Kajian Ekonomi dan Akuntansi Terapan
Volume. 2 Nomor. 1 Maret 2025
e-ISSN : 3046-9414, p-ISSN : 3046-8736, Hal. 224-234
DOI: https://doi.org/10.61132/keat.v2i2.1208
Available online at: https://ejournal.areai.or.id/index.php/KEAT

Analisis Sentimen Masyarakat terhadap Aktivitas Pertambangan di Raja Ampat Menggunakan *Support Vector Machine* dan *Naïve Bayes* dengan Teknik SMOTE

Fitri Dwianasari[1*], Rohmah Diah Yani[2], Karlina Novianto Laksono[3], Nurhafillah Mujaliza[4], Riza Fahlapi[5]
[1-5]Program Studi Akuntansi, Fakultas Ekonomi dan Bisnis, Universitas Bina Sarana Informatika
fdwiana6@gmail.com[1], rhmdyanii@gmail.com[2], karlinnalaksono@gmail.com[3], nmujaliza13@gmail.com[4], riza.rzf@bsi.ac.id[5]

Alamat: Jl. Keramat Raya No. 98, Senen, Jakarta Pusat
*Korespondensi penulis: fdwiana6@gmail.com**

Abstract*: Mining activities in the Raja Ampat area have sparked various public reactions, both supportive and critical, particularly on social media platforms such as Twitter. This study aims to analyze public sentiment regarding the mining operations by employing two classification algorithms. A total of 500 tweets related to Raja Ampat were collected from the X platform, and after data cleaning, 168 were identified as positive sentiments and 303 as negative. Sentiment analysis was conducted using text mining techniques by comparing two algorithms: Support Vector Machine (SVM) and Naïve Bayes. To address the issue of data imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The analysis results showed that SVM achieved an accuracy of 80%, outperforming Naïve Bayes, which reached only 68%. This indicates that SVM performed better in classifying sentiment. Additionally, the application of SMOTE effectively enhanced both algorithms' abilities to detect positive sentiment, as reflected in the precision, recall, and F1-score metrics. For SVM, precision reached 85%, recall 80%, and F1-score 80%, while Naïve Bayes recorded a precision and recall of 69%, and an F1-score of 68%.*

Keywords: *sentiment analysis, Raja Ampat, Twitter, SVM, Naïve Bayes, SMOTE*

Abstrak: Kegiatan pertambangan yang terjadi di wilayah Raja Ampat menimbulkan berbagai tanggapan dari masyarakat, baik yang mendukung maupun yang menentang, terutama terlihat dalam diskusi di media sosial seperti Twitter. Penelitian ini bertujuan untuk menganalisis persepsi publik terhadap aktivitas tersebut melalui pendekatan analisis sentimen menggunakan dua jenis algoritma. Sebanyak 500 tweet terkait Raja Ampat dikumpulkan melalui platform X, dan setelah melalui tahap pembersihan data, sebanyak 168 di antaranya diklasifikasikan sebagai sentimen positif dan 303 sebagai sentimen negatif. Proses analisis dilakukan dengan metode Text Mining, menggunakan perbandingan antara algoritma Support Vector Machine (SVM) dan Naïve Bayes. Untuk mengatasi ketidakseimbangan data, digunakan teknik optimasi Synthetic Minority Over-sampling Technique (SMOTE). Berdasarkan hasil evaluasi, algoritma SVM menunjukkan akurasi sebesar 80%, sementara Naïve Bayes hanya mencapai 68%. Ini menunjukkan bahwa SVM memiliki performa yang lebih unggul. Selain itu, penggunaan SMOTE terbukti meningkatkan efektivitas kedua algoritma dalam mendeteksi sentimen positif, sebagaimana ditunjukkan oleh nilai precision, recall, dan F1-Score. Untuk SVM, precision tercatat sebesar 85%, recall 80%, dan F1-Score 80%; sedangkan Naïve Bayes memperoleh precision dan recall sebesar 69%, serta F1-Score sebesar 68%.

Kata kunci: Analisis Sentimen, Raja Ampat, Twitter, SVM, Naive Bayes, SMOTE

1. LATAR BELAKANG

Raja Ampat, yang selama ini dikenal sebagai salah satu wilayah konservasi laut paling bernilai secara global, kini tengah menghadapi ancaman nyata akibat aktivitas pertambangan nikel yang kontroversial. Berdasarkan laporan terbaru, kegiatan penambangan di beberapa pulau seperti Gag, Kawe, dan Manuran telah menyebabkan kerusakan ekosistem lebih dari 500

Received: Mei 24, 2025; Revised: Juni 05, 2025; Accepted: Juni 19, 2025; Published: Juni 21, 2025

hektar hutan dan tutupan vegetasi. Kondisi ini memicu kekhawatiran dari berbagai pihak, termasuk masyarakat lokal dan kelompok pemerhati lingkungan.

