

BAB I

PENDAHULUAN

A. Latar Belakang

Perkembangan teknologi informasi dan komunikasi dalam beberapa tahun terakhir memberikan perubahan bagi masyarakat dalam mengemukakan pendapat dan berinteraksi, terutama penggunaan media sosial dan *platform* daring. Pada era digital ini, media sosial menjadi ruang publik yang aktif digunakan masyarakat untuk menyampaikan opini terhadap berbagai isu sosial, politik, ekonomi, hingga kebijakan pemerintah. Salah satu isu yang banyak mendapatkan perhatian publik adalah kebijakan kenaikan tarif Pajak Pertambahan Nilai (PPN).

Pajak merupakan sumber pemasukan negara yang sangat penting dalam mendukung program pembangunan dan pemulihan ekonomi nasional [1]. Kementerian Keuangan mencatat realisasi penerimaan Pajak Pertambahan Nilai Dalam Negeri pada bulan Oktober 2024 mencapai Rp. 373,34 triliun [2]. Dengan adanya penyesuaian tarif PPN 11% menjadi 12% diproyeksikan berpotensi menambah penerimaan PPN sebanyak Rp. 75,29 triliun [3]. Salah satu kebijakan pajak yang ditetapkan pemerintah Indonesia adalah kenaikan tarif PPN sebagaimana diatur dalam Undang-Undang Nomor 7 Tahun 2021 tentang Harmonisasi Peraturan Perpajakan (UU HPP). Dalam Bab 4 Pasal 7 ayat (2) disebutkan bahwa tarif PPN akan naik dari 11% menjadi 12% dan mulai diberlakukan pada 1 Januari 2025 [4].

Pada awalnya, kebijakan ini direncanakan akan berlaku secara menyeluruh untuk semua barang dan jasa atau diberlakukan secara umum. Namun, setelah melalui berbagai pertimbangan dan respons publik, pemerintah akhirnya memutuskan bahwa kenaikan PPN 12% hanya akan diterapkan terhadap barang dan jasa mewah. Adapun barang dan jasa mewah yang dikenakan PPN 12% berdasarkan PMK 141/2021 s.t.d.d. PMK 42/2022 dan PMK 96/2021 s.t.d.d. PMK 15/2023 meliputi kendaraan bermotor, hunian mewah, pesawat, balon udara, dan helikopter selain untuk keperluan negara, kapal laut dan kapal pesiar selain untuk keperluan negara, dan *yacht* selain untuk keperluan pariwisata serta senjata api dan amunisi selain untuk keperluan negara [5].

Perubahan kebijakan ini menimbulkan respons baru dari masyarakat. Sebelumnya, ketika isu kenaikan PPN 12% masih menyasar secara umum, terjadi gelombang opini publik yang cukup besar, mulai dari dukungan hingga penolakan. Namun, setelah pemerintah memfokuskan kebijakan ini hanya pada barang dan jasa mewah, terjadi pergeseran opini publik yang menarik untuk dianalisis lebih lanjut. Dalam konteks ini, analisis sentimen menjadi metode yang tepat untuk mengkaji bagaimana respons masyarakat terhadap kebijakan tersebut. Analisis sentimen adalah kegiatan yang digunakan untuk menganalisis pendapat atau opini seseorang tentang suatu topik. Tugas dasar analisis sentimen adalah mengklasifikasikan beberapa teks dari dokumen, kalimat atau fitur, kalimat dari fitur tersebut bisa bersifat positif, negatif atau netral [6].

Penelitian serupa terkait analisis sentimen pernah dilakukan sebelumnya. Penelitian ini berfokus pada analisis sentimen terhadap rencana kenaikan PPN menjadi 12%, dengan menggunakan data yang diperoleh dari media sosial X. Metode yang digunakan pada penelitian ini adalah *Naïve Bayes* dengan akurasi sebesar 83% [7]. Penelitian lainnya juga melakukan analisis sentimen terhadap rencana kenaikan PPN 12% di Indonesia, data diperoleh melalui media sosial X. Penelitian ini menggunakan metode *Decision Tree* memperoleh akurasi sebesar 81,34% [8]. Selain itu, penelitian lainnya menganalisis sentimen publik di X terhadap rencana kenaikan PPN 12%. Metode yang digunakan adalah *Bert* dan memperoleh hasil akurasi sebesar 83% [9]. Penelitian-penelitian ini masih dalam tahap awal dari diterbitkannya rencana kebijakan kenaikan PPN 12% secara luas atau secara umum.

