

Analisis Sentimen Komentar YouTube tentang Pengolahan Sampah Plastik Menggunakan Support Vector Machine dan Random Forest

Dina Lestari Putri^{*1}, Dwi Bayu Laksana Rahmadani²

^{1,2}Universitas Multi Data Palembang, Palembang, Indonesia

E-mail: ^{*1}dinalestariputri_2327250030@mhs.mdp.ac.id,

²dwibayulaksanarahmadani_2327250033@mhs.mdp.ac.id

Abstrak

Perkembangan media sosial menyediakan ruang bagi pengguna untuk menyampaikan tanggapan terhadap berbagai isu, termasuk pengolahan sampah plastik menjadi bahan bakar. Komentar pada YouTube dapat dimanfaatkan sebagai sumber data untuk menganalisis kecenderungan sentimen pengguna terhadap suatu topik. Penelitian ini bertujuan untuk menganalisis sentimen komentar YouTube tentang pengolahan sampah plastik serta membandingkan kinerja algoritma Support Vector Machine (SVM) dan Random Forest. Data penelitian diperoleh melalui proses scraping komentar pada video berjudul “Sulap Sampah Plastik Jadi Bensin RON 92 & Solar. Solusi Cerdas Atasi Limbah” pada kanal Tanilink TV. Sebanyak 1.413 komentar dianalisis melalui tahapan pre-processing, pelabelan sentimen menggunakan pendekatan berbasis leksikon (lexicon-based), ekstraksi fitur teks, pembagian data latih dan data uji dengan rasio 70:30, serta klasifikasi menggunakan SVM dan Random Forest. Hasil pelabelan menunjukkan bahwa 939 komentar (66,45%) dikategorikan sebagai sentimen negatif dan 474 komentar (33,55%) sebagai sentimen positif. Hasil evaluasi menunjukkan bahwa SVM memperoleh akurasi sebesar 85,1% dan weighted F1-score sebesar 0,848, sedangkan Random Forest memperoleh akurasi sebesar 82,5% dan weighted F1-score sebesar 0,818. Berdasarkan hasil pengujian pada dataset yang digunakan, SVM menunjukkan kinerja lebih baik daripada Random Forest. Temuan ini terbatas pada komentar dari video yang dianalisis dan belum dapat digeneralisasikan untuk menggambarkan opini masyarakat secara keseluruhan.

Kata kunci— Analisis Sentimen; Sampah Plastik; YouTube; Support Vector Machine; Random Forest

Abstract

Social media provides a platform for users to express their views on various issues, including the conversion of plastic waste into fuel. YouTube comments can serve as a data source for analyzing users' sentiment toward a particular topic. This study aims to analyze sentiment in YouTube comments concerning plastic waste processing and compare the performance of the Support Vector Machine (SVM) and Random Forest algorithms. The data were collected by scraping comments from a YouTube video entitled “Sulap Sampah Plastik Jadi Bensin RON 92 & Solar. Solusi Cerdas Atasi Limbah” published on the Tanilink TV channel. A total of 1,413 comments were analyzed through several stages, including text preprocessing, lexicon-based sentiment labeling, text feature extraction, splitting the dataset into training and testing sets at a ratio of 70:30, and classification using SVM and Random Forest. The labeling results indicated that 939 comments (66.45%) were classified as negative sentiment, while 474 comments (33.55%) were classified as positive sentiment. The evaluation results showed that SVM achieved an accuracy of 85.1% and a weighted F1-score of 0.848, whereas Random Forest achieved an

accuracy of 82.5% and a weighted F1-score of 0.818. Based on the evaluation of the dataset used in this study, SVM demonstrated better performance than Random Forest. These findings are limited to comments from the analyzed video and cannot be generalized to represent public opinion as a whole.

Keywords— Sentiment Analysis; Plastic Waste; YouTube; Support Vector Machine; Random Forest

1. PENDAHULUAN

Perkembangan teknologi informasi dan meningkatnya penggunaan internet telah mendorong masyarakat untuk lebih aktif menyampaikan pendapat melalui berbagai platform media sosial. Salah satu media yang banyak digunakan adalah YouTube, yang tidak hanya berfungsi sebagai sarana berbagi video, tetapi juga sebagai ruang diskusi melalui fitur komentar. Komentar pengguna memuat berbagai tanggapan dan penilaian terhadap isu tertentu sehingga dapat dimanfaatkan sebagai sumber data untuk menganalisis kecenderungan sentimen. Namun, banyaknya komentar membuat analisis secara manual menjadi kurang efisien. Oleh karena itu, analisis sentimen diperlukan untuk mengidentifikasi dan mengelompokkan kecenderungan sentimen dalam data teks secara otomatis ke dalam kategori tertentu, seperti positif, negatif, atau netral [1].

Salah satu isu lingkungan yang menjadi perhatian adalah permasalahan sampah plastik. Sampah plastik sulit terurai secara alami dan dapat menimbulkan dampak negatif terhadap lingkungan apabila tidak dikelola dengan baik. Di Indonesia, berbagai upaya telah dilakukan untuk mengurangi penggunaan plastik sekali pakai dan meningkatkan pengelolaan sampah plastik. Penelitian Suryani menunjukkan bahwa masyarakat memiliki kesadaran mengenai pentingnya pengelolaan sampah plastik, meskipun terdapat perbedaan pandangan terhadap kebijakan kantong plastik berbayar [2]. Selain itu, penelitian Putri et al. menemukan bahwa pengetahuan dan sikap masyarakat berpengaruh terhadap partisipasi dalam mendukung upaya pengurangan penggunaan kantong plastik [3]. Temuan tersebut menunjukkan bahwa tanggapan masyarakat terhadap persoalan sampah plastik dapat beragam. Oleh karena itu, analisis terhadap komentar di media sosial dapat digunakan untuk mengidentifikasi kecenderungan sentimen pengguna terhadap topik pengelolaan sampah plastik, khususnya pada konteks pengolahan sampah menjadi bahan bakar.

Analisis sentimen merupakan salah satu penerapan *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi polaritas sentimen dalam data teks. Pendekatan ini memungkinkan kumpulan komentar yang berjumlah besar diolah secara sistematis untuk memperoleh gambaran mengenai kecenderungan sentimen pengguna. Koto dan Rahmaningtyas membahas pengembangan dan evaluasi leksikon sentimen bahasa Indonesia yang dapat dimanfaatkan dalam analisis sentimen pada media sosial [4]. Selain pendekatan berbasis leksikon, analisis sentimen juga dapat dilakukan menggunakan algoritma *machine learning* yang mempelajari pola dari data teks untuk mengklasifikasikan sentimen. Pemilihan metode dan kualitas data menjadi aspek penting karena dapat memengaruhi hasil klasifikasi yang diperoleh.

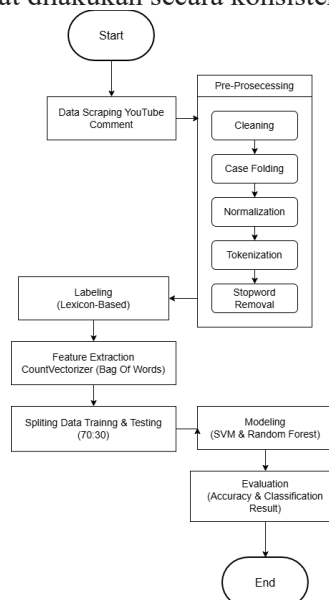
Sejumlah penelitian telah menerapkan algoritma *machine learning* untuk analisis sentimen pada media sosial. Arsyah et al. menerapkan *Support Vector Machine* (SVM) pada data Twitter untuk mengklasifikasikan sentimen [5]. Saputra et al. mengkaji penggunaan SVM yang dikombinasikan dengan normalisasi kata tidak baku pada komentar YouTube untuk meningkatkan performa klasifikasi [6]. Sementara itu, penelitian Satriya et al. membahas

sentimen masyarakat mengenai sampah plastik menggunakan SVM dan melaporkan akurasi sebesar 82,27% [7]. Penelitian-penelitian tersebut menunjukkan bahwa SVM telah digunakan pada beragam konteks analisis sentimen. Di sisi lain, *Random Forest* merupakan algoritma berbasis *ensemble* yang menggabungkan prediksi sejumlah pohon keputusan dan dapat digunakan sebagai pendekatan alternatif dalam klasifikasi teks. Chairat et al. mengkaji analisis sentimen komentar YouTube terkait kebijakan insentif kendaraan listrik di Indonesia menggunakan pendekatan klasifikasi teks [8]. Meskipun demikian, relevansi dan perbedaan hasil antarpelitian perlu ditinjau dengan mempertimbangkan perbedaan objek, sumber data, teknik pelabelan, representasi fitur, dan skema evaluasi.

