

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang Masalah**

Pajak adalah kontribusi wajib kepada negara yang harus dipenuhi oleh individu atau badan yang bersifat memaksa berdasarkan Undang-Undang, dengan tidak mendapatkan imbalan secara langsung dan digunakan untuk keperluan negara bagi sebesar-besarnya kemakmuran rakyat. Definisi pajak ini diatur dalam Undang-Undang Republik Indonesia Nomor 28 Tahun 2007 perubahan ketiga atas Undang-Undang Nomor 6 Tahun 1983 Tentang Ketentuan Umum Dan Tata Cara Perpajakan. Salah satu fungsi pajak adalah untuk membiayai pengeluaran negara (*budgeter*), seperti pembangunan infrastruktur, kesehatan, pendidikan, dan lain sebagainya. Pajak juga berperan dalam mengatur distribusi pendapatan dan kekayaan di masyarakat (*regulerend*), serta mendorong pertumbuhan ekonomi dan investasi (Niru & Sinaga, 2016).

Pajak merupakan sumber pendapatan penting bagi negara dalam meningkatkan kesejahteraan masyarakat. Sebagai sumber pendapatan utama, pajak digunakan untuk membiayai berbagai program pembangunan dan layanan publik, seperti kesehatan, pendidikan, infrastruktur, dan keamanan. Selain itu, pajak juga berperan sebagai alat pengatur kebijakan ekonomi. Pemerintah dapat mengatur kebijakan fiskal untuk mengendalikan inflasi dan menjaga stabilitas ekonomi negara. Dengan jumlah penerimaan pajak yang memadai, negara dapat meminimalkan ketergantungan pada utang luar negeri yang berdampak pada pengurangan beban bunga dan risiko ketidakstabilan ekonomi. Oleh karena itu, kesadaran dan partisipasi masyarakat sebagai wajib pajak sangat penting bagi keberhasilan pengumpulan pajak yang efektif dan efisien serta pembangunan ekonomi yang berkelanjutan (Ismayadie, 2019).

Menteri Keuangan Sri Mulyani Indrawati melaporkan bahwa pada tahun 2022, penerimaan pajak negara tumbuh sebesar 34,3% secara tahunan hingga mencapai Rp.1.716,8 triliun, presentase ini lebih tinggi dibandingkan dengan presentase pada tahun sebelumnya (Rahayu, 2023). Kenaikan penerimaan pajak tersebut tidak

terlepas dari indeks kepercayaan masyarakat terhadap otoritas perpajakan yang mengalami kenaikan sekitar 43% pada tahun 2021. Namun, pada tanggal 24 Februari 2023, kasus penganiayaan yang dilakukan oleh Mario Dandy Satriyo, anak dari pejabat Direktorat Jenderal Pajak, Rafael Alun Trisambodo telah menimbulkan turunnya tingkat kepercayaan masyarakat terhadap otoritas perpajakan. Selain itu, peristiwa pemungutan pajak atas piala yang dimenangkan oleh Fatimah Zahratunnisa juga ikut mempengaruhi persepsi masyarakat terhadap kebijakan perpajakan di Indonesia. Kedua peristiwa besar ini mengakibatkan pengaruh negatif pada serangkaian peristiwa lain yang secara signifikan menurunkan kepercayaan masyarakat terhadap instansi perpajakan. Menurut pakar *big data*, Danny Darussalam, sentimen negatif masyarakat terhadap Ditjen Pajak Kemenkeu melonjak sekitar 47.5% pada february 2023. Hal ini berpotensi mempengaruhi kepatuhan masyarakat dalam memenuhi kewajiban pajak mereka dan mengurangi kenaikan penerimaan pajak yang dicapai pada tahun sebelumnya. Sebab, seperti yang diungkapkan oleh Ibrahim, dkk (2020) kepatuhan wajib pajak dipengaruhi oleh tingkat kepercayaan publik. Semakin tinggi tingkat kepercayaan publik, maka semakin besar juga tingkat kepatuhan wajib pajak. Oleh karena itu, penting bagi otoritas perpajakan untuk membangun kembali kepercayaan masyarakat agar kepatuhan wajib pajak tetap tinggi dan penerimaan pajak negara dapat terus meningkat.

