

BAB I PENDAHULUAN

1.1 Latar Belakang

Trans Padang merupakan angkutan dalam Kota Padang berbasis *Bus Rapid Transit* (BRT) yang mulai beroperasi pada tahun 2014 dengan menyediakan layanan mobilitas yang cepat, nyaman, dan berbiaya rendah (Husna & Asri, 2022). Berbeda dengan angkutan kota lainnya, Trans Padang beroperasi sesuai koridor dan rute yang telah ditetapkan serta hanya dapat menaikkan dan menurunkan penumpang pada titik pemberhentian yang telah ditentukan. Terhitung pada tanggal 12 Desember 2023, Trans Padang telah memiliki enam koridor setelah Pemerintah Kota Padang resmi membuka jalur koridor 2 dan 3 (Jefrimon, 2023). Penambahan koridor yang berdampak pada penambahan jumlah pengguna menjadikan Trans Padang harus memberikan perhatian lebih pada kondisi layanan yang ada.

Berbagai kondisi layanan Trans Padang yang ada dapat menimbulkan respon dari pengguna Trans Padang. Apabila pengguna merasa terdapat kekurangan pada layanan Trans Padang, mereka cenderung akan mengajukan keluhan ke pihak Perumda Padang Sejahtera Mandiri, selaku pengelola Trans Padang. Sebaliknya, apabila pengguna merasa puas terhadap layanan Trans Padang, mereka akan cenderung memberikan pujian atau apresiasi. Respon yang pengguna berikan, baik keluhan ataupun pujian serta apresiasi seringkali disampaikan melalui media sosial.

Trans Padang memiliki beberapa sosial media resmi, yaitu Instagram, TikTok, Facebook, dan Threads. Dari empat sosial media tersebut, Instagram menjadi media sosial utama yang digunakan oleh Trans Padang untuk berinteraksi dengan penggunanya. Hal ini disebabkan oleh jumlah pengikut yang jauh lebih banyak serta pengelola Trans Padang yang secara reguler memberikan pembaharuan informasi melalui Instagram dibandingkan sosial media lainnya. Berdasarkan kondisi tersebut, pihak pengelola Trans Padang secara aktif menggunakan media sosial Instagram sebagai wadah untuk memberikan informasi seputar layanan Trans Padang, melakukan tanya jawab dengan pengguna, dan menerima respon atau opini pengguna terhadap layanan Trans Padang.

Namun, opini yang diberikan oleh pengguna di Instagram terkait dengan Trans Padang belum dikelola dengan baik. Hal ini dapat dilihat dari belum adanya proses pengumpulan dan ekstraksi data opini tersebut untuk memahami sentimen pengguna tentang layanan Trans Padang. Padahal, data-data opini tersebut sangat berguna bagi Perumda Padang Sejahtera Mandiri sebagai bahan evaluasi untuk memperbaiki dan atau meningkatkan pelayanan Trans Padang kedepannya.

Dalam memahami sentimen pengguna tentang layanan Trans Padang tersebut, dapat digunakan metode analisis sentimen. Analisis sentimen merupakan salah satu aplikasi dari *text mining* yang bertujuan untuk menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi masyarakat terhadap suatu produk, layanan, organisasi, individu, topik, dan atributnya (Liu, 2012). Komentar yang disampaikan pengguna Instagram terhadap layanan Trans Padang memiliki aspek layanan yang sangat beragam. Sehingga, diperlukan analisis sentimen berbasis aspek yang mampu memahami dan mengelompokkan sentimen pengguna sesuai dengan aspek yang disampaikan.

Berdasarkan penelitian yang dilakukan oleh (Putra, 2023 & Sakti, 2024) diketahui bahwa terdapat kesenjangan nilai antara harapan penumpang dengan kinerja pelayanan yang telah dirasakan pengguna ketika menggunakan layanan Trans Padang. Putra (2023) menyatakan nilai rata-rata kesenjangan antara harapan dan realisasi kinerja layanan Trans Padang berada pada angka -0.36 , yang berarti bahwa secara keseluruhan, kinerja Trans Padang sedikit di bawah dari harapan pengguna, sehingga perlu diperbaiki dan ditingkatkan. Sakti (2024) pada penelitiannya juga menyatakan nilai tersebut berada di angka $-0,19$, yang menandakan perlunya perhatian dan peningkatan terhadap layanan Trans Padang yang telah terlaksana. Dengan melihat angka kesenjangan antara harapan dan realisasi kinerja tersebut, pengelolaan data-data opini yang disampaikan pengguna di Instagram dapat digunakan oleh Perumda Padang Sejahtera Mandiri untuk mengetahui bagaimana sentimen masyarakat terhadap layanan Trans Padang serta aspek-aspek layanan mana saja yang perlu ditingkatkan dalam rangka memperkecil angka kesenjangan antara harapan dan realisasi kinerja Trans Padang ke depannya.

