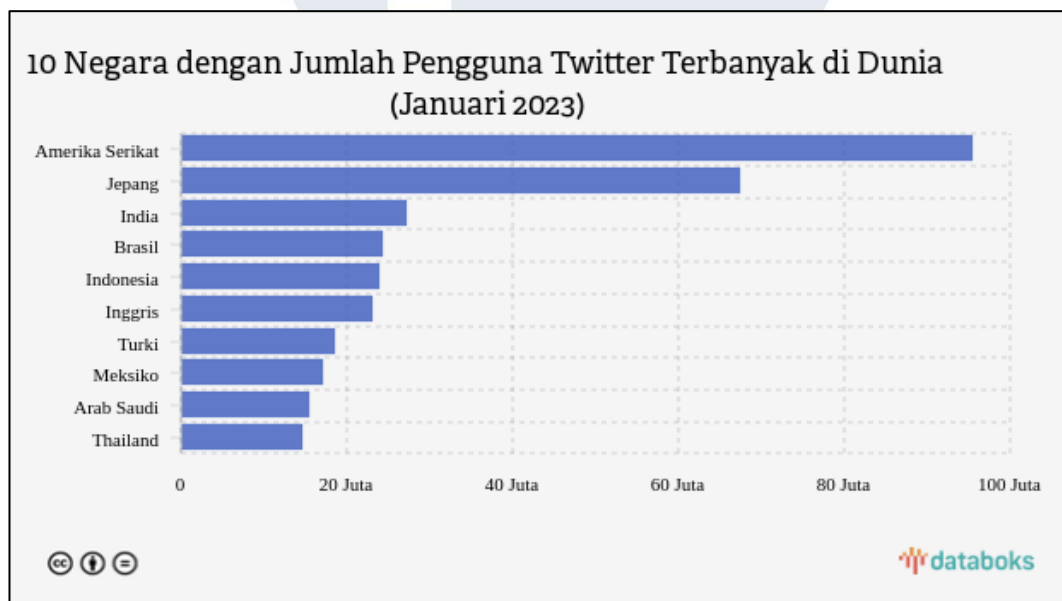


BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi informasi dan media sosial telah berkembang pesat dan juga mengubah cara masyarakat dalam menyampaikan opini terhadap berbagai isu publik, termasuk isu lingkungan. Media sosial memungkinkan pengguna mengekspresikan pandangan secara terbuka dan real-time sehingga menghasilkan data opini publik dalam jumlah besar [1], [2]. Salah satu platform media sosial yang banyak digunakan dalam penelitian opini publik berbasis teks adalah Twitter/X karena karakteristiknya berupa pesan singkat, terbuka, dan aktual serta banyak dimanfaatkan dalam analisis data sosial [3]. Indonesia termasuk dalam peringkat ke lima pengguna twitter terbanyak di dunia pada tahun 2023.



Gambar 1.1 Negara dengan pengguna twitter terbanyak (2023) [4]

Gambar 1.1 merupakan grafik besarnya jumlah pengguna Twitter di berbagai negara, termasuk Indonesia [4]. Gambar tersebut menunjukkan distribusi pengguna Twitter secara global. Tingginya jumlah pengguna tersebut menunjukkan bahwa

Twitter merupakan platform yang aktif digunakan oleh masyarakat untuk menyampaikan opini publik, sehingga berpotensi menjadi sumber data yang relevan dalam analisis sentimen.

