

ANALISIS SENTIMEN KEBIJAKAN EKSPOR PASIR LAUT PADA SOSIAL MEDIA TWITTER MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE*

Raka Ardiansyah Hasibuan*¹, Dian Eka Ratnawati², Rizal Setya Perdana³

^{1,2,3} Universitas Brawijaya, Malang

Email: ¹rakahasibuan@student.ub.ac.id, ²dian_ilkom@ub.ac.id, ³rizalespe@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 25 Maret 2024, diterima untuk diterbitkan: 7 Agustus 2024)

Abstrak

Larangan ekspor pasir laut sudah diterapkan di Indonesia sejak 2003 melalui Surat Keputusan Menperindag No 117/MPP/Kep/2/2003 tentang Penghentian Sementara Ekspor Pasir Laut. Pada 15 Mei 2023, Presiden Joko Widodo melalui Peraturan Pemerintah No 26 Tahun 2023 kembali membolehkan kegiatan ekspor pasir laut keluar negeri. Penambangan pasir adalah proses pengambilan pasir dari alam. Pasir merupakan komoditas yang diperlukan dalam banyak proyek seperti reklamasi lahan dan pembangunan pulau. Proyek-proyek ini memiliki keuntungan ekonomis dan sosial, tapi penambangan pasir juga menghasilkan masalah lingkungan jika dilakukan terus-menerus. Kebijakan ekspor pasir laut ini ramai dibahas berbagai sosial media, salah satunya Twitter. Twitter merupakan media sosial yang cukup populer di Indonesia. Sosial media seperti Twitter dapat menghasilkan ribuan data dalam waktu yang singkat. Opini masyarakat di sosial media Twitter dapat digunakan oleh pemerintah untuk mengkaji ulang apakah kebijakan pembolean ekspor pasir laut ini sudah tepat atau tidak. Dalam penelitian ini, dilakukan analisis sentimen untuk mengetahui opini masyarakat terkait kebijakan ekspor pasir laut. Fitur dan metode yang digunakan adalah SentiWordNet 3.0 untuk labelling dan metode klasifikasi Support Vector Machine. Penelitian ini membandingkan kernel linear, polynomial, sigmoid, dan radial basis function. Kernel polynomial adalah kernel terbaik dan mendapatkan hasil akurasi 80.94%, hasil recall 80.94% dan hasil presisi 80.91%, sedangkan kernel terburuk adalah sigmoid dan mendapatkan hasil akurasi 79.73%, hasil recall 79.73% dan hasil presisi 80.29%. Teknik root cause analysis dengan Fishbone Diagram memberikan beberapa rekomendasi terhadap sentiment negatif antara lain: peningkatan investasi pada teknologi pengerukan ramah lingkungan, penerapan regulasi yang ketat terkait ekspor pasir laut dan kampanye penyuluhan terhadap pentingnya ekosistem serta penjagaan lingkungan.

Kata kunci: analisis sentimen, ekspor pasir laut, *support vector machine*, *fishbone diagram*

SENTIMENT ANALYSIS OF SEA SAND EXPORT POLICY ON TWITTER USING SUPPORT VECTOR MACHINE ALGORITHM

Abstract

The sea sand export ban has been implemented in Indonesia since 2003 through the Decree of the Minister of Industry and Trade No. 117/MPP/Kep/2/2003 concerning Temporary Suspension of Sea Sand Exports. On May 15 2023, President Joko Widodo through Government Regulation No. 26 of 2023 again allowed sea sand export activities abroad. Sand mining is the process of extracting sand from nature. Sand is a commodity required in many projects such as

land reclamation and island development. These projects have economic and social benefits, but sand mining also produces environmental problems if carried out continuously. This sea sand export policy has been widely discussed on various social media, one of which is Twitter. Twitter is a social media that is quite popular in Indonesia. Social media like Twitter can generate thousands of data in a short time. Public opinion on Twitter social media can be used by the government to review whether the policy of allowing sea sand exports is appropriate or not. In this research, sentiment analysis was carried out to determine public opinion regarding sea sand export policies. The features and methods used are SentiWordNet 3.0 for labeling and the Support Vector Machine classification method. This research compares linear, polynomial, sigmoid and radial basis function kernels. The polynomial kernel is the best kernel and gets 80.94% accuracy results, while the worst kernel is sigmoid and gets 79.73% accuracy results. The root cause analysis technique using the Fishbone Diagram provides several recommendations for negative sentiment, including: increasing investment in environmentally friendly dredging technology, implementing strict regulations regarding sea sand exports and outreach campaigns regarding the importance of ecosystems and environmental protection.

