

BAB I PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan teknologi informasi telah membawa perubahan signifikan di berbagai sektor industri, termasuk bidang perpajakan. Di Indonesia, Direktorat Jenderal Pajak (DJP) terus berupaya melakukan inovasi dan digitalisasi sistem guna meningkatkan kualitas pelayanan kepada wajib pajak. Salah satu langkah nyata adalah pengembangan aplikasi CoreTax, yang bertujuan memudahkan proses pelaporan dan pembayaran pajak secara *online* tanpa harus datang langsung ke kantor pajak. Melalui penerapan sistem ini, diharapkan tercipta proses administrasi perpajakan yang lebih efisien, transparan, dan nyaman, sekaligus meningkatkan tingkat kepatuhan pajak di kalangan masyarakat [1].

Namun, meskipun CoreTax diharapkan menjadi solusi modern bagi pelayanan pajak, beragam tanggapan muncul dari para pengguna. Sebagian pengguna merasa terbantu dengan kemudahan akses dan efisiensi waktu yang ditawarkan, sedangkan sebagian lainnya mengalami kendala dalam penggunaan aplikasi, seperti antarmuka yang kurang intuitif, gangguan sistem, atau kesulitan teknis saat pelaporan. Kondisi ini menunjukkan bahwa meskipun digitalisasi telah membawa kemajuan, pengalaman pengguna masih beragam dan memerlukan evaluasi lebih lanjut [2].

Untuk memahami sejauh mana penerapan sistem ini diterima oleh masyarakat, analisis sentimen menjadi metode yang relevan untuk digunakan. Analisis sentimen adalah pendekatan dalam *text mining* dan *Natural Language Processing (NLP)* yang digunakan untuk mengidentifikasi opini publik terhadap suatu objek atau layanan. Melalui teknik ini, komentar atau ulasan pengguna di media sosial dan platform digital dapat diklasifikasikan menjadi sentimen positif, negatif, atau netral. Hasil analisis tersebut dapat memberikan gambaran objektif mengenai kepuasan dan persepsi masyarakat, serta membantu pengambil kebijakan dalam merancang perbaikan layanan publik yang lebih efektif [3].

Dalam ranah pembelajaran mesin (*machine learning*), salah satu algoritma yang banyak digunakan untuk klasifikasi teks adalah *Naive Bayes*. Algoritma ini merupakan metode berbasis probabilistik yang bekerja dengan menghitung kemungkinan suatu data

masuk ke dalam kelas tertentu berdasarkan fitur-fiturnya. Meskipun sederhana, *Naive Bayes* dikenal memiliki kecepatan komputasi tinggi, efisiensi, dan performa akurasi yang baik dalam pengolahan data teks, termasuk pada kasus analisis sentimen [4].

Berdasarkan latar belakang tersebut, penelitian ini berfokus pada penerapan algoritma *Naive Bayes* untuk menganalisis sentimen pengguna terhadap aplikasi CoreTax. Melalui pendekatan ini, diharapkan dapat diketahui bagaimana persepsi publik terhadap layanan digital perpajakan, serta sejauh mana efektivitas algoritma *Naive Bayes* dalam melakukan klasifikasi sentimen terhadap data teks yang bersumber dari platform youtube.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, digitalisasi sistem perpajakan melalui aplikasi CoreTax oleh Direktorat Jenderal Pajak (DJP) menimbulkan beragam persepsi dari pengguna terkait kemudahan dan keandalannya. Oleh karena itu, analisis sentimen diperuntukan untuk mengetahui pandangan masyarakat terhadap aplikasi tersebut menggunakan algoritma *Naive Bayes*. Dari permasalahan tersebut, maka rumusan masalah dalam penelitian ini dapat dijabarkan sebagai berikut :

1. Bagaimana sentimen masyarakat terhadap aplikasi CoreTax berdasarkan ulasan yang diperoleh dari platform digital?
2. Bagaimana penerapan dan kinerja algoritma *Naive Bayes* dalam mengklasifikasikan sentimen pengguna aplikasi CoreTax?