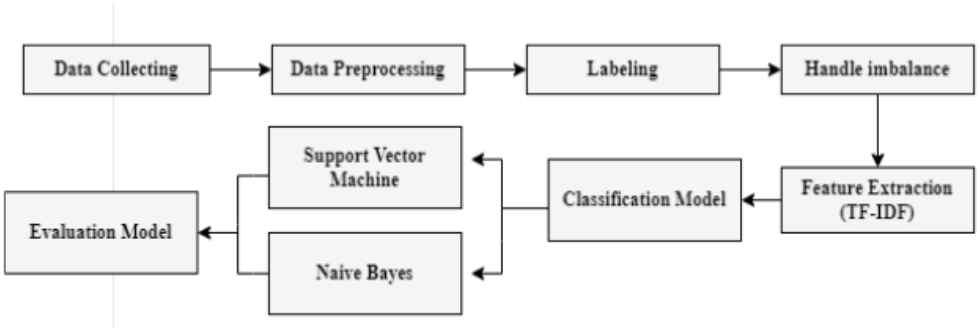
Media sosial, khususnya platform Twitter, menjadi saluran utama bagi publik untuk menyampaikan opini dan sikap mereka terkait isu ini. Melalui pendekatan analisis sentimen terhadap unggahan di Twitter, dapat diperoleh gambaran yang lebih komprehensif mengenai pandangan masyarakat terhadap penambangan di wilayah konservasi tersebut. Berbagai studi sebelumnya telah memanfaatkan algoritma seperti Support Vector Machine (SVM) dan Naive Bayes dalam melakukan klasifikasi sentimen di berbagai konteks. Misalnya, penelitian oleh Armanda dan Tobing (2024) menerapkan SVM untuk mengkaji sentimen publik terhadap pembatalan Piala Dunia U-20 di Indonesia, sementara Tullah dan Akbar (2023) menggunakan Naive Bayes untuk menganalisis sentimen terhadap CPNS di media sosial.

Salah satu tantangan utama dalam analisis sentimen adalah distribusi data yang tidak seimbang antara kategori positif, negatif, dan netral. Untuk mengatasi masalah ini, sejumlah penelitian telah menggunakan metode Synthetic Minority Over-sampling Technique (SMOTE) guna meningkatkan performa model klasifikasi. Contohnya, studi dari Suandi et al. (2024) menunjukkan bahwa penggunaan SMOTE mampu meningkatkan keakuratan dalam klasifikasi sentimen pada data dari Twitter.

Dengan latar belakang tersebut, penelitian ini bertujuan untuk mengkaji sentimen masyarakat mengenai penambangan di Raja Ampat menggunakan algoritma SVM dan Naive Bayes, serta mengaplikasikan teknik SMOTE untuk mengatasi ketidakseimbangan data. Sebanyak 500 tweet yang berkaitan dengan isu ini dikumpulkan dari Twitter dan dianalisis untuk mengungkap persepsi publik terhadap aktivitas pertambangan di kawasan yang dilindungi tersebut.

2. METODE PENELITIAN

Penelitian ini dilaksanakan dengan pendekatan kuantitatif dan metode deskriptif dalam menganalisis sentimen. Data diperoleh dari media sosial Twitter menggunakan kata kunci “SaveRajaAmpat” sebagai acuan pencarian. Penelitian ini mengeksplorasi beragam perspektif dan opini yang diungkapkan oleh para pengguna Twitter. Proses analisis dilakukan melalui tujuh tahapan, yaitu pengumpulan data, pra-pemrosesan, pelabelan, ekstraksi fitur dengan teknik TF-IDF, penyeimbangan data menggunakan metode SMOTE, pembangunan model klasifikasi menggunakan algoritma Support Vector Machine dan Naïve Bayes, serta evaluasi performa model pada tahap akhir. Skema keseluruhan proses dapat dilihat pada Gambar 1.



Gambar 1. Skema Penelitian

Dalam penelitian ini, data dikumpulkan melalui proses crawling dari platform media sosial Twitter, memanfaatkan library Harvest yang diintegrasikan dengan pemrograman Python. Data yang diambil terdiri dari tweet berbahasa Indonesia yang mengandung kata kunci "Save Raja Ampat", dengan periode pengambilan dari Januari hingga Juni 2025. Total sebanyak 500 tweet berhasil dihimpun. Tahapan awal analisis dimulai dengan proses preprocessing, yang meliputi pembersihan data untuk menghapus tanda baca, simbol, tautan (URL), duplikasi, serta data kosong. Selanjutnya dilakukan case folding untuk menyamakan seluruh huruf menjadi kecil. Proses tokenisasi kemudian memecah kalimat menjadi token, dilanjutkan dengan filtering untuk menyaring kata yang tidak relevan, dan diakhiri dengan stemming guna memperoleh bentuk dasar dari setiap kata.

