

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Dalam era transparansi informasi, opini ataupun persepsi dari masyarakat sangatlah penting bagi lembaga tertentu, karena menjadi kompas tak terlihat yang memberikan arah legitimasi dan juga patokan suatu lembaga dalam memberikan layanan kepada publik. Salah satu cara agar pengguna layanan dapat memberikan masukan secara langsung adalah melalui survei yang mengumpulkan tanggapan secara lisan. Adanya teknologi *speech-to-text* kini mempermudah transkripsi data tanggapan lisan menjadi teks, yang membuka peluang untuk memanfaatkan data ini dalam analisis sentimen secara otomatis yang lebih komprehensif. Analisis sentimen atau opinion mining memungkinkan pemahaman terhadap opini, emosi, dan sikap pengguna layanan, yang sangat bermanfaat bagi pengambilan keputusan yang lebih responsive [1]. Sebagai contoh, hasil survei acak terhadap 57 mahasiswa UIN Sunan Gunung Djati Bandung hanya 29 responden atau mungkin sekitar 50,8% mahasiswa yang cukup puas terhadap layanan pendaftaran mahasiswa baru. Temuan ini menunjukkan bahwa nilai kepuasan yang diperoleh masih berada di bawah patokan (*benchmark*) kepuasan layanan pendidikan, yaitu kurang dari 64%, sehingga mengindikasikan perlunya perbaikan pada mekanisme pelayanan pendaftaran. Kondisi ini sekaligus menunjukkan adanya potensi penerapan teknologi *speech-to-text* dan analisis sentimen untuk mengidentifikasi serta memahami opini dan persepsi pengguna layanan secara lebih efisien dan sistematis [2].

Dalam konteks pendaftaran mahasiswa baru, survei layanan yang diberikan oleh institusi pendidikan, seperti kemudahan akses informasi, kelancaran proses pendaftaran online, dan responsivitas petugas pendaftaran, memberikan persepsi atau tanggapan langsung dari mahasiswa sebagai *end-user*. Akan tetapi, data teks yang dihasilkan dari transkripsi suara membutuhkan proses analisis yang cermat agar dapat diklasifikasikan sesuai dengan sentimen yang terkandung di dalamnya, baik positif, negatif, maupun sentimen netral. Penelitian ini bertujuan untuk

mengembangkan model klasifikasi sentimen menggunakan pendekatan metode *Support Vector Machine (SVM)* yang telah terbukti efektif dalam klasifikasi teks. Berdasarkan studi terdahulu, algoritma ini menunjukkan performa yang baik dalam klasifikasi teks, baik pada data media sosial maupun ulasan produk [3].

Untuk memastikan penelitian berjalan sistematis dan terstruktur, pendekatan Cross-Industry Standard Process for Data Mining (*CRISP-DM*) digunakan sebagai kerangka kerja utama. Metode *CRISP-DM* digunakan karena metodologi pengembangan ini terstruktur yang efektif untuk pengembangan proyek *data mining* dan juga *machine learning* [4]. Metode *CRISP-DM* menyediakan enam tahapan yang mencakup pemahaman bisnis (*business understanding*), pemahaman data (*data understanding*), persiapan data (*data preparation*), pemodelan (*modeling*), evaluasi (*evaluation*), dan implementasi (*deployment*) [5]. Tahapan ini memungkinkan penelitian dilakukan dengan alur yang jelas dari awal hingga akhir.

Dalam penelitian ini, data akan diambil dari tanggapan survei suara mahasiswa yang kemudian ditranskripsikan menjadi teks melalui teknologi *Automatic Speech Recognition (ASR)*. Pendekatan yang digunakan akan merujuk pada penelitian terdahulu, seperti penggunaan model *whisper* untuk bahasa Indonesia yang dapat mengoptimalkan akurasi transkripsi [6]. Dengan data hasil transkripsi tersebut, algoritma *Support Vector Machine* akan digunakan untuk mengklasifikasikan sentimen pada tiap tanggapan. Analisis sentimen juga bisa digunakan untuk menunjukkan opini positif (menguntungkan) atau negatif (tidak menguntungkan) terhadap subjek [7]. Melalui penelitian ini, diharapkan dapat diperoleh wawasan yang mendalam mengenai persepsi mahasiswa terhadap layanan pendaftaran mahasiswa baru di institusi Pendidikan.