Proses analisis sentimen berbasis aspek pada layanan Trans Padang terdiri dari beberapa tahapan. Pertama, melakukan klasifikasi opini dan non opini menggunakan algoritma *Random Forest Classifier* (RFC) untuk memisahkan data yang bersifat informasi (non opini) dari data opini yang dapat mempengaruhi hasil analisis sentimen. Kedua, melakukan *topic modelling* untuk mengidentifikasi aspek layanan Trans Padang dan memahami aspek dalam tiap kelas sentimen yang ada menggunakan algoritma *Gibbs Sampling Dirichlet Multinomial Mixture* (GSDMM). Ketiga, melakukan klasifikasi sentimen menggunakan algoritma *Random Forest Classifier*. Pemilihan algoritma RFC dikarenakan dapat meminimalisir terjadinya *overfitting* sehingga dapat meningkatkan performa klasifikasi (Mahing dkk., 2023). Sementara itu, algoritma GSDMM dipilih karena memiliki asumsi bahwa setiap data teks hanya mencakup satu topik atau aspek saja, dimana cocok untuk data teks pendek seperti *tweets*, komentar pada sosial media, dan ulasan pengguna sebagaimana data yang digunakan pada penelitian ini (Abdelmotalieb dkk., 2023). Terakhir, hasil dari ketiga tahapan sebelumnya perlu diintegrasikan dalam sebuah sistem agar dapat digunakan dengan mudah oleh penggunanya. Dengan demikian, penelitian ini mengusulkan pembangunan sistem analisis sentimen berbasis aspek menggunakan algoritma *Random Forest Classifier* dan *Gibbs Sampling Dirichlet Multinomial Mixture* untuk mengatasi permasalahan yang telah diuraikan di atas.

Penelitian mengenai analisis sentimen berbasis aspek telah dilakukan sebelumnya. Ribeiro (2020) meneliti pengembangan alat yang mampu menganalisis sentimen berdasarkan ulasan pelanggan di media sosial, yang menggambarkan produk atau jasa dari perusahaan untuk membantu perusahaan-perusahaan dalam memahami sentimen pelanggan dengan lebih baik. Dalam pengembangan alat tersebut, penelitian ini mencoba menggunakan berbagai algoritma seperti *Logistic Regression*, *Naive Bayes*, *Decision Trees*, *Support Vector Machines*, *Random Forest*, dan *Multi-Layer Perceptron* untuk memprediksi sentimen dari komentar di sosial media serta algoritma *Latent Dirichlet Allocation* (LDA) dan *Gibbs Sampling Dirichlet Multinomial Mixture* (GSDMM) untuk membangun *topic modelling* agar perusahaan dapat mengetahui kapan dan dimana harus mengambil tindakan berdasarkan sentimen pengguna. Hasil penelitian

menunjukkan bahwa kinerja klasifikasi sentimen terbaik dicapai ketika menerapkan model MLP dengan nilai perhitungan *AUC ROC score* 74,84%. Kemudian, *topic modelling* dengan model GSDMM menghasilkan akurasi dan *coherence score* yang lebih unggul dibandingkan dengan LDA, yakni secara berurut di angka 45,89% dan -2,62.