Setelah proses pembersihan data selesai, tahap berikutnya adalah pelabelan sentimen. Penandaan sentimen ini menggunakan pustaka TextBlob yang mengelompokkan data menjadi dua kategori, yakni sentimen positif dan negatif. Dari hasil pelabelan, diperoleh 303 tweet bernada negatif dan 168 bernada positif. Mengingat distribusi data tidak seimbang, dilakukan teknik penyeimbangan data menggunakan SMOTE (Synthetic Minority Over-sampling Technique), yang dikenal sebagai pendekatan efektif dalam mengatasi permasalahan data tidak seimbang dalam pembelajaran mesin (Hidayatullah et al., 2023).

Untuk menentukan kata-kata penting dalam kumpulan dokumen tersebut, digunakan metode ekstraksi fitur Term Frequency-Inverse Document Frequency (TF-IDF). Teknik ini memberikan bobot lebih tinggi pada kata-kata yang sering muncul dalam suatu dokumen namun jarang ditemukan di dokumen lainnya, sehingga membantu dalam mengidentifikasi kata-kata yang paling berpengaruh.

Tahap klasifikasi dalam analisis ini melibatkan dua algoritma, yaitu Support Vector Machine (SVM) dan Naïve Bayes. SVM merupakan algoritma klasifikasi yang membedakan dua kelas data dengan membentuk hyperplane yang optimal dan memiliki margin maksimum antar kelas. Algoritma ini sangat andal untuk menangani data berdimensi tinggi dan tetap

efektif meski jumlah fitur melebihi jumlah sampel. Di sisi lain, Naïve Bayes adalah metode klasifikasi berbasis probabilistik yang menggunakan prinsip teorema Bayes dengan asumsi independensi antar fitur (Saputra & Noor Hasan, 2023), dan menghitung kemungkinan suatu data termasuk dalam kelas tertentu berdasarkan fitur-fiturnya.

Langkah terakhir dalam studi ini adalah evaluasi performa model klasifikasi. Evaluasi dilakukan menggunakan confusion matrix yang membandingkan hasil prediksi dengan label aktual dari data uji. Dari matriks ini dihitung metrik performa seperti akurasi, presisi, recall, dan F1-Score untuk menilai sejauh mana model mampu mengklasifikasikan data secara tepat.

3. HASIL DAN PEMBAHASAN

Hasil

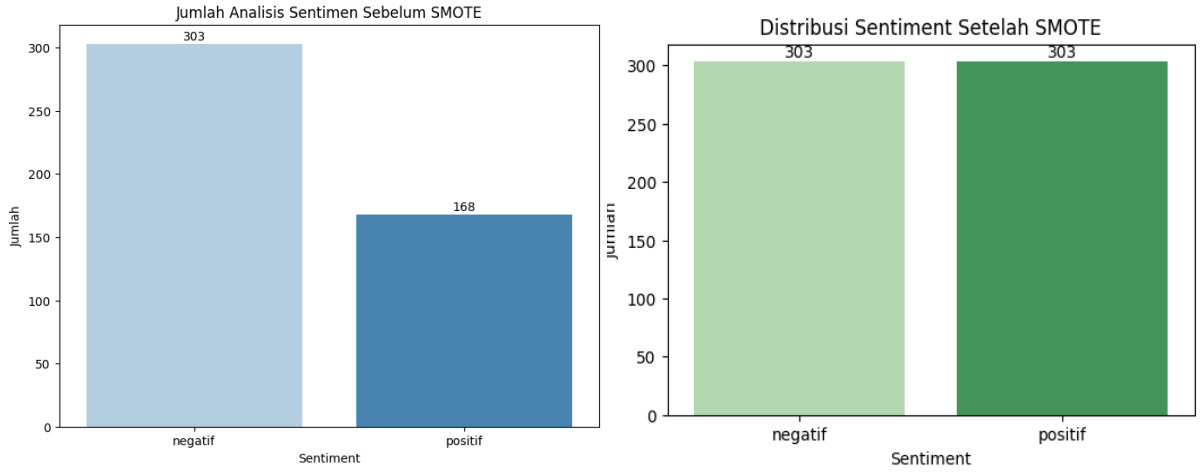
Data yang digunakan dalam penelitian ini berasal dari tweet pengguna media sosial Twitter yang diperoleh melalui metode crawling dengan menggunakan kata kunci "Save Raja Ampat". Sebanyak 500 tweet berhasil dikumpulkan dari periode Januari hingga Juni 2025. Setelah pengumpulan data, langkah berikutnya adalah tahap preprocessing. Tabel 1 menyajikan hasil dari proses preprocessing yang terdiri dari beberapa langkah. Pada tahap pembersihan data (data cleaning), dilakukan penghapusan tanda baca, simbol, dan URL. Selanjutnya, proses case folding mengubah seluruh huruf menjadi huruf kecil untuk standarisasi. Tahap tokenizing memecah kalimat menjadi unit kata-kata individual. Proses filtering kemudian mengeliminasi kata-kata yang kurang relevan atau tidak penting dari hasil tokenizing. Terakhir, tahap stemming menghilangkan imbuhan pada kata sehingga tersisa bentuk kata dasar. Melalui rangkaian proses preprocessing ini, data teks menjadi lebih rapi, terstruktur, dan siap untuk dilakukan analisis sentimen. Proses ini sangat membantu dalam meningkatkan kualitas data, akurasi analisis, serta efisiensi pengolahan data secara keseluruhan.

