *construction*, dan *cutover*. Sistem yang dihasilkan mampu melakukan klasifikasi keluhan secara otomatis berdasarkan kategori yang telah dibentuk, sehingga dapat membantu proses pengelolaan pengaduan pengguna aplikasi MyPelindo menjadi lebih cepat, konsisten, dan efisien.

HASIL DAN PEMBAHASAN

Tahap analisis diawali dengan penerapan metode **Latent Dirichlet Allocation (LDA)** terhadap 8.603 data keluhan pengguna aplikasi MyPelindo yang telah melalui tahapan *preprocessing*, yaitu *cleaning*, *case folding*, *tokenizing*, *stemming*, dan *stopword removal*. Proses ini bertujuan untuk mengidentifikasi pola topik laten yang merepresentasikan jenis-jenis permasalahan yang sering dialami pengguna. Seluruh data kemudian dikonversi ke dalam bentuk *Bag-of-Words* (BoW) dan digunakan sebagai masukan pada model LDA. Hasil *topic modelling* menjadi dasar dalam pembentukan kategori keluhan yang selanjutnya digunakan sebagai label pada proses klasifikasi.

 ISSN 3110-8245 (E)



Gambar 3. Hasil Evaluasi Model LDA dengan Cohorence Score

Penentuan jumlah topik optimal dilakukan menggunakan pendekatan **dynamic topic search** dengan evaluasi **Coherence Score (C_v)**. Pengujian dilakukan pada beberapa variasi jumlah topik, kemudian nilai *coherence* dihitung pada setiap model. Proses pencarian dihentikan secara otomatis menggunakan mekanisme *early stopping* apabila tidak terjadi peningkatan nilai *coherence* dalam tiga iterasi berturut-turut. Pendekatan ini dipilih untuk memperoleh jumlah topik yang mampu memberikan representasi terbaik terhadap keseluruhan data keluhan.

Tabel 1. Hasil Pengujian Coherence Score

Jumlah Topik (K)	Coherence Score (Cv)
5	0,281
6	0,309
7	0,336
8	**0,353**
9	0,348
10	0,344

Berdasarkan Tabel 1, model dengan **8 topik** menghasilkan nilai **Coherence Score** tertinggi sebesar **0,353**, sehingga dipilih sebagai model terbaik. Nilai tersebut menunjukkan bahwa delapan topik yang terbentuk memiliki tingkat keterkaitan semantik paling baik dibandingkan variasi jumlah topik lainnya. Setiap topik kemudian diinterpretasikan berdasarkan kata-kata dominan sehingga diperoleh beberapa tema keluhan, seperti masalah login, pergantian perangkat, kesalahan absensi, sinkronisasi data, hingga gangguan sistem aplikasi.

Tahap berikutnya adalah melakukan **standarisasi topik** menggunakan **Cosine Similarity**. Pengukuran ini bertujuan untuk mengetahui tingkat kemiripan antar topik berdasarkan distribusi probabilitas kata (*topic-term distribution*). Topik-topik yang memiliki tingkat kemiripan tinggi dianggap merepresentasikan permasalahan yang sama sehingga dapat digabungkan menjadi satu kategori yang lebih representatif. Dengan pendekatan ini, struktur kategori menjadi lebih sederhana tanpa menghilangkan informasi penting yang terkandung dalam data.

Tabel 2. Hasil Standarisasi Topik Menggunakan Cosine Similarity

Topik LDA	Kategori Hasil Standarisasi
Topik 1, 3, 5	Ganti Perangkat & Error Umum
Topik 2, 4	Masalah Login & Akun
Topik 6, 7, 8	Absensi & Koreksi Data

Hasil pada Tabel 2 menunjukkan bahwa delapan topik hasil LDA berhasil disederhanakan menjadi **tiga kategori utama**, yaitu **Ganti Perangkat & Error Umum**, **Masalah Login & Akun**, serta **Absensi & Koreksi Data**. Proses standarisasi ini menghasilkan struktur kategori yang lebih ringkas dan mudah dipahami, sehingga dapat digunakan sebagai acuan dalam proses klasifikasi otomatis. Selain itu, penggabungan topik juga mampu mengurangi redundansi akibat kemiripan kata yang muncul pada beberapa topik berbeda.

Setelah proses standarisasi selesai, setiap data keluhan diberi label sesuai kategori hasil penggabungan. Dataset berlabel tersebut kemudian ditransformasikan menggunakan metode **TF-IDF** dan digunakan sebagai data pelatihan pada beberapa algoritma klasifikasi. Berdasarkan hasil pengujian, algoritma **Logistic Regression** memberikan performa terbaik dibandingkan model lain yang diuji. Hal ini menunjukkan bahwa representasi TF-IDF mampu menangkap karakteristik kata yang membedakan masing-masing kategori keluhan sehingga menghasilkan model klasifikasi dengan tingkat akurasi yang tinggi.

