

Analisis Sentimen Komentar Youtube terhadap Pembangunan Tol Kediri–Tulungagung dengan Pendekatan Lexicon–Based

^{1*}Mochamad Bressandy Syach Putra, ²Danang Wahyu Widodo, ³Patmi Kasih

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹putrasyach2@gmail.com, ²danangwahyuwidodo@unpkediri.ac.id,
³patmikasih@unpkediri.ac.id

Penulis Korespondens : Mochamad Bressandy Syach Putra

Abstrak—Pembangunan infrastruktur jalan tol Kediri–Tulungagung menimbulkan beragam tanggapan dari masyarakat yang banyak disampaikan melalui media sosial. Perbedaan pendapat tersebut perlu dianalisis untuk mengetahui kecenderungan sentimen masyarakat terhadap proyek pembangunan tersebut. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat pada media sosial terhadap proyek pembangunan Jalan Tol Kediri–Tulungagung menggunakan pendekatan *lexicon-based*. Data yang digunakan berupa komentar masyarakat yang diperoleh dari Youtube dan telah melalui tahap pra-proses teks, meliputi pembersihan data, tokenisasi, normalisasi, dan penghapusan stopword. Proses analisis sentimen dilakukan dengan menghitung skor sentimen berdasarkan kamus kata positif dan negatif, kemudian mengklasifikasikan komentar ke dalam kategori sentimen positif, netral, dan negatif. Hasil penelitian menunjukkan bahwa sentimen positif lebih dominan dibandingkan sentimen netral dan negatif. Evaluasi kinerja metode *lexicon-based* menghasilkan nilai akurasi sebesar 81%, yang menunjukkan bahwa metode ini cukup tepat untuk analisis sentimen awal. Meskipun memiliki keterbatasan dalam memahami konteks bahasa yang kompleks, pendekatan *lexicon-based* tetap dapat memberikan gambaran umum mengenai opini masyarakat terhadap proyek pembangunan infrastruktur.

Kata Kunci— Analisis Sentimen, Media Sosial, *Lexicon-Based*, Pembangunan Jalan Tol, Opini Masyarakat

Abstract— The construction of the Kediri–Tulungagung toll road has elicited a variety of responses from the public, many of which have been expressed on social media. These differing opinions need to be analyzed to identify trends in public sentiment toward the construction project. This study aims to analyze public sentiment on social media regarding the Kediri–Tulungagung Toll Road construction project using a *lexicon-based* approach. The data used consists of public comments obtained from social media that have undergone text preprocessing, including data cleaning, tokenization, normalization, and stopword removal. The sentiment analysis process was conducted by calculating sentiment scores based on a dictionary of positive and negative words, then classifying the comments into positive, neutral, and negative sentiment categories. The results of the study indicate that positive sentiment is more dominant than neutral and negative sentiment. Performance evaluation of the *lexicon-based* method yielded an accuracy rate of 86%, indicating that this method is sufficiently accurate for preliminary sentiment analysis. Although it has limitations in understanding complex linguistic contexts, the *lexicon-based* approach can still provide a general overview of public opinion regarding infrastructure development projects.

Keywords— *Lexicon-Based*, Public Opinion, Sentiment Analysis, Social Media, Toll Road Construction



I. PENDAHULUAN

Pembangunan Jalan Tol Kediri–Tulungagung merupakan salah satu proyek infrastruktur yang bertujuan untuk meningkatkan konektivitas antarwilayah serta memperlancar mobilitas masyarakat dan distribusi ekonomi. Kehadiran infrastruktur jalan tol diharapkan mampu mendorong pertumbuhan ekonomi daerah melalui peningkatan aksesibilitas dan efisiensi transportasi. Namun, proyek infrastruktur pada umumnya tidak hanya memberikan manfaat, tetapi juga dapat menimbulkan berbagai tanggapan dari masyarakat karena adanya dampak sosial maupun lingkungan yang dirasakan oleh masyarakat sekitar [1]. Oleh karena itu, penting untuk mengetahui bagaimana persepsi dan tanggapan masyarakat terhadap pembangunan Jalan Tol Kediri–Tulungagung.

