

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Perkembangan teknologi internet yang sangat pesat telah membawa perubahan mendasar dalam berbagai aspek kehidupan manusia, termasuk dalam cara individu dan kelompok berkomunikasi. Internet tidak hanya berfungsi sebagai sarana akses informasi, tetapi juga telah mendorong munculnya berbagai platform digital yang menyediakan ruang interaksi baru. Salah satu wujud nyata dari transformasi ini adalah hadirnya media sosial, yang kini memainkan peran strategis dalam menyebarkan informasi, menjalin komunikasi, serta membangun kolaborasi antar pengguna (Ramdhanian *et al.*, 2024).

Di Indonesia, penggunaan media sosial terus menunjukkan peningkatan yang signifikan. Berdasarkan laporan *We Are Social* dan *Meltwater* tahun 2025, tercatat 143 juta pengguna media sosial, dengan 25,2 juta di antaranya aktif di Twitter. Platform ini memiliki fitur *retweet* yang memungkinkan informasi tersebar dengan cepat dan luas, sehingga menjadikannya ruang yang dinamis untuk menangkap opini publik. Fokus Twitter pada konten teks juga mendorong pengguna lebih leluasa mengekspresikan opini, sikap, dan emosi dibandingkan media sosial berbasis visual. Keunggulan tersebut menjadikan data Twitter lebih mudah diolah melalui *text mining* (Wongkar & Angdresey, 2020).

Selain itu, keberadaan fitur *hashtag* pada Twitter memungkinkan pengelompokan dan pencarian topik tertentu secara sistematis, sehingga memudahkan pelacakan isu dan tren sentimen masyarakat secara *real-time*. Dengan struktur data yang terbuka, dominasi teks sebagai format utama, dan sistem pelabelan topik yang efisien, Twitter menjadi salah satu media sosial yang paling ideal untuk analisis persepsi publik terhadap kebijakan pemerintah. Selain itu, keterbukaan akses data melalui API (*Application Programming Interface*) semakin mendukung proses *web crawling* dan analisis data untuk tujuan penelitian (Villavicencio *et al.*, 2021).

Dalam beberapa tahun terakhir, pemerintah Indonesia terus berupaya melakukan reformasi sistem perpajakan, salah satunya dengan memperkenalkan Coretax DJP. Coretax DJP merupakan sistem administrasi perpajakan terbaru yang diterapkan oleh Direktorat Jenderal pajak (DJP) mulai 1 Januari 2025. Sistem ini bertujuan untuk memodernisasi sistem administrasi perpajakan yang ada saat ini. Coretax DJP menyediakan layanan digital yang *modern* untuk mendukung berbagai kebutuhan perpajakan, mulai dari pendaftaran wajib pajak, pelaporan Surat Pemberitahuan Tahunan (SPT), pembayaran pajak, hingga pemeriksaan dan penagihan pajak. Dengan Coretax DJP, pemerintah berusaha meningkatkan transparansi dalam pengelolaan pajak, memperbaiki akuntabilitas, dan meningkatkan kualitas pelayanan kepada wajib pajak (Panjaitan & Yuna, 2024).



Gambar 1.1. Login Page Coretax DJP

Pada akhir tahun 2024, DJP mencatat jumlah wajib pajak terdaftar sebanyak 86,7 juta. Jumlah wajib pajak ini terdiri atas wajib pajak orang pribadi, wajib pajak badan, dan instansi pemerintah. Angka ini meningkat sebesar 17,23% dari tahun sebelumnya (Kurniati, 2025). Pertumbuhan jumlah wajib pajak ini menunjukkan semakin tingginya kebutuhan akan sistem administrasi perpajakan yang lebih modern, efisien, dan mampu mengakomodasikan volume pengguna yang terus meningkat.

Penerimaan masyarakat terhadap Coretax DJP memegang peranan penting dalam menentukan keberhasilan implementasinya. Masyarakat perlu merasa nyaman dan percaya terhadap sistem baru ini agar dapat berkontribusi secara

optimal (Juwita & Qadri, 2024). Namun, dalam beberapa waktu terakhir, muncul berbagai keluhan dari pengguna Twitter terkait penggunaan Coretax DJP. Beberapa isu yang kerap disampaikan mencakup kesulitan dalam mengakses sistem, terjadinya *error* pada *platform*, dan kesulitan adaptasi karena dinilai kurang *user-friendly* dibandingkan sistem sebelumnya. Selain itu, sebagian pengguna berpendapat bahwa perubahan dari sistem lama ke Coretax DJP justru membuat proses administrasi pajak menjadi lebih sulit. Di sisi lain, layanan pelanggan dianggap tidak responsif dan belum mampu memberikan solusi yang jelas atas permasalahan yang dialami pengguna. Minimnya informasi dan sosialisasi mengenai cara penggunaan sistem juga turut menyebabkan kurangnya pemahaman masyarakat terhadap fitur-fitur yang tersedia di Coretax DJP.