Penelitian terdahulu melakukan perbandingan metode *K-Nearest Neighbor* dan *Naïve Bayes* untuk klasifikasi sentimen pengguna aplikasi DANA. Dalam penelitian ini menggunakan data yang diperoleh dari hasil proses *web scraping*. Data yang diperoleh sebanyak 700 data. Data yang dimiliki diberi label berdasarkan ahli bahasa dan kamus *lexicon*. Data dibagi menjadi beberapa rasio untuk melatih model, dimana pembagian data yaitu 90:10, 80:20, dan 70:30. Berdasarkan hasil evaluasi untuk kedua model, data dengan rasio 70:30 memperoleh akurasi 0.9619 atau 96.19%, maka dari itu rasio tersebut digunakan untuk pengujian model. Hasil

yang diperoleh adalah model *K-Nearest Neighbor* mencapai akurasi terbaik di $K=1$ hingga $K=20$ yaitu 0.962. Data yang diberi label oleh ahli bahasa atau *expert* memperoleh akurasi sebesar 96%, presisi 96% dan *recall* 100%. Sedangkan data yang diberi label menggunakan kamus *lexicon* memperoleh akurasi sebesar 78%, presisi 46% dan *recall* 32%. Untuk model *Naïve Bayes* yang diberi label oleh ahli bahasa atau *expert* memperoleh akurasi sebesar 96%, presisi 96% dan *recall* 100%. Sedangkan data yang diberi label menggunakan kamus *lexicon* memperoleh akurasi sebesar 74%, presisi 100% dan *recall* 0% [10].

Penelitian lainnya melakukan perbandingan algoritma SVM dan *Naïve Bayes* untuk analisis sentimen penggunaan mobil listrik di Indonesia. Berdasarkan 1.517 data, nilai akurasi algoritma *naïve bayes* sebesar 63,02% dan akurasi SVM sebesar 70,82%. Dapat disimpulkan bahwa algoritma SVM menjadi algoritma terbaik [11]. Penelitian lainnya melakukan analisis sentimen opini publik terhadap pembelajaran daring di Indonesia dengan menggunakan metode *Support Vector Machine* (SVM). Data dalam penelitian ini diperoleh dari *website developer.twitter.com*, dimana jumlah data yang diperoleh sebanyak 1.257 data dari rentang waktu 18-02-2021 sampai 25-02-2021. Hasil pada penelitian ini adalah iterasi ke-1 mendapatkan nilai akurasi tertinggi yaitu sebesar 83%, sedangkan pada iterasi ke-5 menjadi akurasi terendah yaitu sebesar 54% [12].

Berdasarkan latar belakang tersebut, maka penelitian ini akan melakukan analisis sentimen terhadap kebijakan kenaikan PPN 12% pasca penerapannya yang terbatas hanya pada barang dan jasa mewah. Penelitian ini akan menggunakan metode klasifikasi *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN). Dengan membandingkan performa kedua metode klasifikasi tersebut, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan model analisis sentimen yang andal, serta memberikan gambaran yang komprehensif mengenai opini publik terhadap kebijakan PPN 12% khususnya dalam konteks penerapannya terhadap barang dan jasa mewah. Selain itu, hasil dari penelitian ini juga dapat dijadikan sebagai bahan pertimbangan bagi pembuat kebijakan dalam memahami reaksi publik dan menyusun strategi komunikasi kebijakan yang lebih efektif di masa mendatang. Serta penelitian ini diharapkan

dapat memberikan kesimpulan terkait hasil klasifikasi komentar berdasarkan pola dari data, distribusi dari masing-masing komentar serta hasil evaluasi performa model *support vector machine* dan *k-nearest neighbor* dalam mengklasifikasikan komentar masyarakat terhadap kebijakan PPN 12% pada barang dan jasa mewah.

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka penulis tertarik untuk melakukan penelitian yaitu dengan judul “Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah”.

B. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, rumusan masalah penelitian ini dapat dirumuskan sebagai berikut: bagaimana perbandingan akurasi algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) dalam analisis sentimen kebijakan PPN 12% pada barang dan jasa mewah?