Keywords: *sentiment analysis, sea sand export, support vector machine, fishbone diagram*

1. PENDAHULUAN

Larangan ekspor pasir laut sudah diterapkan di Indonesia sejak 2003 melalui Surat Keputusan Menperindag No 117/MPP/Kep/2/2003 tentang Penghentian Sementara Ekspor Pasir Laut. Pada 15 Mei 2023, Presiden Joko Widodo melalui Peraturan Pemerintah No 26 Tahun 2023 kembali memperbolehkan kegiatan ekspor pasir laut keluar negeri. Penambangan pasir adalah proses pengambilan pasir dari alam. Pasir merupakan komoditas yang diperlukan dalam banyak proyek seperti reklamasi lahan dan pembangunan pulau. Proyek-proyek ini memiliki keuntungan ekonomis dan sosial, tapi penambangan pasir juga menghasilkan masalah lingkungan jika dilakukan terus-menerus (Ashraf et al., 2011).

Dampak penambangan pasir meliputi pengurangan kualitas air dan destabilisasi dasar tanah. Dampak kepada sumber daya hayati berupa pemusnahan infauna dan epifauna. Proses penambangan ini juga dapat menyebabkan erosi dan pencemaran air. Proses penambangan ini menyebabkan polemik di masyarakat. Pasir laut dapat memberikan manfaat ekonomi baik dari lapangan pekerjaan dan peningkatan pendapatan, tetapi penambangan pasir laut menimbulkan masalah lingkungan yang cukup signifikan.

Twitter merupakan media sosial yang cukup populer di Indonesia. Twitter dapat digunakan untuk menyatakan opini dan pendapat seseorang. Sosial media seperti Twitter dapat menghasilkan ribuan data dalam waktu yang singkat. Opini masyarakat di sosial media Twitter dapat digunakan oleh pemerintah untuk mengkaji ulang apakah kebijakan pembolehan ekspor pasir laut ini sudah tepat atau tidak. Teknik analisis sentimen dapat mempergunakan data dari sosial media twitter untuk dilakukan studi analisis sentimen terkait topik kebijakan ekspor pasir laut, dengan harapan dapat dijadikan evaluasi oleh pemerintah dalam melakukan kajian pembolehan ekspor pasir laut di Indonesia (Wagh & Punde, 2018).

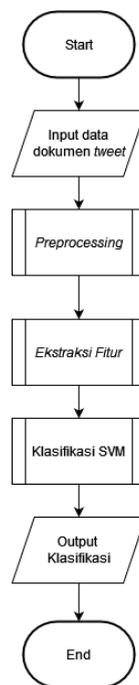
Algoritma *Support Vector Machine* (SVM) diterapkan dalam melakukan pengklasifikasian opini positif dan negatif di Twitter. SVM adalah algoritme pembelajaran mesin yang kuat dan serbaguna yang dapat digunakan untuk mayoritas masalah pembelajaran mesin. Algoritma SVM efisien dan dapat digunakan untuk menangani data yang besar. SVM juga tahan terhadap noise dan outliers (Cortes & Vapnik, 1995). SVM dipilih karena memiliki nilai akurasi yang baik dan cocok digunakan untuk dataset yang bersumberkan dari Twitter, hal ini didukung dengan hasil dari penelitian Neethu & Rajasree yang menunjukkan SVM mendapatkan skor akurasi 90% sedangkan NaiveBayes mendapatkan 89.5% dalam percobaan klasifikasi (Neethu

& Rajasree, 2013). Penelitian selanjutnya yang dilakukan oleh Ahuja et al. pada tahun 2017 mendapatkan hasil akurasi 90% untuk SVM dalam klasifikasi data text Twitter.