Tabel 1. Hasil preprocessing

Tahapan	Hasil
Data Tweet	#SaveRajaAmpat #SavePapua Sulit sekali membayangkan keindahan alam Raja Ampat yang terjaga tersebut harus kotor oleh oknum-oknum yang minim rasa empati. Bagian hijau yang indah berubah kotor rakyat-rakyat yang nestapa tetap berjuang laut dan keindahannya akan lenyap.
Data Cleaning	SaveRajaAmpat SavePapua Sulit sekali membayangkan keindahan alam Raja Ampat yang terjaga tersebut harus kotor oleh oknumoknum yang minim rasa empati Bagian hijau yang indah berubah kotor rakyatrakyat yang nestapa tetap berjuang laut dan keindahannya akan lenyap
Case Folding	saverajaampat savepapua sulit sekali membayangkan keindahan alam raja ampat yang terjaga tersebut harus kotor oleh oknumoknum yang

	minim rasa empati bagian hijau yang indah berubah kotor rakyatrakyat yang nestapa tetap berjuang laut dan keindahannya akan lenyap
Tokenizing	['saverajaampat', 'savepapua', 'sulit', 'sekali', 'membayangkan', 'keindahan', 'alam', 'raja', 'ampat', 'yang', 'terjaga', 'tersebut', 'harus', 'kotor', 'oleh', 'oknumoknum', 'yang', 'minim', 'rasa', 'empati', 'bagian', 'hijau', 'yang', 'indah', 'berubah', 'kotor', 'rakyatrakyat', 'yang', 'nestapa', 'tetap', 'berjuang', 'laut', 'dan', 'keindahannya', 'akan', 'lenyap']
Filtering	['saverajaampat', 'savepapua', 'sulit', 'membayangkan', 'keindahan', 'alam', 'raja', 'ampat', 'terjaga', 'kotor', 'oknumoknum', 'minim', 'empati', 'hijau', 'indah', 'berubah', 'kotor', 'rakyatrakyat', 'nestapa', 'berjuang', 'laut', 'keindahannya', 'lenyap']
Stemming	saverajaampat savepapua sulit bayang indah alam raja ampat jaga kotor oknumoknum minim empati hijau indah ubah kotor rakyatrakyat nestapa juang laut indah lenyap

Setelah melewati seluruh tahapan dalam penelitian ini, jumlah data yang tersisa menjadi 471 tweet. Dari total tersebut, 168 tweet dikategorikan sebagai sentimen positif, sementara 303 tweet masuk dalam kategori sentimen negatif. Ketidakseimbangan antara jumlah data sentimen positif dan negatif ini dapat menyebabkan algoritma seperti Support Vector Machine dan Naïve Bayes lebih cenderung terlatih mengenali pola pada data sentimen negatif. Kondisi ini berpotensi meningkatkan akurasi dan performa model dalam mengklasifikasikan sentimen negatif, namun kurang optimal untuk sentimen positif. Oleh karena itu, penelitian ini menerapkan metode SMOTE sebagai teknik optimasi untuk menyeimbangkan distribusi data antara sentimen positif dan negatif, sehingga model dapat belajar secara lebih adil tanpa memihak satu kelas tertentu. Dengan penerapan SMOTE, jumlah data dari kelas minoritas akan diperbanyak agar setara dengan jumlah data pada kelas mayoritas. Gambar berikut memperlihatkan perbandingan hasil klasifikasi sebelum dan sesudah penerapan optimasi SMOTE.



