

Analisis Sentimen dan Karakteristik Linguistik Komentar Publik terhadap Kebijakan Militer Menggunakan Model RoBERTa

Cindy Setyowati^{1*}, Ariska Nur Anggraini², Endah Kurnia Fitri³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Indonesia

Dikirimkan: 01-04-2026
Diterbitkan: 02-04-2026

Keywords:

Sentiment Analysis;
RoBERTa;
Public Opinion;
Text Classification;
Deep Learning.

E-mail Penulis

korespondensi:
cindyssetyo@gmail.com

Abstrak. Penelitian ini bertujuan untuk menganalisis sentimen serta karakteristik linguistik dalam komentar publik terhadap kebijakan militer menggunakan model berbasis transformer, yaitu RoBERTa. Dataset yang digunakan terdiri dari 32.184 komentar yang terbagi ke dalam tiga kategori sentimen, yaitu positif sebanyak 15.518 data, netral 9.064 data, dan negatif 7.602 data. Tahapan penelitian meliputi prapemrosesan teks, pelabelan, pelatihan model, serta evaluasi menggunakan metrik precision, recall, f1-score, dan confusion matrix. Hasil evaluasi menunjukkan bahwa model RoBERTa memiliki kinerja yang sangat baik dengan tingkat akurasi sebesar 95%. Secara rinci, kelas negatif memperoleh precision 0,81 dan recall 1,00, kelas netral mencapai precision 0,96 dan recall 1,00, sementara kelas positif menunjukkan precision 1,00 dan recall 0,92. Nilai f1-score masing-masing kelas berada pada kisaran tinggi, dengan macro average sebesar 0,94 dan weighted average sebesar 0,95. Confusion matrix menunjukkan bahwa sebagian besar prediksi dilakukan dengan tepat, dengan kesalahan klasifikasi yang relatif kecil, terutama pada sebagian data positif yang terklasifikasi sebagai negatif atau netral. Selain itu, analisis karakteristik linguistik menunjukkan bahwa komentar negatif didominasi oleh ekspresi emosional dan kritik langsung, komentar netral bersifat informatif, sedangkan komentar positif mencerminkan dukungan terhadap kebijakan. Distribusi data yang tidak seimbang juga menunjukkan dominasi sentimen positif dalam opini publik. Penelitian ini menegaskan bahwa RoBERTa efektif dalam menangkap nuansa bahasa dan dapat dimanfaatkan sebagai alat analisis opini publik untuk mendukung perumusan kebijakan yang lebih adaptif dan berbasis data.

Abstract. *This study aims to analyze sentiment and linguistic characteristics in public comments on military policies using a transformer-based model, RoBERTa. The dataset consists of 32,184 comments categorized into three sentiment classes: positive (15,518), neutral (9,064), and negative (7,602). The research process includes text preprocessing, data labeling, model training, and evaluation using precision, recall, f1-score, and confusion matrix metrics. The evaluation results indicate that the RoBERTa model achieves excellent performance with an accuracy of 95%. Specifically, the negative class attains a precision of 0.81 and recall of 1.00, the neutral class achieves a precision of 0.96 and recall of 1.00, while the positive class reaches a precision of 1.00 and recall of 0.92. The f1-scores for all classes are high, with a macro average of 0.94 and a weighted average of 0.95. The confusion matrix shows that most predictions are correct, with minimal misclassification, particularly where a small portion of positive data is incorrectly classified as negative or neutral. Furthermore, linguistic analysis reveals that negative comments are dominated by emotional expressions and*

direct criticism, neutral comments are primarily informative, and positive comments reflect explicit support for the policy. The imbalanced data distribution also indicates a dominance of positive sentiment in public opinion. These findings demonstrate that RoBERTa is effective in capturing linguistic nuances and can be utilized as a reliable tool for public opinion analysis to support more adaptive and data-driven policymaking.

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi telah mendorong meningkatnya partisipasi publik dalam menyampaikan opini terhadap berbagai kebijakan pemerintah, termasuk kebijakan militer [1]. Media sosial dan platform digital menjadi ruang utama bagi masyarakat untuk mengekspresikan pandangan, kritik, maupun dukungan secara terbuka [2]. Fenomena ini menghasilkan volume data tekstual yang sangat besar dan beragam, yang jika dianalisis dengan tepat dapat memberikan wawasan berharga mengenai persepsi publik.

