

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Indonesia menjadi salah satu negara yang pendapatan terbanyaknya berasal dari pajak. Hal ini dapat dilihat dari pernyataan Sri Mulyani Indrawati selaku Menteri Keuangan pada Konferensi Pers Realisasi Anggaran Pendapatan dan Belanja Negara (APBN) 2022 pada tanggal 3 Januari 2023 yang mengatakan bahwa APBN tahun 2022 terealisasi Rp 2.626,4 triliun atau 115,9% dari target, dengan penerimaan pajak berhasil mencapai Rp 1.717,8 triliun atau 115,6% dari target pajak dan menjadi pemasuk paling besar dibandingkan dari penerimaan kepabeanan dan cukai maupun penerimaan negara dari bukan pajak [1]. Namun, tercapainya target penerimaan pajak tersebut memerlukan penantian yang lama. Sejak Tahun 2008, penerimaan pajak tidak mampu mencapai target yang telah ditargetkan dalam setiap APBN [2]. Bahkan target penerimaan pajak pada kuartal kedua di tahun 2019 masih jauh dari yang ditargetkan [3]. Hingga penerimaan pajak mampu mencapai target pada tahun 2021 yang tercatat mencapai Rp 1.277,5 triliun atau setara dengan 103,9% dari target penerimaan pajak, sesuai dari penerimaan yang telah dihimpun oleh Direktorat Jenderal Pajak Kementerian keuangan [2]. Banyaknya pemasukan dari pajak tentunya tidak lepas dari kepatuhan masyarakat wajib pajak itu sendiri. Secara garis besar terdapat dua faktor yang mempengaruhi kepatuhan pajak, yaitu faktor ekonomi seperti tarif pajak, denda, serta pemeriksaan pajak dan faktor non ekonomi yang lebih cenderung berpengaruh terhadap kepatuhan wajib pajak seperti kesadaran wajib pajak, kepercayaan pemerintah, maupun persepsi keadilan pajak [4].

Menjelang berakhirnya bulan Februari 2023, masyarakat Indonesia digemparkan oleh kasus penganiayaan yang dilakukan oleh anak pejabat Direktorat Jenderal Pajak (DJP) kepada anak pengurus pusat GP Ansor. Kasus ini pertama kali diviralkan pada tanggal 21 Februari 2023 oleh salah satu pengguna twitter di media sosial tersebut [5]. Seiring terkuaknya kasus penganiayaan tersebut, masyarakat pengguna media sosial beramai-ramai mengungkapkan akun Tiktok pelaku penganiayaan yang dianggap gemar pamer harta. Pada tanggal 22 Februari 2023, identitas ayah dari pelaku serta harta kekayaannya terkuak yang dalam LHKPN 2021 mencapai Rp 56,1 miliar. Kemudian ayah pelaku penganiayaan yang

merupakan pejabat Direktorat Jendral Pajak (DJP) terseret dipanggil oleh Komisi Pemberantasan Korupsi (KPK) guna dimintai penjelasan terhadap harta kekayaan tersebut pada tanggal 1 Maret 2023. Kasus yang melibatkan pejabat pajak ini berujung pada pemecatan pada 8 Maret 2023 dan disangkakan melanggar Undang-Undang Republik Indonesia Nomor 20 tahun 2001 tentang Perubahan Atas Undang-Undang Nomor 31 tahun 1999 tentang Pemberantasan Tindak Pidana Korupsi pada tanggal 3 April 2023 [6].

Munculnya kasus korupsi yang dilakukan oleh pegawai pajak tersebut dapat mengancam institusi pengumpul pajak maupun pendapatan negara dari pajak. Hal ini dikarenakan berdasarkan penelitian yang telah dilakukan oleh Ario Irawan, persepsi korupsi pajak serta keadilan sistem perpajakan dapat menurunkan kepatuhan wajib pajak dikarenakan ketika wajib pajak memiliki persepsi yang baik terhadap fiskus atau pejabat pajak maka kepatuhan wajib pajak dapat meningkat. Hal ini membuktikan bahwasannya kepercayaan menjadi faktor utama guna meningkatkan kepatuhan wajib pajak karena ketika wajib pajak telah percaya terhadap fiskus tentunya penerimaan pajak akan meningkat [3].