Berdasarkan penelitian terdahulu tersebut, penelitian ini difokuskan pada analisis sentimen komentar YouTube mengenai pengolahan sampah plastik menjadi bahan bakar dalam video “Sulap Sampah Plastik Jadi Bensin RON 92 & Solar. Solusi Cerdas Atasi Limbah” pada kanal Tanilink TV. Penelitian ini membandingkan kinerja SVM dan Random Forest pada dataset yang sama dengan menggunakan skema pembagian data latih dan data uji sebesar 70:30. Perbandingan dilakukan untuk mengetahui perbedaan kinerja kedua algoritma dalam mengklasifikasikan sentimen pada konteks pembahasan tersebut. Dengan demikian, kontribusi penelitian ini terletak pada evaluasi komparatif kedua algoritma terhadap komentar video dengan topik pengolahan sampah plastik menjadi bahan bakar, bukan semata-mata pada penggunaan algoritma yang telah dikenal sebelumnya. Hasil penelitian diharapkan dapat memberikan gambaran mengenai kinerja kedua model pada dataset yang dianalisis.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode komparatif untuk menganalisis sentimen komentar YouTube mengenai pengolahan sampah plastik menjadi bahan bakar. Penelitian dilakukan melalui beberapa tahapan, yaitu pengumpulan data, *pre-processing*, pelabelan sentimen, ekstraksi fitur teks, pembagian data latih dan data uji, klasifikasi menggunakan algoritma *Support Vector Machine* (SVM) dan *Random Forest*, serta evaluasi kinerja model. Kedua algoritma diuji menggunakan dataset dan skema pembagian data yang sama agar hasil perbandingan dapat dilakukan secara konsisten. Alur penelitian pada Gambar 1.



Gambar 1. Alur Penelitian

2.1 Data Scraping

Data penelitian berupa komentar pengguna pada video YouTube berjudul “Sulap Sampah Plastik Jadi Bensin RON 92 & Solar. Solusi Cerdas Atasi Limbah” yang dipublikasikan pada kanal Tanilink TV. Pengumpulan data dilakukan melalui proses *scraping* untuk memperoleh komentar yang selanjutnya digunakan dalam analisis sentimen.

Data yang berhasil dikumpulkan berjumlah 1.456 komentar. Setelah melalui proses *pre-processing* dan seleksi data, diperoleh 1.413 komentar untuk analisis lebih lanjut. Dengan demikian, terdapat pengurangan sebanyak 43 komentar dari dataset awal. Pengurangan tersebut perlu dijelaskan berdasarkan hasil pemeriksaan data dan kriteria seleksi yang diterapkan.

Untuk menjamin keterulangan penelitian, informasi mengenai tanggal pengambilan data, perangkat atau pustaka yang digunakan, cakupan komentar utama dan komentar balasan, serta kriteria inklusi dan eksklusi perlu didokumentasikan. Penanganan komentar duplikat, kosong, dan spam juga perlu dijelaskan sesuai dengan prosedur yang benar-benar diterapkan dalam penelitian.

2.2 Pre-Processing

Pre-processing merupakan tahap persiapan data teks untuk mengurangi unsur yang tidak relevan dan menyeragamkan bentuk teks sebelum dilakukan pelabelan dan klasifikasi. Tahapan yang digunakan dalam penelitian ini meliputi *cleaning*, *case folding*, *normalization*, *tokenization*, dan *stopword removal*.

2.2.1 Cleaning

Cleaning dilakukan untuk membersihkan teks dari unsur yang tidak diperlukan dalam analisis, seperti URL, karakter khusus, dan tanda baca tertentu. Penghapusan angka dan emoji dilakukan sesuai dengan aturan pembersihan yang ditetapkan dalam penelitian. Namun, unsur yang berpotensi memengaruhi makna sentimen perlu dipertahankan apabila relevan terhadap konteks komentar.

2.2.2 Case Folding

Case folding merupakan proses mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*). Tahapan ini bertujuan menyeragamkan penulisan kata sehingga kata yang sama tidak diperlakukan sebagai bentuk berbeda hanya karena penggunaan huruf kapital.

2.2.3 Normalization

Normalization dilakukan untuk menyeragamkan kata tidak baku, singkatan, dan bahasa gaul ke dalam bentuk yang lebih konsisten. Proses ini dapat menggunakan kamus normalisasi yang sesuai dengan karakteristik bahasa pada komentar YouTube. Sumber kamus dan aturan normalisasi yang digunakan perlu dijelaskan agar proses dapat direplikasi.

2.2.4 Tokenization

Tokenization merupakan proses memecah teks menjadi unit-unit kata atau token. Hasil tokenisasi digunakan sebagai masukan untuk tahapan pengolahan teks berikutnya. Dalam pelaksanaannya, pemisahan token perlu memperhatikan kata negasi, seperti “tidak”, “bukan”, dan “belum”, karena kata-kata tersebut dapat mengubah polaritas sentimen.

2.2.5 Stopword Removal

Stopword removal dilakukan dengan menghapus kata-kata umum yang dinilai kurang memberikan informasi bagi proses klasifikasi, seperti “dan”, “yang”, “di”, dan “ke”. Daftar kata yang dihapus perlu disesuaikan dengan karakteristik data. Kata negasi dan kata lain yang

memiliki peran penting dalam menentukan sentimen tidak dihapus secara otomatis tanpa mempertimbangkan dampaknya terhadap makna komentar.

2.3 Data Labeling dan Ekstraksi Fitur

2.3.1 Data Labeling

Pelabelan sentimen dilakukan untuk memberikan kategori sentimen pada setiap komentar sehingga data dapat digunakan dalam proses klasifikasi. Penelitian ini menggunakan pendekatan berbasis leksikon (*lexicon-based*) untuk menentukan kecenderungan sentimen berdasarkan kata-kata yang memiliki nilai polaritas.

Untuk memastikan keterulangan proses pelabelan, perlu dicantumkan nama dan versi leksikon yang digunakan, aturan perhitungan skor sentimen, serta kriteria penentuan label positif, negatif, dan netral. Penanganan komentar yang memperoleh skor nol, skor yang sama untuk beberapa kategori, atau tidak memiliki kata yang tercakup dalam leksikon juga perlu dijelaskan berdasarkan implementasi sebenarnya.

Label yang dihasilkan secara otomatis digunakan sebagai target klasifikasi apabila sesuai dengan prosedur eksperimen. Oleh karena itu, hasil evaluasi model perlu ditafsirkan berdasarkan kualitas label tersebut. Validasi menggunakan anotasi manual dapat digunakan untuk menilai kesesuaian label otomatis dengan makna sentimen yang sebenarnya.

