

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat telah mengubah cara masyarakat menyampaikan opini dan pengalaman terhadap layanan, produk, atau institusi (1). Platform digital seperti Google Maps tidak hanya berfungsi sebagai peta atau panduan lokasi, tetapi juga sebagai media bagi masyarakat untuk menuliskan ulasan mengenai pengalaman mereka, baik berupa kepuasan maupun kritik terhadap kualitas fasilitas dan layanan (2). Ulasan-ulasan tersebut menjadi sumber informasi yang kaya dan dapat memberikan gambaran nyata mengenai persepsi publik terhadap suatu institusi (3). Dengan memanfaatkan data ulasan ini, pihak institusi dapat melakukan evaluasi dan perbaikan secara terukur (4). Namun, secara umum, data ulasan yang tersedia di platform ini bersifat teks bebas dan jumlahnya sangat banyak, sehingga pengolahan secara manual menjadi tidak efisien, rawan subjektivitas, dan sulit untuk mengetahui kecenderungan opini publik secara tepat (5). Hal ini menimbulkan kesulitan bagi institusi, termasuk universitas, untuk menganalisis ulasan yang sering menggunakan bahasa sehari-hari, singkatan, atau bahasa campuran, sehingga analisis manual memerlukan waktu lama dan kurang akurat (6).

Kondisi ini khususnya berlaku untuk Universitas Muhammadiyah Gresik (UMG), di mana ulasan pengguna di Google Maps mencerminkan persepsi publik terhadap layanan dan fasilitas universitas secara umum, seperti kualitas pendidikan, infrastruktur, dan interaksi dengan mahasiswa atau masyarakat (7). Masalah yang dihadapi adalah kesulitan dalam mengolah data ulasan yang besar dan beragam, yang dapat mengakibatkan ketidakmampuan untuk mengetahui sentimen publik secara objektif, sehingga berdampak pada citra dan kualitas pelayanan universitas (8). Urgensi penelitian ini muncul karena tanpa analisis sentimen otomatis, UMG berisiko kehilangan kesempatan untuk evaluasi cepat dan perbaikan layanan, yang penting untuk meningkatkan daya saing dan kepercayaan publik dalam konteks pendidikan tinggi yang kompetitif (9).

Metode yang umum digunakan untuk mengatasi tantangan analisis sentimen pada data teks adalah algoritma *Multinomial Naïve Bayes* (MNB), yang efektif dalam klasifikasi sentimen berdasarkan frekuensi kata (1). Pada dasarnya, Naïve Bayes memiliki beberapa varian, di antaranya *Gaussian Naïve Bayes*, *Bernoulli Naïve Bayes*, dan *Multinomial Naïve Bayes*. *Naïve Bayes* standar seperti *Bernoulli* lebih cocok untuk data biner (misalnya ada/tidak ada kata), sedangkan *Gaussian* digunakan untuk data kontinyu. Sementara itu, *Multinomial Naïve Bayes* dirancang khusus untuk data teks karena mempertimbangkan frekuensi kemunculan kata pada dokumen, sehingga lebih akurat dalam memodelkan representasi TF-IDF maupun *Bag of Words*. Penelitian sebelumnya juga menunjukkan bahwa MNB memiliki performa lebih baik dibandingkan varian NB lainnya dalam tugas klasifikasi teks (10,11). Oleh karena itu, MNB dianggap paling relevan untuk digunakan pada penelitian analisis sentimen berbasis ulasan Google Maps yang memiliki karakteristik data teks bebas.

Penelitian sebelumnya telah membuktikan efektivitasnya, misalnya yang menggunakan NLTK *TextBlob* dan *Naïve Bayes* untuk menganalisis *review* lokasi di Google Maps (7); termasuk yang menerapkan *Naïve Bayes Classifier* untuk analisis ulasan pelanggan di Plus Coffee and Space (12); serta yang mengembangkan aplikasi web analisis sentimen dengan *Multinomial Naïve Bayes* (11); Di antaranya, ada yang menggunakan algoritma *Naïve Bayes* pada ulasan pantai (13); dan seperti halnya yang menerapkan *Multinomial Naïve Bayes* dengan seleksi fitur untuk analisis layanan Indihome (4). Hasil penelitian-penelitian ini menunjukkan bahwa MNB sangat efektif dalam mengolah teks berbasis frekuensi, mampu menghasilkan akurasi tinggi, dan dapat diterapkan dalam berbagai konteks analisis sentimen (14).