Analisis sentimen menjadi salah satu pendekatan yang efektif untuk memahami opini publik secara sistematis. Dengan mengklasifikasikan teks ke dalam kategori sentimen seperti positif, netral, dan negatif, peneliti dapat mengidentifikasi kecenderungan sikap masyarakat terhadap suatu isu [3]. Pendekatan ini sangat relevan dalam konteks kebijakan militer yang seringkali memicu respons emosional dan perdebatan publik yang kompleks.

Seiring dengan perkembangan kecerdasan buatan, khususnya dalam bidang Natural Language Processing (NLP), metode analisis sentimen mengalami peningkatan signifikan [4]. Model berbasis deep learning, terutama arsitektur *transformer*, telah menunjukkan kinerja yang unggul dalam memahami konteks bahasa yang kompleks [5]. Salah satu model yang menonjol adalah *RoBERTa (Robustly Optimized BERT Approach)*, yang merupakan pengembangan dari *BERT* dengan optimasi pada proses pelatihan dan penggunaan data yang lebih besar [6].

RoBERTa memiliki kemampuan untuk menangkap hubungan semantik dan sintaksis dalam teks secara lebih mendalam dibandingkan metode tradisional [7]. Hal ini menjadikannya sangat cocok untuk menganalisis komentar publik yang seringkali mengandung variasi bahasa, ironi, maupun ekspresi emosional [8]. Dengan memanfaatkan model ini, analisis sentimen dapat dilakukan dengan tingkat akurasi yang lebih tinggi dan hasil yang lebih reliabel.

Selain aspek sentimen, karakteristik linguistik dalam komentar publik juga penting untuk dikaji [9]. Analisis ini memungkinkan identifikasi pola bahasa yang digunakan oleh masyarakat dalam menyampaikan opini, seperti penggunaan kata emosional, struktur kalimat, maupun gaya komunikasi [10]. Pemahaman terhadap karakteristik linguistik ini dapat memberikan konteks tambahan yang memperkaya hasil analisis sentimen.

Dalam penelitian ini, digunakan dataset komentar publik yang berjumlah 32.184 data yang telah diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Distribusi data menunjukkan dominasi sentimen positif, yang mengindikasikan kecenderungan dukungan terhadap kebijakan yang dianalisis [11]. Namun demikian, keberadaan sentimen negatif dan netral tetap memberikan perspektif kritis yang penting untuk dipahami.

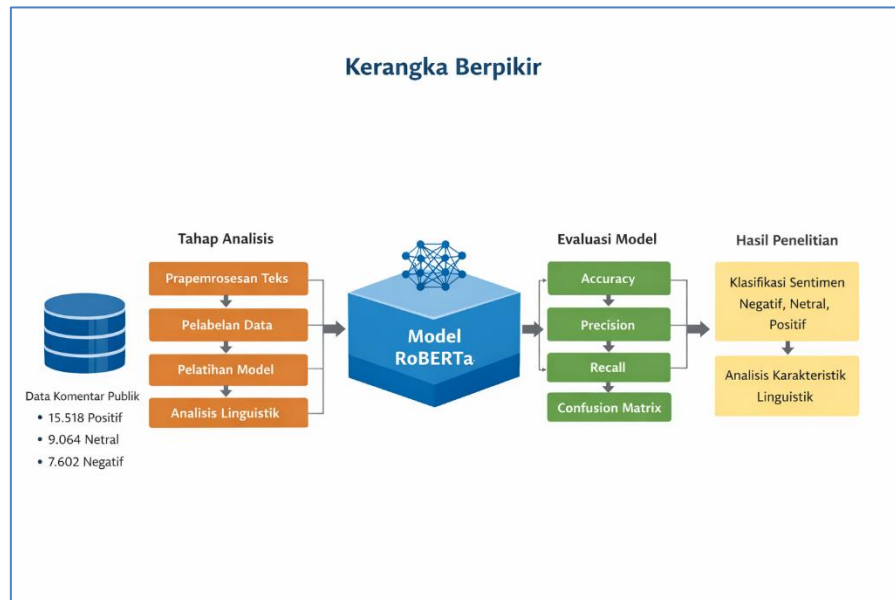
Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen serta karakteristik linguistik komentar publik terhadap kebijakan militer menggunakan model *RoBERTa*. Hasil penelitian diharapkan dapat memberikan kontribusi dalam pengembangan metode analisis opini publik berbasis kecerdasan buatan serta menjadi referensi bagi pengambil kebijakan dalam memahami respons masyarakat secara lebih komprehensif dan berbasis data.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis *Natural Language Processing (NLP)* untuk mengkaji opini publik terhadap kebijakan militer. Pendekatan ini dipilih karena mampu mengolah data tekstual dalam jumlah besar secara sistematis dan objektif [12]. Fokus utama penelitian adalah mengklasifikasikan sentimen serta mengidentifikasi karakteristik linguistik yang muncul dalam komentar publik.