Penelitian lain yang diteliti oleh Apriantoni dkk. (2021) dilakukan untuk mengetahui efek pandemi COVID-19 terhadap sentimen dan emosi masyarakat di tiga area berbeda di kota Surabaya, Indonesia. Penelitian ini melibatkan *topic modelling* dengan analisis sentimen untuk mengetahui informasi tersembunyi serta pola sentimennya dengan lebih baik menggunakan beberapa algoritma berikut: GSDMM, LDA, dan LDA Mallet. Pemodelan analisis sentimen dan deteksi emosi pada penelitian ini dilakukan dengan membandingkan enam algoritma klasifikasi, yaitu *Naïve Bayes*, *Linear Regression*, *Random Forest*, *Support Vector Machine*, *Multilayer Perceptron*, dan *AdaBoost*. *Topic modelling* menggunakan algoritma LDA Mallet menunjukkan nilai *coherence score* tertinggi jika dibandingkan dengan GSDMM dan LDA di masing-masing area yang diteliti, yakni 0.57 untuk area I (Institut Teknologi Sepuluh Nopember), 0.69 untuk area II (Tanjung Plaza Mall), dan 0.5 untuk area III (Rumah Sakit Angkatan Laut). Sementara itu, model deteksi emosi terbaik untuk area I adalah *AdaBoost* dengan nilai akurasi 0,75 dan *f1-score* 0,73, sedangkan untuk area II adalah *Linear Regression* dengan akurasi 0,67 dan *f1-score* 0,70, serta untuk area III nilai akurasi tertinggi yang diperoleh adalah 0,74 menggunakan metode *Linear Regression* dimana *f1-score* tertinggi di angka 0,69 menggunakan metode *AdaBoost*. Hasil penelitian ini juga menunjukkan bahwa selama masa pandemi COVID-19, emosi marah menjadi sentimen paling dominan terkait aktivitas ekonomi masyarakat di Surabaya.

Penelitian serupa juga dilakukan oleh Anuratha dkk. (2021) untuk mengungkapkan perasaan, keprihatinan, dan topik-topik yang dibicarakan masyarakat di India selama pandemic COVID-19 melalui proses analisis sentimen dan identifikasi topik. Penelitian ini menggunakan *dataset* berbahasa Tamil dan bahasa Inggris dari cuitan masyarakat India di Twitter untuk dapat mencapai tujuan tersebut. Proses identifikasi topik yang dibicarakan masyarakat India di Twitter menggunakan algoritma GSDMM pada masing-masing *dataset*. Proses analisis

sentimen pada *dataset* berbahasa Inggris menggunakan algoritma *CrystalFeel*, sedangkan analisis sentimen pada *dataset* berbahasa Tamil dicobakan pada enam algoritma klasifikasi berbeda, yaitu *AdaBoost*, *Naïve Bayes*, *Random Forest*, *Multi-Layer Perceptron*, *Linear SVM*, dan *Linear Regression*, dimana algoritma dengan kinerja terbaik adalah *Linear SVM*. Hasil penelitian ini mengungkapkan bahwa masing-masing *dataset* terdapat tujuh topik yang dibicarakan yang cenderung memiliki sentimen negatif dengan bentuk emosi marah, sedih, dan cemas.

Penelitian terdahulu mengenai analisis sentimen berbasis aspek pada layanan transportasi menggunakan algoritma RFC dan GSDMM belum ditemukan. Berbeda dengan penelitian terdahulu yang disampaikan di atas (Anuratha dkk., 2021; Apriantoni dkk., 2021; Ribeiro, 2020), penelitian menerapkan proses klasifikasi opini dan non opini agar hasil analisis sentimen berbasis aspek difokuskan pada data yang berupa opini saja. Dengan demikian, dilakukanlah penelitian tugas akhir yang berjudul “Pembangunan Sistem Analisis Sentimen Berbasis Aspek Pada Layanan Trans Padang Berdasarkan Komentar Pengguna Instagram menggunakan Algoritma *Random Forest Classifier* dan *Gibbs Sampling Dirichlet Multinomial Mixture*”. Melalui pembangunan sistem analisis sentimen berbasis aspek ini, diharapkan dapat membantu proses pengelolaan data opini pengguna layanan Trans Padang di Instagram sehingga membantu proses pengambilan keputusan oleh Perumda Padang Sejahtera Mandiri dalam meningkatkan pelayanan Trans Padang.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dipaparkan sebelumnya, dapat dirumuskan masalah pada penelitian ini yaitu bagaimana membangun sistem analisis sentimen berbasis aspek terhadap layanan Trans Padang berdasarkan komentar pengguna di Instagram dengan menggunakan algoritma *Random Forest Classifier* dan *Gibbs Sampling Dirichlet Multinomial Mixture* sehingga dapat menghasilkan informasi yang berharga bagi pihak Perumda Padang Sejahtera Mandiri, selaku pengelola Trans Padang, sebagai bahan pengambilan keputusan dalam rangka meningkatkan pelayanan Trans Padang.

