

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan salah satu negara dengan prevalensi perokok yang sangat tinggi di tingkat global. Data survei kesehatan Indonesia (SKI) tahun 2023 menunjukkan bahwa pada tahun 2024 terdapat 74,5% laki-laki dan 3,0% perempuan yang tercatat sebagai perokok aktif, sebanyak 22,46% masyarakat Indonesia usia kurang lebih 10 tahun tercatat merokok setiap hari. Dari total tersebut, sekitar 3,2% menggunakan rokok elektronik, dengan estimasi pengguna sebanyak 181.278 jiwa.[1]. Pada tahun 2025 jumlah perokok diprediksi akan meningkat dari 31,7%, diproyeksi menjadi 37,5%. Peningkatan prevalensi merokok di Indonesia berpotensi memperbesar risiko penyakit kronis hingga kematian. WHO menyebutkan bahwa konsumsi rokok menjadi penyebab sekitar 300.000 kematian setiap tahunnya di Indonesia. [2]. Sebagai perbandingan, Associate Professor Ryan Courtney dari Nasional Drug and Alcohol Research Centre (NDARC) UNSW Sydney pada tahun 2023 melaporkan hasil penelitian yang berkaitan dengan topik ini. Penggunaan tembakau alternatif berupa rokok elektrik (*vape*) memiliki tingkat efektivitas yang relatif lebih tinggi dalam mendukung upaya perokok untuk menghentikan kebiasaan merokok dibandingkan pendekatan konvensional. [3].

Dalam beberapa tahun terakhir, penggunaan rokok elektrik (*vape*) semakin populer di berbagai kalangan masyarakat Indonesia dan menjadi sorotan publik

pada media sosial, terutama di kalangan remaja dan dewasa muda. Muncul beragam pandangan bahwa *Vape* dianggap sebagai alternatif dari rokok konvensional karena diyakini memiliki risiko kesehatan yang lebih rendah dan tidak menimbulkan bau asap dan karena dianggap aman, Pandangan ini menyebabkan banyak pengguna baru yang beralih ke *vape* tanpa memahami secara mendalam dampak jangka panjang dari penggunaannya. Seiring pesatnya perkembangan teknologi informasi, masyarakat kini semakin aktif mengekspresikan pandangannya terkait rokok elektrik (*vape*) melalui platform media sosial. Di media sosial X, perdebatan ini tampak jelas melalui unggahan dan komentar masyarakat yang mencerminkan sentimen positif, negatif, dan netral. Data Reportal “Digital 2025 Indonesia” menyebut bahwa diawal 2025 X mencatat 25,2 juta pengguna, di mana X menjadi salah satu platform yang relevan karena berbasis teks [4].

Meskipun produk ini sering dipromosikan sebagai alat bantu berhenti merokok, berbagai penelitian kesehatan dan lembaga medis dunia masih memperdebatkan tingkat keamanannya salah satu laporan dari perpustakaan BPOM yaitu “Bahaya Rokok Elektronik” menjelaskan bahwa kandungan nikotin dan zat berbahaya yang dapat merusak sistem pernapasan, iritasi paru, serta gangguan jantung [5]. Penelitian kesehatan tentang “karakterisasi polutan logam dan potensi radioaktif dalam *E-liquid* rokok elektrik” menjelaskan penemuan logam berat seperti aluminium (Al), titanium (Ti), tembaga (Cu), nickel (Ni), arsenik (As), mangan (Mn) dalam kadar berbeda antar sampel. Meskipun aktivitas alpha dan beta radioaktif berada di bawah batas deteksi,

keberadaan logam berat dianggap indikasi risiko. Poin kontroversi penelitian ini menunjukkan bahwa meskipun produk diklaim lebih aman tapi masih terdapat kontaminasi yang berpotensi membahayakan [6]. Penelitian selanjutnya yaitu “kajian efek rokok elektrik terhadap kesehatan” menjelaskan kandungan berbahaya yang menyebabkan kanker kandung kemih, iritasi mulut, batuk, vertigo, gangguan saluran cerna, dan mual. Namun ada juga yang melaporkan efek positif yaitu seperti melegakan pernafasan, mengurangi batuk, dan tidur lebih nyenyak. Akan tetapi kesimpulan tetap menunjukkan bahwa rokok elektrik didalamnya terdapat nikotin dan sejumlah komponen kimia lain yang memiliki risiko merugikan terhadap kesehatan tubuh[7].

