

Implementation of the Multilingual Sentiment Method in Public Sentiment Perspectives Toward Education Policy Directions in Indonesia

Fajar Widiyanto*

Universitas Pamulang, Tangerang Selatan, Indonesia

*Corresponding author: fajarwidiyanto95@gmail.com

Abstract

The current development of information technology has increased the dynamics of public opinion regarding the direction of education policy in Indonesia, which is spread across various digital platforms such as social media and news portals. The linguistic diversity used by the public, including Indonesian, regional languages, and mixed foreign languages, poses a challenge for conventional sentiment analysis, which is generally monolingual. This study aims to implement the Multilingual Sentiment Method to analyze and map public sentiment perspectives toward education policies in Indonesia more comprehensively and accurately. The research methodology encompasses data collection utilizing web scraping and APIs, data preprocessing (cleaning, case folding, tokenization, stopword removal, and normalization), automatic labeling, and the application of Natural Language Processing (NLP) using multilingual pre-trained models. Model performance evaluation is conducted using accuracy, distribution, precision, and prediction metrics. Data visualization analysis is also performed to illustrate sentiment distribution and emerging issue trends. The results indicate that the multilingual approach successfully improves classification performance compared to monolingual methods, particularly in detecting sentiments in mixed-language texts. The distribution of public sentiment is dominated by neutral and positive categories, whereas negative sentiments emerge regarding issues of technical implementation and equitable access to education. In conclusion, the implementation of the Multilingual Sentiment Method is effective in providing real-time mapping of public opinion, thereby supporting the formulation of education policies that are more responsive and data-driven.

Keywords: Education, Sentiment Analysis, Multilingual Sentiment, Orange

Article History:

Received : 18 January, 2026

Revised : 03 March, 2026

Accepted : 31 August, 2026

© 2026 Widiyanto. Published by Program
Studi Statistika Fakultas MIPA Universitas
Lambung Mangkurat.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



How to Cite:

F. Widiyanto, "Implementation of the Multilingual Sentiment Method in Public Sentiment Perspectives Toward Education Policy Directions in Indonesia," *RAGAM: Journal of Statistics and Its Application*, vol. 5, no. 2, pp. 82-94, 2026, doi: 10.20527/ragam.v5i2.18091.

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah membawa perubahan signifikan dalam cara masyarakat menyampaikan opini dan pandangannya terhadap kebijakan pemerintah, termasuk di bidang pendidikan. Media sosial seperti twitter, facebook, dan instagram dan youtube menjadi wadah bagi publik untuk mengekspresikan pendapat, baik dalam bentuk dukungan maupun kritik terhadap arah kebijakan pendidikan yang diterapkan oleh pemerintah. Besarnya volume data opini publik ini membuka peluang besar bagi penelitian berbasis data mining dan machine learning untuk menggali informasi dan pola sentimen masyarakat secara sistematis.

Kebijakan pendidikan yang diambil oleh pemerintah, seperti implementasi Kurikulum Merdeka, sistem zonasi sekolah, hingga digitalisasi pembelajaran pasca pandemi COVID-19, sering kali menimbulkan beragam tanggapan publik. Analisis terhadap sentimen publik ini penting dilakukan agar pemerintah dan pemangku kebijakan dapat memahami sejauh mana kebijakan tersebut diterima, serta mengidentifikasi aspek yang perlu diperbaiki. Namun, opini publik di media sosial memiliki sifat yang tidak terstruktur dan beragam, sehingga dibutuhkan metode analisis yang efektif untuk mengklasifikasikan sentimen tersebut ke dalam kategori positif, negatif, atau netral.

Meskipun kajian mengenai analisis sentimen terhadap kebijakan pemerintah telah banyak dilakukan, terdapat celah krusial yang belum terjamah secara optimal dalam literatur saat ini. Sebagian besar penelitian terdahulu cenderung bersifat monolingual, di mana model hanya dilatih untuk mengenali teks dalam Bahasa Indonesia baku. Faktanya, masyarakat Indonesia dalam mengekspresikan opini mengenai Pendidikan terutama di media sosial sering kali menggunakan campur kata antara Bahasa Indonesia dan Inggris serta penggunaan dialek bahasa daerah.

Salah satu metode yang dipakai dalam analisis sentimen ini adalah *multilingual sentiment*, merupakan suatu proses dalam analisis sentimen yang dirancang untuk mengidentifikasi, mengklasifikasikan opini atau emosi (positif, negatif, netral) dalam teks yang ditulis dalam lebih dari satu Bahasa [1], research gap dalam penelitian ini sebagian besar analisis sentimen kebijakan di Indonesia hanya berfokus pada Monolingual (Bahasa Indonesia saja). Padahal, opini publik di media sosial sering kali bersifat *code-switching* (campur kode) atau menggunakan dialek daerah, Oleh karena itu, penelitian ini hadir untuk mengisi ruang kosong tersebut dengan mengimplementasikan metode yang mampu mengolah keberagaman bahasa guna mendapatkan potret persepsi publik yang lebih inklusif dan representatif.