Analisis metode SVM pada kebijakan ekspor pasir laut akan diimplementasikan pada analisis akar permasalahan. Root Cause Analysis (RCA) yang diterapkan pada penelitian ini menggunakan diagram Fishbone. Diagram Fishbone adalah sebuah alat untuk mengidentifikasi akar penyebab kualitas suatu masalah yang dipelopori oleh Kaoru Isikawa pada tahun 1960-an. (Watson, 2004)

2. METODE PENELITIAN

Penelitian ini menggunakan algoritma *support vector machine* dalam proses klasifikasi terhadap dataset sentiment kebijakan ekspor pasir laut. Terdapat 7 tahapan metodologi penelitian yaitu studi literatur, pengumpulan data, *preprocessing* data, pembobotan kata, klasifikasi SVM, analisis, dan penarikan kesimpulan serta pemberian saran. Untuk diagram alur sistem dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alur Sistem

2.1. Studi Literatur

Studi literatur merupakan pengumpulan informasi mengenai literatur yang mendukung jalannya penelitian. Landasan-landasan kepustakaan ini didapatkan dari jurnal mengenai penelitian terdahulu. Pada penelitian ini, studi literatur difokuskan pada pembahasan analisis sentimen twitter dan algoritma *Support Vector Machine* beserta hasil performa analisisnya

2.2. Pengumpulan Data

Pada penelitian, data yang dikumpulkan berupa *tweet* yang mengandung kata kunci "ekspor pasir laut". Data yang diambil merupakan data dari rentang 14 Juni 2023 - 8 September 2023. Data diambil dengan menggunakan *twitter crawler* bernama "*tweet-harvest*" menggunakan *javascript* dan *auth token* twitter. Data yang berhasil dikumpulkan berjumlah 856. Penelitian ini hanya akan menggunakan 661 data dikarenakan 195 dari data yang didapatkan tidak memiliki kecenderungan positif/negatif.

2.3. Preprocessing Data

Preprocessing data yang akan dilakukan adalah *case folding*, formalisasi, *tokenization*, *lemmatization* dan *stopwords removal*. Seluruh proses ini bertujuan untuk menghilangkan *noise* pada *dataset* sehingga data menjadi relevan dan mudah untuk diolah oleh sistem (Singh & Kumari, 2016). Pada Tabel 1 dapat dilihat data hasil dari *preprocessing* data.

Tabel 1. Hasil *Preprocessing*

Sebelum proses preprocessing	Setelah proses preprocessing
@Suhana_web_id Harus di kecam itu pasir laut ko' di keruk di ekspor..la terus yg jadi pertanyaan uang dari hasil jualan pasir itu uang nya buat apa ya? 😊	['condemned', 'sea', 'sand', 'dredged', 'export', 'question', 'money', 'used', 'selling', 'sand']

2.4. Ekstraksi Fitur

Ekstraksi fitur yang akan dilakukan adalah pengambilan *part-of-speech* (PoS). PoS akan digunakan untuk mendapatkan skor sentimen suatu kalimat serta labelnya. Skor sentimen akan menentukan apakah suatu kalimat bersentimen positif atau negatif. Pada Tabel 2 dapat dilihat data hasil akhir dari Ekstraksi Fitur.

Tabel 2. Hasil Ekstraksi Fitur

Text	Label
@Suhana_web_id Harus di kecam itu pasir laut ko' di keruk di ekspor..la terus yg jadi pertanyaan uang dari hasil jualan pasir itu uang nya buat apa ya? 😊	Negative

2.5. Term-Weighting

Metode Term Weighting atau pembobotan yang dipilih pada penelitian ini adalah Term Frequency – Inverse Document Frequency (TF-IDF). Term Frequency suatu kata tertentu (t) ditentukan dengan menghitung seberapa banyak suatu istilah muncul dalam suatu dokumen dengan jumlah total kata dalam dokumen tersebut. IDF (Frekuensi Dokumen Invers) adalah penghitungan seberapa penting suatu kata (Ahuja et al., 2019). Semakin sering suatu kata muncul pada dokumen maka nilainya menjadi semakin tidak berarti. Nilai TF-IDF didapatkan dengan melakukan perkalian nilai TF dengan IDF.