Upaya awal dalam membangun kembali kepercayaan publik terhadap otoritas perpajakan adalah dengan melakukan analisis sentimen. Analisis sentimen dapat dimanfaatkan untuk memantau opini dan sentimen masyarakat terhadap pelayanan pajak, sehingga pemerintah dapat mengambil tindakan yang tepat untuk meningkatkan kepercayaan masyarakat terhadap lembaga pajak. Menurut penjelasan dari Finandra & Hamami (2021), analisis sentimen adalah proses menganalisis teks digital untuk menentukan nada emosional positif, negatif, atau netral pada sebuah pesan. Dalam konteks ini, Analisis sentimen dapat dilakukan berdasarkan data tweet dari pengguna Twitter di Indonesia untuk memantau opini dan sentimen masyarakat terhadap pajak.

Berdasarkan data *We Are Social* (2023) dan *Hootsuite* (2023), terdapat peningkatan jumlah pengguna Twitter di Indonesia pada awal tahun 2023 yang mencapai 24 juta, dibandingkan dengan jumlah pengguna pada Januari 2022 yang hanya mencapai 18,45 juta. Dalam hal penggunaan Twitter, Indonesia menduduki peringkat kelima terbesar di dunia setelah Amerika Serikat (95,4 juta), Jepang (67,5 juta), India (27,3 juta) dan Brasil (24,3 juta). Menurut Dwi Adriansah (*Country Industry Head Twitter Indonesia*), *Data Twitter Trends* selama tiga tahun terakhir menunjukkan bahwa pembicaraan terkait keuangan meningkat sebesar 38%, termasuk pembahasan tentang menabung, produk investasi, pasar *kripto*, dan dompet *digital*. Twitter dapat menjadi sumber data yang penting untuk analisis sentimen terkait pajak karena pengguna Twitter di Indonesia cukup banyak dan terus meningkat, serta pembicaraan terkait keuangan yang cukup signifikan di platform tersebut. Selain itu, Twitter juga memiliki fitur yang memungkinkan pengguna untuk menggunakan tagar atau hashtag, sehingga memudahkan dalam memantau sentimen atau opini terkait topik tertentu, seperti pajak.

Dalam penelitian ini, data *tweet* pengguna twitter dapat dikumpulkan menggunakan metode *scraping* dengan bantuan *Snsrape* dan *API Twitter*. Kemudian, data yang diambil dapat dianalisis untuk mengidentifikasi sentimen masyarakat terhadap pajak menggunakan algoritma *machine learning* seperti TF-IDF dan *Naive Bayes*. Algoritma TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah metode statistik yang digunakan untuk mengevaluasi seberapa penting suatu kata dalam sebuah dokumen dengan memperhatikan frekuensi kemunculan kata tersebut dalam dokumen tersebut dan dalam seluruh dokumen yang ada (Kim & Gil, 2019). Sementara itu, algoritma *Naive Bayes* adalah algoritma yang digunakan untuk melakukan klasifikasi teks berdasarkan probabilitas dari setiap kata yang terdapat dalam dokumen (Silaen et al., 2022).

Selain itu, terdapat pula beberapa algoritma *machine learning* lain yang dapat digunakan untuk analisis sentimen, seperti *Random Forest*, *Support Vector Machine (SVM)*, dan *Neural Network* (Miftahusalam et al., 2022). Namun, dalam penelitian ini penulis memilih menggunakan algoritma TF-IDF dan *Naive Bayes* karena kedua algoritma tersebut sudah banyak digunakan dalam analisis sentimen