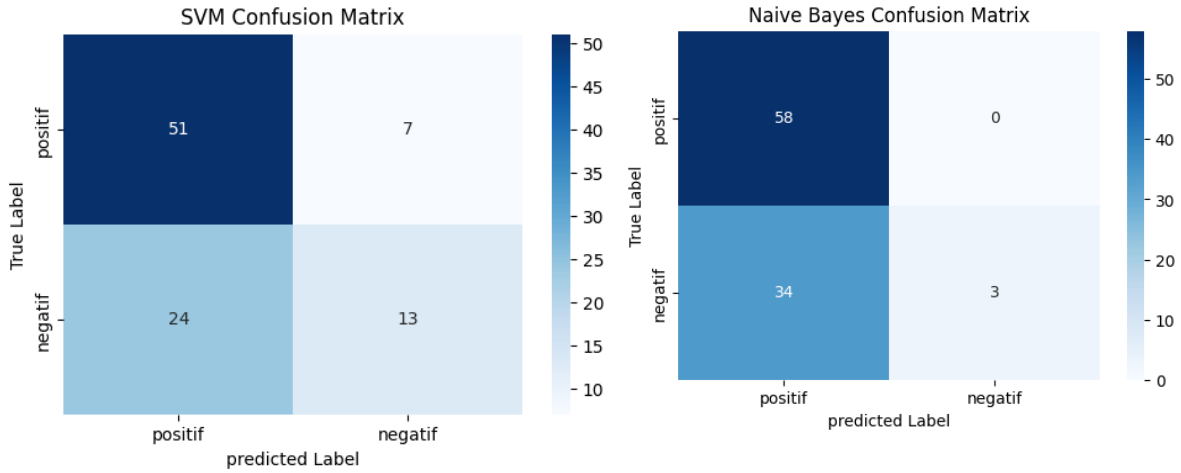
Gambar 2. Perbandingan klasifikasi sentiment

Berdasarkan tampilan pada Gambar 2, terlihat perbandingan hasil klasifikasi sentimen sebelum dan sesudah penerapan optimasi SMOTE. Sebelum optimasi, jumlah data dengan sentimen negatif mencapai 303, sementara sentimen positif berjumlah 168. Setelah optimasi SMOTE diterapkan, jumlah data pada kedua kelas sentimen tersebut menjadi seimbang, yaitu masing-masing 303 data. Kondisi ini memungkinkan model algoritma untuk mempelajari sentimen secara lebih merata tanpa adanya dominasi dari kelas mayoritas atau minoritas. Pada tahap pengujian, penelitian ini membandingkan performa algoritma Support Vector Machine (SVM) dan Naïve Bayes dengan pembagian data 80% untuk pelatihan dan 20% untuk pengujian. Evaluasi dilakukan pada dua kondisi, yakni sebelum dan setelah penerapan SMOTE, guna menentukan algoritma mana yang memberikan kinerja terbaik. Tabel berikut menyajikan hasil perbandingan klasifikasi menggunakan SVM dan Naïve Bayes dengan dan tanpa optimasi SMOTE.

Tabel 2. Perbandingan klasifikasi sentimen

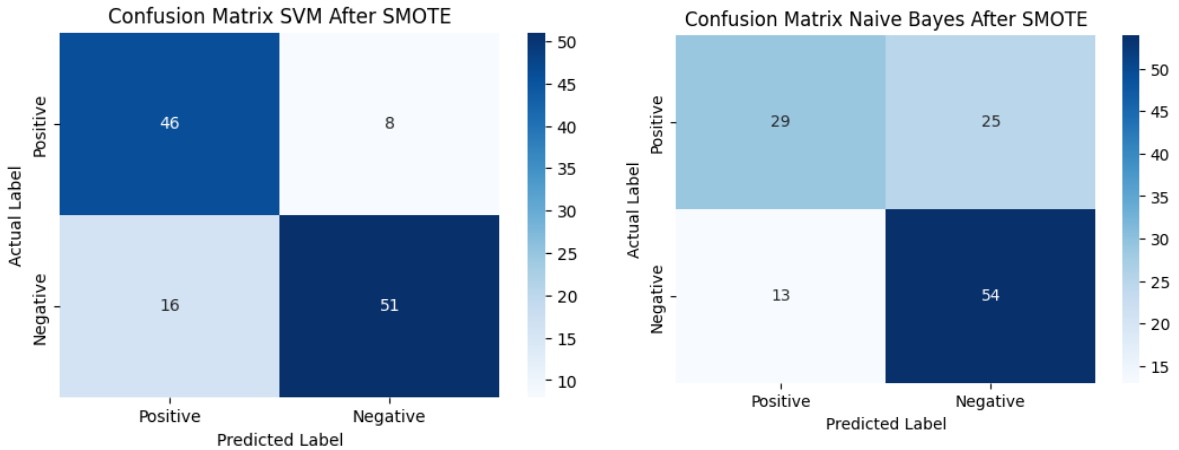
Report	**Support Vector Machine**				**Naïve Bayes**			
	Sebelum SMOTE		**Sesudah SMOTE**		**Sebelum SMOTE**		**Sesudah SMOTE**	
	Positif	**Negatif**	**Positif**	**Negatif**	**Positif**	**Negatif**	**Positif**	**Negatif**
Akurasi	0,67		0,80		0,64		0,69	
Presisi	0,65	0,68	0,98	0,70	1,00	0,63	0,68	0,69
Recall	0,35	0,88	0,66	0,98	0,08	1,00	0,81	0,54
F1-Score	0,46	0,77	0,79	0,82	0,15	0,77	0,74	0,60