Tabel 3. Hasil Evaluasi Model Klasifikasi Terbaik

Algoritma	Akurasi	Precision	Recall	F1-Score
Logistic Regression	96,05%	0,96	0,96	0,96

Berdasarkan Tabel 3, algoritma **Logistic Regression** memperoleh nilai **akurasi sebesar 96,05%** dengan nilai **precision**, **recall**, dan **F1-score** masing-masing sebesar **0,96**. Nilai tersebut menunjukkan bahwa model mampu mengklasifikasikan keluhan pengguna dengan tingkat ketepatan yang sangat baik. Tingginya performa model dipengaruhi oleh kualitas proses *preprocessing*, representasi fitur menggunakan TF-IDF, serta proses standarisasi kategori yang berhasil menghasilkan label yang lebih konsisten sebelum dilakukan proses klasifikasi.

Model klasifikasi terbaik selanjutnya diimplementasikan ke dalam sistem berbasis website menggunakan framework **Flask**. Sistem memungkinkan pengguna memasukkan teks keluhan secara langsung, kemudian melakukan proses *preprocessing*, ekstraksi fitur, dan prediksi kategori secara otomatis menggunakan model yang telah dilatih. Hasil prediksi ditampilkan secara real-time sehingga petugas dapat segera mengetahui jenis keluhan yang diajukan pengguna. Implementasi ini menunjukkan bahwa integrasi metode LDA, *Cosine Similarity*, dan **Logistic Regression** mampu mendukung proses pengelolaan keluhan menjadi lebih cepat, konsisten, dan efisien dibandingkan proses klasifikasi manual.

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan pada penelitian mengenai Pengambangan standarisasi kategori keluhan aplikasi MyPelindo menggunakan Latent Dirichlet Allocation dan Klasifikasi.diperoleh simpulan sebagai berikut

1. Proses ekstraksi topik keluhan pengguna aplikasi MyPelindo menggunakan metode Latent Dirichlet Allocation (LDA) berhasil mengidentifikasi struktur topik laten dari data teks keluhan. Melalui evaluasi coherence score, diperoleh jumlah topik optimal sebanyak 8 topik. Selanjutnya, dilakukan proses standarisasi menggunakan pendekatan cosine similarity untuk mengukur kemiripan antar topik. Hasilnya, delapan topik tersebut berhasil disederhanakan menjadi tiga kategori utama yang lebih representatif, yaitu Ganti Perangkat & Error Umum, Masalah Login & Akun, serta Absensi & Koreksi Data. Proses ini terbukti mampu mengurangi kompleksitas hasil topic modelling sekaligus meningkatkan interpretabilitas kategori keluhan.
2. Model klasifikasi dibangun dengan memanfaatkan hasil standarisasi topik sebagai label kategori. Proses pelabelan data dapat mentransformasi dari pendekatan unsupervised learning (LDA) menjadi supervised learning. Berdasarkan hasil perbandingan beberapa algoritma klasifikasi, model Logistic Regression menunjukkan performa terbaik dengan nilai akurasi sebesar 96,05% dan f1-score sebesar 0,96. Setelah dilakukan optimasi menggunakan GridSearchCV, model memperoleh parameter terbaik dengan akurasi validasi silang sebesar 0,9823. Hasil ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang tinggi dan stabil dalam mengidentifikasi kategori keluhan pengguna.
3. Model yang telah dibangun kemudian berhasil diimplementasikan dalam sistem berbasis website menggunakan framework Flask. Sistem mampu menerima input teks keluhan dari pengguna, melakukan proses ekstraksi fitur menggunakan TF-IDF, serta menghasilkan prediksi kategori secara otomatis menggunakan model Logistic Regression. Hasil pengujian menunjukkan bahwa sistem dapat berjalan dengan baik, mampu menangani berbagai variasi input, serta memberikan hasil prediksi secara cepat dan konsisten. Dengan demikian, implementasi ini tidak hanya membuktikan keberhasilan model secara teoritis, tetapi juga menunjukkan bahwa model dapat diaplikasikan secara praktis dalam skenario penggunaan nyata.

Saran

Berdasarkan hasil penelitian yang diperoleh,terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan penelitian selanjutnya

1. Penelitian selanjutnya disarankan untuk menggunakan dataset ulasan yang lebih besar dan beragam agar model machine learning dapat menangkap variasi pola bahasa pengguna secara lebih komprehensif.
2. Pada tahap klasifikasi, penggunaan model yang lebih kompleks seperti *Deep Learning* atau *pre-trained language model* (misalnya BERT) dapat dipertimbangkan untuk meningkatkan kemampuan model dalam memahami konteks bahasa.
3. Dari sisi implementasi, sistem dapat dikembangkan lebih lanjut agar siap digunakan dalam lingkungan nyata (*production-ready*), termasuk penambahan fitur manajemen pengguna, peningkatan keamanan data, serta optimasi performa sistem.

Pengembangan sistem berbasis *multimodal* juga dapat dipertimbangkan dengan mengintegrasikan data non-teks seperti gambar atau tangkapan layar, sehingga mampu menangkap konteks keluhan secara lebih komprehensif.