Berdasarkan pemaparan latar belakang diatas, telah dijelaskan bahwa penelitian ini bertujuan untuk mengembangkan model klasifikasi sentimen pada data survei suara mahasiswa yang ditranskripsikan menggunakan teknologi *speech-to-text*. Dengan menggunakan algoritma *Support Vector Machine* dan metode *CRISP-DM* sebagai kerangka kerja, diharapkan penelitian ini dapat memberikan wawasan yang lebih dalam tentang persepsi mahasiswa terhadap berbagai aspek layanan pendaftaran mahasiswa baru. Oleh karena itu, judul yang relevan untuk penelitian ini adalah "Analisis Sentimen Pendaftaran Mahasiswa Baru Berbasis

Speech-to-text dengan Menggunakan Metode *Support Vector Machine*", yang mencerminkan fokus penelitian pada analisis sentimen berbasis transkripsi suara untuk meningkatkan kualitas layanan pendaftaran mahasiswa baru.

1.2 Rumusan Masalah

Berdasarkan dengan deskripsi latar belakang yang telah disebutkan sebelumnya, maka ada permasalahan yang bisa dirumuskan dalam penelitian ini yaitu sebagai berikut:

1. Bagaimana penerapan algoritma *Support Vector Machine* dalam mengklasifikasikan sentimen pada data survei suara mahasiswa yang telah ditranskripsikan menggunakan teknologi *speech-to-text*?
2. Bagaimana performa algoritma *Support Vector Machine* dalam mengklasifikasikan sentimen (positif, negatif, atau netral) dari hasil transkripsi survei suara mahasiswa tentang pendaftaran mahasiswa baru?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah melakukan pengembangan dan pengujian sistem analisis sentimen otomatis pada survei layanan pendaftaran mahasiswa baru berbasis teknologi *speech-to-text* dan algoritma *Support Vector Machine*, yaitu:

1. Menerapkan algoritma *Support Vector Machine* dalam mengklasifikasikan sentimen pada data survei suara mahasiswa yang telah ditranskripsikan menggunakan teknologi *speech-to-text*.
2. Mengetahui performa (*precision*, *recall*, *f1-score* dan *accuracy*) dari algoritma *Support Vector Machine* dalam mengklasifikasikan sentimen (positif, negatif dan netral) dari data hasil transkripsi survei suara mahasiswa tentang pendaftaran mahasiswa baru.

1.4 Batasan Masalah

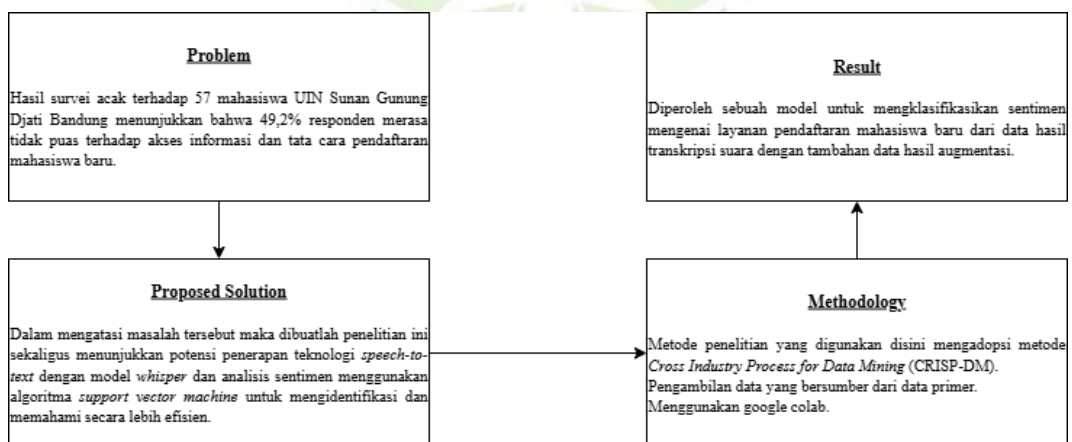
Dalam upaya mencapai tujuan penelitian yang diharapkan, penelitian ini memiliki batasan-batasan tertentu. Berikut adalah beberapa batasan yang relevan yang menjadi fokus dalam penelitian ini:

1. Penelitian ini terbatas pada analisis sentimen terhadap tanggapan mahasiswa mengenai pendaftaran mahasiswa baru di UIN Sunan Gunung Djati Bandung,

seperti pada kemudahan akses informasi, kelancaran proses pendaftaran online, dan responsivitas petugas pendaftaran.

2. Dataset dikumpulkan dari 47 mahasiswa yang masuk UIN Sunan Gunung Djati Bandung Jurusan Teknik Informatika angkatan 2024 menggunakan Jalur Mandiri. Adapun untuk jumlah 47 mahasiswa yang masuk UIN Sunan Gunung Djati Bandung Jurusan Teknik Informatika angkatan 2024 didapat dari Bagian Penerimaan Mahasiswa Baru UIN Sunan Gunung Djati Bandung.
3. Data yang digunakan dalam penelitian ini berasal dari survei suara mahasiswa dengan format opus yang akan ditranskripsi menjadi teks menggunakan teknologi *Automatic Speech Recognition* (ASR), dengan fokus pada penggunaan model *whisper*.
4. Penelitian ini terbatas pada analisis sentimen yang dibagi menjadi tiga kategori: positif, negatif, dan netral.