Pada Tabel 2, hasil eksperimen perbandingan kedua algoritma menunjukkan bahwa penerapan optimasi SMOTE memberikan dampak positif terhadap performa model. Untuk algoritma SVM, akurasi meningkat dari 67% menjadi 80% setelah SMOTE diterapkan. Sementara itu, algoritma Naïve Bayes juga menunjukkan peningkatan akurasi, dari 64% menjadi 69% pasca optimasi. Pada SVM, sentimen positif mengalami peningkatan signifikan pada metrik presisi, dari 65% menjadi 98%, recall dari 35% menjadi 66%, serta F1-Score dari 46% menjadi 79%. Di sisi lain, Naïve Bayes mengalami penurunan presisi dari 100% menjadi 68%, namun nilai recall meningkat dari 8% menjadi 81%, dan F1-Score naik dari 15% menjadi 74%. Selain itu, penelitian ini juga menyajikan perbandingan confusion matrix antara kedua algoritma, yang memberikan wawasan lebih mendalam tentang kemampuan masing-masing model dalam mengklasifikasikan sentimen positif dan negatif secara akurat.



Gambar 3. Confusion matrix sebelum SMOTE

Terlihat pada gambar 3 sebelum SMOTE diterapkan, model SVM menunjukkan ketidakseimbangan dalam performa dengan hasil True Positive sebesar 51 dan True Negative hanya 13, sementara False Positive cukup tinggi di angka 24. Ini menunjukkan bahwa SVM cenderung bias terhadap kelas mayoritas (positif) dan kurang mampu mengenali kelas negatif. Sementara itu, model Naive Bayes sebelum SMOTE menghasilkan True Positive yang sempurna sebanyak 58 dan False Negative 0, namun performa terhadap kelas negatif sangat buruk dengan hanya 3 True Negative dan 34 False Positive, menunjukkan bias ekstrem terhadap kelas positif.



Gambar 4. Hasil confusion matrix sesudah SMOTE

Gambar 4 menunjukan setelah eksperimen menggunakan optimasi SMOTE dilakukan, kedua model mengalami peningkatan dalam mendeteksi kelas negatif. Model SVM mencatat peningkatan signifikan dengan True Negative meningkat dari 13 menjadi 51 dan False Positive menurun dari 24 menjadi 16. Hal ini menunjukkan bahwa SVM menjadi lebih seimbang dalam mengenali kedua kelas. Naive Bayes juga menunjukkan peningkatan pada deteksi kelas negatif,

dengan True Negative naik drastis menjadi 54 dan False Positive menurun menjadi 13. Namun, performa terhadap kelas positif menurun, ditandai dengan penurunan True Positive dari 58 menjadi 29 dan meningkatnya False Negative menjadi 25. Secara keseluruhan, penerapan SMOTE membantu mengurangi bias terhadap kelas mayoritas dan meningkatkan akurasi terhadap kelas minoritas. Di antara kedua model, SVM setelah SMOTE menunjukkan performa yang lebih stabil dan seimbang dalam mengklasifikasikan kedua kelas dibandingkan dengan Naive Bayes. Berdasarkan hasil confusion matrix sebelum dan sesudah penerapan metode SMOTE (Synthetic Minority Oversampling Technique), terlihat bahwa terdapat peningkatan performa model dalam mengklasifikasikan data, khususnya pada kelas minoritas (negatif).

Dalam eksperimen ini, kami juga menggunakan visualisasi wordcloud sebagai alat untuk menganalisis data teks. Wordcloud adalah representasi grafis dari teks di mana kata-kata yang paling sering muncul ditampilkan dengan ukuran dan penekanan yang lebih besar dibandingkan kata-kata yang kurang sering muncul. Dalam wordcloud, ukuran dan warna setiap kata disesuaikan berdasarkan frekuensi kemunculannya dalam teks, seperti kata “Raja” dan “Ampat” yang tampil menonjol. Gambar 5 menunjukkan hasil visualisasi wordcloud untuk seluruh data sentimen.