1.3 Batasan Masalah

Adapun batasan masalah yang peneliti tetapkan dalam penelitian ini yaitu:

1. Data yang digunakan bersumber dari data komentar dari postingan akun Instagram resmi milik Trans Padang yaitu @official_transpadang.psm, data komentar dari postingan Instagram dengan *hashtag* (#) transpadang, data komentar dari postingan akun Instagram @infosumar mengenai Trans Padang, dan data komentar dari postingan akun Instagram @infopadang_ mengenai Trans Padang.
2. Data yang diambil adalah data komentar dari postingan yang di unggah di Instagram pada 1 Januari 2021 hingga 11 Januari 2024.
3. Penelitian ini tidak memperhatikan latar belakang atau demografi dari pemilik akun Instagram.
4. Hasil analisis sentimen diklasifikasikan dalam sentimen positif dan sentimen negatif.
5. Aspek-aspek pelayanan yang digunakan pada penelitian ini bersumber dari hasil *topic modelling* menggunakan algoritma *Gibbs Sampling Dirichlet Multinomial Mixture*.
6. Sistem analisis sentimen berbasis aspek pada layanan Trans Padang dibangun dengan mengikuti metode pembangunan sistem *waterfall* hingga tahap *integration* dan *system testing*.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk melakukan pembangunan sistem analisis sentimen berbasis aspek terhadap layanan Trans Padang menggunakan algoritma *Random Forest Classifier* dan *Gibbs Sampling Dirichlet Multinomial Mixture* sehingga sistem tersebut mampu mengklasifikasikan sentimen pengguna terkait layanan Trans Padang berdasarkan aspek-aspek pelayanan yang dikomentari di Instagram. Sistem yang dibangun akan menyajikan informasi tentang sentimen pengguna, sehingga membantu pihak pengelola Trans Padang, yakni Perumda Padang Sejahtera Mandiri, dalam proses pengambilan keputusan terkait peningkatan layanan Trans Padang di masa mendatang.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi pihak Padang Sejahtera Mandiri, selaku pengelola operasional Trans Padang, dalam memahami sentimen penggunaannya mengenai pelayanan transportasi Trans Padang berdasarkan aspek-aspek pelayanan yang sering dibicarakan pada media sosial Instagram. Selain itu, hasil analisis sentimen berbasis aspek yang disajikan pada sistem yang dibangun dapat berguna sebagai bahan evaluasi pelayanan Trans Padang oleh Padang Sejahtera Mandiri sehingga membantu proses pengambilan keputusan dalam rangka meningkatkan pelayanan transportasi Trans Padang

1.6 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini adalah sebagai berikut:

BAB I: PENDAHULUAN

Bab ini terdiri atas beberapa sub bab yang berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan laporan tugas akhir.

BAB II: TINJAUAN PUSTAKA

Bab ini menguraikan teori dan informasi pendukung yang relevan dengan masalah yang akan diselesaikan pada penelitian tugas akhir ini. Beberapa teori yang diuraikan yaitu tentang Trans Padang, Analisis Media Sosial Instagram, *Text Mining*, Sentimen Analisis, *Topic Modelling*, *FastText*, Algoritma *Random Forest Classifier*, Teknik Evaluasi Model Klasifikasi, Algoritma *Gibbs Sampling Dirichlet Multinomial Mixture*, *Dashboard*, Visualisasi, *SQLite*, dan *Framework Flask*.

BAB III: METODE PENELITIAN

Bab ini menjelaskan langkah-langkah yang akan ditempuh untuk melaksanakan penelitian. Berisikan informasi mengenai objek penelitian, metode pengumpulan data, dan diagram alir penelitian yang diuraikan secara rinci.

BAB IV: ANALISIS DAN RANCANGAN SISTEM

Bab ini menjelaskan rancangan model analisis sentimen berbasis aspek yang akan dilakukan pada penelitian ini. Berisikan uraian pembahasan setiap langkah penelitian sesuai diagram alir penelitian yang dimulai dari proses akuisisi data, membuat model klasifikasi data opini dan non opini menggunakan algoritma

Random Forest Classifier, *topic modelling* dengan algoritma *Gibbs Sampling* *Dirichlet Multinomial Mixture*, dan model analisis sentimen berbasis aspek menggunakan algoritma *Random Forest Classifier*. Diuraikan juga rancangan sistem analisis sentimen berbasis aspek yang akan dibangun.

BAB V: IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan proses implementasi model analisis sentimen berbasis aspek yang telah dikembangkan sebelumnya ke dalam sebuah sistem. Berisikan uraian kebutuhan perangkat keras dan lunak yang digunakan dalam implementasi sistem, implementasi antar muka sistem, dan pengujian sistem.

BAB VI: PENUTUP

Bab ini berisi tentang kesimpulan dari hasil penelitian yang telah dilaksanakan dan saran untuk pengembangan penelitian kedepannya.