DAFTAR PUSTAKA

Chen, W. K., Riantama, D., & Chen, L. S. (2020). Using a text mining approach to hear voices of customers from social media toward the fast-food restaurant industry. Sustainability, 13(1), 268.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences, 101(suppl 1), 5228-5235

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. Journal of the American society for information science, 41(6), 391-407.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Alkaff, M., Baskara, A. R., & Maulani, I. (2021). Klasifikasi Laporan Keluhan Pelayanan Publik Berdasarkan Instansi Menggunakan Metode LDA-SVM. Jurnal Teknologi Informasi dan Ilmu Komputer, 8(6).

Marthasari, G. I., Hayatin, N., & Yuniarti, M. (2022). Content Classification based-on Latent Semantic Analysis and Support Vector Machine (LSA-SVM). Jurnal Transformatika, 19(2), 144-150.

Saikin, S., Fadli, S., Akbar, J., & Fahmi, H. (2025). TOPIC MODELING JUDUL PENELITIAN MENGGUNAKAN METODE LATENT DIRICHLET ALLOCATION (LDA) DAN FAKTORISASI MATRIKS NON-NEGATIF (NMF). JATI (Jurnal Mahasiswa Teknik Informatika), 9(2), 3439-3445.

Abdurohman, M., Husna, R., Ali, I., Dwilestari, G., & Rahaningsih, N. (2022). Penerapan Model Klasifikasi Dalam Tingkat Kepuasan Layanan Publik Kelurahan Karyamulya Dengan Menggunakan Algoritma Decision Tree. INFORMATION MANAGEMENT

FOR EDUCATORS AND PROFESSIONALS: Journal of Information Management, 6(1), 81-90.

Diansyah, S. (2022). Klasifikasi tingkat kepuasan pengguna dengan menggunakan metode k-nearest neighbors (knn). Jurnal Sistim Informasi Dan Teknologi, 7-12.

Puspitasari, D. (2025). Analisis Sentimen Berbasis Aspek Pada Ulasan Multibahasa Dari Aplikasi Gobis Suroboyo Menggunakan Lda dan Svm (Doctoral dissertation, UPN Veteran Jawa Timur).

Fatmawati, F., & Affandes, M. (2018). Klasifikasi Keluhan Menggunakan Metode Support Vector Machine (SVM) Pada Akun Facebook Group iRaise Helpdesk. Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi, 3(1), 24-30.

Abdurrazzaq, M. A. (2023). Analisis Ulasan Aplikasi MyPertamina Menggunakan Topic Modeling dengan Latent Dirichlet Allocation. KALBISCIENTIA Jurnal Sains dan Teknologi, 10(1), 1-6.

Kustiyahningsih, Y., & Permana, Y. (2024). Penggunaan Latent Dirichlet Allocation (LDA) dan Support-Vector Machine (SVM) Untuk Menganalisis Sentimen Berdasarkan Aspek Dalam Ulasan Aplikasi EdLink. Teknika, 13(1), 127-136.

Saputra, H. T., Damanik, A., Shaquille, M. M., Rusydi, M. A., Riziq, M. H., & Zulva, N. S. (2025). Analisis Modelling Pada Reviewes Lazada Indonesia Menggunakan Latent Dirichlet Allocation (LDA) Untuk Optimalisasi Strategi Bisnis. JEKIN-Jurnal Teknik Informatika, 5(1), 361-371.

Puspita, E., Shiddieq, D. F., & Roji, F. F. (2024). Pemodelan Topik pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc): Topic Modeling on Online News Media Using Latent Diriclet Allocation (Case Study Somethinc Brand). MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4(2), 481-489.

IdCloudHost. (2025). Mengenal Rapid Application Development

Metodologi Fleksibel. https://idcloudhost.com/rapidapplication-development/

Martin, J. (1991). Rapid Application Development. Macmillan

Publishing Company.

Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

Hofmann, T. (1999). Probabilistic latent semantic analysis. In Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI), 289-296.

Jurafsky, D., & Martin, J. H. (2021). Speech and Language Processing (3rd ed. draft). Stanford University.

Wallach, H. M. (2008). Structured Topic Models for Language (Doctoral dissertation, University of Cambridge).

Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann/Elsevier.

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. Informatica, 31, 249-268.

Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In Mining Text Data, 163-222. Springer.

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). Introduction to Data Mining (2nd ed.). Pearson.

Quinlan, J. R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttler, D. (2012). Exploring topic coherence over many models and many topics. Proceedings of the 2012 Joint Conference on EMNLP and CoNLL, 952-961.

Roder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 399-408.

Aggarwal, C. C. (2018). Machine Learning for Text. Springer.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI Magazine, 17(3), 37-54.

Dunham, M. H. (2003). Data Mining: Introductory and Advanced Topics. Prentice Hall.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.

Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. Proceedings of the First Instructional Conference on Machine Learning, 133-142.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), 2, 1137-1143.