2.3.2 Ekstraksi Fitur Teks

Ekstraksi fitur dilakukan untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Teknik ekstraksi fitur yang digunakan harus disesuaikan dengan implementasi program penelitian, yaitu *CountVectorizer* dengan pendekatan *Bag of Words* (BoW), *Term Frequency-Inverse Document Frequency* (TF-IDF), atau kombinasi keduanya apabila memang diterapkan.

Pada pendekatan BoW, teks direpresentasikan berdasarkan frekuensi kemunculan kata dalam dokumen. Sementara itu, TF-IDF memberikan bobot pada kata berdasarkan frekuensi kemunculannya dalam dokumen dan distribusinya pada keseluruhan dokumen. Teknik yang benar-benar digunakan beserta parameter pentingnya perlu dicantumkan secara konsisten pada metode, hasil, dan pembahasan.

2.4 Klasifikasi dan Evaluasi Model

Klasifikasi sentimen dilakukan menggunakan dua algoritma, yaitu *Support Vector Machine* (SVM) dan *Random Forest*. Dataset yang sama digunakan untuk kedua algoritma dengan pembagian data latih dan data uji sebesar 70:30. Dari total 1.413 komentar, sebanyak 989 komentar digunakan sebagai data latih dan 424 komentar sebagai data uji.

2.4.1 Support Vector Machine (SVM)

Support Vector Machine merupakan algoritma klasifikasi yang menentukan batas pemisah optimal (*hyperplane*) untuk membedakan kelas data. Pada klasifikasi teks, SVM dapat digunakan untuk mempelajari pola representasi numerik komentar dan memprediksi kategori sentimennya. Jenis kernel dan parameter yang digunakan dalam penelitian perlu dilaporkan berdasarkan konfigurasi model yang sebenarnya [10].

2.4.2 Random Forest

Random Forest merupakan algoritma klasifikasi berbasis *ensemble* yang menggabungkan prediksi dari sejumlah pohon keputusan (*decision trees*) untuk menentukan kelas akhir. Algoritma ini digunakan sebagai pembandingan SVM dalam mengklasifikasikan sentimen

komentar. Jumlah pohon, kedalaman pohon, dan parameter lain yang digunakan perlu dicantumkan berdasarkan implementasi penelitian [11].

2.4.3 Evaluasi Kinerja Model

Evaluasi dilakukan menggunakan *confusion matrix* dan metrik akurasi, *precision*, *recall*, serta *F1-score*. Metrik tersebut digunakan untuk mengukur kinerja klasifikasi secara keseluruhan dan pada setiap kelas sentimen.

Akurasi menunjukkan proporsi prediksi yang benar terhadap seluruh data uji. *Precision* mengukur proporsi prediksi suatu kelas yang benar, sedangkan *recall* mengukur proporsi data aktual suatu kelas yang berhasil dikenali oleh model. *F1-score* merupakan rata-rata harmonis antara *precision* dan *recall*. Adapun *support* menunjukkan jumlah data aktual pada masing-masing kelas.

Nilai metrik kedua algoritma dibandingkan untuk mengetahui perbedaan kinerja SVM dan Random Forest pada dataset yang digunakan. Interpretasi hasil mempertimbangkan distribusi kelas dan kualitas label sentimen sehingga keunggulan model tidak hanya ditentukan berdasarkan nilai akurasi.

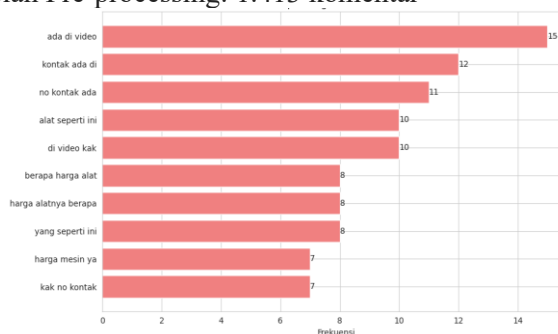
3. HASIL DAN PEMBAHASAN

3.1 Hasil Pre-processing

Dataset awal yang diperoleh melalui proses pengumpulan komentar YouTube berjumlah 1.456 komentar. Setelah melalui tahapan *pre-processing*, jumlah data yang digunakan dalam analisis menjadi 1.413 komentar. Dengan demikian, terdapat pengurangan sebanyak 43 komentar. Pengurangan ini perlu dijelaskan berdasarkan proses yang benar-benar dilakukan, misalnya penghapusan komentar kosong, duplikat, atau komentar yang tidak memenuhi kriteria analisis. Penjelasan mengenai penyebab pengurangan data diperlukan agar proses pengolahan dataset dapat ditelusuri dan direplikasi.

Jumlah Dataset Sebelum Pre-processing: 1.456 komentar

Jumlah Dataset Setelah Pre-processing: 1.413 komentar



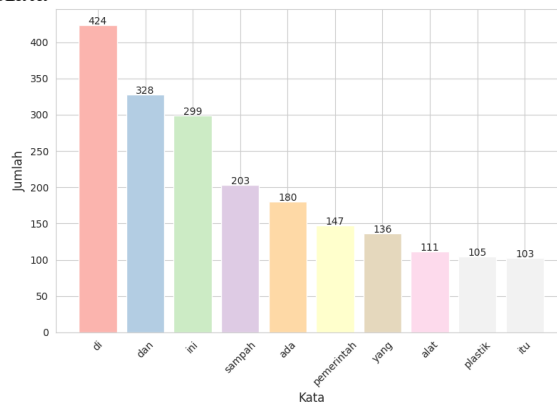
Gambar 2. Visualisasi Top 3-gram Sebelum Pre-processing

Berdasarkan visualisasi *top 3-gram* sebelum *pre-processing*, frasa yang paling sering muncul adalah “ada di video”, “kontak ada di”, dan “no kontak ada”. Kemunculan frasa-frasa tersebut mengindikasikan bahwa sebagian komentar membahas informasi kontak narasumber atau informasi yang berkaitan dengan alat yang ditampilkan dalam video. Namun, interpretasi ini perlu dipahami berdasarkan konteks komentar secara keseluruhan karena frekuensi kemunculan frasa saja belum cukup untuk memastikan maksud setiap pengguna. Visualisasi ini memberikan gambaran awal mengenai pola penggunaan kata dalam komentar sebelum dilakukan pembersihan data.



membantu menyederhanakan representasi teks dan mengurangi kemunculan sejumlah kata umum. Meskipun demikian, visualisasi *wordcloud* hanya menunjukkan frekuensi relatif kata, bukan polaritas sentimen. Dengan demikian, kemunculan kata tertentu tidak secara langsung menunjukkan bahwa komentar yang memuat kata tersebut bersifat positif atau negatif.

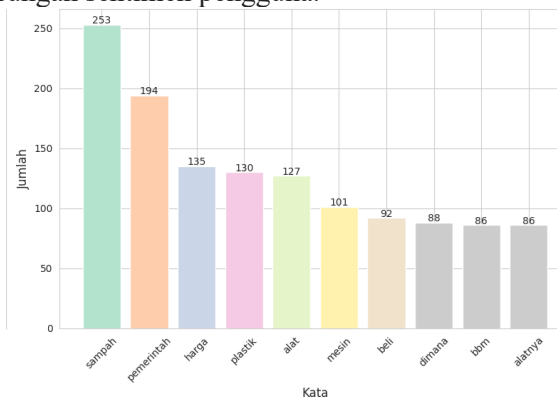
Analisis Frekuensi Kata



Gambar 6. Frekuensi Kata Setelah Pre-processing

Berdasarkan hasil penghitungan frekuensi kata setelah *pre-processing*, kata “sampah” memiliki frekuensi tertinggi, yaitu sebanyak 253 kali. Kata berikutnya adalah “pemerintah” sebanyak 194 kali, “harga” sebanyak 135 kali, “plastik” sebanyak 130 kali, dan “alat” sebanyak 127 kali. Distribusi tersebut memperlihatkan bahwa kata-kata yang berkaitan dengan sampah, pemerintah, harga, plastik, serta alat atau teknologi cukup sering muncul dalam kumpulan komentar yang dianalisis.