dan terbukti memberikan hasil yang cukup baik dalam klasifikasi teks. Penelitian yang sebelumnya dilakukan oleh Farah Zhafira, dkk (2021) memperlihatkan bahwa penggunaan TF-IDF dan *Naive Bayes* memiliki akurasi terbaik sebesar 97% dalam analisis sentimen kebijakan kampus merdeka berdasarkan komentar pada youtube. Selain itu, Putu Gita Naraswati, dkk (2021) juga pernah melakukan studi Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan COVID-19 di Indonesia dengan *Naive Bayes Classification*. Hasil penelitian tersebut Menunjukkan bahwa kebijakan penanggulangan COVID-19 di Jakarta belum memperoleh hasil yang maksimal, serta evaluasi yang dilakukan pada *Naive Bayes Classification* menunjukkan bahwa tingkat akurasi klasifikasi mencapai 87,34%, sensitivitas 93,43%, dan spesifisitas 71,76%, yang berarti metode ini sudah cukup efektif.

Berdasarkan latar belakang tersebut, penulis termotivasi untuk melakukan Analisis Sentimen Pengguna Twitter Terhadap Topik Perpajakan di Indonesia Menggunakan TF-IDF dan *Naive Bayes*. Penulis berharap dapat memberikan informasi penting bagi pemerintah atau lembaga terkait untuk memperbaiki kebijakan dan tindakan terkait perpajakan di Indonesia.

## 1.2. Identifikasi Masalah

Adapun identifikasi masalah yang dilakukan sebelum melakukan penelitian adalah sebagai berikut:

1. Kasus penganiayaan yang dilakukan oleh Mario Dandy, anak pejabat Ditjen Pajak Kemenkeu terhadap David Ozora mengakibatkan turunnya tingkat kepercayaan masyarakat terhadap otoritas perpajakan.
2. Pada Maret 2023, Fatimah Zahratunnisa yang memenangkan perlombaan bernyanyi di TV Jepang mengirimkan piala ke Indonesia dan harus membayar pajak sebesar 4 juta rupiah. Peristiwa ini dapat mempengaruhi persepsi masyarakat terhadap kebijakan perpajakan di Indonesia dan dapat menurunkan kepercayaan masyarakat terhadap instansi perpajakan Indonesia. Selain itu, peristiwa ini juga dapat menimbulkan ketidakpuasan masyarakat terhadap kebijakan perpajakan yang dianggap tidak adil.

3. Sentimen pengguna Twitter terhadap topik perpajakan di Indonesia belum terdokumentasi secara komprehensif dan baik. Hal ini dapat menjadi kendala dalam pengambilan kebijakan yang berhubungan dengan perpajakan di Indonesia. Oleh karena itu, diperlukan penelitian yang dapat mengumpulkan dan menganalisis sentimen pengguna Twitter terhadap topik perpajakan di Indonesia secara komprehensif.
4. Dibutuhkan sebuah metode analisis sentimen yang dapat membantu memahami pandangan pengguna Twitter terhadap topik perpajakan di Indonesia. Salah satu metode analisis sentimen yang dapat digunakan adalah pembobotan TF-IDF dan klasifikasi *Naive Bayes*.

### **1.3. Ruang Lingkup**

Penelitian ini menggunakan teknik pembobotan TF-IDF (*Term Frequency–Inverse Document Frequency*) dan metode klasifikasi *Naive Bayes* untuk mengidentifikasi perasaan, pandangan, serta pendapat dari pengguna twitter terhadap kebijakan dan layanan perpajakan di Indonesia.

### **1.4. Batasan Masalah**

Dalam penyusunan penelitian ini diperlukan beberapa batasan masalah untuk menghindari meluasnya permasalahan yang akan dibahas, yaitu:

1. Masalah yang diteliti pada penelitian ini adalah efektivitas pembobotan TF-IDF dan metode *Naive Bayes* dalam klasifikasi data sentimen pengguna twitter terhadap topik perpajakan di Indonesia.
2. Data sentimen dalam penelitian ini dikumpulkan berdasarkan kumpulan *tweet* pengguna twitter terhadap topik perpajakan pasca kasus penganiayaan Mario Dandy dan pemungutan pajak atas piala yang dimenangkan oleh Fatimah Zahratunnisa.
3. Data *tweet* dikumpulkan mulai dari tanggal 21 Februari 2023 sampai tanggal 31 maret 2023.
4. Jenis *tweet* yang diambil merupakan *tweet* yang berisi kritik atau saran.

