

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media digital, yang mencakup platform-platform seperti media sosial (misalnya Facebook, Twitter/X, Instagram, TikTok), situs berita online, blog, dan podcast, telah merevolusi cara masyarakat memperoleh, mendistribusikan, dan bahkan memproduksi informasi politik. Platform yang bersifat dinamis dan interaktif ini menyajikan sarana yang lebih partisipatif dan langsung bagi pengguna. Salah satu keunggulan utamanya adalah kemampuan untuk menyajikan informasi secara real-time, yang memungkinkan berita politik, isu kebijakan, dan peristiwa penting diakses oleh publik dengan cepat. Akibatnya, individu dapat merespons dan terlibat dalam diskusi politik dengan segera. Selain itu, platform ini menciptakan ruang komunikasi dua arah yang lebih terbuka dan inklusif, menghilangkan hambatan geografis dan kelas sosial, serta memperkuat suara individu dalam dialog politik[1].

Salah satu isu yang menarik perhatian publik di media sosial Twitter(X) adalah kebijakan pemerintah mengenai kenaikan Pajak Pertambahan Nilai (PPN) menjadi 12%. PPN merupakan salah satu penyumbang pajak terbesar karena peningkatan faktor produksi seperti membuat, menyalurkan hingga memperdagangkan barang atau jasa kepada konsumen [2]. Kebijakan ini menghadapi tantangan tersendiri, terutama karena diterapkan di tengah situasi tekanan global yang turut mendorong naiknya harga bahan baku. Menyikapi tekanan dari berbagai pihak, pemerintah akhirnya memutuskan bahwa kenaikan sebesar 12% hanya diberlakukan untuk PPnBM [3].

Analisis sentimen terhadap data yang bersumber dari platform Twitter(X) dapat memberikan representasi umum mengenai persepsi masyarakat terhadap suatu kebijakan. Melalui penerapan teknik analisis sentimen, data tidak terstruktur seperti cuitan dapat diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral, sehingga memungkinkan pemahaman opini publik secara lebih

terstruktur dan sistematis [4]. Penelitian ini menggunakan metode pelabelan dengan memanfaatkan kamus InSet Lexicon, yang merupakan pendekatan berbasis lexicon-based. Kamus ini berisi kumpulan kata yang telah diberikan skor polaritas tertentu, yang digunakan untuk menilai dan menentukan bobot sentimen dalam tanggapan masyarakat terhadap suatu topik [5]. Berdasarkan skor polaritas tersebut, tanggapan kemudian dikelompokkan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Namun demikian, penggunaan bahasa yang bersifat informal dalam unggahan media sosial kerap menyulitkan proses pemahaman secara mendalam. Selain itu, keberagaman opini yang diungkapkan oleh pengguna dapat menjadi tantangan dalam mengidentifikasi pola sentimen secara akurat dan konsisten [6]. Oleh karena data yang diperoleh dari platform Twitter(X) memiliki karakteristik yang tidak terstruktur, diperlukan penerapan teknik *text mining* untuk mengolah dan menganalisis data tersebut secara efektif. Beberapa metode yang umum digunakan dalam analisis sentimen di antaranya adalah algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* [7].

Penelitian ini membandingkan algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* guna mengidentifikasi metode klasifikasi yang memberikan performa paling optimal. Metode *Naïve Bayes* dianggap unggul dalam proses klasifikasi dibandingkan metode klasifikasi lainnya, terutama dari segi akurasi dan efisiensi kinerja [8]. Di sisi lain, algoritma *Support Vector Machine (SVM)* dikenal memiliki kemampuan yang baik dalam memisahkan data ke dalam kelas-kelas tertentu serta menghasilkan tingkat akurasi yang tinggi, sehingga sangat efektif untuk diterapkan dalam analisis sentimen [9]. SVM bekerja dengan menentukan *hyperplane* atau bidang pemisah optimal dalam ruang fitur yang bertujuan untuk memisahkan satu kelas dari kelas lainnya secara maksimal [10].

Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Naïve Bayes* dan SVM dalam analisis sentimen terhadap cuitan Twitter(X) yang membahas kenaikan PPN 12%. Dengan membandingkan kedua algoritma ini, diharapkan dapat diketahui algoritma mana yang lebih efektif dalam

mengklasifikasikan sentimen positif, negatif, atau netral terkait isu tersebut. Penelitian ini tidak hanya memberikan kontribusi ilmiah dalam bidang analisis sentimen, tetapi juga dapat menjadi acuan bagi pembuat kebijakan untuk memahami persepsi masyarakat dan merancang strategi komunikasi yang lebih efektif.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja algoritma *Naive Bayes* dalam analisis sentimen Twitter tentang kenaikan PPN 12%?
2. Bagaimana kinerja algoritma SVM dalam analisis sentimen Twitter tentang kenaikan PPN 12%?
3. Algoritma mana yang lebih efektif untuk analisis sentimen terhadap data Twitter tentang kenaikan PPN 12%?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah, terdapat beberapa batasan masalah yang ditetapkan, yaitu:

1. Menggunakan bahasa pemrograman Python.
2. Data yang digunakan adalah postingan dan komentar terkait isu kenaikan PPN 12% di platform sosial media Twitter(X) yang di crawling pada tanggal 8 sampai 10 April 2025.
3. Algoritma yang dibandingkan adalah *Naive Bayes* dan SVM.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Menganalisis kinerja algoritma *Naive Bayes* dalam klasifikasi sentimen Twitter(X) tentang kenaikan PPN 12% untuk mengetahui sejauh mana algoritma ini mampu mengidentifikasi opini publik secara tepat.
2. Menganalisis kinerja algoritma Support Vector Machine (SVM) dalam klasifikasi sentimen Twitter(X) tentang kenaikan PPN 12% guna mengevaluasi keakuratannya dalam menangkap persepsi masyarakat.

3. Membandingkan efektivitas algoritma *Naive Bayes* dan SVM dalam analisis sentimen terhadap data Twitter(X) tersebut untuk menentukan algoritma yang paling optimal sebagai acuan dalam pengembangan sistem monitoring opini publik berbasis media sosial, khususnya untuk mendukung pengambilan keputusan kebijakan fiskal oleh pemerintah atau instansi terkait.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan kontribusi ilmiah dalam bidang analisis sentimen, khususnya dalam membandingkan algoritma *Naive Bayes* dan SVM.
2. Membantu pemerintah atau pemangku kebijakan untuk memahami persepsi masyarakat terhadap kebijakan kenaikan PPN 12%.
3. Menjadi referensi bagi peneliti lain yang ingin melakukan penelitian serupa di bidang analisis sentimen.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini adalah sebagai berikut:

BAB I Pendahuluan Berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II Tinjauan Pustaka Membahas teori-teori yang mendukung penelitian, seperti teori analisis sentimen, algoritma Naive Bayes, algoritma SVM, serta studi terdahulu yang relevan.

BAB III Metode Penelitian Menguraikan metode yang digunakan dalam penelitian, meliputi pengumpulan data, preprocessing, penerapan algoritma, dan evaluasi kinerja model.

BAB IV Hasil dan Pembahasan Menyajikan hasil penelitian serta analisis dan pembahasan terhadap hasil yang diperoleh.

BAB V Kesimpulan dan Saran Berisi kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.