Dominasi kata “sampah” dan “plastik” menunjukkan keterkaitan data dengan topik utama penelitian. Sementara itu, kemunculan kata “pemerintah” dan “harga” mengindikasikan bahwa pembahasan dalam komentar juga mencakup aspek kebijakan dan ekonomi. Kata “alat” dapat berkaitan dengan teknologi pengolahan yang ditampilkan dalam video. Namun, interpretasi tersebut tetap perlu memperhatikan konteks kalimat karena satu kata dapat digunakan dalam berbagai maksud. Frekuensi kata juga tidak dapat digunakan secara langsung untuk menyimpulkan kecenderungan sentimen pengguna.



Gambar 7. Frekuensi Kata Sebelum Pre-processing

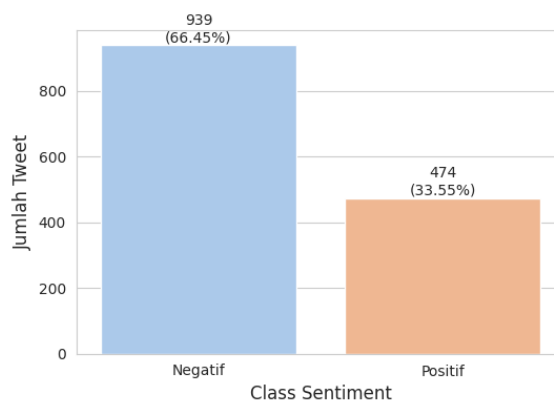
Berdasarkan Gambar 7, kata “sampah” memiliki frekuensi kemunculan tertinggi, yaitu sebanyak 253 kali, diikuti kata “pemerintah” sebanyak 194 kali, “harga” sebanyak 135 kali, “plastik” sebanyak 130 kali, dan “alat” sebanyak 127 kali. Selain itu, kata “mesin” muncul sebanyak 101 kali, “BBM” sebanyak 92 kali, “dimana” sebanyak 88 kali, serta kata “bahan” dan “adanya” masing-masing sebanyak 86 kali. Hasil visualisasi ini menunjukkan distribusi frekuensi kata yang paling sering muncul dalam kumpulan komentar YouTube yang dianalisis.

Dominasi kata “sampah” dan “plastik” memperlihatkan keterkaitan komentar dengan topik utama penelitian, yaitu pengolahan sampah plastik. Kemunculan kata “pemerintah” menunjukkan bahwa pembahasan dalam komentar juga berkaitan dengan peran pemerintah, sedangkan kata “harga” mengindikasikan adanya pembahasan mengenai aspek ekonomi. Sementara itu, kata “alat”, “mesin”, “bbm”, dan “bahan” menunjukkan bahwa komentar juga memuat istilah yang berkaitan dengan teknologi dan proses pengolahan sampah menjadi bahan bakar.

Apabila dibandingkan dengan visualisasi frekuensi kata setelah *pre-processing*, grafik ini dapat digunakan untuk mengamati perubahan distribusi kata sebelum dan sesudah pembersihan data. Perbandingan tersebut perlu menggunakan kriteria penghitungan yang sama agar perubahan frekuensi dapat ditafsirkan secara tepat. Selain itu, frekuensi kata hanya menunjukkan seberapa sering kata muncul dalam dataset dan tidak secara langsung menunjukkan sentimen positif atau negatif. Oleh karena itu, interpretasi mengenai dukungan atau penolakan pengguna terhadap pengolahan sampah plastik menjadi bahan bakar perlu didasarkan pada konteks komentar dan hasil klasifikasi sentimen.

3.2 Hasil Pelabelan Data (Lexicon-based)

Hasil pelabelan menggunakan pendekatan berbasis leksikon menunjukkan bahwa dataset terdiri atas dua kategori sentimen, yaitu negatif dan positif. Sebanyak 939 komentar dikategorikan sebagai sentimen negatif, sedangkan 474 komentar dikategorikan sebagai sentimen positif.



Gambar 8. Jumlah Analisis Sentimen

Berdasarkan hasil pelabelan sentimen, terdapat 939 komentar (66,45%) yang termasuk dalam kategori sentimen negatif dan 474 komentar (33,55%) yang termasuk dalam kategori sentimen positif. Hasil tersebut menunjukkan bahwa label negatif lebih banyak ditemukan daripada label positif dalam dataset yang dianalisis.

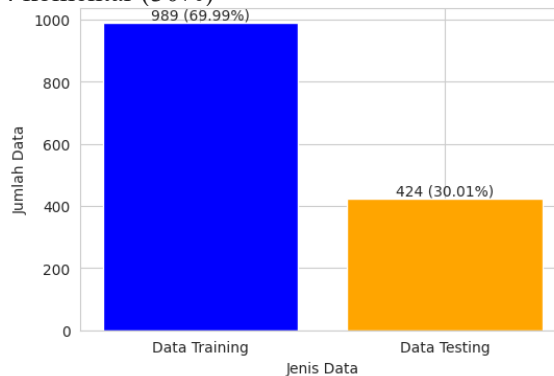
Dominasi label negatif ini menggambarkan distribusi hasil pelabelan berbasis leksikon pada komentar dari video yang menjadi objek penelitian. Namun, hasil tersebut tidak dapat langsung ditafsirkan sebagai bukti bahwa mayoritas masyarakat menolak pengolahan sampah plastik menjadi bahan bakar. Pelabelan otomatis berbasis leksikon berpotensi mengalami kesalahan ketika menghadapi negasi, ironi, sarkasme, atau penggunaan kata yang maknanya bergantung pada konteks. Oleh karena itu, hasil pelabelan perlu dipahami sesuai dengan metode yang digunakan dan tidak disamakan dengan penilaian sentimen oleh anotator manusia, kecuali telah dilakukan validasi terhadap label tersebut. Selain itu, metode penelitian perlu menjelaskan nama dan versi leksikon, cara menghitung skor sentimen, ambang batas penentuan kelas, serta perlakuan terhadap komentar dengan skor nol atau skor yang tidak menunjukkan kecenderungan sentimen yang jelas. Penjelasan tersebut diperlukan agar proses pelabelan dapat direplikasi dan kualitas label yang digunakan dalam pelatihan model dapat dinilai.

3.2.1 Pembagian Data (Splitting Data)

Jumlah data latih dan data uji dalam penelitian ini adalah sebagai berikut.

Jumlah Data Latih: 989 komentar (70%)

Jumlah Data Uji: 424 komentar (30%)



Gambar 9. Jumlah Data Training dan Testing

Dataset yang berjumlah 1.413 komentar dibagi menjadi 989 data *training* (70%) dan 424 data *testing* (30%). Data *training* digunakan untuk membangun model klasifikasi, sedangkan data *testing* digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan data yang tidak digunakan pada proses pelatihan. Pembagian data tersebut memungkinkan perbandingan kinerja algoritma Support Vector Machine (SVM) dan Random Forest pada data uji yang sama. Agar hasil evaluasi dapat direplikasi, naskah perlu menyebutkan metode pembagian data yang digunakan, nilai *random state* atau *random seed* apabila diterapkan, serta apakah pembagian dilakukan secara *stratified* untuk mempertahankan proporsi kelas sentimen. Informasi tersebut tidak boleh ditambahkan sebagai fakta sebelum diperiksa pada implementasi penelitian.