Studi mengenai kajian analisis sentimen telah diterapkan secara luas pada berbagai bidang penelitian sebelumnya mengenai rokok elektrik. Salah satunya adalah penelitian yang dilakukan oleh Rosalinda[8] pada tahun 2017 berjudul “Analisis sentimen masyarakat terhadap rokok pada twitter menggunakan algoritma *k-nearest Neighbor* (K-NN) dan *Support Vector Machine* (SVM)” penelitian menggunakan 6042 dataset yang diambil dari *twitter* dengan kata kunci yaitu Rokok. Proses klasifikasi berdasarkan hasil pengujian menggunakan confusion matrix, metode *K-Nearest Neighbor* (K-NN) dengan parameter $K = 1$ menghasilkan tingkat akurasi sebesar 87,95% sedangkan algoritma *Support Vector Machine* (SVM) memperoleh nilai akurasi sebesar 81,93%. Penelitian yang lain pada dilakukan oleh Dio [9] pada tahun 2022 berjudul “Analisis sentimen pengguna Twitter Terhadap rokok elektrik (*vape*) di Indonesia menggunakan *naive bayes*” dengan 1000 data tweet model yang

dihasilkan menunjukkan tingkat ketepatan klasifikasi sebesar 77,5%. Distribusi hasil analisis memperlihatkan bahwa proporsi sentimen negatif mencapai 11,7%, sentimen netral sebesar 77,3%, serta sentimen positif sebesar 11%. Penelitian ini menunjukkan bahwa opini masyarakat masih menjadi pro-kontra. Penelitian selanjutnya oleh Hani[10] berjudul “Perbandingan Algoritma NBC dan SVM dalam analisis sentimen Dampak kesehatan rokok elektrik” penelitian ini menelaah pandangan masyarakat mengenai dampak kesehatan dari rokok elektrik dengan memanfaatkan dua metode klasifikasi, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM), pada data yang bersumber dari media sosial X. Total data yang digunakan dalam analisis berjumlah 1.464. Hasil pengujian menunjukkan bahwa *Naive Bayes Classifier* (NBC) mencapai tingkat akurasi sebesar 80,5%, dengan nilai *precision* 36%, *recall* 80,6%, serta *F1-Score* 50,4%. Sementara itu, algoritma SVM menampilkan kinerja yang lebih optimal dengan akurasi 95,2%, *precision* 82,4%, *recall* 77,8%, dan *F1-Score* 79,9%. Temuan tersebut mengindikasikan bahwa metode SVM lebih unggul dibandingkan bahwa metode SVM lebih unggul dibandingkan Naive Bayes dalam mengklasifikasikan sentimen terkait dampak kesehatan rokok elektrik.

Penelitian perbandingan sentimen analisis menggunakan metode *Support Vector Machine* sudah banyak dilakukan selain yang berhubungan mengenai rokok tembakau ataupun rokok elektrik. Salah satu penelitian oleh khoirul [11] Salah satu penelitian terkait mengangkat terkait mengangkat judul “Perbandingan Metode *Support Vector Machine* dan *Decision Tree* untuk analisis sentimen review komentar pada aplikasi transportasi online” hasil

pengujian menunjukan bahwa algoritma *Support Vector Machine* (SVM) memperoleh tingkat akurasi sebesar 90,20%, sedikit lebih tinggi dibandingkan *Decision Tree* yang mencapai 89,80%. Temuan tersebut menegaskan bahwa SVM memiliki kinerja yang lebih baik dalam melakukan sentimen. Berikutnya dilakukan oleh Gishella [12] berjudul “perbandingan metode *Naive Bayes* dan *Support Vector Machine* untuk analisis sentimen pada ulasan pengguna aplikasi LinkedIn” pengumpulan data dilakukan melalui teknik *web scraping*, kemudian dilanjutkan dengan tahap *text pre-processing* sebelum proses klasifikasi. Hasil pengujian menunjukan bahwa algoritma *Naive Bayes* memperoleh nilai akurasi sebesar 88%, *precision* 88%, *recall* 85%, serta *F1-Score* 88%. Selanjutnya penelitian dilakukan oleh kelvin [13] berjudul “analisis perbandingan sentimen *Corona Virus Disease-2019* (Covid-19) pada *twitter* menggunakan metode *Logistic Regression* dan *Support Vector Machine* (SVM)” akurasi yang dihasilkan SVM 91,15% dan *Logistic Regression* mendapatkan akurasi 87,68%.