Pascakasus yang melibatkan pejabat pajak tersebut menimbulkan banyak seruan di masyarakat untuk tidak membayar pajak yang dinilai oleh Tauhid Ahmad selaku Direktur Eksekutif Institute for Developments of Economics and Finance (Indef) sebagai bentuk penurunan kepercayaan terhadap institusi pengumpul pajak [7]. Salah satu seruan yang ramai muncul di masyarakat adalah gerakan boikot bayar pajak di media sosial twitter, dikarenakan masyarakat merasa uang pajak disalahgunakan untuk foya-foya oknum pejabat atau keluarga pejabat pajak itu sendiri [8]. Namun, tidak sedikit juga masyarakat yang menentang boikot bayar pajak karena sama saja tidak mendukung pembangunan nasional di berbagai bidang [9].

Ramainya pro dan kontra terhadap minat masyarakat dalam membayar pajak di twitter tentunya tidak lepas dari penjelasan dan kekuatan media sosial tersebut. Twitter merupakan salah satu fasilitas *microblogging* secara *realtime* yang menjadi tempat berbagi pengalaman secara bebas tanpa adanya halangan. *Tweet* yang diunggah oleh pengguna juga terekam secara publik. Sehingga twitter dapat dijadikan sebuah sumber data opini masyarakat yang bermanfaat dalam berbagai aspek keperluan [10]. Selain itu, twitter memiliki kekuatan yang dapat mendorong sentimen publik serta mengatur kemarahan publik, simpati, sukacita, dan ketakutan [11]. Dengan banyaknya *tweet* di media sosial twitter terhadap pro dan kontra minat masyarakat dalam membayar pajak serta kekuatan sosial media tersebut, unggahan

yang tersebar didalamnya dapat dijadikan bahan penelitian kali ini.

Analisa sentimen adalah sebuah cabang keilmuan dari penelitian penambangan teks (*text mining*) atau *data mining* [12]. Analisa sentimen merupakan suatu metode komputasi otomatis dengan memahami, mengekstrak, dan mengelola data teks guna memperoleh informasi sentimental yang terkandung dalam sebuah kalimat pendapat, perilaku, dan emosi seseorang dengan suatu entitas dari kumpulan data yang tidak terstruktur [13]. Oleh karena itu, analisis sentimen diperlukan karena data yang diambil dari twitter masih banyak mengandung *noise* dan memiliki karakteristik yang tidak terstruktur [14]. Selain itu, analisa sentimen juga dilakukan untuk melihat opini atau kecenderungan opini seseorang terhadap sebuah masalah atau objek dengan mengkategorikan opini tersebut menjadi positif, netral, atau negatif [12].

Salah satu teknik yang dapat digunakan dalam menganalisis sentimen adalah *Naive Bayes*. *Naive Bayes* merupakan metode klasifikasi yang berasal dari gagasan untuk memprediksi peluang kejadian di masa depan berdasarkan pengalaman yang telah terjadi di masa lalu. Gagasan ini pertama kali diutarakan oleh Thomas Bayes yang kemudian dikenal dengan Teorema Bayes [15]. *Naive Bayes* memiliki beberapa keunggulan dalam pengimplementasiannya. Kesederhanaan dalam perhitungan, kecepatan, dan presisi yang tinggi menjadi keunggulan dari metode *Naive Bayes* [16]. Dalam penggunaannya terdapat 3 variasi *Naive Bayes* yang biasa digunakan, yaitu *Gaussian Naive Bayes*, *Multinomial Naive Bayes* dan *Bernoulli Naive Bayes* [17].

Penelitian serupa pernah dilakukan oleh Eni Tri Handayani bersama Ari Sulistiyawati yang membahas mengenai analisa sentimen terkait respon masyarakat dari kabar harian Covid-19 dari twitter Kementerian Kesehatan Masyarakat Republik Indonesia yang kemudian diklasifikasikan menjadi positif, netral, dan negatif menggunakan metode *Naive Bayes*. Dalam penelitiannya mereka menggunakan 2397 data yang didapati dari hasil pengumpulan data atau *crawling* menggunakan python serta memanfaatkan fasilitas *twitterscraper*. Dari penelitian tersebut didapati hasil skor akurasi sebesar sebesar 78%, *precision* 92%, *recall* 85%, dan *F1-score* 88% [18].