Model yang digunakan dalam penelitian ini adalah *RoBERTa (Robustly Optimized BERT Approach)*, yaitu model berbasis arsitektur *transformer* yang merupakan pengembangan dari *BERT*. *RoBERTa* dipilih karena memiliki kemampuan yang lebih baik dalam memahami konteks bahasa melalui pelatihan yang lebih optimal dan penggunaan data yang lebih besar [13]. Model ini telah terbukti efektif dalam berbagai tugas klasifikasi teks.

Sumber data dalam penelitian ini berasal dari komentar publik yang dikumpulkan dari platform digital. Data yang digunakan berjumlah 32.184 komentar yang mencerminkan beragam opini masyarakat terhadap kebijakan militer. Data tersebut telah diberi label sentimen yang terdiri dari tiga kategori, yaitu positif, netral, dan negatif.



Gambar 1. Kerangka Berpikir

Distribusi data menunjukkan bahwa terdapat 15.518 komentar positif, 9.064 komentar netral, dan 7.602 komentar negatif. Ketidakseimbangan distribusi ini menunjukkan dominasi sentimen positif dalam dataset. Kondisi ini perlu diperhatikan karena dapat mempengaruhi performa model dalam mengenali kelas minoritas.

Tahap awal penelitian adalah prapemrosesan data untuk memastikan kualitas teks yang digunakan. Proses ini meliputi pembersihan data dari *noise* seperti tanda baca berlebih, *URL*, emoji, serta karakter *non-alfabet* [14]. Langkah ini penting untuk mengurangi gangguan dalam proses analisis. Selanjutnya dilakukan normalisasi teks untuk menyeragamkan bentuk kata, termasuk konversi huruf menjadi huruf kecil dan perbaikan kata tidak baku [15]. Proses ini bertujuan untuk meningkatkan konsistensi data sehingga model dapat mengenali pola bahasa dengan lebih baik. Tahap berikutnya adalah tokenisasi, yaitu proses memecah teks menjadi unit-unit kata atau subword sesuai dengan *tokenizer* yang digunakan oleh *RoBERTa* [16]. Tokenisasi ini memungkinkan model memahami struktur teks secara lebih detail. Selain itu, dilakukan padding dan *truncation* untuk menyeragamkan panjang input [17].

Dataset kemudian dibagi menjadi data latih dan data uji dengan proporsi tertentu untuk memastikan evaluasi model yang objektif [18]. Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk mengukur performa model pada data yang belum pernah dilihat sebelumnya [19]. Proses pelatihan dilakukan dengan metode *fine-tuning* terhadap model *RoBERTa* yang telah dipra-latih. *Fine-tuning* memungkinkan model menyesuaikan parameter internalnya sesuai dengan karakteristik dataset penelitian [20]. Parameter seperti *learning rate*, *batch size*, dan jumlah *epoch* diatur untuk mendapatkan hasil yang optimal [21]. Selama proses pelatihan, digunakan fungsi *loss* untuk mengukur kesalahan prediksi model dan *optimizer* untuk memperbarui bobot model [22]. Proses ini dilakukan secara iteratif hingga model mencapai konvergensi atau performa yang stabil. Teknik regularisasi juga dapat digunakan untuk mencegah *overfitting*.

Setelah pelatihan selesai, dilakukan evaluasi model menggunakan beberapa metrik kinerja. Metrik yang digunakan meliputi *accuracy* untuk mengukur tingkat ketepatan keseluruhan, serta *precision*, *recall*, dan *f1-score* untuk mengevaluasi performa pada masing-masing kelas sentimen [23]. Selain itu, *confusion matrix* digunakan untuk memberikan gambaran distribusi prediksi model terhadap kelas sebenarnya [24]. Melalui *confusion matrix*, dapat diketahui pola kesalahan klasifikasi yang terjadi [25], seperti kesalahan pada kelas positif yang terklasifikasi sebagai netral atau negatif.