Gambar 1. 2. Contoh Keluhan Masyarakat terhadap Coretax DJP di Twitter

Keluhan-keluhan ini menunjukkan bahwa masih terdapat tantangan dalam implementasi Coretax DJP, yang perlu dianalisis lebih lanjut agar dapat diatasi dengan solusi yang tepat. Salah satu pendekatan yang dapat digunakan untuk memahami opini masyarakat secara lebih sistematis adalah analisis sentimen. Analisis sentimen atau yang juga dikenal sebagai penambangan opini merupakan salah satu cabang *Text Mining* yang digunakan untuk mempelajari opini, perasaan,

dan emosi yang terkandung dalam kata-kata atau kalimat (Villavicencio *et al.*, 2021). Metode yang dapat digunakan dalam analisis sentimen yaitu metode *machine learning*, yang mengandalkan algoritma untuk mengklasifikasikan sentimen berdasarkan pola yang ditemukan dalam data teks.

Dalam pendekatan *machine learning*, pelatihan model dilakukan dengan memberikan data teks berlabel sentimen agar algoritma dapat mempelajari hubungan antara teks dan kategorinya. Model kemudian dievaluasi menggunakan data uji untuk mengukur akurasi dan kemampuan generalisasinya. Jika hasil evaluasi baik, model dapat digunakan untuk memprediksi sentimen pada data baru yang belum pernah dianalisis (Purnamasari *et al.*, 2023).

Beberapa algoritma *machine learning* yang umum digunakan dalam analisis sentimen yaitu Naive Bayes, *Support Vector Machine (SVM)*, *Random Forest*, dan *K-Nearest Neighbors (KNN)*. Dari berbagai algoritma tersebut, dalam penelitian ini digunakan algoritma Naive Bayes, karena memiliki keunggulan dalam menangani data teks dengan baik serta efisien dalam proses komputasi. Naive Bayes juga dikenal karena kesederhanaan dan keefektifannya dalam menangani klasifikasi teks dalam jumlah besar (Negara *et al.*, 2024).

Para peneliti banyak yang menggunakan metode Naive Bayes untuk melakukan analisis sentimen terhadap berbagai isu. Penelitian dengan metode tersebut pernah dilakukan oleh (Negara *et al.*, 2024) yang menganalisis sentimen publik terhadap penunjukkan staf kepresidenan. Penelitian ini memberikan hasil bahwa metode Naive Bayes efektif digunakan untuk mengklasifikasikan 3.000 komentar di media sosial (Twitter, Facebook, Youtube) yang dibagi menjadi 80% data latih dan 20% data uji ke dalam tiga polaritas, yaitu positif, negatif, dan netral.

Penelitian oleh (Qahar *et al.*, 2024) mengenai analisis faktor yang mempengaruhi pengalaman pengguna *Mobile JKN* pernah dilakukan dengan pendekatan analisis sentimen. Dalam penelitian tersebut digunakan model klasifikasi Naive Bayes (NB) dan *Support Vector Machine (SVM)* yang dikombinasikan dengan *Synthetic Minority Over-sampling Technique (SMOTE)* untuk mengatasi ketidakseimbangan data. Hasil yang diperoleh menunjukkan

performa model yang sangat baik, dengan nilai akurasi sebesar 93,33%, presisi 93,76%, *recall* 93,33%, dan *F1-score* 93,43%.

Penelitian lain oleh (Ressan & Hassan, 2022) yang melakukan analisis sentimen terkait pandemi COVID-19 menggunakan beberapa variasi dari algoritma Naive Bayes, seperti Multinomial Naive Bayes dan Bernouli Naive Bayes. Hasil Penelitian ini menunjukkan bahwa model Multinomial Naive Bayes berhasil mengklasifikasikan 3057 *tweets* ke dalam empat polaritas (gembira, sedih, takut dan marah) dan 10.000 *tweets* ke dalam dua polaritas (positif dan negatif) dengan akurasi yaitu 91,6%. Hasil ini menunjukkan bahwa algoritma Naive Bayes dapat secara efektif mengidentifikasi respons publik terhadap kebijakan kesehatan.

Analisis sentimen masyarakat Indonesia terhadap alat pendeteksi COVID-19 (*Genose*) menggunakan algoritma Naive Bayes juga pernah dilakukan oleh (Jonathan & Nataliani, 2022). Hasil penelitian berupa tabel *confussion matrix* untuk menghitung nilai *precision*, *recall*, dan *accuracy*. Nilai *precision* yang didapat adalah sebesar 93%, *recall* sebesar 88%, dan *accuracy* sebesar 91%.