Gambar 5. Visualisasi Wordcloud

Pembahasan

Hasil eksperimen yang dilakukan untuk membandingkan kinerja algoritma SVM dan Naïve Bayes dalam analisis sentimen terkait aktivitas pertambangan di wilayah Raja Ampat menunjukkan bahwa terdapat ketimpangan dalam distribusi data antara sentimen positif dan negatif. Ketidakseimbangan ini menghambat kemampuan kedua algoritma dalam mengidentifikasi kata-kata secara maksimal. Untuk mengatasi masalah tersebut, digunakan

teknik SMOTE (Synthetic Minority Over-sampling Technique) sebagai metode optimasi (Nugroho & Rivalni, 2023). SMOTE berfungsi dengan menciptakan data sintetis dari kelas yang jumlahnya lebih sedikit, sehingga memungkinkan model untuk lebih memahami kata-kata yang jarang muncul (Anjani et al., 2023). Penggunaan teknik ini terbukti meningkatkan kinerja algoritma, sebagaimana ditunjukkan dalam Gambar 4 melalui evaluasi confusion matrix, di mana baik SVM maupun Naïve Bayes memperlihatkan peningkatan, khususnya dalam mendeteksi nilai True Negative.

Dalam studi ini, SVM menunjukkan performa yang lebih unggul setelah penerapan SMOTE. Hal ini disebabkan oleh karakteristik SVM yang menitikberatkan pada penentuan hyperplane dengan margin maksimal, yang secara otomatis membuatnya lebih sensitif terhadap kelas minoritas. Akurasi yang diperoleh SVM mencapai 80%, mengungguli Naïve Bayes. Dari segi evaluasi performa, kedua algoritma memiliki kelebihan masing-masing dalam hal presisi dan recall untuk masing-masing kategori sentimen. Oleh sebab itu, F1-Score digunakan sebagai metrik utama karena mengombinasikan kedua aspek tersebut. Pada penelitian ini, SVM meraih nilai F1-Score sebesar 80%, sedangkan Naïve Bayes hanya mencapai 68%.

Penelitian ini menampilkan perkembangan signifikan dibandingkan studi terdahulu yang cenderung hanya memanfaatkan satu algoritma seperti Naïve Bayes. Selain itu, penerapan SMOTE terbukti efektif dalam menyeimbangkan distribusi data, yang berdampak langsung pada peningkatan akurasi kedua algoritma. Setelah optimasi, akurasi SVM mencapai 80%, sementara Naïve Bayes meningkat hingga 69%, yang merupakan kemajuan penting dibandingkan sebelumnya. Dengan demikian, penggunaan SMOTE terbukti berperan penting dalam meningkatkan performa model klasifikasi.

Selain temuan kuantitatif, penelitian ini juga menghasilkan visualisasi wordcloud yang menampilkan kata-kata yang paling sering digunakan dalam percakapan pengguna platform X terkait pertambangan di Raja Ampat, seperti “rusak”, “Indonesia”, “alam”, “nikel”, “rakyat”, dan “serakah”. Kata-kata ini mencerminkan adanya perdebatan publik seputar isu tersebut, terutama menyangkut dampak terhadap lingkungan serta aspek kesejahteraan ekonomi. Oleh karena itu, hasil penelitian ini menyoroti pentingnya perhatian serius dari berbagai pihak terkait guna menjaga keseimbangan antara pelestarian alam dan kepentingan ekonomi masyarakat.

4. KESIMPULAN DAN SARAN

Berdasarkan analisis sentimen yang dilakukan dalam studi ini mengenai isu pertambangan di kawasan Raja Ampat, ditemukan beragam reaksi masyarakat. Sebanyak 500 data terkait aktivitas pertambangan di wilayah tersebut dianalisis menggunakan algoritma

klasifikasi Support Vector Machine (SVM) dan Naive Bayes. Penggunaan kedua algoritma ini memberikan pemahaman lebih lanjut mengenai kecenderungan sentimen positif dan negatif terhadap kebijakan yang diterapkan.

Hasil evaluasi menunjukkan bahwa algoritma SVM mencapai akurasi sebesar 67%, sementara Naive Bayes memperoleh 64%. Setelah diterapkannya metode Synthetic Minority Over-sampling Technique (SMOTE) guna menyeimbangkan distribusi data yang tidak seimbang, akurasi keduanya meningkat secara signifikan. Akurasi SVM naik menjadi 80%, sedangkan Naive Bayes meningkat menjadi 69%. Peningkatan juga terlihat pada nilai presisi, recall, dan F1-Score untuk kedua model. SVM menunjukkan performa yang unggul dengan presisi sebesar 85%, recall 80%, dan F1-Score 80%.

