

BAB I

PENDAHULUAN

1.1 Latar Belakang

Mobil merupakan bagian tak terpisahkan dari industri otomotif. Saat ini, beragam jenis mobil dengan keunggulan masing-masing dari setiap merek mudah ditemukan, baik dari segi fitur, harga, maupun performa mesin. Peningkatan penggunaan kendaraan berbahan bakar minyak sejalan dengan meningkatnya polusi udara akibat emisi gas buang [1]. Situasi ini memunculkan gagasan inovatif untuk mengembangkan mesin mobil yang ramah lingkungan sebagai respons terhadap kekhawatiran masyarakat yang semakin peduli terhadap polusi udara. Dalam beberapa tahun terakhir, teknologi mesin mobil telah berkembang pesat, tidak hanya berbasis bahan bakar minyak tetapi juga dengan hadirnya teknologi *hybrid*. Teknologi *hybrid* menjadi pilihan menarik karena mampu menggabungkan motor pembakaran internal dengan motor listrik. Sistem kerja *hybrid* pada mesin mobil menggunakan prinsip pengisian ulang regeneratif saat kendaraan beroperasi. Mobil dengan mesin *hybrid* ini dikenal sebagai HEV (*Hybrid Electric Vehicle*).[2].

Kemajuan teknologi telah mempermudah masyarakat dalam mengakses informasi dengan sangat cepat, dan internet menjadi salah satu bukti nyata dari perkembangan ini. Saat ini, internet memiliki peran penting dalam proses distribusi informasi. Banyak orang menyampaikan pendapatnya melalui media sosial, terutama X (dulu dikenal sebagai Twitter), yang memungkinkan mereka untuk berbagi informasi serta berkomunikasi dengan cepat dan efektif [3]. X merupakan salah satu media sosial yang sangat populer di kalangan pengguna internet, karena kesederhanaannya serta kemudahan dalam penggunaannya. Pengguna dapat dengan bebas menyampaikan pendapat atau opini mereka, menjadikan X sebagai platform yang sangat efektif untuk berbagi informasi dan berinteraksi dengan orang lain [4].

Analisis sentimen adalah metode yang digunakan untuk mengukur dan mengevaluasi suatu kasus atau objek tertentu melalui teks dalam bentuk kalimat atau dokumen. Melalui analisis ini, emosi dan pendapat yang terkandung dalam teks dapat diidentifikasi, sehingga memungkinkan untuk menarik kesimpulan dan membuat keputusan berdasarkan hasil evaluasi tersebut. [5]. Analisis sentimen

adalah instrumen yang sangat efektif untuk menentukan apakah evaluasi suatu teks bersifat positif atau negatif. Hasil analisis ini diperoleh dari keterkaitan antara aspek kognitif, psikologis, dan sosial. Pengguna tidak hanya menggambarkan isi konten mengenai suatu peristiwa, tetapi juga mengungkapkan perasaan mereka. Hampir semua perusahaan, baik profesional maupun non-teknis, menggunakan teknik analisis sentimen untuk memahami maksud dari umpan balik dan keinginan pelanggan terkait kualitas produk atau layanan yang mereka tawarkan. [6].

Untuk melaksanakan tugas analisis sentimen ini, diperlukan metode yang dapat mendukung dalam proses klasifikasi. Ada beberapa algoritma yang dapat digunakan diantaranya *Support Vector Machine* (SVM), *Naïve Bayes Classifier*, *k-Nearest Neighbor* (KNN), *Artificial Neural Network* (ANN) dan *Decision Tree*. Dalam, ditunjukkan bahwa algoritma *Support Vector Machine* memiliki ketahanan dan kemampuan generalisasi yang lebih tinggi serta akurasi klasifikasi yang lebih stabil dibandingkan dengan algoritma yang lain [7]. Salah satu metode yang dapat digunakan untuk melakukan analisis ini adalah *Support Vector Machine* atau SVM dalam konteks analisis sentimen. SVM merupakan salah satu metode klasifikasi dengan menggunakan metode *machine learning (supervised learning)* yang memprediksi kelas berdasarkan pola dari hasil proses training yang diciptakan oleh Vladimir Vapnik. Klasifikasi dilakukan dengan garis pembatas (*hyperlane*) yang memisahkan antara kelas opini positif dan opini negatif [8].

