

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital saat ini, perkembangan teknologi informasi membuat produksi data meningkat sangat tajam di berbagai bidang [1]. Menurut laporan World Bank, volume data secara global telah mencapai sekitar 120 zettabyte dan diproyeksikan akan terus melonjak hingga melampaui 180 zettabyte pada tahun 2025 [2]. Di dunia pendidikan sendiri, sistem juga belajar semakin canggih dengan maraknya penggunaan teknologi dan platform pembelajaran online [3], [4].

Dalam ekosistem pendidikan tinggi, umpan balik (feedback) dari mahasiswa telah diakui sebagai instrumen utama dan krusial untuk mengevaluasi serta meningkatkan kualitas pengajaran dosen secara berkelanjutan [5]. Berdasarkan penelitian *Assessing Teacher Competencies in Higher Education*, komentar mahasiswa ini sangat berguna untuk menilai bagaimana metode pengajaran dan interaksi dosen di dalam kelas, pihak kampus juga bisa memahami apa yang dirasakan oleh mahasiswa dengan lebih jelas [6].

Namun masalahnya, Data komentar mahasiswa ini biasanya berbentuk kalimat teks yang tidak terstruktur [6]. Karena jumlahnya yang banyak, sangat sulit dan memakan waktu jika harus dibaca dan dianalisis satu per satu secara manual [7]. Seiring dengan masifnya akumulasi data digital yang diproduksi oleh ekosistem pendidikan modern, pendekatan komputasi otomatis seperti machine learning menjadi sangat krusial. Teknologi ini dibutuhkan untuk memproses, menganalisis, dan mengekstraksi informasi strategis dari tumpukan data tersebut secara efektif dan efisien [8].

Untuk memproses volume teks yang tidak terstruktur secara otomatis, pemanfaatan teknologi *Natural Language Processing* (NLP) sangat diperlukan. Di dalam payung NLP, teknik analisis sentimen menjadi pendekatan yang paling ideal untuk mengekstraksi dan mengklasifikasikan polaritas ulasan mahasiswa ke dalam kategori positif maupun negatif [9].

Di antara berbagai algoritma machine learning yang tersedia untuk klasifikasi teks, Naïve Bayes sering kali menjadi pilihan utama. Meskipun memiliki landasan komputasi yang relatif sederhana dan efisien untuk diterapkan, algoritma ini terbukti sangat tangguh dalam mempertahankan tingkat akurasi yang tinggi saat memproses dan mengklasifikasikan kumpulan data opini berskala besar [10].

Untuk menghasilkan wawasan yang lebih komprehensif, klasifikasi tidak cukup hanya sebatas pelabelan positif atau negatif secara biner. Melalui pendekatan sentiment scoring, setiap ulasan teks akan diberikan bobot nilai numerik untuk mengukur intensitas atau kekuatan sentimen yang terkandung di dalamnya [11]. Dan untuk menguji metode tersebut, penelitian ini menggunakan dataset ulasan mahasiswa publik berbahasa Inggris bernama EduRABSA, yang berisi evaluasi terhadap proses pembelajaran dan kinerja dosen [12], [13].

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang dan identifikasi celah penelitian, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana kinerja algoritma Naïve Bayes yang dipadukan dengan sentiment scoring dalam menganalisis sentimen umpan balik mahasiswa terhadap dosen pada dataset teks ulasan mahasiswa publik berbahasa Inggris?

1.3 Batasan Masalah

Untuk menjaga fokus penelitian dan memastikan hasil yang diperoleh sesuai dengan tujuan yang telah ditetapkan, maka batasan masalah dalam penelitian ini adalah sebagai berikut:

- a. Data yang digunakan dalam penelitian ini terbatas pada dataset teks ulasan mahasiswa publik berbahasa Inggris yang digunakan sebagai *benchmark* untuk pengujian metode analisis sentimen.
- b. Penelitian ini hanya menerapkan algoritma Naïve Bayes dengan ekstraksi fitur TF-IDF untuk klasifikasi sentimen positif dan negatif, serta mengintegrasikan sentiment scoring sebagai pendekatan pendukung untuk mengukur intensitas sentimen.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disusun, tujuan dari penelitian ini adalah sebagai berikut:

- a. Menganalisis kinerja algoritma Naïve Bayes dalam mengklasifikasikan sentimen positif dan negatif pada dataset teks ulasan mahasiswa publik menggunakan ekstraksi fitur TF-IDF.
- b. Mengevaluasi integrasi sentiment scoring dalam memberikan nilai intensitas sentimen untuk memperkuat interpretasi hasil klasifikasi sentimen.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam bidang analisis sentimen melalui penerapan metode Naïve Bayes yang dipadukan dengan sentiment scoring pada evaluasi pendidikan tinggi, dengan memperkaya kajian pengolahan umpan balik mahasiswa berbasis teks sekaligus menghasilkan sistem evaluasi kinerja dosen yang lebih objektif, informatif, dan berbasis data sebagai bahan pertimbangan pengambilan keputusan di institusi pendidikan.