3. Hasil dan Pembahasan

Dataset yang digunakan dalam penelitian ini terdiri dari 32.184 komentar publik yang diperoleh dari *platform* digital dan mencerminkan berbagai opini masyarakat terhadap kebijakan militer. Data tersebut telah melalui proses pelabelan ke dalam tiga kategori sentimen, yaitu positif sebanyak 15.518 data, netral sebanyak 9.064 data, dan negatif sebanyak 7.602 data. Distribusi ini menunjukkan adanya ketidakseimbangan kelas, di mana sentimen positif mendominasi dibandingkan dua kelas lainnya. Dataset ini kemudian digunakan sebagai dasar dalam proses pelatihan dan evaluasi model *RoBERTa* untuk mengidentifikasi pola sentimen serta karakteristik linguistik dalam teks secara lebih mendalam.

Tabel 1. *Dataset* Penelitian

No	comment_id	created_at	profile_pic_url	text	user_id	username
0	18115065256473652	1746621032	https://instagram.fcgk4-5.fna.fbcdn.net/...	@dopla_7 pa iye sok kabur wae sakolana	37865439330	ral_537
1	18042536120211439	1746621099	https://instagram.fcgk4-5.fna.fbcdn.net/...	@ral_537 tah sok di ajak si iyeu @lpand_	8121815340	dopla_7
2	17850542778449686	1746621693	https://instagram.fcgk4-5.fna.fbcdn.net/...	@dopla_7 Kirim KA dua Nana KA barak TNI pak 🙏	37865439330	ral_537
3	18032698235368796	1746622239	https://scontent-cgk1-1.cdninstagram.com/...	Anakku pak tolong jemput.. biar soleh Dan disiplin	64429045897	firandraal
...
10	18063566329846381	1746623148	https://instagram.fcgk4-4.fna.fbcdn.net/...	@ral_537 yeuhhh sok nanaon teh	6034761534	lpand_

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data teks sebelum digunakan dalam proses pelatihan model. Proses ini meliputi pembersihan data (*cleaning*) dengan menghapus *URL*, tanda baca, emoji, dan karakter khusus yang tidak relevan, serta menghilangkan mention seperti *@username* yang tidak memiliki kontribusi terhadap analisis sentimen. Selanjutnya dilakukan normalisasi teks dengan mengubah seluruh huruf menjadi huruf kecil (*case folding*) dan memperbaiki kata tidak baku atau singkatan agar lebih konsisten. Setelah itu, dilakukan tokenisasi menggunakan *tokenizer RoBERTa* untuk memecah teks menjadi unit *subword* yang sesuai dengan model. Selain itu, diterapkan padding dan truncation untuk menyeragamkan panjang input teks. Tahapan ini bertujuan agar data yang dihasilkan lebih bersih, terstruktur, dan optimal untuk diproses oleh model dalam menghasilkan klasifikasi sentimen yang akurat.

Tabel 2. *Dataset* Penelitian Setelah *Preprocessing*

No	text	cleaned_komentar
0	@dopla_7 pa iye sok kabur wae sakolana	dopla pa iye sok kabur wae sakolana
1	@ral_537 tah sok di ajak si iyeu @lpand_	ral tah sok ajak si iyeu lpand
2	@dopla_7 Kirim KA dua Nana KA barak TNI pak 🙏	dopla kirim ka dua nana ka barak tni pak
3	Anakku pak tolong jemput.. biar soleh Dan disiplin	anak pak jemput biar soleh disiplin insyaallah
4	@ral_537 yeuhhh sok nanaon teh	ral yeuhhh sok nanaon teh
5	Kdm ok 🙏🙏	kdm
6	😄😄😄	(kosong)
7	#konnasham #kpai mah diam aja ngga usah kerja...	konnasham kpai mah diam aja ngga usah kerja
8	Gubernur yang lain baru ketahuan GAK ADA KERJA...	gubernur baru tahu gak kerja nya
9	Ayo pak..ke kalsel...biar maju indonesia meny...	ayo pak ke kalsel biar maju indonesia seluruh
10	Hebat, semoga anak anak generasi sekarang yg ...	hebat moga anak anak generasi sekarang yg pimp
11	@sarahaaulia2088 abi mah kapungkur baong kitu t...	sarahaaulia abi mah kapungkur baong kitu teh
12	Anakku pak usia 2,5 tahun kayaknya juga perlu ...	anak pak usia tahun kayak perlu bina sini
...
...
32183	Pa, anak saya boleh gk ikut pelatihan di tempat...	pa anak gk ikut latih tempat titip anak gak mas

Penelitian ini menghasilkan model klasifikasi sentimen berbasis *RoBERTa* yang mampu mengidentifikasi opini publik terhadap kebijakan militer ke dalam tiga kategori, yaitu positif, netral, dan negatif. Berdasarkan hasil

pengujian terhadap 100 data uji, model menunjukkan performa yang sangat baik dengan tingkat akurasi sebesar 95%. Hal ini mengindikasikan bahwa model memiliki kemampuan yang tinggi dalam memahami konteks bahasa pada komentar publik.