C. Batasan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang telah diuraikan, adapun batasan masalah dalam penelitian “Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah” adalah sebagai berikut:

1. Data yang digunakan adalah ulasan atau komentar publik terhadap kebijakan PPN 12% pada barang dan jasa mewah yang diperoleh dari platform Youtube. Data yang diperoleh mulai tanggal 15 November 2024 hingga 26 Maret 2025. Data diperoleh dari 67 video youtube yang berbeda-beda dengan total data yang dikumpulkan sebanyak 8.706 data.
2. Analisis sentimen pada penelitian ini difokuskan pada tiga kelas utama yaitu sentimen positif, sentimen negatif dan sentimen netral.
3. Evaluasi performa yang digunakan dalam perbandingan model *Support Vector Machine* dan *K-Nearest Neighbor* yaitu menggunakan nilai *accuracy*, *precision*, *recall* dan *f1-score*. Metrik utama yang digunakan untuk perbandingan adalah nilai *accuracy*.

4. *InSet Lexicon* yang digunakan untuk pelabelan sentimen komentar atau ulasan opini publik terhadap kebijakan PPN 12% pada barang dan jasa mewah menggunakan *InSet Lexicon* yang dibuat oleh Fajri Koto dan Gemala Y. Rahmaningtyas.
5. Pelabelan sentimen dilakukan untuk mendapatkan label aktual atau label asli berdasarkan *inset lexicon*. Pelabelan ini akan dilakukan diawal dan kamus *inset lexicon* digunakan hanya untuk mendapatkan pelabelan sentimen terhadap *dataset* yang mana akan digunakan dalam pelatihan dan pengujian metode *Support Vector Machine* dan *K-Nearest Neighbor*.

D. Tujuan Penelitian

Penelitian ini bertujuan untuk membandingkan hasil akurasi antara algoritma *Support Vector Machine* dan *K-Nearest Neighbor* dalam melakukan analisis sentimen kebijakan PPN 12% pada barang dan jasa mewah.

E. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat baik secara praktis maupun konseptual dengan beberapa hasil penelitian yang dapat diterapkan. Manfaat dari penelitian ini sebagai berikut:

1. Bagi Penulis : Penelitian ini diharapkan dapat memperkaya literatur mengenai perbandingan algoritma klasifikasi yaitu *support vector machine* dan *k-nearest neighbor*, khususnya dalam konteks analisis sentimen kebijakan publik.
2. Bagi Peneliti : Penelitian ini dapat dijadikan sumber dan referensi bagi peneliti di masa yang akan datang, terutama jika mengambil permasalahan perbandingan metode lainnya dengan studi kasus dan metode yang berbeda.
3. Bagi Pembaca : Penelitian dan penulisan karya ilmiah ini memberikan kontribusi dan dampak yang cukup signifikan dalam memperluas ilmu pengetahuan pada pengembangan pembelajaran mesin dan pemrosesan

bahasa alami atau teks khususnya penggunaan algoritma *Support Vector Machine* dan *K-Nearest Neighbor*.

4. Bagi Instansi Terkait : Penelitian ini dapat digunakan sebagai bahan pertimbangan untuk memahami reaksi publik terhadap kebijakan PPN 12% pada barang dan jasa mewah serta dapat membantu menyusun strategi komunikasi publik dan evaluasi kebijakan berbasis data sentimen masyarakat. Penelitian ini juga dapat dijadikan dasar evaluasi berdasarkan data bagi pemerintah untuk mengukur tingkat penerimaan masyarakat terhadap kebijakan PPN 12% pada barang dan jasa mewah. Dengan adanya informasi dari model SVM dan KNN dapat digunakan dalam membangun sistem pemantauan opini publik otomatis untuk memitigasi gejala penolakan atau meluruskan disinformasi yang beredar di media sosial.
5. Bagi Masyarakat Umum dan Kalangan Atas: Hasil penelitian ini dapat mencerminkan suara dan opini publik secara objektif, apakah bersifat positif (mendukung), negatif (menolak) atau netral (tidak keduanya) terhadap kebijakan PPN 12% yang dikenakan pada barang dan jasa mewah berdasarkan data bukan hanya sekedar asumsi. Penelitian ini memberikan gambaran nyata mengenai respons masyarakat atas kebijakan tersebut setelah diterapkan secara terbatas hanya untuk barang dan jasa mewah.

Bagi masyarakat umum, penelitian ini menjadi ruang agar pendapat mereka dapat diakui dan dipertimbangkan dalam evaluasi kebijakan publik.

Bagi masyarakat kalangan atas penelitian ini dapat menjadi refleksi tentang bagaimana publik melihat kontribusi mereka melalui pajak barang dan jasa mewah serta mendorong rasa tanggung jawab sosial.

Secara keseluruhan, penelitian ini membantu menjadi dialog antara masyarakat umum, kalangan atas dan pemerintah agar kebijakan yang dibuat bisa lebih adil dan diterima oleh semua pihak.