Berdasarkan kondisi yang ada dan hasil penelitian sebelumnya, penelitian ini menawarkan solusi berupa penerapan algoritma *Multinomial Naïve Bayes* untuk menganalisis sentimen ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik (13). Penelitian ini akan melalui beberapa tahapan, mulai dari data *cleaning*, normalisasi kata, tokenisasi, penghapusan *stopwords*, *stemming*, pelabelan sentimen, pelatihan model, hingga pengujian sistem menggunakan input

teks secara manual (4). Dengan metode ini, diharapkan pihak universitas dapat memperoleh informasi sentimen publik secara cepat dan akurat, sehingga dapat menjadi dasar bagi pengambilan keputusan, evaluasi layanan, dan pengembangan strategi komunikasi yang lebih efektif untuk meningkatkan kualitas pelayanan dan citra institusi secara menyeluruh (3).

1.2 Rumusan Masalah

Permasalahan yang diangkat dalam penelitian ini adalah bagaimana menganalisis sentimen ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik secara otomatis dan akurat, khususnya terkait fasilitas kampus, kualitas pendidikan, dan interaksi dengan mahasiswa, guna mengidentifikasi kecenderungan opini publik terhadap kualitas layanan dan citra universitas tersebut. Secara spesifik, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana penerapan algoritma *Multinomial Naïve Bayes* dalam melakukan klasifikasi sentimen pada data ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik?
2. Bagaimana tingkat akurasi model yang dihasilkan berdasarkan hasil evaluasi menggunakan *confusion matrix*?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan fokus pada tujuan yang ingin dicapai, maka batasan masalah dalam penelitian ini ditetapkan sebagai berikut:

1. Data yang digunakan merupakan ulasan pengguna dari Google Maps yang berkaitan dengan Universitas Muhammadiyah Gresik.
2. Algoritma yang digunakan untuk klasifikasi adalah *Multinomial Naïve Bayes*, tanpa membandingkannya dengan algoritma lain.
3. Kategori sentimen yang digunakan hanya terdiri dari positif, negatif, dan netral.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menganalisis sentimen ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik dengan menggunakan algoritma *Multinomial Naïve Bayes*, khususnya terkait fasilitas kampus, kualitas pendidikan, dan interaksi dengan mahasiswa, sehingga dapat diketahui kecenderungan opini publik terhadap universitas tersebut dan mendukung

peningkatan citra serta pelayanan UMG. Secara spesifik, tujuan penelitian ini adalah:

1. Menerapkan algoritma *Multinomial Naïve Bayes* untuk melakukan klasifikasi sentimen pada ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik.
2. Mengevaluasi tingkat akurasi model yang dihasilkan melalui *confusion matrix*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi beberapa pihak, antara lain:

1. Bagi Peneliti

Penelitian ini memberikan pengalaman dan pemahaman langsung kepada peneliti dalam menerapkan metode *Multinomial Naïve Bayes* pada analisis sentimen berbasis teks. Selain itu, penelitian ini juga memperdalam kemampuan dalam melakukan pengolahan data teks (*text preprocessing*) menggunakan *Python* di platform Google Colab, mulai dari *cleaning data*, *tokenizing*, *stopword removal*, *stemming*, hingga proses evaluasi model menggunakan *confusion matrix*.

2. Bagi Universitas Muhammadiyah Gresik

Hasil penelitian ini diharapkan dapat memberikan informasi mengenai persepsi publik terhadap Universitas Muhammadiyah Gresik berdasarkan ulasan pengguna di Google Maps. Informasi tersebut dapat menjadi bahan evaluasi bagi pihak universitas untuk meningkatkan kualitas pelayanan, fasilitas, serta citra kampus di mata masyarakat.

3. Bagi Pembaca dan Peneliti Selanjutnya

Penelitian ini dapat dijadikan referensi bagi mahasiswa atau peneliti lain yang ingin melakukan penelitian serupa di bidang *text mining* dan *analisis sentimen*. Selain itu, hasil penelitian ini juga dapat menjadi acuan untuk pengembangan sistem analisis sentimen yang lebih kompleks dengan penerapan algoritma lain atau perluasan sumber data.

1.6 Metodologi Penelitian

Metodologi penelitian ini menggunakan model *Waterfall* sebagai pendekatan utama untuk menganalisis sentimen ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik. Model *Waterfall* adalah pendekatan linier yang terdiri dari fase-fase berurutan: pengumpulan kebutuhan (*requirements gathering*), perancangan sistem (*system design*), implementasi (*implementation*), pengujian (*testing*), dan penyebaran/pemeliharaan (*deployment/maintenance*). Setiap fase harus diselesaikan sebelum melanjutkan ke fase berikutnya, yang memastikan proses yang terstruktur dan dapat dikontrol. Adapun tahapan penelitian ini sesuai dengan fase-fase model *Waterfall* adalah sebagai berikut:

1. Pengumpulan Kebutuhan (*Requirements Gathering*)

Pada fase ini, dilakukan pengumpulan data dan studi literatur. Pengumpulan data dilakukan dengan mengambil ulasan pengguna Universitas Muhammadiyah Gresik di Google Maps secara semi-otomatis menggunakan metode *web scraping* dengan bahasa pemrograman *Python*. Data yang diperoleh meliputi komentar pengguna dan rating bintang, kemudian disimpan dalam format CSV untuk keperluan pengolahan lebih lanjut. Selain itu, dilakukan pengumpulan dan kajian literatur terkait analisis sentimen, algoritma *Multinomial Naïve Bayes*, *text preprocessing*, serta hasil penelitian terdahulu yang relevan. Literatur ini menjadi landasan teori dan pedoman dalam merancang penelitian.