Seiring dengan perkembangan teknologi informasi, YouTube menjadi salah satu platform digital yang banyak dimanfaatkan sebagai sarana penyebaran informasi dan penyampaian opini publik [2]. Melalui fitur komentar, masyarakat dapat memberikan tanggapan terhadap berbagai isu yang sedang berkembang, termasuk pembangunan infrastruktur. Komentar yang dihasilkan mengandung beragam pandangan yang dapat mencerminkan persepsi dan sentimen masyarakat terhadap suatu peristiwa atau kebijakan. Dengan jumlah komentar yang terus bertambah, diperlukan suatu metode yang mampu mengolah dan menganalisis data tersebut sehingga dapat menghasilkan informasi yang bermanfaat bagi pengambilan keputusan.

Salah satu metode yang dapat digunakan untuk memahami opini masyarakat adalah analisis sentimen. Beberapa penelitian sebelumnya menunjukkan bahwa analisis sentimen mampu digunakan untuk mengidentifikasi dan memahami tanggapan masyarakat terhadap kebijakan publik maupun proyek pemerintah dengan memanfaatkan data dari media sosial [3]. Salah satu pendekatan yang sering digunakan adalah metode *lexicon-based*, yaitu metode yang mengklasifikasikan sentimen berdasarkan kamus kata positif dan negatif [4]. Penelitian terdahulu menunjukkan bahwa pendekatan *lexicon-based* mampu memberikan gambaran umum mengenai kecenderungan sentimen masyarakat terhadap suatu isu yang dibahas di media sosial [5]. Selain itu, metode ini dinilai cukup efektif untuk data berbahasa Indonesia karena tidak memerlukan data latih dalam jumlah besar dan dapat langsung diterapkan pada data teks yang tersedia [6]. Beberapa penelitian juga menyebutkan bahwa pendekatan *lexicon-based* digunakan untuk menganalisis data media sosial yang umumnya mengandung bahasa informal, singkatan, dan berbagai ekspresi emosional [7].

Meskipun penelitian mengenai analisis sentimen telah banyak dilakukan, sebagian besar penelitian sebelumnya masih berfokus pada isu kebijakan nasional, layanan publik, atau topik umum lainnya. Penelitian yang secara khusus membahas sentimen masyarakat terhadap proyek pembangunan infrastruktur daerah, terutama pembangunan Jalan Tol Kediri–Tulungagung. Selain itu, karakteristik komentar pada YouTube yang beragam menjadikan analisis sentimen terhadap topik ini menarik untuk diteliti lebih lanjut.

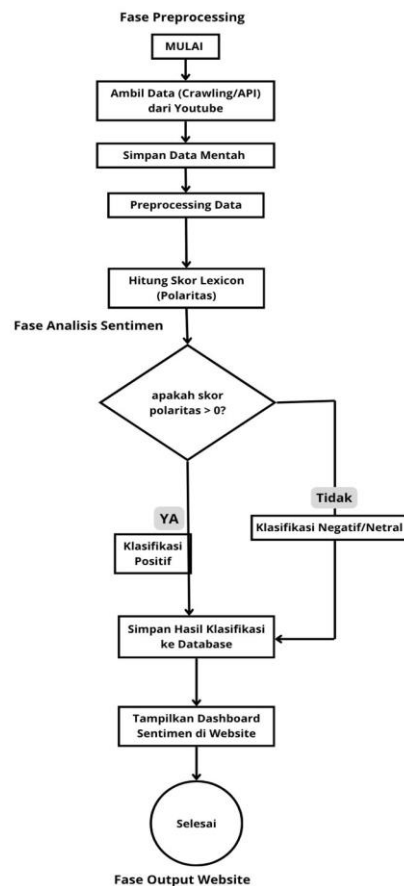
Berdasarkan permasalahan tersebut, penelitian ini menerapkan metode *lexicon-based* untuk menganalisis sentimen masyarakat terhadap pembangunan Jalan Tol Kediri–Tulungagung berdasarkan komentar yang diperoleh dari YouTube. Metode ini dipilih karena sederhana, mudah