3.3 Evaluasi Analisis Sentimen

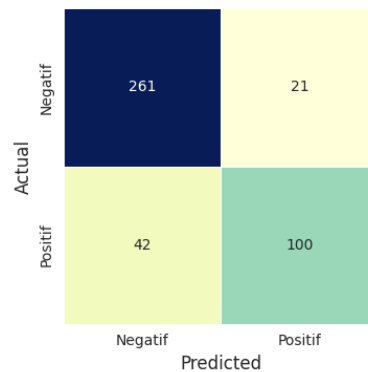
Evaluasi dilakukan untuk membandingkan kinerja algoritma Support Vector Machine (SVM) dan Random Forest dalam mengklasifikasikan komentar ke dalam kategori sentimen negatif dan positif. Evaluasi menggunakan *confusion matrix*, akurasi, *precision*, *recall*, dan F1-score. Penggunaan beberapa metrik diperlukan karena akurasi saja belum sepenuhnya menggambarkan kemampuan model dalam mengenali masing-masing kelas sentimen.

3.3.1 Hasil Evaluasi Support Vector Machine (SVM)

Berdasarkan *confusion matrix*, model SVM berhasil mengklasifikasikan 261 data negatif dan 100 data positif dengan benar. Sementara itu, terdapat 21 data negatif yang salah diklasifikasikan sebagai positif dan 42 data positif yang salah diklasifikasikan sebagai negatif. Dengan demikian, model mengklasifikasikan 361 dari 424 data uji secara benar.

Hasil tersebut menghasilkan akurasi sebesar 85,1%. Berdasarkan perhitungan dari *confusion matrix*, model memperoleh *precision* sekitar 0,861 dan *recall* sekitar 0,926 pada kelas negatif. Pada kelas positif, model memperoleh *precision* sekitar 0,826 dan *recall* sekitar 0,704. Nilai *weighted F1-score* yang dilaporkan adalah 0,848.

Hasil tersebut memperlihatkan bahwa model SVM memiliki kemampuan yang lebih tinggi dalam mengenali kelas negatif dibandingkan kelas positif. Dari 142 data positif aktual pada data uji, sebanyak 100 data berhasil diklasifikasikan dengan benar, sedangkan 42 data lainnya salah diklasifikasikan sebagai negatif. Dengan demikian, meskipun akurasi keseluruhan mencapai 85,1%, kemampuan model dalam mendeteksi sentimen positif masih perlu diperhatikan.



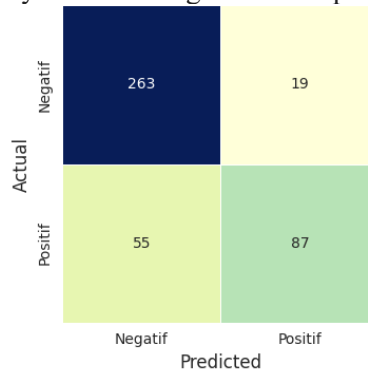
Gambar 10. SVM Confusion Matrix

3.3.2 Hasil Evaluasi Random Forest

Model Random Forest berhasil mengklasifikasikan 263 data negatif dan 87 data positif dengan benar. Namun, terdapat 19 data negatif yang salah diklasifikasikan sebagai positif dan 55 data positif yang salah diklasifikasikan sebagai negatif. Secara keseluruhan, model mengklasifikasikan 350 dari 424 data uji secara benar.

Hasil evaluasi menunjukkan bahwa Random Forest memperoleh akurasi sebesar 82,5% dan *weighted F1-score* sebesar 0,818. Berdasarkan perhitungan dari *confusion matrix*, model memperoleh *precision* sekitar 0,827 dan *recall* sekitar 0,933 pada kelas negatif. Pada kelas positif, model memperoleh *precision* sekitar 0,821 dan *recall* sekitar 0,613.

Hasil tersebut menunjukkan bahwa Random Forest mampu mengenali kelas negatif dengan baik, tetapi masih mengalami kesulitan dalam mendeteksi sebagian komentar positif. Dari 142 data positif aktual pada data uji, sebanyak 87 data diklasifikasikan dengan benar, sedangkan 55 data salah diklasifikasikan sebagai negatif. Oleh karena itu, penilaian kinerja Random Forest tidak cukup hanya berdasarkan jumlah data negatif yang berhasil diklasifikasikan, tetapi juga perlu mempertimbangkan kemampuannya dalam mengenali kelas positif.



Gambar 11. Random Forest Confusion Matrix

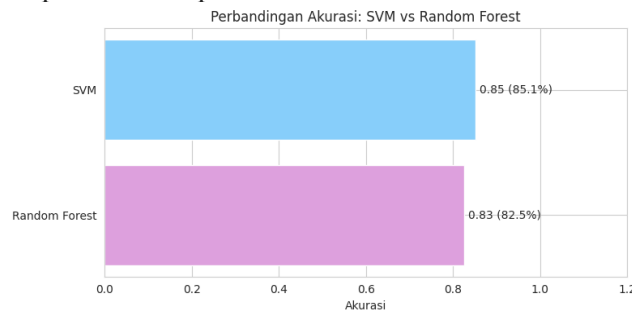
3.3.3 Perbandingan Akurasi SVM dan Random Forest

Hasil perbandingan menunjukkan bahwa SVM memperoleh akurasi sebesar 85,1%, sedangkan Random Forest memperoleh akurasi sebesar 82,5%. Dengan demikian, pada pembagian data dan dataset yang digunakan dalam penelitian ini, SVM memiliki akurasi lebih tinggi daripada Random Forest. Selisih akurasi kedua algoritma adalah sekitar 2,6 poin persentase.

Perbedaan tersebut menunjukkan bahwa SVM menghasilkan jumlah prediksi benar yang lebih banyak pada data uji yang digunakan. SVM mengklasifikasikan 361 komentar secara benar,

sedangkan Random Forest mengklasifikasikan 350 komentar secara benar. Dengan demikian, terdapat selisih 11 prediksi benar antara kedua model.

Meskipun demikian, hasil ini hanya menunjukkan perbandingan kinerja pada konfigurasi eksperimen dan pembagian data yang digunakan. Hasil tersebut belum cukup untuk menyimpulkan bahwa SVM selalu lebih unggul daripada Random Forest pada seluruh dataset komentar YouTube atau pada seluruh permasalahan analisis sentimen.



Gambar 12. Perbandingan Akurasi: SVM vs Random Forest

3.3.4 Classification Report SVM

Model SVM menghasilkan akurasi sebesar 85,1% dengan nilai *weighted F1-score* sebesar 0,848. Berdasarkan *confusion matrix*, kelas negatif memiliki *precision* sekitar 0,861, *recall* sekitar 0,926, dan F1-score sekitar 0,892. Sementara itu, kelas positif memiliki *precision* sekitar 0,826, *recall* sekitar 0,704, dan F1-score sekitar 0,761.

Nilai tersebut memperlihatkan bahwa kemampuan SVM dalam mengenali kelas negatif lebih tinggi dibandingkan kelas positif. Nilai *recall* kelas negatif sebesar 0,926 menunjukkan bahwa sebagian besar komentar yang berlabel negatif berhasil dikenali oleh model. Sebaliknya, *recall* kelas positif sebesar 0,704 menunjukkan bahwa masih terdapat komentar positif yang diklasifikasikan sebagai negatif. Oleh karena itu, performa SVM perlu dipahami berdasarkan karakteristik masing-masing kelas, bukan hanya berdasarkan akurasi keseluruhan.

	precision	recall	f1-score	support
Negatif	0.861	0.926	0.892	282
Positif	0.826	0.704	0.760	142
accuracy	0.851	0.851	0.851	1
macro avg	0.844	0.815	0.826	424
weighted avg	0.850	0.851	0.848	424

Gambar 13. Classification Report for SVM

3.3.5 Classification Report Random Forest

Model Random Forest memperoleh akurasi sebesar 82,5% dengan nilai *weighted F1-score* sebesar 0,818. Berdasarkan *confusion matrix*, kelas negatif memiliki *precision* sekitar 0,827, *recall* sekitar 0,933, dan F1-score sekitar 0,877. Pada kelas positif, model memiliki *precision* sekitar 0,821, *recall* sekitar 0,613, dan F1-score sekitar 0,703.