Nilai TF-IDF didapatkan dengan melakukan perkalian nilai TF pada persamaan 1 dengan nilai IDF pada persamaan 2.

$$tf_{t,d} = \frac{N_{t,d}}{N_d} \quad (1)$$

$$idf_t = \log \frac{n}{n_k} \quad (2)$$

$tf_{t,d}$ merupakan nilai *term frequency* t di dokumen d , $N_{t,d}$ adalah jumlah munculnya *term* t di dokumen d , N_d merupakan total *term* pada dokumen d , idf_t adalah nilai IDF dari *term* t , n adalah jumlah dokumen dan n_k jumlah dokumen yang memuat *term* t .

2.6. Klasifikasi SVM

Support Vector Machine (SVM) adalah algoritme pembelajaran mesin yang kuat dan serbaguna yang dapat digunakan untuk berbagai masalah pembelajaran mesin. SVM adalah algoritme yang efisien dan dapat digunakan untuk berbagai masalah pembelajaran mesin. SVM

adalah algoritme yang efisien dan dapat digunakan untuk menangani data yang besar. SVM juga tahan terhadap noise dan outliers.

Metode SVM dimulai dengan mencari garis pembatas (hyperplane) yang memisahkan antara satu kelas dengan kelas lainnya. Hyperplane terbaik didapatkan dengan memberikan margin maksimal. Terdapat kernel yang digunakan pada Support Vector Machine yaitu kernel linear, kernel polynomial, kernel radial basis function (RBF) dan kernel sigmoid. Perhitungan hyperplane dapat dilihat pada persamaan 3 dan perhitungan margin dapat dilihat pada persamaan 4.

$$h(x) = w \cdot x + b \quad (3)$$

$$\text{Margin} = |d_{h_1} - d_{h_2}| = \frac{2}{\|w\|} \quad (4)$$

2.7. Confusion Matrix

Confusion Matrix adalah teknik untuk mengevaluasi klasifikasi suatu model. Cara kerja metode ini adalah dengan membandingkan matriks dari prediksi dengan kelas asli serta prediksi nilai klasifikasi. *Confusion matrix* dalam penelitian ini menggunakan empat kondisi yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

3. HASIL DAN PEMBAHASAN

Bab ini menganalisis dan menguji hasil implementasi penerapan sistem. Pengujian ini dibuat berdasarkan rancangan yang dirancang pada bab Perancangan dan menampilkan hasil akurasi dan evaluasi kinerja dengan menguji 10 kali *fold* dan menghitung rata-rata untuk setiap skenario.

3.1. Skenario Pertama-Pengujian *Degree* dan *Class-Balancing*

Pada scenario pertama digunakan seluruh data sebanyak 404 negatif dan 257 positif. Hasil pengujian scenario pertama dipaparkan pada Tabel 3 dan Tabel 4.

Tabel 3. Hasil Rerata *Accuracy* Skenario Pertama *Non-Balanced*

Degree	Akurasi
1	80.03%
2	74.13%
3	71.86%
4	70.50%
5	70.35%
6	70.20%
7	70.35%
8	70.35%
9	69.9%
10	69.9%

Tabel 4. Hasil Rerata *Accuracy* Skenario Pertama *Balanced*

Degree	Akurasi
1	79.73%
2	75.64%
3	72.77%
4	71.26%
5	70.35%
6	70.65%

Degree	Akurasi
7	70.35%
8	70.35%
9	70.35%
10	70.2%

3.2. Skenario Kedua-Pengujian 4 kernel Terhadap Parameter C

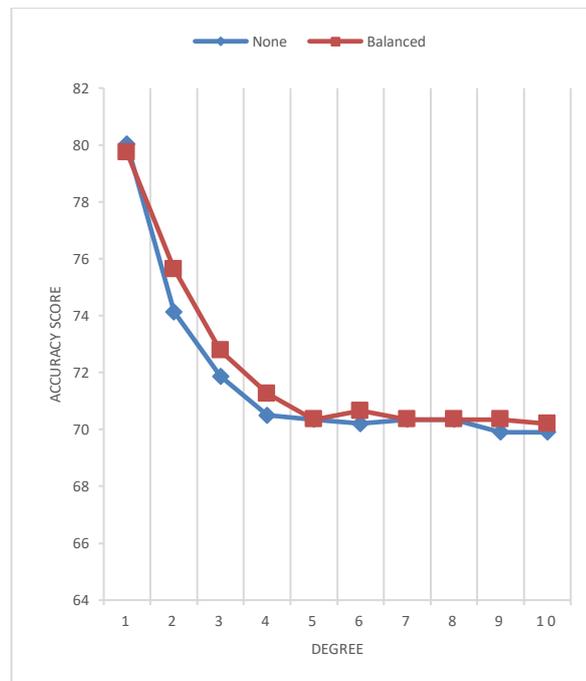
Pada scenario kedua menggunakan 4 kernel berbeda yaitu linear, polynomial, sigmoid dan rbf terhadap 3 nilai parameter C, yaitu 0.1, 1 dan 10. Hasil pengujian dapat dilihat pada Tabel 5.