diterapkan, serta mampu memberikan hasil analisis sentimen secara cepat tanpa memerlukan data latih dalam jumlah besar. Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen masyarakat ke dalam kategori positif, netral, dan negatif terhadap pembangunan Jalan Tol Kediri–Tulungagung. Hasil penelitian diharapkan dapat memberikan gambaran mengenai kecenderungan sentimen masyarakat terhadap proyek tersebut serta menjadi bahan evaluasi bagi pemerintah dan pihak terkait dalam menyusun strategi komunikasi publik.

II. METODE

Pada penelitian ini digunakan pendekatan yang komprehensif untuk menganalisis sentimen masyarakat terhadap proyek pembangunan Jalan Tol Kediri–Tulungagung. Proses penelitian meliputi beberapa tahapan utama, yaitu pengumpulan data komentar dari media sosial, praproses data teks, penerapan metode analisis sentimen berbasis *lexicon-based*, serta evaluasi kinerja model menggunakan pengujian akurasi.

Penelitian ini menggunakan metode *Lexicon-Based Sentiment Analysis* untuk menganalisis sentimen masyarakat terhadap pembangunan Tol Kediri–Tulungagung berdasarkan komentar pada platform YouTube. Tahapan penelitian dimulai dari pengambilan data komentar, *preprocessing* teks, perhitungan skor sentimen menggunakan kamus leksikon, klasifikasi sentimen, hingga penyajian hasil dalam bentuk dashboard website. Metode *Lexicon-Based* merupakan pendekatan yang menentukan sentimen berdasarkan pencocokan kata dengan kamus sentimen yang telah memiliki bobot polaritas tertentu [8].



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1, proses penelitian diawali dengan pengambilan data komentar YouTube menggunakan teknik *crawling*. Data yang berhasil diperoleh kemudian disimpan sebagai data mentah untuk diproses lebih lanjut. Setelah itu dilakukan *preprocessing* data untuk membersihkan komentar sebelum dilakukan analisis sentimen. Data hasil *preprocessing* selanjutnya dicocokkan dengan kamus leksikon positif dan negatif untuk memperoleh skor polaritas. Berdasarkan skor tersebut sistem menentukan kategori sentimen positif, negatif, atau netral. Hasil klasifikasi kemudian disimpan ke dalam database dan ditampilkan melalui dashboard website.

A. Pengumpulan Data

Data penelitian diperoleh dari komentar pada video YouTube yang membahas pembangunan Jalan Tol Kediri–Tulungagung. Pengumpulan data dilakukan menggunakan teknik *crawling* pada beberapa video yang relevan dengan topik penelitian. Adapun kriteria data yang digunakan dalam penelitian ini adalah komentar yang ditulis dalam bahasa Indonesia, berkaitan dengan pembangunan Jalan Tol Kediri–Tulungagung, serta tidak mengandung spam atau komentar duplikat. Data komentar dikumpulkan dari video yang diunggah pada periode Oktober 2025 – April 2026. Komentar yang berhasil dikumpulkan kemudian disimpan dalam bentuk dataset dan digunakan sebagai data utama dalam proses analisis sentimen [9].

B. Preprocessing

Tahap *preprocessing* bertujuan untuk membersihkan dan menyiapkan data teks agar dapat diproses oleh metode *Lexicon-Based*. Tahap ini sangat penting karena kualitas data yang baik akan mempengaruhi hasil klasifikasi sentimen yang diperoleh [10].

1. Tahapan *Case Folding*

Case folding merupakan proses mengubah seluruh huruf pada komentar menjadi huruf kecil (*lowercase*). Tujuan dari proses ini adalah untuk menyeragamkan bentuk kata sehingga tidak terjadi perbedaan antara huruf kapital dan huruf kecil saat proses analisis.