Hasil tersebut menunjukkan bahwa Random Forest memiliki *recall* kelas negatif yang tinggi, tetapi *recall* kelas positif lebih rendah dibandingkan SVM. Nilai *recall* kelas positif sebesar 0,613 menunjukkan bahwa model hanya berhasil mengenali 87 dari 142 komentar positif pada data uji. Kondisi ini menjelaskan mengapa Random Forest memperoleh akurasi dan *weighted F1-score* yang lebih rendah daripada SVM pada eksperimen ini.

	precision	recall	f1-score	support
Negatif	0.827	0.933	0.877	282
Positif	0.821	0.613	0.702	142
accuracy	0.825	0.825	0.825	1
macro avg	0.824	0.773	0.789	424
weighted avg	0.825	0.825	0.818	424

Gambar 14. Classification Report untuk Random Forest

3.3.6 Ringkasan Perbandingan Kinerja Model

Untuk memperjelas perbandingan, hasil evaluasi kedua algoritma disajikan dalam satu tabel. Nilai *precision*, *recall*, dan *F1-score* per kelas berikut dihitung dari jumlah pada *confusion matrix*. Nilai yang dihitung ditampilkan hingga tiga angka desimal, sedangkan akurasi dan *weighted F1-score* mengikuti hasil yang dilaporkan dalam naskah.

Tabel 1. Perbandingan Kinerja Algoritma SVM dan Random Forest

Metrik	Kelas	SVM	Random Forest
Akurasi	Keseluruhan	85,1%	82,5%
<i>Precision</i>	Negatif	0,861	0,827
<i>Recall</i>	Negatif	0,926	0,933
<i>F1-score</i>	Negatif	0,892	0,877
<i>Precision</i>	Positif	0,826	0,821
<i>Recall</i>	Positif	0,704	0,613
<i>F1-score</i>	Positif	0,761	0,703
<i>Weighted F1-score</i>	Keseluruhan	0,848	0,818
Prediksi benar	Keseluruhan	361	350

Berdasarkan tabel tersebut, SVM memperoleh akurasi, *weighted F1-score*, dan *F1-score* kelas positif yang lebih tinggi daripada Random Forest. Sementara itu, Random Forest memiliki *recall* kelas negatif sedikit lebih tinggi daripada SVM. Temuan ini memperlihatkan bahwa keunggulan suatu algoritma bergantung pada metrik yang digunakan untuk menilai kinerjanya. Dalam konteks penelitian ini, SVM menghasilkan kinerja keseluruhan yang lebih baik berdasarkan akurasi dan *weighted F1-score*, serta lebih baik dalam mengenali kelas positif berdasarkan *recall* dan *F1-score*.

3.4 Pembahasan

Hasil visualisasi *top n-gram* dan *wordcloud* menunjukkan bahwa komentar yang dianalisis memuat kata-kata yang berkaitan dengan topik sampah plastik, seperti “sampah”, “plastik”, “BBM”, dan “Pertamina”. Visualisasi sebelum *pre-processing* juga memperlihatkan kemunculan kata-kata umum, seperti “yang”, “dan”, serta “di”. Setelah dilakukan pembersihan data, visualisasi memperlihatkan kata-kata yang lebih terfokus pada topik penelitian. Hasil ini menunjukkan pentingnya tahap *pre-processing* dalam mempersiapkan data teks sebelum proses klasifikasi. Temuan ini sejalan dengan penelitian Dadgar et al. yang menyatakan bahwa *pre-processing* merupakan tahapan penting dalam *text mining* untuk meningkatkan kualitas fitur yang digunakan dalam proses klasifikasi [12]. Selain itu, Tanesab et al. menjelaskan bahwa proses *tokenizing*, *cleansing*, dan *filtering* merupakan tahapan penting dalam analisis sentimen komentar YouTube [13].

Meskipun demikian, perbedaan visualisasi sebelum dan setelah *pre-processing* perlu ditafsirkan secara hati-hati. Penggunaan 3-gram sebelum *pre-processing* dan 4-gram setelah *pre-processing* menyebabkan kedua visualisasi tidak sepenuhnya sebanding. Oleh karena itu, klaim mengenai perubahan pola frasa sebaiknya didasarkan pada perbandingan dengan parameter *n*-

gram yang sama. Demikian pula, *wordcloud* dan frekuensi kata hanya memberikan informasi mengenai kemunculan kata, bukan secara langsung mengungkap polaritas sentimen. Dengan demikian, hasil visualisasi berfungsi sebagai gambaran karakteristik teks, sedangkan penentuan sentimen dilakukan melalui proses pelabelan dan klasifikasi.

Pada tahap ekstraksi fitur, naskah perlu memastikan konsistensi antara metode yang dijelaskan dan implementasi program. Bagian metode pada naskah awal menyebutkan penggunaan CountVectorizer dengan pendekatan *Bag of Words*, sedangkan bagian pembahasan menyebutkan TF-IDF. Kedua teknik tersebut memiliki mekanisme pembobotan fitur yang berbeda. CountVectorizer merepresentasikan teks berdasarkan jumlah kemunculan kata, sedangkan TF-IDF memberikan bobot berdasarkan frekuensi kata dalam dokumen dan distribusinya pada kumpulan dokumen. Oleh karena itu, metode yang dituliskan sebagai metode penelitian harus disesuaikan dengan kode program yang benar-benar digunakan. Pembahasan mengenai efektivitas TF-IDF dan penelitian terkait [14], [15], [16] tetap relevan apabila TF-IDF memang digunakan dalam eksperimen. Jika implementasi sebenarnya menggunakan CountVectorizer, uraian tersebut harus disesuaikan dengan metode yang diterapkan agar tidak menimbulkan ketidakkonsistenan metodologis.

Berdasarkan hasil pelabelan, sentimen negatif mendominasi dataset dengan persentase 66,45%, sedangkan sentimen positif mencapai 33,55%. Distribusi ini menunjukkan bahwa label negatif lebih banyak dihasilkan daripada label positif pada kumpulan komentar yang dianalisis. Hasil tersebut dapat digunakan untuk menggambarkan kecenderungan label sentimen dalam dataset, tetapi tidak secara otomatis membuktikan bahwa sebagian besar masyarakat menolak pengolahan sampah plastik menjadi bahan bakar. Hal ini karena data hanya berasal dari satu video YouTube dan label sentimen ditentukan menggunakan pendekatan berbasis leksikon. Hasil pelabelan otomatis juga berpotensi dipengaruhi oleh konteks, negasi, ironi, atau sarkasme yang tidak selalu dapat dikenali oleh leksikon. Temuan mengenai penggunaan komentar YouTube untuk mempelajari respons pengguna sejalan dengan penelitian Bhuiyan et al. [17] dan Jadhav [18]. Namun, interpretasi hasil penelitian ini tetap dibatasi pada dataset yang digunakan.

Hasil evaluasi menunjukkan bahwa SVM memperoleh akurasi sebesar 85,1%, sedangkan Random Forest memperoleh akurasi sebesar 82,5%. SVM menghasilkan 361 prediksi benar dari 424 data uji, sedangkan Random Forest menghasilkan 350 prediksi benar. Hasil tersebut memperlihatkan bahwa SVM memberikan kinerja keseluruhan yang lebih baik pada eksperimen ini. Temuan tersebut sejalan dengan penelitian Alkaff et al. yang melaporkan kinerja SVM dalam analisis sentimen komentar YouTube [16]. Hasil penelitian ini juga dapat dibandingkan dengan penelitian Tanesab et al. yang memperoleh akurasi sebesar 84% menggunakan SVM pada data komentar YouTube [13], serta penelitian Jamil et al. mengenai penerapan algoritma *machine learning* dalam klasifikasi sentimen komentar YouTube [19]. Namun, perbandingan nilai akurasi antarpelelitian harus memperhatikan perbedaan dataset, proses pelabelan, metode ekstraksi fitur, pembagian data, dan konfigurasi model. Oleh karena itu, perbedaan angka akurasi tidak dapat langsung dijadikan bukti bahwa satu penelitian memiliki metode yang secara umum lebih baik daripada penelitian lainnya.