Berdasarkan dari hasil penelitian sebelumnya yaitu pada jurnal dengan judul Perbandingan Model Svm, Knn Dan *Naïve Bayes* Untuk Analisis Sentiment Pada Data X: Studi Kasus Calon Presiden 2024. Pada penelitian tersebut, akurasi yang paling tinggi diperoleh metode SVM dengan tingkat akurasi 88%, dilanjut dengan *Naïve Bayes* sebesar 70%, dan yang terakhir KNN sebesar 64% [9]. Penelitian analisis sentimen yang lainnya yang dilakukan dengan membandingkan hasil akurasi dari metode SVM dan *Naïve Bayes* dengan judul penelitian Perbandingan Metode Klasifikasi *Support Vector Machine* dan *Naïve Bayes* untuk Analisis Sentimen pada Ulasan Tekstual di Google *Play Store*. Berdasarkan penelitian tersebut didapatkan SVM memiliki akurasi yang lebih tinggi daripada *Naïve Bayes* dengan nilai akurasinya sebesar 81,46%, pada *Naïve Bayes* akurasi yang didapatkan rata – rata sebesar 75,41% [10].

TF-IDF adalah salah satu metode pembobotan fitur yang paling populer dan sering digunakan, dengan akurasi serta *recall* yang cukup tinggi. TF-IDF digunakan terutama dalam *information retrieval* untuk menilai seberapa penting suatu kata dalam sebuah dokumen relatif terhadap koleksi dokumen lainnya. Pembobotan ini penting karena jika suatu kata sering muncul dalam satu dokumen, kontribusinya akan semakin besar. Namun, jika kata tersebut sering muncul di banyak dokumen, kontribusinya akan menjadi lebih kecil. TF-IDF terdiri dari dua komponen utama: *Term Frequency* (TF), yang mengukur frekuensi kemunculan suatu kata dalam dokumen, dan *Inverse Document Frequency* (IDF), yang mengurangi bobot kata jika sering muncul di banyak dokumen [11]. Di mana TF merupakan banyaknya kemunculan sebuah kata dalam suatu dokumen dan jumlahnya sebanding dengan besaran bobot kata tersebut karena semakin sering muncul, maka bobotnya semakin besar. Sedangkan, pada konsep IDF, suatu kata yang sering muncul memiliki bobot kecil [12]. TF-IDF sejauh ini dikenal menjadi metode ekstraksi fitur terbaik untuk analisis teks [13].

Dari beberapa penelitian terdahulu yang telah dilakukan, metode *Support Vector Machine* (SVM) menjadi salah satu pilihan yang digunakan pada penelitian analisis sentimen ini. Penelitian ini menjabarkan proses yang menghasilkan analisis sentimen dengan metode klasifikasi SVM dan pembobotan TF-IDF. SVM dan TF-IDF dapat digunakan untuk menganalisis sentimen terkait opini masyarakat terhadap mobil listrik pada media sosial X. Selain itu, penelitian ini juga mengkaji efektivitas metode SVM dalam menganalisa sentimen opini masyarakat, yang dapat menjadi kontribusi penting dalam pengembangan analisis sentimen di berbagai bidang. Dengan demikian, penelitian ini memberi kontribusi positif dalam rangka mengoptimalkan kelayakan mobil *hybrid* dan pengembangan kendaraan yang lebih baik.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Mobil *hybrid* semakin populer sebagai kendaraan ramah lingkungan, yang memadukan mesin berbahan bakar fosil dengan motor listrik. Masyarakat di berbagai belahan dunia, termasuk pengguna Sosial Media X, secara aktif

membagikan pengalaman, pendapat, dan ulasan mereka mengenai mobil hybrid. Ulasan ini mencakup berbagai aspek, seperti performa, efisiensi bahan bakar, kenyamanan, teknologi, dan dampaknya terhadap lingkungan. Namun, ulasan ini bersifat masif dan beragam mulai dari yang sangat positif hingga sangat negatif dan tersebar dalam bentuk teks yang tidak terstruktur.

1.2.2 Metode penelitian

Penelitian ini mengusulkan metode Support Vector Machine (SVM) dan pendekatan representasi teks dengan model TF-IDF serta menggunakan penyeimbang data SMOTE.

