

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring pertumbuhan jumlah penduduk, permintaan akan sarana transportasi ikut meningkat. Transportasi sendiri terbagi menjadi dua jenis, yakni transportasi pribadi, seperti sepeda, sepeda motor, dan mobil pribadi, serta transportasi umum, seperti bus dan kereta, yang bertujuan memberikan layanan aman, nyaman, cepat, dan terjangkau. Namun, banyak masyarakat masih memilih kendaraan pribadi karena faktor kenyamanan, aksesibilitas, dan jadwal transportasi umum yang kurang fleksibel, sehingga menyebabkan kemacetan. Untuk mengatasi hal ini, pemerintah perlu meningkatkan kualitas transportasi umum dengan memperbaiki kenyamanan, memperluas rute, dan menyediakan fasilitas pendukung agar masyarakat beralih ke transportasi umum [1].

Transportasi umum menjadi solusi efektif untuk mengurangi kemacetan, terutama di daerah perkotaan dengan mobilitas yang tinggi. Berdasarkan data BPS, pada tahun 2024, jumlah penduduk di Jawa Timur mencapai 41,64 juta jiwa, yang berkontribusi pada peningkatan volume lalu lintas, terutama kendaraan pribadi. Sebagai pusat industri dan keuangan, Jawa Timur menghadapi tantangan transportasi yang semakin kompleks. Salah satu cara untuk mengatasinya adalah dengan mengoptimalkan transportasi umum seperti Bus Trans Jatim, yang memiliki kapasitas besar, fasilitas yang memadai, serta jadwal yang teratur. Beberapa kota besar di Indonesia telah berhasil mengembangkan sistem transportasi massal guna mengurangi kepadatan lalu lintas. Trans Jatim, yang merupakan bagian dari program Bus Rapid Transit (BRT), dapat menjadi pilihan utama masyarakat dengan layanan yang efisien, nyaman, dan terjangkau. Seiring dengan pertumbuhan populasi dan ekonomi, kebutuhan akan transportasi publik yang lebih baik semakin meningkat, sehingga pengembangan layanan bus dan moda transportasi umum lainnya menjadi langkah penting dalam mendukung mobilitas masyarakat [2].

Bus Trans Jatim merupakan layanan transportasi publik berbasis Bus Rapid Transit (BRT) yang beroperasi di wilayah aglomerasi perkotaan di Provinsi Jawa Timur, termasuk kawasan Gerbangkertosusila. Layanan Bus Trans Jatim diresmikan oleh Gubernur Jawa Timur, Khofifah Indar Parawansa, pada 19 Agustus

2022. Pada awal pelaksanaannya, bus ini hanya melayani Koridor I, dan kini telah berkembang hingga Koridor V, yaitu : Koridor I (Sidoarjo-Surabaya-Gresik), Koridor II (Mojokerto-Surabaya), Koridor III (Mojokerto-Gresik), Koridor IV (Gresik-Lamongan), dan Koridor V (Surabaya-Bangkalan). Terkait tarif, harga tiket telah ditetapkan oleh penyelenggara, sebagaimana disampaikan oleh Bu Susi, yaitu Rp 5.000 untuk masyarakat umum dan Rp 2.500 untuk pelajar/mahasiswa. Peluncuran Bus Trans Jatim bertujuan meningkatkan kualitas layanan transportasi publik dengan melibatkan pihak swasta sebagai pengelola. Program ini juga menjadi solusi inovatif serta wujud visi dalam menyediakan alternatif mobilitas bagi masyarakat. Pada minggu pertama operasionalnya, jumlah pengguna layanan ini mencapai 2.800 hingga 3.000 penumpang per hari, dan terus meningkat hingga 3.500–4.000 penumpang per hari [4].