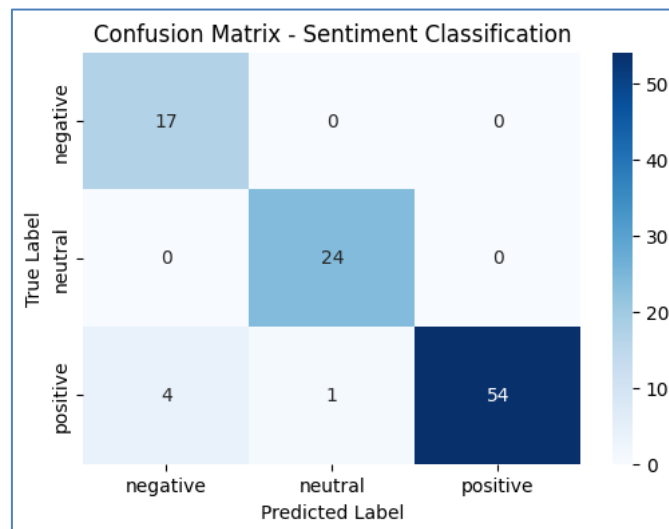
Tabel 3. Classification Report

Kelas	Precision	Recall	F1-Score	Support
Negative	0.81	1.00	0.89	17
Neutral	0.96	1.00	0.98	24
Positive	1.00	0.92	0.96	59
Accuracy			0.95	100
Macro Avg	0.92	0.97	0.94	100
Weighted Avg	0.96	0.95	0.95	100

Secara lebih rinci, hasil evaluasi menunjukkan bahwa kelas negatif memiliki nilai precision sebesar 0,81 dan recall sebesar 1,00, yang berarti seluruh data negatif berhasil terdeteksi dengan baik, meskipun masih terdapat beberapa prediksi yang kurang tepat dari sisi precision. Kelas netral menunjukkan performa yang sangat optimal dengan precision 0,96 dan recall 1,00, menandakan bahwa model sangat konsisten dalam mengenali komentar yang bersifat informatif atau tidak berpihak.

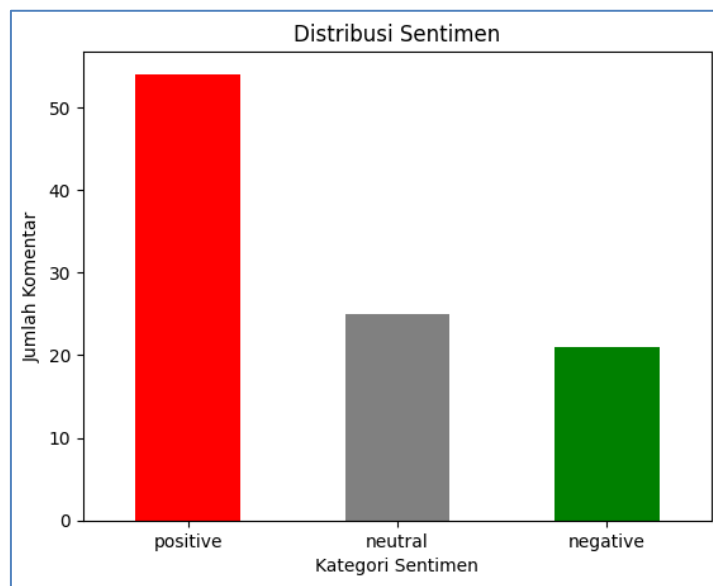
Sementara itu, kelas positif memperoleh precision sebesar 1,00 dan recall sebesar 0,92. Hal ini menunjukkan bahwa hampir seluruh prediksi positif yang dihasilkan model adalah benar, namun masih terdapat sebagian kecil data positif yang tidak terklasifikasi dengan tepat. Nilai f1-score pada ketiga kelas berada pada rentang tinggi, yaitu antara 0,89 hingga 0,98, yang menunjukkan keseimbangan antara precision dan recall.