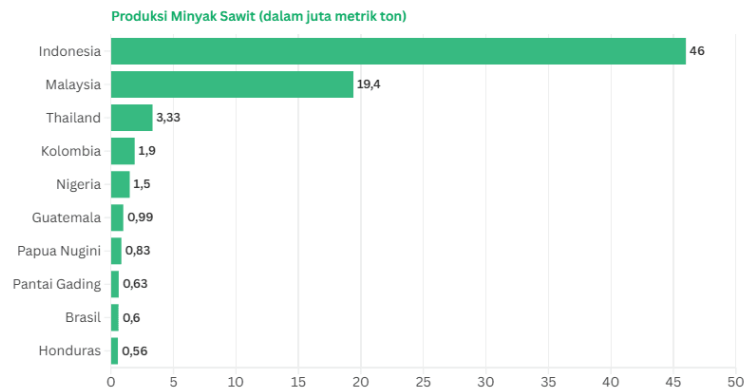
Isu lingkungan kelapa sawit merupakan topik yang banyak dibahas dalam pembahasan nasional maupun internasional. Kelapa sawit merupakan komoditas strategis Indonesia yang berkontribusi terhadap perekonomian nasional dan pembangunan pedesaan dengan menyediakan sekitar 16,5 juta tenaga kerja pada tahun 2024 [5]. Selain itu, Indonesia juga dikenal sebagai negara penghasil kelapa sawit terbesar di dunia, sehingga peran komoditas ini sangat penting dalam sektor ekspor dan pendapatan negara.

Di balik kontribusi terhadap perekonomian nasional, industri kelapa sawit juga menghadapi berbagai tantangan terkait keberlanjutan lingkungan. Perluasan lahan perkebunan sering dikaitkan dengan terjadinya deforestasi[6], degradasi lahan, kebakaran hutan dan lahan, hilangnya habitat satwa liar, serta meningkatnya emisi gas rumah kaca [7][8]. Selain itu, perubahan fungsi kawasan hutan menjadi lahan perkebunan juga dapat mengurangi kemampuan tanah dalam menyerap air sehingga meningkatkan risiko terjadinya banjir dan bencana ekologis lainnya. Oleh karena itu, isu lingkungan kelapa sawit menjadi salah satu topik yang banyak mendapat perhatian dari masyarakat, pemerintah, maupun organisasi internasional.

Permasalahan tersebut tidak hanya berdampak terhadap lingkungan, tetapi juga menimbulkan kerugian sosial dan ekonomi. Banjir yang diduga dipengaruhi oleh perubahan tutupan lahan mengakibatkan kerugian hingga Rp68,67 triliun [9]. Selain itu, pemerintah juga melakukan evaluasi terhadap tata kelola kawasan hutan dengan mencabut jutaan hektare izin usaha perkebunan dan kehutanan sebagai upaya memperbaiki pengelolaan lingkungan [10]. Kondisi tersebut menunjukkan bahwa isu lingkungan kelapa sawit memiliki dampak yang luas sehingga perlu mendapatkan perhatian lebih lanjut.

10 Produsen Minyak Sawit Terbesar di Dunia

(Periode 2024/2025)



Sumber: USDA Foreign Agricultural Service

GoodStats

Gambar 1.2 Penghasil minyak sawit di dunia periode 2024/2025 [11]

Gambar 1.2 menunjukkan bahwa Indonesia merupakan produsen minyak sawit terbesar di dunia pada periode 2024/2025 dengan total produksi mencapai 46 juta metrik ton. Jumlah tersebut jauh lebih tinggi dibandingkan negara produsen lainnya seperti Malaysia yang menghasilkan 19,4 juta metrik ton. Data ini menunjukkan bahwa industri kelapa sawit memiliki peran yang sangat penting bagi perekonomian Indonesia [11]. Besarnya produksi kelapa sawit tersebut menyebabkan isu lingkungan yang berkaitan dengan industri ini semakin banyak menjadi perhatian masyarakat. Perkembangan informasi yang cepat melalui media sosial membuat berbagai opini mengenai dampak lingkungan kelapa sawit terus bermunculan dan menjadi bahan diskusi publik.

Pembahasan global mengenai dampak lingkungan industri kelapa sawit turut memengaruhi persepsi masyarakat, termasuk yang tercermin dalam percakapan di media sosial [12], [13]. Media sosial menjadi sarana utama bagi masyarakat untuk menyampaikan pandangan, opini, serta tanggapan terhadap isu lingkungan yang berkaitan dengan industri kelapa sawit. Penelitian sebelumnya menunjukkan bahwa konten media sosial bertema lingkungan memiliki variasi sentimen yang dapat dianalisis menggunakan teknik machine learning [14]. Penelitian lain juga

menunjukkan bahwa Twitter efektif digunakan untuk menganalisis opini publik terhadap isu lingkungan di Indonesia [15].