Bus Trans Jatim merupakan moda transportasi aglomerasi yang berfungsi menghubungkan berbagai pusat aktivitas seperti kawasan industri, perdagangan, dan pemukiman di wilayah Jawa Timur. Untuk meningkatkan kualitas layanan transportasi publik, Dinas Perhubungan Provinsi Jawa Timur meluncurkan TransJatim Ajaib (Aplikasi Jatim Informasi Bus) sebagai inovasi digital yang menyediakan informasi *realtime* terkait jadwal keberangkatan, posisi bus, serta rute dan halte terdekat. Aplikasi ini juga berfungsi sebagai media pembayaran elektronik dan sarana informasi pelayanan publik yang mempermudah masyarakat dalam mengakses layanan Bus Trans Jatim [6]. Tren penggunaan Bus Trans Jatim menunjukkan peningkatan signifikan; misalnya, pada periode Januari–Maret 2025 tercatat jumlah penumpang mencapai 627.946 pada Koridor I, 284.305 pada Koridor II, 210.636 pada Koridor III, 246.892 pada Koridor IV, 233.775 pada Koridor V [7]. Hal ini menunjukkan bahwa masyarakat mulai beralih ke moda transportasi publik berbasis digital, meskipun masih diperlukan peningkatan operasional dan kenyamanan agar pertumbuhan pengguna dapat berkelanjutan.

Meskipun jumlah pengguna Bus Trans Jatim terus menunjukkan peningkatan, layanan ini masih menghadapi sejumlah kendala yang dikeluhkan oleh penumpang. Sejak mulai beroperasi pada Agustus 2022, berbagai permasalahan muncul, seperti kondisi halte yang mengalami kebocoran, akses keluar masuk halte yang sempit sehingga menyebabkan penumpukan penumpang, serta waktu tunggu bus yang

terlalu lama terutama pada jam sibuk. Selain itu, kebersihan di dalam bus juga masih dinilai kurang optimal [1]. Pengguna juga menyoroti keterbatasan jumlah dan penempatan halte, jam operasional yang belum fleksibel, serta minimnya armada dan rute perjalanan yang tersedia [4]. Fasilitas pendukung pun masih belum memadai, misalnya papan informasi rute di halte maupun di dalam bus yang belum terpasang secara menyeluruh, sehingga menyulitkan penumpang yang belum menggunakan aplikasi TransJatim Ajaib. Bahkan, beberapa halte belum dilengkapi dengan stop kontak untuk pengisian daya, dan ada pula halte yang tidak memiliki aliran listrik, sehingga kondisi menjadi gelap pada malam hari [3].

Ulasan serta opini penumpang yang tersebar di berbagai platform media sosial merupakan sumber informasi berharga untuk mengevaluasi tingkat kepuasan dan mengidentifikasi kendala dalam penggunaan layanan Bus Trans Jatim. Informasi ini penting karena dapat menjadi dasar bagi penyedia layanan dalam meningkatkan kualitas transportasi publik sesuai kebutuhan masyarakat. Namun, keberagaman sudut pandang dan besarnya volume ulasan membuat proses analisis menjadi semakin kompleks. Oleh karena itu, diperlukan metode yang mampu mengolah informasi tersebut secara sistematis dan terstruktur. Dalam penelitian ini digunakan pendekatan *Topic Modeling* dengan *Latent Dirichlet Allocation* (LDA), yang dinilai efektif untuk menganalisis kumpulan dokumen berskala besar dan mengidentifikasi pola tersembunyi dalam opini publik [5].

Penelitian terdahulu yang relevan dengan penelitian [8] menggunakan *Latent Dirichlet Allocation* (LDA) sebagai metode pemodelan topik untuk mengidentifikasi isu-isu terkait pinjaman online. Penelitian tersebut menerapkan pendekatan *Knowledge Discovery in Databases* (KDD), yang mencakup tahap pengumpulan data, *preprocessing*, transformasi, *data mining*, dan evaluasi hasil. Hasil penelitian menunjukkan bahwa skenario terbaik diperoleh dengan iterasi 10, $\alpha = 0.1$, dan $\beta = 0.1$, yang menghasilkan coherence score tertinggi sebesar 0,391.