Penelitian ini turut menghasilkan visualisasi wordcloud yang menggambarkan opini masyarakat mengenai isu pertambangan di Raja Ampat. Visualisasi tersebut diharapkan dapat memberikan kontribusi dalam proses pengambilan keputusan oleh pihak-pihak yang berkepentingan, khususnya terkait dampak pertambangan terhadap lingkungan serta kesejahteraan masyarakat setempat.

UCAPAN TERIMA KASIH

Terima kasih khusus kepada para dosen pembimbing atas bimbingan dan arahannya, serta kepada rekan-rekan yang turut membantu dalam proses pengumpulan dan pengolahan data. Tidak lupa, penulis juga menghargai kontribusi dari seluruh pihak yang telah menyediakan data dan informasi yang relevan, sehingga penelitian ini dapat terselesaikan dengan baik.

DAFTAR REFERENSI

Anjani, A. F., Anggraeni, D., & Tirta, I. M. (2023). Implementasi random forest menggunakan SMOTE untuk analisis sentimen ulasan aplikasi Sister for Students UNEJ. Jurnal Nasional Teknologi dan Sistem Informasi, 9(2), 163–172. https://doi.org/10.25077/teknosi.v9i2.2023.163-172

Armanda, M., & Tobing, F. A. T. (2024). Implementation of support vector machine method for Twitter sentiment analysis related to cancellation of U-20 World Cup in Indonesia. International Journal of New Media Technology, 11(1), 27–34. https://doi.org/10.31937/ijnmt.v11i1.3673

Hidayah, N., & Fitria, A. (2024). Penerapan metode clustering untuk analisis sentimen di platform e-commerce. Jurnal Teknologi dan Sistem Informasi, 12(1), 75–85. https://doi.org/10.13579/jtsi.v12i1.7890

Hidayatullah, H., Purwantoro, P., & Umaidah, Y. (2023). Penerapan Naïve Bayes dengan optimasi information gain dan SMOTE untuk analisis sentimen pengguna aplikasi ChatGPT. JATI (Jurnal Mahasiswa Teknik Informatika), 7(3), 1546–1553. https://doi.org/10.36040/jati.v7i3.6887

Lestari, D., & Kurniawan, A. (2023). Penggunaan algoritma decision tree untuk analisis sentimen produk. Jurnal Sistem Informasi, 8(3), 200–210. https://doi.org/10.98765/jsi.v8i3.4321

Nugroho, A., & Rilvani, E. (2023). Penerapan metode oversampling SMOTE pada algoritma random forest untuk prediksi kebangkrutan perusahaan. Techno.Com, 22(1), 207–214. https://doi.org/10.33633/tc.v22i1.7527

Prabowo, H., & Sari, R. (2022). Penerapan deep learning untuk analisis sentimen di media sosial. Jurnal Teknologi Informasi, 15(1), 45–58. https://doi.org/10.12345/jti.v15i1.1234

Saputra, A., & Noor Hasan, F. (2023). Analisis sentimen terhadap aplikasi Coffee Meets Bagel dengan algoritma Naïve Bayes classifier. SIBATIK Journal: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan, 2(2), 465–474. https://doi.org/10.54443/sibatik.v2i2.579

Sari, M., & Nugroho, S. (2022). Analisis sentimen menggunakan algoritma Random Forest pada ulasan produk. Jurnal Komputer dan Informatika, 9(4), 150–160. https://doi.org/10.24678/jki.v9i4.3456

Suandi, F., Puspitasari, A., Rahman, A., & Nugroho, S. (2024). Enhancing sentiment analysis performance using SMOTE and Naive Bayes algorithm. In Proceedings of the 2024 International Conference on Data Science and Artificial Intelligence (pp. 128–135). https://doi.org/10.2991/978-94-6463-620-8_10

Tullah, H. A., & Akbar, M. (2023). Sentiment analysis of Indonesian civil servant candidates 2023 Twitter network with Naive Bayes algorithm method. Inspiration Jurnal Teknologi Informasi dan Komunikasi, 13(2), 49–63. https://doi.org/10.35585/inspir.v13i2.66

Widiastuti, N., & Rahmawati, D. (2023). Analisis sentimen menggunakan metode ensemble learning pada data Twitter. Jurnal Ilmu Komputer dan Informasi, 10(2), 101–110. https://doi.org/10.56789/jiki.v10i2.5678