	1_Data_Aslri	2_Case_Folding
0	Remajaku dstu omahku.. 🤔 Industri kraft aku dsik n...	remajaku dstu omahku industri kraft aku dsik n...
1	Sing WIS cair ngguya ngguyuu...mesam mesem ..si...	sing wis cair ngguya ngguyumesam mesem sing du...
2	Wilayah kota kediri lambat ,terutama wilayah m...	wilayah kota kediri lambat terutama wilayah mr...

Gambar 2. Tahapan *Case Folding*

2. Tahapan *Tokenizing*

Tokenizing merupakan proses pemecahan kalimat menjadi kumpulan kata atau token. Tahapan ini memudahkan sistem dalam melakukan pencocokan kata dengan kamus leksikon [9].

	2_Case_Folding	3_Tokenizing
0	remajaku dstu omahku industri kraft aku dsik n...	[remajaku, dstu, omahku, industri, kraft, aku, ...
1	sing wis cair ngguya ngguyumesam mesem sing du...	[sing, wis, cair, ngguya, ngguyumesam, mesem, ...
2	wilayah kota kediri lambat terutama wilayah mr...	[wilayah, kota, kediri, lambat, terutama, wila...

Gambar 2. Tahapan *Tokenizing*

3. Tahapan *Stopword Removal*

Stopword removal dilakukan untuk menghapus kata-kata umum yang tidak memiliki nilai sentimen, seperti "dan", "yang", "di", dan "ke", sehingga proses analisis menjadi lebih fokus pada kata yang memiliki makna sentimen [10].

	3_Tokenizing	4_Stopword_Removal
0	[remajaku, dstu, omahku, industri, kraft, aku,...	[remajaku, dstu, omahku, industri, kraft, aku,...
1	[sing, wis, cair, ngguya, ngguyumesam, mesem, ...	[sing, wis, cair, ngguya, ngguyumesam, mesem, ...
2	[wilayah, kota, kediri, lambat, terutama, wila...	[wilayah, kota, kediri, lambat, terutama, wila...

Gambar 3. Tahapan Stopword Removal

C. Modeling

Analisis sentimen dilakukan menggunakan pendekatan *lexicon-based*, yaitu metode yang menentukan sentimen berdasarkan kamus kata positif dan negatif [11]. Setiap kata positif diberikan bobot +1, sedangkan kata negatif diberikan bobot -1. Skor sentimen suatu komentar dihitung dengan menjumlahkan seluruh bobot kata yang terdapat di dalam komentar.

$$\text{Skor Sentimen} = \sum_{i=1}^n w_i \quad (1)$$

Dengan w_i merupakan bobot kata ke- i dan n adalah jumlah kata dalam komentar [8].

$$\text{Sentimen} = \begin{cases} \text{Positif, jika skor} > 0 \\ \text{Netral, jika skor} = 0 \\ \text{Negatif, jika skor} < 0 \end{cases} \quad (2)$$

Berdasarkan nilai skor sentimen yang diperoleh, komentar kemudian diklasifikasikan menggunakan aturan keputusan sebagai berikut:

Pendekatan *lexicon-based* dipilih karena memiliki keunggulan dalam kesederhanaan implementasi dan tidak memerlukan data latih berlabel dalam jumlah besar. Metode ini juga banyak digunakan dalam penelitian analisis sentimen berbahasa Indonesia karena keterbatasan dataset berlabel serta kemampuannya dalam memberikan gambaran awal opini publik secara cepat [12], [13].