Penelitian lainnya [9], penelitian ini melakukan analisis data dan pemodelan topik menggunakan LDA pada 33.896 ulasan pengguna Shopee di PlayStore, dengan mempertimbangkan parameter *passes* (5, 10, 15, 20) dan *iterations* (50, 100) serta *random state* 142. Hasil pengujian menunjukkan bahwa kombinasi *passes* 10 dan *iterations* 50 menghasilkan *coherence score* tertinggi, parameter

tersebut dipilih sebagai konfigurasi akhir untuk model LDA, yang menghasilkan tujuh topik utama, yakni kepuasan pengguna, pengiriman dan kesesuaian barang, kepercayaan pengguna, pengalaman pengguna, layanan dan bantuan, aplikasi, serta promosi dan penawaran. Hasil analisis menunjukkan bahwa topik aplikasi merupakan yang paling kritis karena menerima jumlah rating 1 terbanyak, dengan kata-kata bernilai bobot tinggi seperti lambat, berat, dan buka, menandakan bahwa isu utama yang perlu diperbaiki berkaitan dengan kinerja sistem dan fungsionalitas aplikasi.

Penelitian lainnya [10], yang dimana penelitian menerapkan LDA sebagai metode pemodelan topik guna menemukan pola yang bermakna dalam korpus medis dan meningkatkan efektivitas konsultasi dalam sistem telemedicine. Penelitian ini dilakukan melalui tahapan pengumpulan data, preprocessing teks, penerapan model LDA, serta evaluasi hasil menggunakan *coherence score*. Hasil penelitian menunjukkan bahwa pemodelan topik menggunakan LDA berhasil memberikan performa yang baik, dengan *coherence score* tertinggi sebesar 0,551075 dan standar deviasi tertinggi sebesar 0,5327286, yang membuktikan efektivitas pendekatan ini dalam mengidentifikasi topik dalam data telemedicine.

Penelitian lainnya [11], penelitian ini menerapkan *Latent Dirichlet Allocation* (LDA) sebagai metode pemodelan topik guna mengelompokkan berita secara otomatis. Tahapan penelitian meliputi pengumpulan data dari portal berita detik.com, *preprocessing* teks, penerapan model LDA, serta evaluasi hasil. Hasil penelitian menunjukkan bahwa metode LDA berhasil mengidentifikasi tiga kategori utama yang sering muncul dalam berita, yaitu topik 1 tentang kejadian bencana alam, topik 2 mengenai tokoh dan permasalahan politik, serta topik 3 yang membahas berita seputar Piala Dunia. Temuan ini membuktikan bahwa pemodelan topik dengan LDA dapat digunakan untuk menyusun sistem keputusan topik berita secara lebih sistematis dan efisien.

Berdasarkan permasalahan di atas, penelitian ini dilakukan untuk menganalisis topik-topik utama dalam ulasan terkait pelayanan Bus Trans Jatim menggunakan metode *Latent Dirichlet Allocation* (LDA). Metode ini dipilih karena mampu mengelompokkan kata-kata dalam ulasan pengguna ke dalam kumpulan topik tersembunyi (*latent topics*) secara otomatis tanpa memerlukan label sentimen

atau data pelatihan sebelumnya. Berbeda dengan analisis sentimen yang hanya mengukur kecenderungan emosi positif atau negatif terhadap suatu layanan, LDA dapat mengidentifikasi topik-topik spesifik seperti kenyamanan, tarif, jadwal, kebersihan, atau keamanan yang sering dibahas publik, sehingga memberikan pemahaman yang lebih mendalam dan terstruktur terhadap opini pengguna. Dengan demikian, hasil penelitian ini dapat memberikan gambaran komprehensif mengenai topik pelayanan yang paling sering dibahas oleh pengguna Bus Trans Jatim. Informasi tersebut dapat dijadikan dasar oleh pengelola dan Dinas Perhubungan Provinsi Jawa Timur dalam meningkatkan kualitas pelayanan berdasarkan umpan balik masyarakat, sekaligus mempertahankan topik-topik layanan yang telah dinilai baik oleh pengguna.