2. Perancangan Sistem (*System Design*)

Fase ini mencakup perancangan data dan analisis sistem. Tahap perancangan dilakukan untuk mengolah data mentah menjadi data berkualitas agar dapat dianalisis secara akurat. Proses ini meliputi data cleaning untuk menghapus karakter atau elemen tidak relevan, normalisasi, tokenisasi, penghapusan *stopwords*, dan *stemming* menggunakan *library Sastrawi*. Berdasarkan hasil pengolahan data, dilakukan analisis kebutuhan sistem untuk menentukan fitur yang digunakan dalam penerapan algoritma *Multinomial Naïve Bayes*. Analisis ini juga mencakup penentuan kategori sentimen (positif, negatif, netral) serta metode evaluasi model menggunakan *confusion matrix*.

3. Implementasi (*Implementation*)

Pada fase ini, dilakukan penerapan algoritma *Multinomial Naïve Bayes* pada data ulasan yang telah diproses. Model dilatih menggunakan Google Colab dengan *library Python*, kemudian diuji untuk menghasilkan prediksi sentimen berdasarkan ulasan baru.

4. Pengujian (*Testing*)

Fase pengujian dilakukan dengan memasukkan teks ulasan secara manual ke dalam sistem. Sistem akan memprediksi sentimen ulasan tersebut, misalnya:

Input teks: “Kampusnya bersih dan nyaman.”

Prediksi sentimen: Positif.

Tahap ini bertujuan untuk mengevaluasi kinerja sistem dalam memberikan prediksi yang sesuai dengan konteks ulasan pengguna.

5. Penyebaran/Pemeliharaan (*Deployment/Maintenance*)

Fase terakhir adalah penyusunan laporan penelitian yang mencakup seluruh tahapan, hasil analisis, pembahasan, serta kesimpulan dan saran sesuai dengan sistematika penulisan skripsi. Laporan ini juga berfungsi sebagai dokumentasi akhir untuk evaluasi dan potensi pengembangan lebih lanjut.

Dengan pendekatan *Waterfall* ini, penelitian dapat dilakukan secara sistematis, meminimalkan risiko kesalahan, dan memastikan bahwa setiap fase didokumentasikan dengan baik untuk keperluan replikasi atau evaluasi.

1.7 Sistematika Penulisan

Sistematika penulisan ini bertujuan untuk memudahkan pembaca dalam memahami materi skripsi secara komprehensif dan memenuhi persyaratan sebagai karya ilmiah. Adapun sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I: PENDAHULUAN

Membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian (berdasarkan model *Waterfall*), dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Membahas dasar teoritis yang menjadi landasan dalam pelaksanaan penelitian, termasuk konsep analisis sentimen pada ulasan Google Maps, pengolahan bahasa alami (*Natural Language Processing*), algoritma *Multinomial Naïve Bayes* , serta penelitian terdahulu yang relevan dengan analisis sentimen berbasis teks.

BAB III: ANALISIS DAN PERANCANGAN SISTEM

Menjelaskan perancangan sistem secara rinci beserta metode yang digunakan dalam penelitian, termasuk pengolahan data (data cleaning, tokenisasi, stemming), desain model klasifikasi *Multinomial Naïve Bayes* sebagai inti sistem, serta evaluasi kinerja model menggunakan *confusion matrix*.

BAB IV: IMPLEMENTASI DAN HASIL PENGUJIAN

Menjelaskan hasil implementasi sistem analisis sentimen yang dibangun (seperti aplikasi atau model klasifikasi berbasis *Multinomial Naïve Bayes*), cara kerja program yang dikembangkan (misalnya, menggunakan *Python* dan Google Colab), serta hasil pengujian model klasifikasi sentimen terhadap data ulasan pengguna di Google Maps mengenai Universitas Muhammadiyah Gresik, termasuk demo pengujian dan akurasi sistem.

BAB V: KESIMPULAN DAN SARAN

Memuat kesimpulan dari penelitian serta saran-saran untuk perbaikan dan pengembangan lebih lanjut, termasuk potensi penerapan algoritma lain atau perluasan data pada sistem yang dibangun.

DAFTAR PUSTAKA

Berisi semua referensi yang digunakan dalam penelitian, termasuk buku, jurnal, artikel, dan sumber lainnya, disusun sesuai gaya sitasi Vancouver.