1.2.3 Solusi Permasalahan

Solusi untuk mendapatkan klasifikasi sentimen pada mobil *hybrid* yang memiliki tingkat akurasi yang tinggi yaitu dengan mengimplementasikan metode *Support Vector Machine* (SVM) dengan pembobotan TF-IDF.

1.2.4 Pertanyaan Penelitian

Berdasarkan pada permasalahan yang ada, penelitian ini menghasilkan pertanyaan penelitian yaitu:

1. Bagaimana pengaruh penggunaan parameter C dengan nilai yang berbeda serta pengaruh metode penyeimbangan data menggunakan SMOTE (Synthetic Minority Over-sampling Technique) terhadap akurasi metode klasifikasi Support Vector Machine (SVM) dengan pembobotan TF-IDF dalam melakukan analisis sentimen terhadap mobil *hybrid* ?
2. Bagaimana hasil sentimen dari implementasi metode *Support Vector Machine* (SVM) dengan pembobotan TF-IDF untuk analisis sentimen terhadap mobil *hybrid* ?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

1. Penelitian ini bertujuan untuk menganalisis dan mengevaluasi pengaruh penggunaan parameter C dengan nilai yang berbeda serta pengaruh metode penyeimbangan data menggunakan SMOTE (*Synthetic Minority Oversampling Technique*) terhadap akurasi klasifikasi sentimen dengan

menggunakan *Support Vector Machine* (SVM) dan pembobotan TF-IDF dalam ulasan mengenai mobil hybrid.

2. Mengetahui hasil sentimen dari implementasi metode *Support Vector Machine* (SVM) dengan pembobotan TF-IDF untuk analisis sentimen terhadap mobil *hybrid*.

1.3.2 Manfaat

Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk menganalisis sentimen pada berbagai konteks, termasuk analisis sentimen terhadap mobil *hybrid* pada social media X. Hasil penelitian ini dapat mengetahui akurasi analisis sentimen sehingga lebih akurat dalam menentukan sentimen masyarakat terhadap mobil *hybrid*.

Meningkatkan Kesadaran Masyarakat: Dengan menganalisis sentimen masyarakat terhadap mobil *hybrid*, penelitian ini dapat meningkatkan kesadaran masyarakat tentang kelebihan dan kekurangan mobil *hybrid*. Hal ini dapat membantu masyarakat dalam membuat keputusan yang lebih informasi dan bijak dalam memilih mobil.

1.4 Batasan Masalah

Agar tujuan dari penelitian ini dapat tercapai dengan maksimal, maka terdapat batasan masalah sebagai berikut:

1. Metode yang digunakan yaitu *Support Vector Machine* (SVM) dengan pembobotan TF-IDF.
2. Penelitian ini menggunakan opini yang disampaikan dengan kata kunci yang digunakan adalah mobil *hybrid*.
3. Data yang digunakan berjumlah 1677 dalam bentuk csv dimulai dari tanggal 1 Januari 2020 hingga 1 November 2024.
4. Fokus pada klasifikasi sentimen ini menjadi 2 kategori utama: positif dan negatif.
5. Penggunaan parameter C (cost) dengan nilai 0.1, 1, 10, dan 100, serta pembagian data dengan rasio 80:20 dan 90:10, dilengkapi dengan penyeimbangan data menggunakan teknik *SMOTE* (*Synthetic Minority Over-sampling Technique*).

1.5 Sistematika Penelitian

Sistematika penulisan penelitian tugas akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan konteks penelitian, permasalahan yang dihadapi, solusi yang diajukan untuk mengatasi masalah tersebut, tujuan serta keuntungan yang dapat diperoleh dari penelitian ini, dan juga struktur penelitian skripsi yang disajikan

BAB II KAJIAN PUSTAKA

Bab ini membahas teori mengenai metode *Support Vector Machine* (SVM), pemobotan TF-IDF serta penelitian terkait dari metode yang diusulkan.

BAB III METODE PENELITIAN

Bab ini memaparkan rancangan penelitian yang diterapkan, serta rencana uji coba penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan terkait hasil dari pengujian yang dilakukan dalam penelitian ini.

BAB V PENUTUP

Bab ini menjelaskan terkait kesimpulan dan hasil yang telah dilakukan selama penelitian berlangsung.

LAMPIRAN

Bab ini membahas tentang lampiran yang didapat selama penelitian berlangsung