Tabel 5. Hasil Rerata *Accuracy Kernel* terhadap C

Kernel Type	C Value	Average Accuracy
<i>Sigmoid</i>	0.1	72.77%
<i>Sigmoid</i>	1	79.73%
<i>Sigmoid</i>	10	78.67%
<i>RBF</i>	0.1	63.24%
<i>RBF</i>	1	78.97%
<i>RBF</i>	10	79.88%
<i>Linear</i>	0.1	75.65%
<i>Linear</i>	1	79.73%
<i>Linear</i>	10	80.78%
<i>Polynomial</i>	0.1	60.97%
<i>Polynomial</i>	1	80.03%
<i>Polynomial</i>	10	80.94%

3.3. Analisa Hasil Pengujian Skenario Pertama

Tren perbandingan dari hasil akurasi terhadap degree pada Tabel 3 Dan Tabel 4 dapat dilihat pada Gambar 2.



Gambar 2. Hasil Tren Perbandingan *None-Balanced* dan *Balanced*

Gambar 2. Menunjukkan bahwa peningkatan nilai degree baik pada fitur none-balanced dan balanced menghasilkan nilai akurasi yang lebih rendah. Penggunaan fitur balanced

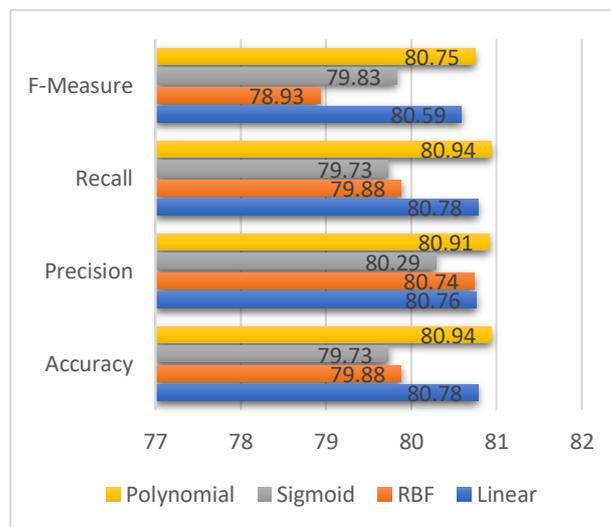
memberikan peningkatan nilai akurasi yang lebih menyebar jika dibandingkan dengan none-balanced. Secara rata-rata, peningkatan nilai akurasi sebesar 0.5%

3.4. Analisa Hasil Pengujian Skenario Kedua

Hasil analisis pada pengujian skenario kedua menunjukkan beberapa hal berikut:

1. Parameter C terbaik untuk kernel sigmoid adalah 1
2. Parameter C terbaik untuk kernel RBF adalah 10
3. Parameter C terbaik untuk kernel linear adalah 10
4. Parameter C terbaik untuk kernel polynomial adalah 10

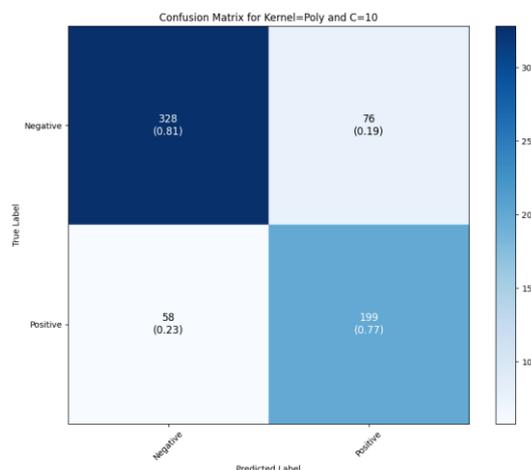
Untuk grafik perbandingan masing-masing *kernel* dengan parameter terbaik dapat dilihat pada Gambar 3.