Penelitian lain oleh (Fauzi *et al.*, 2023) juga menunjukkan bahwa algoritma Naive Bayes, yang menghasilkan *accuracy* 76,25%, *F1 score* 78%, dan *recall* 80%, memiliki kinerja yang lebih baik dibandingkan algoritma *Random Forest* dalam mengklasifikasikan opini publik terhadap isu perubahan iklim di Twitter. Dengan mempertimbangkan penelitian-penelitian sebelumnya, dapat disimpulkan bahwa algoritma Naive Bayes telah terbukti efektif dalam analisis sentimen berbagai isu. Namun, penelitian yang secara khusus menganalisis sentimen masyarakat Indonesia terhadap sistem Coretax DJP masih sangat terbatas. Oleh karena itu peneliti tertarik untuk melakukan penelitian yang berjudul “Analisis Sentimen Masyarakat Indonesia Terhadap Coretax DJP di Twitter Menggunakan Algoritma Naive Bayes”.

Urgensi penelitian ini terletak pada pentingnya memahami persepsi masyarakat terhadap sistem perpajakan yang baru, mengingat pajak merupakan penyumbang terbesar dalam anggaran pendapatan di Indonesia. Oleh karena itu,

keberlanjutan dan efektivitas sistem perpajakan sangat krusial bagi perekonomian negara. Sistem perpajakan yang baik tidak hanya harus efisien secara administratif, tetapi juga harus dapat diterima dengan baik oleh masyarakat agar kepatuhan wajib pajak meningkat. Dengan memahami opini masyarakat melalui analisis sentimen, pemerintah dapat mengevaluasi penerimaan masyarakat terhadap sistem perpajakan yang baru serta mengambil langkah-langkah yang diperlukan untuk meningkatkan efektivitas implementasi Coretax DJP.

1.2. Identifikasi Masalah

1. Tingkat penerimaan masyarakat terhadap sistem Coretax DJP masih rendah, ditandai dengan banyaknya keluhan yang muncul di media sosial, khususnya Twitter.
2. Aksesibilitas sistem yang buruk, seperti kesulitan dalam mengakses *platform* dan seringnya terjadi *error*; menyebabkan gangguan dalam proses administrasi perpajakan.
3. Perubahan dari sistem lama ke Coretax DJP dianggap memperumit proses administrasi pajak.
4. Layanan *helpdesk* DJP cenderung kurang responsif dan belum memberikan solusi yang optimal, sehingga menyulitkan wajib pajak dalam mengatasi kendala.
5. Minimnya informasi dan sosialisasi menyebabkan kurangnya pemahaman masyarakat dalam menggunakan Coretax DJP.

1.3. Rumusan Masalah

1. Bagaimana sentimen masyarakat Indonesia terhadap Coretax DJP di twitter menggunakan algoritma Naive Bayes?
2. Bagaimana kinerja algoritma Naive Bayes dalam mengklasifikasikan sentimen masyarakat Indonesia terhadap Coretax DJP di Twitter?

1.4. Tujuan Penelitian

1. Menganalisis sentimen masyarakat Indonesia terhadap Coretax DJP di Twitter menggunakan algoritma Naive Bayes.

2. Mengevaluasi kinerja algoritma Naive Bayes dalam mengklasifikasi sentimen masyarakat Indonesia terhadap Coretax DJP di Twitter.

1.5. Batasan Masalah

1. Data yang digunakan dalam penelitian ini diambil dari *platform* media sosial Twitter, dengan fokus pada unggahan yang mengandung kata kunci “Coretax”.
2. Penelitian ini hanya menganalisis opini masyarakat dalam bentuk teks berbahasa Indonesia yang diunggah selama bulan Februari, Maret, dan April.
3. Algoritma yang digunakan untuk analisis sentimen adalah Naive Bayes, dengan proses klasifikasi dibatasi pada tiga polaritas, yaitu positif, negatif, dan netral.
4. Data penelitian dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji.
5. Proses pelabelan data dilakukan secara otomatis dengan memanfaatkan *library* Python *TextBlob*.

1.6. Manfaat Penelitian

1. Bagi Direktorat Jenderal Pajak, penelitian ini memberikan gambaran kepada Direktorat Jenderal Pajak tentang persepsi masyarakat terhadap Coretax DJP, sehingga dapat meningkatkan kualitas layanan perpajakan serta sebagai bahan evaluasi terhadap sistem yang diterapkan.
2. Bagi peneliti, penelitian ini memberikan kesempatan untuk mengembangkan pemahaman teoritis dan keterampilan teknis dalam penerapan algoritma Naive Bayes untuk analisis sentimen, khususnya pada data media sosial Twitter. Selain itu, peneliti juga dapat memahami opini masyarakat terhadap kebijakan publik dan berkontribusi langsung dalam memberikan masukan untuk evaluasi serta pengembangan sistem administrasi perpajakan di Indonesia.
3. Bagi Pembaca, penelitian ini dapat menambah wawasan dan literatur ilmiah mengenai penerapan algoritma Naive Bayes dalam analisis sentimen, khususnya di bidang perpajakan. Hasil penelitian ini juga dapat menjadi rujukan bagi pembaca yang ingin melakukan kajian serupa di masa mendatang.