Analisis sentimen merupakan teknik dalam text mining yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini dalam teks ke dalam kategori sentimen tertentu menggunakan pendekatan machine learning [14], [16]. Dalam konteks media sosial, analisis sentimen memungkinkan pengukuran persepsi publik terhadap suatu isu berdasarkan data teks yang dihasilkan pengguna [17]. Metode machine learning yang banyak digunakan dalam klasifikasi sentimen teks adalah Support Vector Machine (SVM) dan Naive Bayes, yang terbukti efektif dalam klasifikasi teks media sosial dan data berdimensi tinggi, mulai dari isu politik, produk, kesehatan, hoax, hingga isu lingkungan [3], [18], [17], [19], [20].

Perkembangan penelitian analisis sentimen menunjukkan bahwa metode machine learning telah banyak digunakan untuk mengklasifikasikan opini publik pada berbagai topik di media sosial. Penelitian yang dilakukan oleh Pusvita dkk. menganalisis opini publik mengenai pembukaan kebun kelapa sawit di Papua menggunakan pendekatan machine learning. Penelitian tersebut menunjukkan bahwa analisis sentimen telah diterapkan pada isu pembukaan kebun kelapa sawit di Papua menggunakan pendekatan machine learning [14].

Beberapa penelitian telah membandingkan algoritma Support Vector Machine (SVM) dan Naïve Bayes pada berbagai domain, seperti penyitaan aset koruptor, Artificial Intelligence, dan obesitas [21]-[23]. Hasil penelitian tersebut menunjukkan bahwa SVM secara konsisten memberikan performa klasifikasi yang lebih baik dibandingkan Naïve Bayes.

Berdasarkan beberapa penelitian terdahulu tersebut, dapat diketahui bahwa analisis sentimen telah banyak diterapkan pada berbagai topik menggunakan algoritma Support Vector Machine (SVM) dan Naïve Bayes. Namun, penelitian yang secara khusus membahas opini publik mengenai isu lingkungan kelapa sawit di Indonesia masih relatif terbatas. Selain itu, belum banyak penelitian yang membandingkan performa kedua algoritma pada dataset yang sama dengan

menerapkan tahapan preprocessing, pelabelan sentimen berbasis lexicon yang divalidasi oleh manusia, pembobotan TF-IDF, penyeimbangan data menggunakan RandomOverSampler, serta optimasi parameter menggunakan Grid Search. Kondisi tersebut menunjukkan bahwa masih terdapat celah penelitian yang perlu dikaji lebih lanjut.

Berdasarkan kondisi tersebut, penelitian ini dilakukan untuk menganalisis sentimen opini publik terhadap isu lingkungan kelapa sawit di Indonesia menggunakan data tweet berbahasa Indonesia yang dikumpulkan dari media sosial Twitter/X pada periode 2021-2025. Periode tersebut dipilih karena mencakup rentang waktu ketika isu lingkungan kelapa sawit masih menjadi salah satu topik yang banyak diperbincangkan oleh masyarakat di media sosial. Penelitian ini menerapkan algoritma Support Vector Machine (SVM) dan Naïve Bayes untuk mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral. Selanjutnya, performa kedua algoritma dibandingkan berdasarkan hasil evaluasi model untuk mengetahui metode yang memberikan performa klasifikasi terbaik. Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai kecenderungan opini publik terhadap isu lingkungan kelapa sawit di Indonesia serta memberikan rekomendasi algoritma klasifikasi yang memiliki performa terbaik. Selain itu, hasil penelitian ini juga diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang membahas analisis sentimen pada isu lingkungan maupun pengembangan metode klasifikasi teks menggunakan pendekatan machine learning.

1.2 Rumusan Masalah

1. Bagaimana hasil analisis sentimen opini publik terhadap isu lingkungan kelapa sawit di Indonesia berdasarkan data media sosial Twitter/X?
2. Bagaimana perbandingan performa algoritma Support Vector Machine (SVM) dan Naïve Bayes dengan menerapkan RandomOverSampler dan Grid Search dalam mengklasifikasikan sentimen tweet terkait isu lingkungan kelapa sawit berdasarkan hasil evaluasi model?