Analisis *confusion matrix* memberikan penjelasan lebih rinci mengenai perbedaan kinerja kedua algoritma. SVM berhasil mengklasifikasikan 100 dari 142 komentar positif dengan benar, sedangkan Random Forest berhasil mengklasifikasikan 87 komentar positif. Dengan demikian, *recall* kelas positif SVM mencapai sekitar 0,704, sedangkan Random Forest mencapai sekitar 0,613. Hasil ini menunjukkan bahwa SVM lebih baik dalam menemukan komentar positif pada data uji yang digunakan. Sebaliknya, Random Forest mengklasifikasikan 263 komentar negatif dengan benar, sedikit lebih banyak daripada SVM yang mengklasifikasikan 261 komentar negatif dengan benar. Nilai *recall* kelas negatif Random Forest sebesar 0,933 juga sedikit lebih tinggi dibandingkan SVM sebesar 0,926.

Perbedaan tersebut menunjukkan bahwa evaluasi model perlu mempertimbangkan kemampuan klasifikasi pada setiap kelas. SVM memiliki F1-score kelas positif sebesar 0,761, sedangkan Random Forest sebesar 0,703. Pada kelas negatif, SVM memperoleh F1-score sebesar 0,892 dan Random Forest sebesar 0,877. Dengan demikian, SVM menunjukkan keseimbangan yang lebih baik berdasarkan F1-score kedua kelas pada dataset ini. Temuan tersebut dapat dibandingkan dengan penelitian Zakiyah et al. yang melaporkan kinerja SVM dalam analisis opini publik di media sosial [20]. Selain itu, penelitian Alkaff et al. membahas penggunaan SVM pada komentar YouTube [16]. Kendati demikian, hasil penelitian ini tidak berarti bahwa SVM akan selalu lebih unggul pada seluruh dataset atau topik analisis sentimen.

Secara keseluruhan, penelitian ini menunjukkan bahwa SVM menghasilkan akurasi dan *weighted F1-score* yang lebih tinggi dibandingkan Random Forest pada dataset komentar YouTube mengenai pengolahan sampah plastik yang digunakan. Hasil tersebut mendukung pembahasan mengenai penggunaan representasi fitur teks dan algoritma klasifikasi dalam analisis sentimen, sebagaimana dibahas dalam penelitian Dadgar et al. [12], Shahmirzadi et al. [14], serta Eminagaoglu [21]. Namun, kesimpulan mengenai kombinasi metode tertentu harus mengikuti teknik ekstraksi fitur yang benar-benar digunakan dalam implementasi penelitian. Jika kode program menggunakan TF-IDF, hasil penelitian dapat dibahas dalam konteks representasi fitur TF-IDF dan klasifikasi SVM. Jika menggunakan CountVectorizer, uraian tersebut perlu disesuaikan dengan representasi *Bag of Words*.

Penelitian ini memiliki keterbatasan yang perlu diperhatikan. Pertama, data hanya berasal dari komentar pada satu video YouTube sehingga hasilnya belum dapat digeneralisasikan untuk menggambarkan opini masyarakat secara keseluruhan. Kedua, label sentimen dihasilkan menggunakan pendekatan berbasis leksikon sehingga validitas label bergantung pada karakteristik leksikon dan konteks bahasa komentar. Ketiga, hasil perbandingan model diperoleh dari satu pembagian data latih dan data uji dengan rasio 70:30. Oleh sebab itu, hasil evaluasi menggambarkan kinerja model pada konfigurasi eksperimen yang digunakan dan belum membuktikan kestabilan kinerja pada pembagian data yang berbeda. Keterbatasan ini dapat menjadi pertimbangan untuk penelitian selanjutnya melalui validasi label, penggunaan data dari beberapa video, analisis kesalahan klasifikasi, serta pengujian dengan skema evaluasi tambahan.

4. KESIMPULAN

1. Penelitian ini telah menganalisis sentimen komentar YouTube pada video “Sulap Sampah Plastik Jadi Bensin RON 92 & Solar. Solusi Cerdas Atasi Limbah” yang dipublikasikan pada kanal Tanilink TV menggunakan algoritma Support Vector Machine (SVM) dan Random Forest. Setelah melalui tahapan *pre-processing*, sebanyak 1.413 komentar digunakan dalam proses analisis dan evaluasi klasifikasi sentimen.
2. Hasil pelabelan berbasis leksikon menunjukkan bahwa sebanyak 939 komentar (66,45%) termasuk dalam kategori sentimen negatif dan 474 komentar (33,55%) termasuk dalam kategori sentimen positif. Hasil ini menunjukkan bahwa label sentimen negatif lebih dominan daripada label positif pada komentar dalam video yang dianalisis. Temuan tersebut tidak dapat digeneralisasikan untuk menggambarkan opini publik secara keseluruhan karena penelitian hanya menggunakan komentar dari satu video YouTube dan label sentimen diperoleh secara otomatis menggunakan pendekatan berbasis leksikon.
3. Hasil evaluasi menunjukkan bahwa SVM memperoleh akurasi sebesar 85,1% dan *weighted F1-score* sebesar 0,848, sedangkan Random Forest memperoleh akurasi sebesar 82,5% dan *weighted F1-score* sebesar 0,818. Berdasarkan kedua metrik tersebut, SVM menunjukkan kinerja keseluruhan yang lebih baik pada dataset yang digunakan.

Meskipun demikian, Random Forest memiliki *recall* kelas negatif sedikit lebih tinggi dibandingkan SVM, sedangkan SVM memiliki *recall* dan F1-score kelas positif yang lebih tinggi. Hasil ini menunjukkan bahwa perbandingan kinerja kedua algoritma perlu mempertimbangkan performa pada masing-masing kelas, bukan hanya akurasi keseluruhan.

4. Penelitian ini memiliki keterbatasan, yaitu sumber data hanya berasal dari satu video YouTube, pelabelan sentimen dilakukan secara otomatis menggunakan pendekatan berbasis leksikon, dan evaluasi model menggunakan pembagian data latih dan data uji dengan rasio 70:30. Apabila evaluasi hanya dilakukan menggunakan satu kali pembagian data, hasil kinerja model belum dapat menggambarkan kestabilan performa pada pembagian data yang berbeda. Selain itu, dataset yang digunakan hanya mencakup dua kelas sentimen, yaitu positif dan negatif. Penelitian selanjutnya disarankan untuk menggunakan komentar dari beberapa video, memvalidasi hasil pelabelan otomatis, serta menerapkan pengujian tambahan seperti *cross-validation* untuk memperoleh gambaran kinerja model yang lebih menyeluruh.