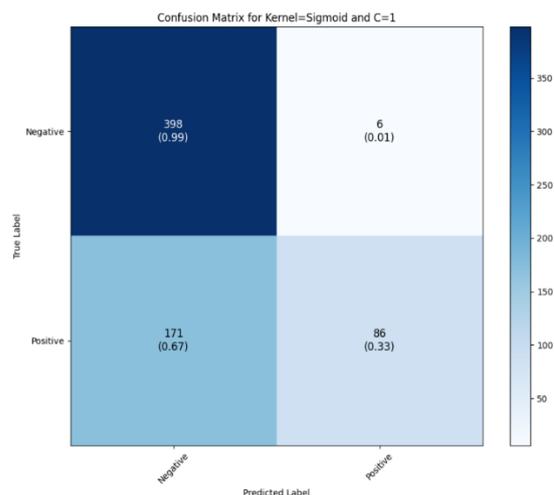
Gambar 3. Perbandingan *kernel* Parameter Terbaik

Gambar 3. Merupakan perbandingan masing-masing *kernel*. Masing-masing kernel memiliki performa yang mirip satu dengan yang lain. Secara umum, kernel polynomial mengungguli seluruh kernel lainnya dari seluruh matrix yang ada dengan nilai accuracy 80.94%, recall 80.94%, precision 80.91% dan F-Measure 80.75%. Kernel dengan performa terburuk adalah sigmoid dengan nilai *accuracy* 79.73%, *recall* 79.73%, *precision* 80.29% dan *F-Measure* 79.83%.

Berikut adalah hasil *confusion matrix* dari kernel polynomial dan kernel sigmoid:



Gambar 4. Hasil *Confusion Matrix* kernel Polynomial

Gambar 5. Hasil *Confusion Matrix* kernel Sigmoid

Dapat dilihat bahwa kernel polinomial merupakan model yang cukup baik untuk mendeteksi kasus positif. Kernel ini juga memiliki akurasi tinggi dalam prediksi secara keseluruhan. Dapat dikatakan bahwa kernel polinomial menawarkan keseimbangan yang lebih baik antara presisi dan recall jika dibandingkan dengan kernel sigmoid.

3.5. Visualisasi Word Cloud Sentimen Negatif

Pada bagian ini akan dilakukan visualisasi *word cloud* pada sentimen negatif. Sentimen negatif yang berjumlah 404 data merupakan data yang paling banyak disampaikan pada platform twitter. Hasil dari *word cloud* untuk kelas negatif dapat dilihat pada Gambar 6.

Gambar 6. Visualisasi *Word Cloud* Kelas Negatif

Word Cloud pada Gambar 6 mengandung kata berukuran besar dan kecil. Kata-kata yang berukuran besar adalah *sea*, *sand*, *export*, *policy*, *regulation*, *island*, dan *maritime*. Kata-kata ini termasuk dalam kata-kata yang paling sering digunakan pada tweet dengan kategori negative. Terdapat juga kata-kata dengan ukuran lebih kecil seperti *profit*, *commerce*, *debt* dan *environment*.

3.6. Root Cause Analysis (RCA)

Analisis RCA merupakan proses terakhir setelah analisis *word cloud* sebelumnya. Analisis RCA pada sentiment *tweet* negatif kebijakan ekspor pasir laut bertujuan untuk mengumpulkan berbagai macam kemungkinan permasalahan, mengidentifikasi akar permasalahan dan menentukan solusi/rekomendasi yang dapat diberikan. Hasil dari akar permasalahan dapat dilihat pada Tabel 7.