1.3 Batasan Masalah

Untuk menjaga fokus penelitian ini, beberapa batasan masalah ditetapkan sebagai berikut:

1. Penelitian ini dibatasi pada analisis sentimen terhadap ulasan atau komentar pengguna mengenai aplikasi CoreTax dengan menggunakan algoritma *Naive Bayes*, tanpa membahas pengembangan atau evaluasi fitur aplikasi CoreTax itu sendiri.
2. Klasifikasi sentimen dalam penelitian ini hanya mencakup dua kategori, yaitu sentimen positif dan sentimen negatif, sehingga sentimen netral tidak dianalisis.
3. Sumber data penelitian dibatasi pada ulasan atau komentar pengguna yang terdapat

pada platform YouTube yang berkaitan dengan aplikasi CoreTax.

4. Proses pengolahan data, pelatihan model, dan evaluasi hasil klasifikasi sentimen dilakukan menggunakan bahasa pemrograman Python dengan bantuan lingkungan Google Colab.

1.4 Tujuan Penelitian

Penelitian ini dilakukan untuk menganalisis dan mengidentifikasi sentimen masyarakat terhadap aplikasi CoreTax dengan memanfaatkan algoritma *Naive Bayes* sebagai metode klasifikasi teks. Adapun tujuan penelitian ini adalah sebagai berikut:

1. Menganalisis sentimen masyarakat terhadap aplikasi CoreTax berdasarkan ulasan pengguna di platform digital.
2. Mengklasifikasikan opini pengguna aplikasi CoreTax ke dalam kategori sentimen positif dan negatif menggunakan algoritma *Naive Bayes*.
3. Mengevaluasi kinerja algoritma *Naive Bayes* dalam analisis sentimen berdasarkan metrik *accuracy*, *precision*, *recall*, dan *f1-score*.
4. Memberikan gambaran objektif mengenai persepsi masyarakat terhadap aplikasi CoreTax sebagai layanan digital perpajakan.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis sebagai berikut:

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi terhadap pengembangan ilmu di bidang *Natural Language Processing (NLP)* dan *machine learning*, khususnya dalam penerapan algoritma *Naive Bayes* untuk analisis sentimen berbasis teks. Dengan menggunakan metode klasifikasi probabilistik, penelitian ini dapat memperkaya pemahaman mengenai cara kerja *Naive Bayes* dalam mengolah data teks dan mengidentifikasi pola opini publik terhadap suatu layanan digital. Selain itu, hasil penelitian ini dapat menjadi referensi bagi peneliti selanjutnya yang ingin mengkaji penerapan algoritma *machine learning* lainnya pada analisis sentimen atau mengembangkan model yang lebih kompleks untuk kasus

serupa

2. Manfaat Praktis

Secara praktis, penelitian ini diharapkan dapat memberikan wawasan kepada Direktorat Jenderal Pajak (DJP) dan pihak pengembang aplikasi CoreTax mengenai persepsi masyarakat terhadap layanan digital perpajakan. Hasil analisis sentimen dapat digunakan sebagai dasar untuk meningkatkan kualitas fitur, memperbaiki antarmuka, serta memperkuat strategi komunikasi digital agar lebih sesuai dengan kebutuhan pengguna. Selain itu, penelitian ini juga dapat menjadi contoh penerapan teknologi analitik dalam evaluasi pelayanan publik, sehingga mendorong penggunaan data berbasis opini masyarakat dalam pengambilan keputusan dan pengembangan kebijakan yang lebih responsif.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Pada bab ini terdiri atas latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan. Bab ini memberikan gambaran umum mengenai alasan, arah, serta tujuan dari penelitian yang dilakukan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi rangkuman teori dan konsep yang relevan dengan penelitian, termasuk teori tentang analisis sentimen, algoritma *Naive Bayes*, serta literatur dan penelitian terdahulu yang mendukung studi ini. Tinjauan pustaka ini bertujuan untuk membangun landasan teoritis yang kuat sebagai dasar analisis.

BAB III METODE PENELITIAN

Bab ini membahas secara rinci mengenai objek penelitian, sumber dan teknik pengumpulan data, tahapan pengolahan data, serta penerapan algoritma *Naive Bayes* untuk klasifikasi sentimen. Selain itu, dijelaskan juga metode evaluasi model untuk menilai tingkat akurasi hasil klasifikasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penerapan algoritma *Naive Bayes* terhadap data ulasan pengguna

aplikasi CoreTax. Pembahasan mencakup analisis hasil klasifikasi sentimen, interpretasi data, serta evaluasi performa model berdasarkan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

BAB V PENUTUP

Bab ini memuat kesimpulan yang diperoleh dari hasil penelitian serta saran untuk pengembangan penelitian selanjutnya. Kesimpulan dirumuskan berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada bab sebelumnya.