1.2 Rumusan Masalah

Berikut merupakan rumusan masalah pada penelitian ini yang terdiri dari permasalahan, solusi permasalahan, dan pertanyaan penelitian.

1.2.1 Permasalahan

Berdasarkan latar belakang yang telah dijelaskan, dapat diidentifikasi permasalahan bahwa penyedia layanan masih mengalami kesulitan dalam mengelompokkan topik utama dari ulasan penumpang secara manual terkait berbagai topik layanan. Selain itu, analisis yang hanya dilakukan secara umum tanpa pemetaan topik yang jelas dapat menyebabkan kurangnya pemahaman terhadap isu-isu spesifik yang sering muncul. Oleh karena itu, diperlukan metode *topic modeling* yang mampu mengidentifikasi dan mengelompokkan berbagai topik dalam ulasan penumpang secara otomatis, sehingga penyedia layanan dapat memperoleh wawasan yang lebih terstruktur untuk meningkatkan kualitas pelayanan Bus Trans Jatim.

1.2.2 Solusi Permasalahan

Berdasarkan permasalahan yang diidentifikasi, maka solusi yang dapat diusulkan dalam penelitian ini adalah penggunaan metode *Latent Dirichlet Allocation* diterapkan dalam *topic modeling* untuk menganalisis ulasan terkait pelayanan Bus Transjatim. Pendekatan ini membuat pengenalan serta pengelompokan topik utama yang kerap dibahas oleh pengguna, sehingga

memberikan wawasan yang lebih mendalam mengenai topik layanan yang paling sering dibahas oleh penggunaannya.

1.2.3 Pertanyaan Penelitian

Berdasarkan permasalahan dan solusi yang diidentifikasi, maka pertanyaan penelitian yang dapat diusulkan dalam penelitian ini adalah sebagai berikut :

1. Berapakah jumlah topik (k) yang menghasilkan nilai *Coherence Score* tertinggi pada pemodelan *Topic Modeling* menggunakan *Latent Dirichlet Allocation* (LDA) terhadap ulasan pelayanan Bus Trans Jatim?
2. Bagaimana pengaruh parameter alpha, beta, dan jumlah iterasi terhadap nilai *Coherence Score* dalam pemodelan topik menggunakan metode LDA pada ulasan pelayanan Bus Trans Jatim?
3. Topik apa saja yang dihasilkan oleh pemodelan *Topic Modeling* menggunakan LDA terhadap ulasan pelayanan Bus Trans Jatim, serta bagaimana distribusi topik dan kata-kata dominan yang merepresentasikan setiap topik tersebut?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

Berdasarkan permasalahan, solusi, dan pertanyaan penelitian, maka tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui jumlah topik (k) yang menghasilkan nilai *Coherence Score* tertinggi pada pemodelan *Topic Modeling* menggunakan *Latent Dirichlet Allocation* (LDA) terhadap ulasan pelayanan Bus Trans Jatim, dilakukan pengujian dengan beberapa variasi jumlah topik.
2. Menganalisis pengaruh parameter alpha, beta, dan jumlah iterasi terhadap nilai *Coherence Score* guna menentukan konfigurasi terbaik dalam menghasilkan topik yang paling relevan dengan pembahasan utama mengenai topik harga, fasilitas, pelayanan, rute, dan keamanan.
3. Mengidentifikasi kata-kata dominan dalam setiap topik hasil pemodelan LDA untuk memahami konteks pembahasan yang paling sering muncul dalam ulasan pengguna, sehingga dapat ditemukan topik

pembahasan yang paling sering muncul dalam ulasan pengguna, sehingga dapat ditemukan topik pembahasan utama yang berkaitan dengan harga, fasilitas, pelayanan, rute, maupun keamanan pada layanan Bus Trans Jatim.