3. Bagaimana karakteristik opini publik terhadap isu lingkungan kelapa sawit di Indonesia berdasarkan distribusi sentimen yang dihasilkan dari algoritma dengan performa terbaik?

1.3 Batasan Masalah

Batasan dalam penelitian ini adalah:

1. Data penelitian berasal dari media sosial Twitter/X yang berkaitan dengan isu lingkungan kelapa sawit di Indonesia pada periode 2021–2025.
2. Data berupa tweet berbahasa Indonesia terkait isu lingkungan kelapa sawit.
3. Klasifikasi sentimen terdiri dari positif, negatif, dan netral.
4. Metode klasifikasi yang digunakan adalah Support Vector Machine (SVM) dan Naive Bayes.
5. Data yang digunakan merupakan data publik dan tidak mencakup informasi privat pengguna.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

1. Menganalisis hasil sentimen opini publik terhadap isu lingkungan kelapa sawit di Indonesia berdasarkan data media sosial Twitter/X.
2. Membandingkan performa algoritma Support Vector Machine (SVM) dan Naïve Bayes dengan menerapkan RandomOverSampler dan Grid Search berdasarkan hasil evaluasi model.
3. Menganalisis distribusi sentimen opini publik terhadap isu lingkungan kelapa sawit berdasarkan hasil klasifikasi menggunakan algoritma yang memiliki performa terbaik.

1.4.2 Manfaat Penelitian

1. Manfaat akademis

Memberikan kontribusi terhadap pengembangan penelitian analisis sentimen berbasis media sosial, khususnya pada isu lingkungan kelapa

sawit melalui penerapan dan perbandingan algoritma Support Vector Machine (SVM) dan Naïve Bayes.

2. **Manfaat praktis**

Memberikan gambaran mengenai kecenderungan opini publik terhadap isu lingkungan kelapa sawit berdasarkan data media sosial Twitter/X sehingga dapat menjadi sumber informasi bagi masyarakat, peneliti, maupun pelaku industri.

3. **Manfaat kebijakan**

Mendukung pengambilan keputusan berbasis data bagi pemerintah atau pemangku kebijakan dalam merespons isu lingkungan terkait industri kelapa sawit.

1.5 **Sistematika Penulisan**

Sistematika penulisan skripsi ini adalah sebagai berikut:

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

BAB II Landasan Teori : Bab ini berisi penelitian terdahulu yang relevan, teori-teori yang mendukung penelitian, meliputi analisis sentimen, *text mining*, *Natural Language Processing* (NLP), *Term Frequency-Inverse Document Frequency* (TF-IDF), algoritma *Support Vector Machine* (SVM), *Naïve Bayes*, *RandomOverSampler*, *Synthetic Minority Over-sampling Technique* (SMOTE), *Grid Search*, framework *Cross-Industry Standard Process for Data Mining* (CRISP-DM), serta metode evaluasi model.

BAB III Metodologi Penelitian : Bab ini menjelaskan metodologi penelitian berdasarkan framework *CRISP-DM*, yang meliputi *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Selain itu, dijelaskan proses pengumpulan data, *preprocessing*, pelabelan sentimen, validasi label, pembobotan TF-IDF, pembagian data, penyeimbangan data menggunakan *RandomOverSampler* dan SMOTE, optimasi parameter menggunakan *Grid Search*, proses pemodelan

menggunakan algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes*, serta metode evaluasi model.

BAB IV Hasil dan Pembahasan : Bab ini berisi hasil implementasi penelitian berdasarkan tahapan *CRISP-DM*, meliputi analisis data, distribusi sentimen, hasil pemodelan menggunakan *Support Vector Machine* (SVM) dan *Naïve Bayes* pada berbagai skenario pengujian, evaluasi performa model menggunakan *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*, implementasi *deployment* berupa aplikasi analisis sentimen, perbandingan hasil dengan penelitian terdahulu, serta pembahasan hasil penelitian.

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