Penelitian serupa mengenai analisa sentimen pada media sosial twitter juga pernah dilakukan oleh Auliya Rahman Isnain bersama teman-temannya. Dalam penelitiannya mereka melakukan analisa sentimen kepada 2000 data opini masyarakat dari twitter menggunakan metode support vector machine dan ekstraksi fitur TF-IDF yang membagi kelas menjadi sentimen positif 68,75% dan negatif

31,25% terhadap kebijakan pembatasan sosial berskala besar yang dilakukan oleh pemerintah. Hasil dari penelitian tersebut adalah skor akurasi sebesar 74%, *precision* 75%, *recall* 92% dan *F1-score* 83% [19].

Berdasarkan latar belakang permasalahan yang telah dijabarkan, dibuatlah penelitian analisa sentimen pengguna media sosial twitter terhadap minat membayar pajak. Penelitian ini dilakukan guna mengetahui sentimen opini masyarakat di media sosial twitter terhadap minatnya membayar pajak pascakasus yang melibatkan pegawai pajak yang dapat menurunkan kepercayaan masyarakat wajib pajak terhadap institusi pengumpul pajak. Metode yang akan digunakan dalam penelitian ini adalah tiga variasi algoritma *Naive Bayes*, yaitu *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Bernoulli Naive Bayes*. Ketiga variasi *Naive Bayes* tersebut akan dievaluasi menggunakan confusion matrix yang kemudian dilakukan perbandingan terhadap hasil evaluasi untuk mencari hasil terbaik diantaranya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara mengimplementasikan *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Benoulli Naive Bayes* untuk menganalisis sentimen pengguna twitter terhadap minatnya membayar pajak?
2. Bagaimana cara mengukur performa *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Benoulli Naive Bayes* yang digunakan untuk menganalisa sentimen mengenai minat membayar pajak para pengguna media sosial twitter?
3. Varian manakah yang menghasilkan hasil performa klasifikasi terbaik?

1.3 Batasan Permasalahan

Berdasarkan latar belakang dan rumusan masalah diatas, maka dapat dibuat batasan masalah sebagai berikut:

1. Data yang digunakan merupakan *tweet* yang mengandung kata kunci "stop bayar pajak", "berhenti bayar pajak", "stop pajak", "tetap bayar pajak", dan "tetep bayar pajak".

2. Data yang digunakan berjumlah 2617 data terhitung sejak tanggal 21 Februari 2023 sampai dengan 3 April 2023.
3. Data diklasifikasi menjadi positif, netral, dan negatif.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dibuat, maka tujuan dari penelitian ini adalah:

1. Mengimplementasikan algoritma *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Benoulli Naive Bayes* guna mengklasifikasi sentimen pengguna twitter terhadap minat membayar pajak.
2. Mengukur hasil performa *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Benoulli Naive Bayes* menggunakan *confusion matrix* dalam analisis sentimen pengguna media sosial Twitter terhadap minatnya membayar pajak.
3. Melakukan perbandingan terhadap hasil performa dari *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, dan *Benoulli Naive Bayes* guna mendapatkan varian mana yang menghasilkan performa terbaik.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dapat diambil dari penelitian kali ini adalah:

1. Dapat membantu pemerintah maupun masyarakat untuk mengetahui ekspresi dari opini yang ramai diperbincangkan oleh pengguna twitter terkait minat mereka dalam membayar pajak.
2. Mengetahui performa serta akurasi algoritma *Naive Bayes* dalam proses analisa sentimen dari topik terkait.

1.6 Sistematika Penulisan

Adapun sistematika penulisan yang akan digunakan dalam laporan skripsi ini guna membantu dalam pencarian informasi yang diperlukan serta menampilkan penyelesaian pekerjaan yang terstruktur sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini menjabarkan isi dari latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.
- Bab 2 LANDASAN TEORI
Bab ini memuat pengertian maupun definisi dari teori-teori yang dijadikan dasar pengetahuan dalam penulisan dan pelaksanaan penelitian kali ini. Seperti pengertian pajak, *Naive Bayes*, confusion matrix, dan lainnya.
- Bab 3 METODOLOGI PENELITIAN
Bab ini menjelaskan mengenai metode yang digunakan selama melakukan penelitian, dari mulai telaah dan analisa literatur hingga evaluasi sistem.
- Bab 4 HASIL DAN DISKUSI
Bab ini menjabarkan mengenai hasil implementasi sistem serta evaluasi dari sistem yang telah dibangun.
- Bab 5 SIMPULAN DAN SARAN
Bab ini berisikan simpulan dari hasil penelitian yang telah dilakukan serta saran yang dapat digunakan sebagai referensi penelitian berikutnya.