III. HASIL DAN PEMBAHASAN

A. Hasil Klasifikasi Sentimen

Setelah proses preprocessing selesai, sistem melakukan klasifikasi sentimen menggunakan metode *Lexicon-Based*. Setiap kata pada komentar dicocokkan dengan kamus leksikon positif dan negatif untuk memperoleh skor sentimen. Berdasarkan skor tersebut, sistem menentukan kategori sentimen berupa positif, negatif, atau netral.

	text_clean	kata_kunci	skor_akhir	label_sentimen	label_manual
0	Remajaku dstu omahku.0000 Industri kraft aku ...		0	Netral	Positif
1	Sing WIS cair ngguya ngguyu...mesam mesem ...si...	cair(5), senyum(-1), sing(5), pusing(-3)	6	Positif	Negatif
2	Wilayah kota kediri lambat ,terutama wilayah m...	lambat(-4), proses(1)	-3	Negatif	Negatif
3	Bismillah Jalan Tol Gresik - Lamongan - Tuban ...	bismillah(3)	3	Positif	Positif
4	Sengdurung pencairan sabar lor...mengko lek way...		0	Netral	Netral
5	Pembebasan semampir sebelah heritage kediri be...		0	Netral	Netral
6	Untuk gayam tahap 2 arah bandara kapan ki pros...	lama(-4), bengkalai(-5), blas(-2)	-11	Negatif	Negatif
7	Jogja bawen kapan lur		0	Netral	Netral
8	Tol kere,tanda tangan tanda tangan tok tp GK d...		0	Netral	Negatif
9	Situbondo kapan bos		0	Netral	Netral

Gambar 5. Hasil Klasifikasi Sentimen

Berdasarkan Gambar 5, sistem berhasil melakukan pelabelan sentimen secara otomatis terhadap komentar yang telah melalui tahap preprocessing. Kolom kata_kunci menunjukkan kata-kata yang dikenali oleh sistem beserta bobot sentimennya, sedangkan kolom skor_akhir menunjukkan hasil penjumlahan bobot yang diperoleh dari setiap komentar. Sebagai contoh, komentar yang mengandung kata dengan bobot positif lebih banyak akan menghasilkan skor positif dan diklasifikasikan sebagai sentimen positif. Sebaliknya, komentar yang mengandung lebih banyak kata negatif akan menghasilkan skor negatif dan dikategorikan sebagai sentimen negatif.

Selain menampilkan hasil klasifikasi sistem pada kolom label_sentimen, tabel juga menampilkan kolom label_manual yang digunakan sebagai pembanding dalam proses evaluasi. Perbandingan antara kedua label tersebut digunakan untuk mengetahui tingkat ketepatan metode *Lexicon-Based* dalam melakukan klasifikasi sentimen. Dari hasil yang diperoleh terlihat bahwa sebagian besar label yang dihasilkan sistem telah sesuai dengan pelabelan manual, sehingga menunjukkan bahwa metode yang digunakan mampu mengidentifikasi kecenderungan sentimen komentar masyarakat dengan cukup baik.

Tahap pelabelan sentimen ini menjadi proses utama dalam penelitian karena hasil klasifikasi yang diperoleh akan digunakan pada tahap evaluasi akurasi serta analisis distribusi sentimen masyarakat terhadap pembangunan Tol Kediri–Tulungagung.

B. Hasil Evaluasi Model

Evaluasi model sistem analisis sentimen dilakukan untuk mengetahui sejauh mana pendekatan *lexicon-based* mampu mengklasifikasikan komentar masyarakat secara tepat. Pada penelitian ini, evaluasi performa sistem dilakukan menggunakan beberapa metrik pengukuran, yaitu akurasi, *precision*, *recall*, dan *F1-score*. Penggunaan beberapa metrik evaluasi bertujuan untuk memberikan gambaran yang lebih komprehensif mengenai performa sistem, tidak hanya dari sisi ketepatan klasifikasi secara keseluruhan, tetapi juga dari kemampuan sistem dalam mengenali masing-masing kelas sentimen.