1.5 Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran

Dalam tugas akhir ini, pertama-tama, dimulai dengan mengidentifikasi sebuah temuan mengenai survei acak yang dilakukan kepada 57 mahasiswa UIN Sunan Gunung Djati Bandung hanya 29 responden atau mungkin sekitar 50,8% mahasiswa yang cukup puas terhadap layanan pendaftaran mahasiswa sedangkan sisa responden lainnya yaitu 15 orang atau 26,3% responden merasa bahwa akses informasi mengenai pendaftaran sulit didapatkan mencakup informasi mengenai jalur masuk yang ada hingga kebutuhan berkas yang harus dilengkapi baik itu saat melakukan pendaftaran maupun pendaftaran ulang. Adapun untuk sisa 13

responden menyebutkan mendapatkan kendala juga namun bisa mengatasi hal tersebut dengan cara *browsing* ataupun mencari ke akun Instagram yang memberikan info tentang kampus seperti akun @infouinsgd. Untuk mengetahui mayoritas pendapat atau tanggapan terhadap layanan tersebut maka dilakukan penelitian ini dengan subjek penelitian mahasiswa Teknik Informatika angkatan 2024 yang masuk menggunakan jalur mandiri, baik itu jalur mandiri tes (CBT), prestasi (seni, olahraga, dll), tahfidz Al-Qur'an dan khusus lulusan madrasah aliyah unggulan. Alasan peneliti memilih subjek penelitian mahasiswa Teknik Informatika angkatan 2024 yang masuk menggunakan jalur mandiri dikarenakan dua hal, yaitu peneliti sendiri merupakan mahasiswa Teknik Informatika juga, alasan kedua ialah dikarenakan angkatan 2024 yang masuk melalui jalur mandiri adalah angkatan mahasiswa paling baru dan juga jalur mandiri adalah jalur dimana para mahasiswa benar-benar melakukan pendaftaran dari awal menggunakan layanan pendaftaran mahasiswa baru dari kampus tanpa adanya pihak ketiga. Sebagai solusi untuk mengatasi permasalahan tersebut maka dibuatlah penelitian ini sekaligus menunjukkan potensi dari penerapan teknologi *speech-to-text* dengan model whisper dan penganalisisan sentimen menggunakan algoritma *support vector machine* untuk mengidentifikasi dan memahami pendapat dari mahasiswa secara lebih efisien. Metodologi yang digunakan adalah CRISP-DM, dengan pengambilan data yang bersumber dari data primer dan pengembangan menggunakan Google Colab. Hasil akhirnya adalah sebuah model untuk mengklasifikasikan sentimen mengenai layanan pendaftaran mahasiswa baru dari data hasil transkripsi suara dengan tambahan data hasil augmentasi.

1.6 Sistematika Penulisan

Sistematika penulisan laporan memuat sistematika penulisan laporan tugas akhir dengan memberikan gambaran kandungan setiap bab, urutan penulisan, serta keterkaitan antara satu bab dengan bab lainnya dalam sebuah laporan tugas akhir. Berikut sistematika penulisan laporan tugas akhir.

BAB I PENDAHULUAN

Bab ini terdiri dari latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, kerangka pemikiran penelitian, dan sistematika penulisan.

BAB II KAJIAN LITERATUR

Pada bab ini membahas terkait literatur atau penelitian terdahulu, konsep konsep, teori-teori, model, dan rumus yang menjadi landasan dalam proses analisis permasalahan dengan topik masalah yang diambil.

BAB III METODOLOGI PENELITIAN

Metodologi penelitian berisi penjelasan langkah-langkah dan teknik yang diterapkan dalam penelitian, diuraikan secara sistematis dan terstruktur.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan dua hal utama, yang pertama adalah pemaparan tentang temuan atau hasil penelitian berdasarkan langkah-langkah penelitian yang telah dilakukan. Selanjutnya adalah pembahasan hasil atau temuan penelitian sebagai jawaban terhadap rumusan masalah penelitian.

BAB V SIMPULAN DAN SARAN

Bab ini berfokus pada penarikan kesimpulan dari hasil penelitian yang diperoleh serta menjawab pertanyaan penelitian atau rumusan masalah. Selain itu, bab ini juga memberikan saran untuk penelitian selanjutnya yang dapat dilakukan agar meningkatkan kualitas dari penelitian tersebut