5. Tahap pra-proses data dibatasi menjadi 7 (tujuh) tahapan, yaitu *Removing Punctuation, Case Folding, Stopword Removal, Normalisasi Kata, Removing url and username, stemming dan Pelabelan Sentimen*.
6. Normalisasi kata-kata dalam penelitian ini dilakukan berdasarkan kamus *Colloquial Indonesian Lexicon* (Kamus Bahasa Alay Indonesia) tanpa melakukan pemeriksaan manual.
7. Proses pelabelan sentimen pada penelitian ini dilakukan menggunakan pustaka InSet (*Indonesia Sentiment Lexicon*).
8. Penelitian ini tidak melakukan analisis konteks antar kata di dalam data sentimen pengguna twitter.
9. Analisis data yang dilakukan pada penelitian ini merupakan analisis teks dan tidak melakukan analisis profil pengguna atau informasi lain.
10. Tingkat ketepatan klasifikasi dibatasi oleh nilai akurasi, presisi dan sensitivitas.

### **1.5. Rumusan Masalah**

Berdasarkan latar belakang yang sudah diuraikan sebelumnya, permasalahan yang dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana tingkat sentimen masyarakat pengguna twitter terhadap topik perpajakan di Indonesia?
2. Bagaimana evaluasi penggunaan metode pembobotan TF-IDF dan Klasifikasi Naive Bayes dalam melakukan analisis sentimen topik perpajakan?

### **1.6. Tujuan Penelitian**

Adapun tujuan dari penelitian ini ialah sebagai berikut :

1. Mengukur tingkat sentimen masyarakat pengguna twitter terhadap topik perpajakan pasca serangkaian peristiwa yang secara signifikan menurunkan kepercayaan masyarakat terhadap instansi perpajakan.

2. Mendeskripsikan evaluasi penggunaan metode pembobotan TF-IDF dan Klasifikasi Naive Bayes dalam melakukan analisis sentimen topik perpajakan.

### 1.7. Manfaat Penelitian

Adapun manfaat yang diharapkan pada penelitian ini yaitu sebagai berikut:

1. Bagi pemerintah Indonesia, Hasil analisis sentimen dapat memberikan wawasan yang berharga bagi pemerintah dalam mengambil tindakan yang lebih tepat dan efektif dalam meningkatkan kesadaran dan kepatuhan perpajakan di Indonesia.
2. Bagi Direktorat Jenderal Pajak (DJP), DJP dapat memanfaatkan hasil penelitian ini untuk meningkatkan efektivitas dan efisiensi dalam layanan perpajakan. Dengan memahami pandangan dan sikap masyarakat terhadap perpajakan, DJP dapat mengambil tindakan yang lebih tepat dan efektif dalam meningkatkan kesadaran dan kepatuhan perpajakan di Indonesia.
3. Bagi penulis, Penulis dapat memperoleh wawasan baru mengenai sentimen pengguna Twitter terhadap topik perpajakan di Indonesia, serta meningkatkan keterampilan dalam melakukan analisis sentimen menggunakan TF-IDF dan *Naive Bayes*.
4. Bagi pembaca, Pembaca dapat memperoleh pemecahan masalah yang relevan terhadap analisis sentimen menggunakan pembobotan TF-IDF dan klasifikasi *Naive Bayes*, khususnya analisis terhadap sentimen pengguna twitter dalam topik layanan pajak di Indonesia.
5. Bagi akademik, Penelitian ini dapat memberikan kontribusi pada pengembangan ilmu pengetahuan di bidang analisis sentimen dan *Natural Language Processing*. Hasil penelitian yang signifikan dapat memperluas pengetahuan dan pemahaman, serta membantu untuk memperbarui teori dan konsep yang ada.