5. SARAN

1. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dengan mengambil komentar dari beberapa video YouTube yang membahas pengolahan sampah plastik menjadi bahan bakar. Penggunaan sumber data yang lebih beragam diharapkan dapat memberikan gambaran yang lebih luas mengenai kecenderungan sentimen pada topik tersebut. Apabila menggunakan platform media sosial lain, karakteristik data dari setiap platform perlu diperhatikan agar hasil analisis dapat dibandingkan secara tepat.
2. Penelitian berikutnya dapat menambahkan kategori sentimen netral sehingga klasifikasi tidak hanya terbatas pada sentimen positif dan negatif. Selain itu, validasi terhadap hasil pelabelan otomatis berbasis leksikon perlu dilakukan, misalnya melalui pemeriksaan manual oleh anotator, untuk mengetahui kesesuaian label dengan konteks komentar. Langkah ini diharapkan dapat meningkatkan keandalan data berlabel yang digunakan dalam pelatihan dan evaluasi model.
3. Penelitian selanjutnya dapat membandingkan algoritma Support Vector Machine (SVM) dan Random Forest dengan algoritma klasifikasi lainnya, seperti Naïve Bayes, Logistic Regression, IndoBERT, atau model *deep learning* lainnya. Perbandingan tersebut perlu menggunakan dataset, pembagian data, dan metrik evaluasi yang konsisten agar perbedaan kinerja antarmodel dapat dianalisis secara lebih objektif.
4. Untuk memperoleh hasil evaluasi yang lebih andal, penelitian mendatang dapat menerapkan optimasi parameter (*hyperparameter tuning*) dan pengujian tambahan, seperti *cross-validation*, guna mengetahui kestabilan kinerja model pada pembagian data yang berbeda. Teknik penyeimbangan data juga dapat dipertimbangkan apabila ditemukan ketidakseimbangan kelas yang memengaruhi hasil klasifikasi. Penerapan teknik tersebut perlu dievaluasi berdasarkan metrik per kelas, khususnya *precision*, *recall*, dan F1-score, agar peningkatan kinerja tidak hanya terlihat pada akurasi keseluruhan.

DAFTAR PUSTAKA

- [1] Adilla, R. E., Huda, M., Aziz, M., & Suadaa, L. H. (2024). Aspect-based sentiment analysis of transportation electrification opinions on YouTube comment data. *Jurnal*

- Aplikasi Statistika & Komputasi Statistik*, 16(2), 140–157.
<https://doi.org/10.34123/jurnalasks.v16i2.790>
- [2] Suryani, A. S. (2016). Persepsi masyarakat dan analisis willingness to pay terhadap kebijakan kantong plastik berbayar studi di Jakarta dan Bandung. *Kajian*, 21(4), 359–376.
- [3] Putri, N. W., Fitriyani, F., Rahmalber, T. V., Falikha, D. D., Tafsia, S. I., & Setiawati, S. D. (2023). Pengetahuan, sikap dan partisipasi masyarakat dalam penerapan kebijakan pengurangan kantong plastik sebagai upaya mengurangi sampah plastik di Kota Padang. *Jurnal Kesehatan Lingkungan Indonesia*, 22(3), 259–267.
<https://doi.org/10.14710/jkli.22.3.259-267>
- [4] Koto, F., & Rahmaningtyas, G. Y. (2017). Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. In *2017 International Conference on Asian Language Processing (IALP)* (pp. 391–394). IEEE.
- [5] Arsyah, U. I., Pratiwi, M., & Muhammad, A. (2024). Twitter sentiment analysis of public space opinions using SVM and TF-IDF methods. *The Indonesian Journal of Computer Science*, 13(1). <https://doi.org/10.33022/ijcs.v13i1.3594>
- [6] Saputra, A. N. A., Saputro, R. E., & Saputra, D. I. S. (2025). Enhancing sentiment analysis accuracy using SVM and slang word normalization on YouTube comments. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 9(2), 687–699.
<https://doi.org/10.33395/sinkron.v9i2.14613>
- [7] Satriya, R. D., Suarjaya, I. M. A. D., & Pratama, I. P. A. E. (2022). Sentiment analisis antusias masyarakat terhadap sampah plastik dengan menggunakan metode support vector machine (SVM). *Jurnal Ilmiah Teknologi dan Komputer*, 3(1), 880–893.
- [8] Chairat, A. S. N., Rizal, R., & Himawan, H. (2025). Comparative sentiment analysis of YouTube comments on Indonesia's electric vehicle incentive policy using TF-IDF and IndoBERTweet models. *Jurnal Teknik Informatika*, 6(6), 5988–6002.
<https://doi.org/10.52436/1.jutif.2025.6.6.5499>
- [9] Doholio, N. P., Hasan, I. K., & Abdussamad, S. N. (2025). Comparison of Word2Vec and CountVectorizer with mutual information in support vector machine (SVM) for public sentiment analysis. *Journal of Mathematics and Computer Statistics*, 8(1), 12–23.
- [10] Parameswari, S. D., Lubis, M., Suakanto, S., Ramadhan, Y. Z., Amanah, R. N., & Dila, R. A. (2025). Studi perbandingan Naïve Bayes dan support vector machine (SVM) dalam analisis sentimen pengguna Metaverse. *Jurnal Teknologi dan Manajemen Industri Terapan*, 4(3), 1059–1065. <https://doi.org/10.55826/jtmit.v4i3.1122>
- [11] Salman, A. H., & Al-Jawher, W. A. M. (2024). Performance comparison of support vector machines, AdaBoost, and random forest for sentiment text analysis and classification. *J. Port Sci. Res.*, 7(3), 300–311.
- [12] Dadgar, S. M. H., Araghi, M. S., & Farahani, M. M. (2016). A novel text mining approach based on TF-IDF and support vector machine for news classification. In *2016*

- IEEE International Conference on Engineering and Technology (ICETECH)* (pp. 112–116). IEEE.
- [13] Tanesab, F. I., Sembiring, I., & Purnomo, H. D. (2017). Sentiment analysis model based on YouTube comment using support vector machine. *International Journal of Computer Science and Software Engineering*, 6(8), 180–185.
- [14] Shahmirzadi, O., Lugowski, A., & Younge, K. (2019). Text similarity in vector space models: A comparative study. In *2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 659–666). IEEE. <https://doi.org/10.1109/ICMLA.2019.00120>
- [15] Samant, S. S., Murthy, N. L. B., & Malapati, A. (2019). Improving term weighting schemes for short text classification in vector space model. *IEEE Access*, 7, 166578–166592. <https://doi.org/10.1109/ACCESS.2019.2953918>
- [16] Alkaff, M., Baskara, A. R., & Wicaksono, Y. H. (2020). Sentiment analysis of Indonesian movie trailer on YouTube using delta TF-IDF and SVM. In *2020 Fifth International Conference on Informatics and Computing (ICIC)* (pp. 1–5). IEEE. <https://doi.org/10.1109/ICIC50835.2020.9288579>
- [17] Bhuiyan, H., Ara, J., Bardhan, R., & Islam, M. R. (2017). Retrieving YouTube video by sentiment analysis on user comment. In M. F. Ahmad Fauzi (Ed.), *Proceedings of the 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA 2017)* (pp. 474–478). IEEE. <https://doi.org/10.1109/ICSIPA.2017.8120658>
- [18] Jadhav, S. S. (2024). *Sentiment analysis of user comments for a YouTube educational videos* [Master's thesis, National College of Ireland]. NORMA@NCI Library. <https://norma.ncirl.ie/8712/>
- [19] Jamil, M., Hadiyanto, H., & Sanjaya, R. (2024). Sentiment analysis: Classifying public comments on YouTube in disaster management simulation in Indonesia using Naïve Bayes and support vector machine. *Ingénierie des Systèmes d'Information*, 29(2), 437–446. <https://doi.org/10.18280/isi.290205>
- [20] Zakiyah, N. P., Umam, K., & Mahfudh, A. A. (2025). Public opinion on the MBG program: Comparative evaluation of InSet and VADER lexicon labeling using SVM on platform X. *Journal of Applied Informatics and Computing*, 9(6), 3937–3944.
- [21] Eminagaoglu, M. (2022). A new similarity measure for vector space models in text classification and information retrieval. *Journal of Information Science*, 48(4), 463–476. <https://doi.org/10.1177/0165551520968055>