Analisis confusion matrix memberikan gambaran lebih lanjut mengenai distribusi kesalahan klasifikasi. Seluruh data negatif (17) dan netral (24) berhasil diklasifikasikan dengan benar tanpa kesalahan. Namun, pada kelas positif terdapat beberapa kesalahan klasifikasi, yaitu 4 data positif yang diprediksi sebagai negatif dan 1 data sebagai netral. Hal ini menunjukkan bahwa model masih mengalami sedikit kesulitan dalam membedakan beberapa komentar positif yang memiliki nuansa mirip dengan sentimen lain.



Gambar 2. Confusion Matrix

Selain performa model, distribusi dataset juga menjadi faktor penting dalam hasil penelitian ini. Data menunjukkan bahwa jumlah komentar positif (15.518) jauh lebih dominan dibandingkan netral (9.064) dan negatif (7.602). Ketidakseimbangan ini berpotensi memengaruhi kecenderungan model untuk lebih mudah mengenali kelas mayoritas, meskipun dalam penelitian ini model tetap mampu mempertahankan performa yang baik pada semua kelas.



Gambar 3. Hasil Evaluasi Model Dalam Distribusi Sentimen Pada Komentar

Dari sisi linguistik, analisis menunjukkan bahwa komentar negatif cenderung menggunakan bahasa yang emosional, kritis, dan langsung, sering kali mengandung kata-kata dengan konotasi kuat. Sebaliknya, komentar netral lebih bersifat deskriptif dan informatif tanpa menunjukkan sikap yang jelas. Komentar positif didominasi oleh ekspresi dukungan, apresiasi, atau persetujuan terhadap kebijakan yang dibahas.

Temuan ini menunjukkan bahwa model *RoBERTa* tidak hanya efektif dalam klasifikasi sentimen, tetapi juga mampu menangkap nuansa bahasa yang kompleks dalam komentar publik. Kemampuan ini menjadi keunggulan utama dibandingkan metode tradisional yang cenderung terbatas dalam memahami konteks.



Gambar 4. Word Cloud

Secara keseluruhan, hasil penelitian ini menegaskan bahwa pendekatan berbasis transformer, khususnya *RoBERTa*, sangat relevan untuk analisis opini publik dalam konteks kebijakan. Model ini dapat digunakan sebagai alat bantu dalam memahami persepsi masyarakat secara lebih sistematis dan berbasis data.

Namun demikian, penelitian ini masih memiliki keterbatasan, terutama terkait dengan distribusi data yang tidak seimbang dan ukuran data uji yang relatif kecil. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan seimbang, serta mengeksplorasi teknik penanganan imbalance seperti oversampling atau class weighting.

Dengan demikian, hasil dan pembahasan ini memberikan gambaran komprehensif mengenai performa model serta dinamika linguistik dalam opini publik, yang dapat menjadi dasar bagi pengembangan penelitian lanjutan maupun implementasi praktis dalam analisis kebijakan.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model RoBERTa mampu memberikan kinerja yang sangat baik dalam melakukan klasifikasi sentimen terhadap komentar publik terkait kebijakan militer. Model ini berhasil mencapai tingkat akurasi sebesar 95%, dengan nilai precision, recall, dan f1-score yang tinggi pada masing-masing kelas sentimen. Hal ini menunjukkan bahwa pendekatan berbasis transformer efektif dalam memahami konteks bahasa yang kompleks dalam data teks. Selain itu, hasil analisis menunjukkan bahwa distribusi sentimen pada data uji didominasi oleh sentimen positif, diikuti oleh netral dan negatif. Meskipun terdapat ketidakseimbangan data, model tetap mampu mengklasifikasikan ketiga kategori sentimen dengan baik. Namun, masih ditemukan sedikit kesalahan klasifikasi pada kelas positif yang menunjukkan adanya kemiripan konteks dengan kelas lain. Dari sisi linguistik, komentar negatif cenderung mengandung ekspresi emosional dan kritik langsung, sementara komentar netral bersifat informatif, dan komentar positif menunjukkan dukungan terhadap kebijakan. Hal ini mengindikasikan bahwa model tidak hanya mampu mengklasifikasikan sentimen, tetapi juga menangkap karakteristik bahasa yang digunakan oleh masyarakat. Dengan demikian, penelitian ini membuktikan bahwa penggunaan model RoBERTa dalam analisis sentimen dapat menjadi solusi yang efektif untuk memahami opini publik secara otomatis dan berbasis data. Ke depan, penelitian dapat dikembangkan dengan menggunakan dataset yang lebih besar, metode penyeimbangan data, serta eksplorasi model lain untuk meningkatkan performa dan generalisasi hasil.