Berdasarkan penjabaran yang ada, maka dalam penelitian ini diusulkan sebuah judul “Analisis Sentimen Masyarakat Indonesia terhadap Rokok Elektrik (*vape*) pada Media Sosial X menggunakan *Support Vector Machine* (SVM)” penelitian ini diharapkan dapat memberikan sumbangan pemikiran pada berbagai bidang terkait. Dari sisi keilmuan, penelitian ini berkontribusi dalam pengembangan kajian terkait penerapan metode *Support Vector Machine* (SVM) pada analisis sentimen teks berbahasa Indonesia. Dari aspek teknis, penelitian ini menghasilkan sebuah model klasifikasi sentimen berbasis SVM yang dapat dimanfaatkan untuk mengelompokkan opini masyarakat dengan

memanfaatkan data yang bersumber dari media sosial X. Sementara itu, dari perspektif sosial, temuan penelitian ini diharapkan mampu menyajikan gambaran yang lebih menyeluruh mengenai pandangan masyarakat Indonesia terhadap penggunaan rokok elektrik (*vape*).

1.2 Rumusan Masalah

Berlandaaskan permasalahan yang dipaparkan dalam latar belakang, penelitian ini memfokuskan kajian pada pemanfaatan algoritma *Support Vector Machine* (SVM) dalam mengelompokkan sentimen masyarakat Indonesia mengenai rokok elektrik (*vape*)?

1.3 Batasan Masalah

Beberapa batasan masalah dalam melakukan penelitian ini yaitu:

1. Penelitian ini memanfaatkan kumpulan data *tweet* yang menggunakan bahasa indonesia.
2. Data *tweet* yang dikumpulkan terdiri dari dua kata kunci yakni rokok elektrik dan *vape*.
3. *Tools* yang digunakan dalam penelitian ini adalah *Google Colab*.
4. Hasil klasifikasi dalam penelitian ini mencakup tiga kategori, yaitu positif, negatif, dan netral.

1.4 Tujuan Penelitian

Penelitian ini difokuskan pada penggunaan algoritma *Support Vector Machine* (SVM) untuk mengelompokkan sentimen masyarakat Indonesia terhadap rokok elektrik (*vape*) berdasarkan data yang bersumber dari platform media sosial *X*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dan manfaat berupa:

1. Penelitian ini dapat memberikan gambaran persepsi masyarakat Indonesia terhadap penggunaan rokok elektrik (*vape*) melalui media sosial *X*.
2. Meningkatkan pemahaman dan keterampilan teknik dalam penerapan algoritma *Support Vector Machine* (SVM) menggunakan *tools Google Colab* serta tahap *text preprocessing*.

1.6 Sistematika Penulisan

Untuk memperjelas alur pembahasan agar mudah dipahami, penulis menyajikan gambaran umum susunan penulisan skripsi sebagai berikut.

BAB I PENDAHULUAN

Bab ini memuat uraian mengenai latar belakang penelitian, rumusan permasalahan, batasan penelitian, tujuan yang ingin dicapai, manfaat penelitian, serta sistematika penulisan

BAB II TINJAUAN PUSTAKA

Bab ini mengkaji penelitian-penelitian terdahulu yang relevan serta landasan teori yang mendukung dan berkaitan dengan topik penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang digunakan, meliputi tahapan pengumpulan data, proses pengolahan data, serta teknik analisis yang diterapkan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisikan hasil yang dicapai serta memberikan penjelasan terkait temuan yang dilakukan.

BAB V PENUTUP

Bab ini menyajikan ringkasan dari keseluruhan hasil penelitian yang telah dilakukan dan memberikan saran untuk pengembangan lebih lanjut terkait topik yang dibahas.