1.3.2 Manfaat

Adapun manfaat yang diperoleh dari penelitian ini adalah sebagai berikut :

1. Teoritis: Penelitian ini dapat memberikan kontribusi terhadap pengembangan ilmu pengetahuan di bidang *Natural Language Processing* (NLP), khususnya dalam penerapan metode Topic Modeling menggunakan *Latent Dirichlet Allocation* (LDA) untuk menganalisis opini publik pada sektor transportasi publik. Hasil penelitian ini juga dapat menjadi referensi bagi penelitian selanjutnya yang berfokus pada analisis teks dan evaluasi layanan publik berbasis data media sosial.
2. Praktis: Memberikan masukan bagi Dinas Perhubungan Jawa Timur dan pengelola Bus Trans Jatim dalam meningkatkan kualitas pelayanan berdasarkan topik utama yang menjadi perhatian pengguna, sehingga hasil penelitian ini dapat dijadikan sebagai acuan dalam pengambilan keputusan maupun evaluasi layanan.

1.4 Batasan Masalah

Batasan masalah yang dapat diterapkan agar penelitian ini tidak terlalu luas dan terarah, adalah sebagai berikut:

1. Data yang akan digunakan dari satu sumber media sosial tertentu (Youtube, Instagram, Tiktok), sehingga hasil analisis hanya merepresentasikan karakteristik pengguna pada platform tersebut.
2. Metode yang digunakan dalam penelitian ini adalah *Latent Dirichlet Allocation* (LDA) untuk pemodelan topik.
3. Evaluasi model dilakukan menggunakan *coherence score* dengan variasi parameter Iterasi, Alpha dan Beta sehingga mendapatkan hasil pemodelan topik terbaik.
4. Visualisasi hasil *topic modeling* menggunakan *wordcloud* dan *pyLDAvis*.

5. Bahasa ulasan tidak semuanya baku, melainkan campuran (bahasa Indonesia, Jawa, dan bahasa informal/slang), sehingga memerlukan tahap prapemrosesan yang lebih kompleks.

1.5 Sistematika Penulisan Proposal

Pada penelitian ini sistematika penulisan skripsi yang digunakan adalah sebagai berikut :

- I. **BAB I PENDAHULUAN**, yang terdiri dari latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan skripsi.
- II. **BAB II KAJIAN PUSTAKA**, yang terdiri dari teori-teori pendukung seperti *Topic Modelling*, *Text Mining*, *Natural Language Processing*, *Machine Learning*, *Preprocessing Data*, Pembobotan Kata TF IDF, *Latent Dirichlet Allocation* dan penelitian terkait.
- III. **BAB III METODELOGI PENELITIAN**, bab ini berisi penjelasan tentang data penelitian, diuraikan arsitektur sistem serta skenario uji coba.
- IV. **BAB IV HASIL DAN PEMBAHASAN**, bab ini menjelaskan hasil dari penerapan metode *Latent Dirichlet Allocation* (LDA) dalam menganalisis komentar terkait pelayanan Bus Trans Jatim, serta pembahasan atas temuan-temuan yang diperoleh. Analisis dilakukan mulai dari proses preprocessing data teks, pemodelan topik, evaluasi koherensi model, hingga visualisasi dan interpretasi topik-topik yang dihasilkan.
- V. **BAB V PENUTUP**, bab ini menjelaskan kesimpulan dari hasil penelitian yang telah dilakukan terkait analisis komentar pengguna terhadap pelayanan Bus Trans Jatim menggunakan metode *Latent Dirichlet Allocation* (LDA), serta saran untuk pengembangan dan penelitian lanjutan. Kesimpulan disusun berdasarkan hasil analisis topik yang diperoleh dari proses pemodelan dengan LDA. Selain itu, bab ini juga memberikan masukan terkait potensi perbaikan metode maupun penggunaan pendekatan lain yang dapat diterapkan dalam penelitian sejenis di masa mendatang.