Daftar Rujukan

- [1] A. F. R. Harahap and A. M. Harahap, "Peran digitalisasi dalam meningkatkan partisipasi publik pada pengambilan keputusan tata negara," *Jurnal Educatio (Jurnal Pendidikan Indonesia)*, vol. 9, no. 2, pp. 769–776, 2023, doi: 10.29210/1202323208.
- [2] M. F. R. Azis, S. Angeline Lusman, and I. A. Indriyany, "Kebebasan Ekspresi Digital Perspektif Gender dalam Kasus Laras Faizati," *JISoH*, vol. 2, no. 1, pp. 939–950, Mar. 2026, doi: 10.63822/rxhtpm74.
- [3] N. Nurcholilah, S. Suherman, and A. Anshor, "Analisis Sentimen Masyarakat Pada Twitter Terhadap Kasus HIV/AIDS Di Indonesia Menggunakan Algoritma Naive Bayes," *jiwp*, vol. 9, no. 14, pp. 584–599, Jul. 2023.
- [4] D. Tribuana, U. Usman, dan D. Dayanti, "Penerapan Natural Language Processing Untuk Analisis Sentimen Terhadap Layanan Publik Di Media Sosial Twitter," *JTBC*, vol. 1, no. 1, hlm. 28–37, Jul 2025.
- [5] S. Butsianto and A. M. Rifa'i, "Analisis Sentimen Ulasan Aplikasi Jamsostek dengan SVM, Random Forest, dan Logistic Regression," *INFEB*, vol. 7, no. 3, pp. 700–706, Sep. 2025.
- [6] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan Lexicon-Based pada Performa IndoBERT: Sentiment Analysis of Coretax: A Comparison of Manual, Transformers-Based, and Lexicon-Based Data Labeling on IndoBERT Performance", *MALCOM*, vol. 5, no. 3, pp. 1037-1048, Jul. 2025.
- [7] J.J. Hidayat, C. Setyowati, M.D.I Amin, K. Bimasakti, and A.P. Werdana, "Deep Learning-based Sentiment Analysis of Public Comments on Military Education Using RoBERTa Algorithm and Rule-Based Hybrid Parameters", *jmcs*, vol. 4, no. 2, pp. 277-292, Jul. 2025.
- [8] U. I. Shabrina, M. I. Java, and S. Rochimah, "OPTIMIZING SENTIMENT ANALYSIS IN EDUCATIONAL YOUTUBE VIDEOS: A COMPARATIVE STUDY OF ROBERTA AND MULTINOMIAL NAIVE BAYES", *JUTI*, vol. 22, no. 2, pp. 83–90, Jul. 2024, doi: 10.12962/j24068535.v22i2.a1204.
- [9] [2] B. O. Lubis, B. Firmansyah, H. Putra, S. Irawan, A. N. Alizah, M. I. Adnan, and D. M. Situmorang, "Public Sentiment Analysis on the Development of the Nusantara Capital City (IKN) using the Naive Bayes Algorithm," *Jurnal Ilmiah Sistem Informasi*, vol. 5, no. 1, Jan. 2026, doi: 10.51903/nx1aza77.
- [10] Fajrul Khaer, Ilmiawan, and A. Aba, "Analyze the Language Used in Political News and Its Implications for Public Opinion: Analisis Bahasa yang Digunakan Dalam Berita Politik, Serta Implikasinya Terhadap Opini Publik", *J. Riseta Soshum*, vol. 2, no. 1, pp. 1–6, May 2025, doi: 10.70392/jrs.v2i1.0106.
- [11] A. Anshor, "Analisa Sentimen Warganet Terhadap KTT G20 Bali Menggunakan Algoritma Naive Bayes," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 11, no. 3, pp. 819–828, 2022. [Online]. Available: <https://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/article/view/1030>
- [12] J. J. Hidayat, C. Setyowati, and A. P. Werdana, "Sentiment Analysis of Instagram User Comments related to the Inauguration of Mr. Prabowo Subianto as President of the Republic of Indonesia Using Natural Language Processing", *Int. J. Data. Science.*, vol. 6, no. 2, pp. 94–102, Dec. 2025.
- [13] F. Zangari, R. Cimmaruta, and G. Nascetti, "Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae)," *Biological Journal of the Linnean Society*, vol. 87, no. 4, pp. 515–536, Apr. 2006, doi: 10.1111/j.1095-8312.2006.00585.x.
- [14] A. Halim and A. Safuwani, "Analisis sentimen opini warganet Twitter terhadap tes screening Genose pendeteksi virus Covid-19 menggunakan metode Naive Bayes berbasis particle swarm optimization," *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 5, no. 1, pp. 170–178, 2023.
- [15] W. Wiyanto and Z. Setyaningsih, "Sentiment Analysis Pemutusan Hubungan Kerja Akibat Pandemi Covid-19 Menggunakan Algoritma Naive Bayes Dan PSO", *SISFOKOM*, vol. 10, no. 3, pp. 426–431, Dec. 2021.
- [16] E. Widodo, E. P. Rachmawati, and P. Buana, "Pengaruh konten YouTube terhadap penerimaan mahasiswa baru: studi analisis sentimen BERT dan korelasi Spearman," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 7, no. 2, pp. 186–194, 2025.
- [17] I. A. Putri, A. Firmansyah, and A. Suprianto, "Analisis sentimen peran artificial intelligence terhadap kreativitas dan efektivitas