=====				
HASIL AKURASI TERBARU: 86.00%				
=====				
	precision	recall	f1-score	support
Negatif	0.96	0.77	0.85	30
Netral	0.84	0.85	0.85	61
Positif	0.84	0.92	0.88	59
accuracy			0.86	150
macro avg	0.88	0.84	0.86	150
weighted avg	0.86	0.86	0.86	150

Gambar 6. Hasil Akurasi

Berdasarkan hasil pengujian, sistem analisis sentimen memperoleh nilai akurasi sebesar 86%. Hasil evaluasi juga menunjukkan bahwa nilai *precision*, *recall*, dan *F1-score* memiliki nilai yang relatif seimbang. Nilai *weighted average precision*, *recall*, dan *F1-score* masing-masing sebesar 0,86. Hal ini mengindikasikan bahwa sistem tidak hanya mampu memberikan prediksi yang tepat, tetapi juga cukup konsisten dalam mengklasifikasikan data pada seluruh kelas sentimen, terutama pada kelas dengan jumlah data yang lebih besar, seperti sentimen positif dan netral.

Namun demikian, nilai *macro average recall* yang sedikit lebih rendah dibandingkan nilai *precision* menunjukkan bahwa masih terdapat keterbatasan sistem dalam mendeteksi seluruh data pada setiap kelas sentimen secara merata. Kondisi ini terutama dipengaruhi oleh ketidakseimbangan jumlah data antar kelas, di mana jumlah komentar dengan sentimen negatif lebih sedikit dibandingkan kelas lainnya.

Perbedaan nilai pada setiap metrik evaluasi tersebut menunjukkan bahwa pendekatan *lexicon-based* masih memiliki keterbatasan dalam memahami konteks kalimat secara mendalam. Komentar yang mengandung bahasa tidak baku, makna ganda, atau unsur sarkasme berpotensi menyebabkan kesalahan klasifikasi, karena metode ini sangat bergantung pada keberadaan kata-kata positif dan negatif yang terdapat dalam kamus sentimen.

IV. KESIMPULAN

Berdasarkan hasil analisis sentimen masyarakat terhadap proyek pembangunan Jalan Tol Kediri–Tulungagung menggunakan pendekatan *lexicon-based*, dapat disimpulkan bahwa metode ini mampu memberikan gambaran umum mengenai opini publik yang berkembang di media sosial. Hasil klasifikasi sentimen menunjukkan bahwa sentimen positif lebih mendominasi dibandingkan sentimen netral dan negatif. Kondisi ini mengindikasikan bahwa mayoritas masyarakat cenderung memberikan tanggapan yang mendukung atau setidaknya bersikap netral terhadap proyek pembangunan jalan tol tersebut. Evaluasi kinerja sistem menunjukkan nilai akurasi sebesar 86%, sehingga pendekatan *lexicon-based* dapat dinilai cukup efektif untuk digunakan sebagai metode analisis sentimen awal dengan proses yang sederhana dan efisien.

Meskipun demikian, hasil penelitian juga menunjukkan bahwa pendekatan *lexicon-based* masih memiliki keterbatasan, khususnya dalam mengklasifikasikan komentar yang mengandung bahasa tidak baku, makna ambigu, atau unsur sarkasme. Keterbatasan ini disebabkan oleh ketergantungan metode terhadap kamus kata positif dan negatif tanpa mempertimbangkan konteks kalimat secara menyeluruh. Namun, dengan mempertimbangkan kemudahan implementasi dan hasil akurasi yang diperoleh, metode *lexicon-based* tetap relevan digunakan sebagai metode dasar dalam analisis sentimen masyarakat berbasis teks.

DAFTAR PUSTAKA

- [1] F. Nurpandi, F. S. Sulaeman, dan A. Hermawan, “Analisis Sentimen Terhadap Kinerja Kepolisian Indonesia Menggunakan Metode Multinomial Naive Bayes, Long Short-Term Memory, dan Lexicon-Based,” *Media Jurnal Informatika*, vol. 16, no. 1, pp. 1–10, 2024, doi: 10.35194/mji.v16i1.4165.
- [2] D. Andini dan D. A. Muthia, “Implementasi Metode Lexicon Based dan Support Vector Machine Pada Analisis Sentimen Ulasan Pengguna ChatGPT,” *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 9, no. 2, 2024, doi: 10.31294/ijcit.v9i2.23948.
- [3] A. H. D. Fatra, N. Hayatin, dan C. S. K. Aditya, “Analisa Sentimen Tweet Berbahasa Indonesia Dengan Menggunakan Metode Lexicon Pada Topik Perpindahan Ibu Kota Indonesia,” *Jurnal Repositor*, vol. 2, no. 11, 2024, doi: 10.22219/repositor.v2i11.30761.
- [4] B. Setiawan, “A Review of Sentiment Analysis Applications in Indonesia Between 2023–2024,” *JIEET (Journal of Information Engineering and Educational Technology)*, vol. 8, no. 2, pp. 71–83, 2025, doi: 10.26740/jieet.v8n2.p71-83.