Tabel 7. Akar Permasalahan Umum

Akar Permasalahan	Penyebab	Rekomendasi
Keterbatasan teknologi pengerukan yang ramah lingkungan	Kurangnya investasi dalam pengembangan teknologi ekstraksi modern	Peningkatan investasi dalam riset dan pengembangan teknologi yang ramah lingkungan
Kerusakan pada ekosistem laut dan pantai karena ekstraksi pasir laut yang tidak terkendali.	Keterbatasan pemahaman akan ekologi ekosistem terkait dan dampak ekstraksi pasir laut.	Lakukan penelitian lebih lanjut untuk memahami dampak ekologis secara mendalam. Terapkan regulasi yang ketat untuk melindungi ekosistem alam saat ekspor pasir laut
Kurangnya data dan metrik untuk menilai dampak ekstraksi pasir laut	Metode pengukuran yang tidak memadai untuk evaluasi efisiensi ekstraksi pasir laut	Mengembangkan indikator kinerja yang lebih baik untuk mengukur dampak lingkungan dan efisiensi operasional
Keterbatasan sumber daya pasir laut	Peningkatan permintaan global terhadap pasir laut	Mendorong diversifikasi material konstruksi yang ramah lingkungan dan berkelanjutan, serta menggali alternatif pengganti pasir laut untuk mengurangi tekanan pada sumber daya tersebut
Kurangnya kesadaran dan kepercayaan masyarakat terkait ekspor pasir laut	Kurangnya keterlibatan stakeholders dan komunitas lokal di pesisir	Lakukan kampanye penyuluhan guna meningkatkan kesadaran masyarakat mengenai regulasi yang relevan

4. KESIMPULAN DAN SARAN PENELITIAN

Hasil analisis dari 661 data Twitter menunjukkan mayoritas sentimen negatif (50.5%), menyoroti kekhawatiran terhadap kebijakan ekspor pasir laut. Sentimen positif (25.25%) dan netral (24.5%) juga tergambar, mencerminkan variasi pendapat masyarakat. Pengujian kernel SVM menunjukkan performa baik, terutama kernel polynomial dengan akurasi 80.94%, precision 80.91%, recall 80.94% dan f-measure 80.75% setelah penanganan ketidakseimbangan kelas. RCA mengidentifikasi sumber masalah sentimen negatif, termasuk teknologi tidak ramah lingkungan, dampak lingkungan, kurangnya data evaluasi, keterbatasan sumber daya pasir laut, dan kurangnya kesadaran masyarakat. Saran penelitian selanjutnya adalah penambahan data dari platform sosial media lain dan penggunaan metode klasifikasi yang berbeda sebagai pembanding dengan kernel pada SVM. Peneliti juga menyarankan untuk menggunakan media sosial lain untuk mendapatkan data yang lebih beragam.

DAFTAR PUSTAKA

- Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science*, 152, 341–348. <https://doi.org/10.1016/j.procs.2019.05.008>
- Ashraf, M. A., Maah, M. J., Yusoff, I., Wajid, A., & Mahmood, K. (2011). Sand mining effects, causes and concerns: A case study from Bestari Jaya, Selangor, Peninsular Malaysia. *Scientific Research and Essays*, 6(6), 1216–1231.

- Neethu, M. S., & Rajasree, R. (2013). Sentiment analysis in twitter using machine learning techniques. *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, 1–5. <https://doi.org/10.1109/ICCCNT.2013.6726818>
- Nurdeni, D. A., Budi, I., & Santoso, A. B. (2021). Sentiment Analysis on Covid19 Vaccines in Indonesia: From The Perspective of Sinovac and Pfizer. *2021 3rd East Indonesia Conference on Computer and Information Technology (EIconCIT)*, 122–127. <https://doi.org/10.1109/EIconCIT50028.2021.9431852>
- Pradana, A. W., & Hayaty, M. (2019). The effect of stemming and removal of stopwords on the accuracy of sentiment analysis on indonesian-language texts. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 375–380.
- Seref, B., & Bostanci, E. (2018). Sentiment Analysis using Naive Bayes and Complement Naive Bayes Classifier Algorithms on Hadoop Framework. *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–7. <https://doi.org/10.1109/ISMSIT.2018.8567243>
- Singh, T., & Kumari, M. (2016). Role of Text Pre-processing in Twitter Sentiment Analysis. *Procedia Computer Science*, 89, 549–554. <https://doi.org/10.1016/j.procs.2016.06.095>
- Watson, G. (2004). The legacy of Ishikawa. *Quality Progress*, 37(4), 54.
- Wagh, R., & Punde, P. (2018). Survey on Sentiment Analysis using Twitter Dataset. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 208–211. <https://doi.org/10.1109/ICECA.2018.8474783>