- mahasiswa dalam penyelesaian tugas akhir menggunakan decision tree berbasis SMOTE,” JATI (Jurnal Mahasiswa Teknik Informatika), vol. 9, no. 3, pp. 4909–4916, 2025.
- [18] J. J. Hidayat, C. Setyowati, and A. P. Werdana, “Perancangan Sistem Prediksi Penyakit pada Tanaman Padi Berbasis Image Processing Menggunakan Algoritma VGG-16 Transfer Learning dan K-Means Segmentation”, *J. Pract. Computer Sci.*, vol. 5, no. 1, pp. 1-15, May 2025.
- [19] M. D. I. Amin, J. J. Hidayat, C. Setyowati, E. K. Fitri, A. N. Anggraini, and A. P. Werdana, “Implementasi Model LSTM Untuk Peramalan Curah Hujan Di Bekasi Dengan Pemanfaatan Data Cuaca BMKG”, *JTID*, vol. 1, no. 2, pp. 90–99, Dec. 2025, Accessed: Apr. 01, 2026. [Online]. Available: <https://jurnal.ipdig.id/index.php/jtid/article/view/200>
- [20] D. S. Rizkiyan, D. A. Kurnia, Y. A. Wijaya, N. Rahaningsih, and W. Prihartono, “Optimalisasi akurasi model deep learning pada klasifikasi sampah plastik menggunakan strategi fine-tuning MobileNetV2,” *Journal of Computer Science and Artificial Intelligence (JCSAI)*, vol. 2, no. 6, Feb. 2026.
- [21] J. J. Hidayat, “Prediksi Diabetes Menggunakan Deep Neural Network dengan Penyesuaian Hiperparameter Berbasis Bayesian Optimization”, *J. Pract. Computer Sci.*, vol. 5, no. 2, pp. 130-143, Jan. 2026.
- [22] A. Z. Kamalia, Choiriyatun Nisa Latansa, and Zaenur Rozikin, “Klasifikasi Kondisi Pasar Harga Emas ANTAM Indonesia Menggunakan Algoritma Decision Tree ”, *JUKTISI*, vol. 4, no. 3, pp. 2087–2098, Jan. 2026.
- [23] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada Twitter,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [24] J. J. Hidayat, M. D. I. Amin, E. K. Fitri, A. N. Anggraini, A. P. Werdana, and C. Setyowati, “Implementasi Model EfficientNetB0 Pada Pembuatan Aplikasi Desktop Untuk Identifikasi Hama Tanaman Sawi Berbasis Deep Learning”, *JTID*, vol. 1, no. 2, pp. 82–89, Nov. 2025, Accessed: Apr. 01, 2026. [Online]. Available: <https://jurnal.ipdig.id/index.php/jtid/article/view/195>
- [25] A. M. Rifai, S. Raharjo, E. Utami, and D. Ariatmanto, “Analysis for diagnosis of pneumonia symptoms using chest X-ray based on MobileNetV2 models with image enhancement using white balance and contrast limited adaptive histogram equalization (CLAHE),” *Biomed. Signal Process. Control*, vol. 90, p. 105857, Apr. 2024, doi: 10.1016/j.bspc.2023.105857.