- [5] G. Z. Rahma dan M. Rosyda, "Perbandingan Kinerja Algoritma Convolutional Neural Network (CNN) dan Recurrent Neural Network (RNN) pada Analisis Sentimen Pemilu Presiden 2024," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 5, 2024, doi: 10.52436/1.jpti.652.
- [6] S. Fitrianti dan A. Yudhistira, "Analisis Sentimen Media Sosial Terhadap Calon Pilkada 2024 Dengan Metode Naïve Bayes," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 4, 2024, doi: 10.52436/1.jpti.610.
- [7] M. Chaerul, Septiadi, dan G. Triyono, "Analisis Sentimen Kebijakan Pembatasan Subsidi Bahan Bakar Minyak di Indonesia Tahun 2024 Menggunakan Algoritma Klasifikasi," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 12, 2025, doi: 10.52436/1.jpti.825.
- [8] B. S. Rintyarna, "Ina-SASet: Dataset for Developing Indonesian Sentiment Lexicon for Extracting Consumer Preference Based on Fine Grained Sentiment Analysis Technique," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 23s, pp. 295–300, 2024.
- [9] R. Syahputri, J. K. Tanubrata, dan S. Trisnawati, "Penerapan Lexicon-Based Untuk Analisis Sentiment Masyarakat Terhadap Kasus Kebocoran Data Di Indonesia," *Jurnal AI2MTech*, vol. 1, no. 2, 2024.
- [10] Asro dan Solihin, "Comparative Evaluation of Preprocessing Techniques in Twitter Sentiment Analysis for Indonesia's 2024 Regional Elections," *INOVTEK Polbeng - Seri Informatika*, 2025, doi: 10.35314/tt65bb54.
- [11] A. A. Firdaus, A. I. Hadiana, dan A. K. Ningsih, "Klasifikasi Sentimen pada Aplikasi Shopee Menggunakan Fitur Bag of Word dan Algoritma Random Forest," *Ranah Research: Journal of Multidisciplinary Research and Development*, vol. 6, no. 5, 2024, doi: 10.38035/rj.v6i5.994.
- [12] N. C. Ramadani, "Analisis Sentimen Untuk Mengukur Ulasan Pengguna Aplikasi Mobile Legend Menggunakan Algoritma Naive Bayes, SVM, Random Forest, Decision Tree, dan Logistic Regression," *JSI: Jurnal Sistem Informasi*, vol. 16, no. 1, pp. 123–138, 2024, doi: 10.18495/jsi.v16i1.152.
- [13] H. Hariyadi, D. Firdo, dan M. H. Al Rafi, "Implementasi Algoritma Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Ulasan Aplikasi Canva," *Jurnal Minfo Polgan*, vol. 13, no. 1, 2024, doi: 10.33395/jmp.v13i1.13568.
- [14] B. Setiawan, "A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024," *Journal of Information Engineering and Educational Technology (JIEET)*, vol. 8, no. 2, pp. 71–83, 2025, doi: i10.26740/jieet.v8n2.p71-83.
- [15] R. I. Syah, H. Hoiriyah, dan M. Walid, "Analisis Sentimen Pengguna Media Sosial Terhadap Aplikasi M-Health Peduli Lindungi Dengan Metode Lexicon Based dan Naïve Bayes," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 6, no. 1, 2023, doi: 10.21927/ijubi.v6i1.3275