Melalui pendekatan ini, diharapkan hasil penelitian dapat memberikan gambaran nyata mengenai persepsi masyarakat terhadap kebijakan pendidikan yang sedang berlangsung. Selain itu, hasil analisis dapat menjadi masukan bagi pemerintah dalam merumuskan kebijakan pendidikan yang lebih adaptif dan responsif terhadap kebutuhan masyarakat. Selain itu, penelitian ini juga secara praktis berkontribusi pada pengembangan penerapan *machine learning* di bidang pendidikan dan kebijakan publik, khususnya dalam konteks analisis data teks. Implementasi *Multilingual Sentiment* dalam konteks sosial ini tidak hanya memberikan manfaat akademik, tetapi juga praktis bagi lembaga pemerintah, akademisi, dan masyarakat umum dalam memahami dinamika opini publik secara komprehensif.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Sentimen analysis atau opinion mining merupakan teknik komputasional untuk menemukan, mengekstrak, dan mengukur opini, emosi, atau sikap dalam teks untuk

menentukan polaritas suatu pernyataan (positif, negatif, atau netral). Menurutnya, sentimen analysis adalah salah satu anak cabang Natural Language Processing (NLP) yang digunakan luas untuk memahami opini publik terhadap masalah sosial maupun kebijakan [2]. Sedangkan analisis sentimen juga dapat diartikan sebagai proses otomatis untuk mengkategorikan teks berdasarkan opini yang terkandung di dalamnya menggunakan algoritma klasifikasi seperti Naïve Bayes, SVM, atau *Deep Learning* fokus utamanya penelitian ini adalah bagaimana model pembelajaran mesin dapat digunakan untuk mengklasifikasikan sentimen dari dataset multibahasa tanpa membangun model terpisah untuk setiap bahasa [3]. Analisis sentimen merupakan proses analisis opini publik untuk mengetahui arah pandangan masyarakat terhadap isu tertentu dengan bantuan algoritma klasifikasi teks [4]

Berdasarkan pandangan para ahli di atas, dapat disimpulkan bahwa analisis sentimen merupakan model metode berbasis Natural Language Processing (NLP) dan machine learning untuk mengidentifikasi dan mengukur opini publik yang terkandung dalam teks. Dalam konteks penelitian ini, analisis sentimen digunakan untuk memahami persepsi masyarakat terhadap arah kebijakan pendidikan dengan pendekatan *Multilingual Sentiment Method*.

2.2 Multilingual Sentiment

Multilingual sentiment analysis merupakan suatu proses dalam analisis sentimen yang dirancang untuk mengidentifikasi, mengklasifikasikan opini atau emosi (positif, negatif, netral) dalam teks yang ditulis dalam lebih dari satu bahasa. Hal ini melibatkan teknik Natural Language Processing (NLP) serta algoritma pembelajaran mesin (machine learning) untuk menangani keragaman bahasa dan budaya dalam data teks [1].

Multilingual sentiment memiliki kelebihan diantaranya model ini mampu menganalisis opini dari berbagai bahasa secara bersamaan sehingga menghasilkan gambaran sentimen yang lebih representatif, adaptif Karena interaksi digital bersifat multibahasa, metode ini lebih sesuai untuk analisis opini publik modern serta fleksibilitas dalam konteks global untuk analisis media sosial yang sering mengandung campuran bahasa (*code-mixing*)[3].

2.3 Kebijakan Pendidikan

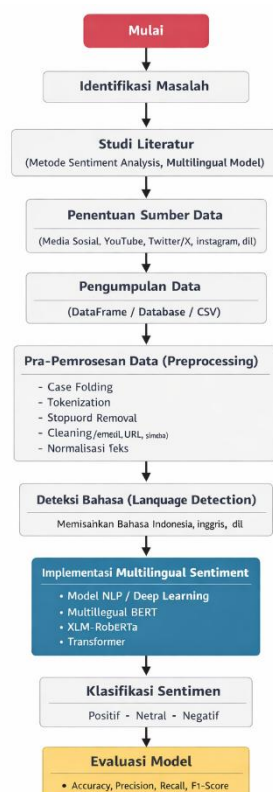
Keberhasilan kebijakan pendidikan dapat diukur melalui persepsi publik terhadap kebijakan tersebut, karena opini masyarakat berpengaruh terhadap implementasi dan legitimasi kebijakan. Oleh karena itu, analisis sentimen publik terhadap kebijakan pendidikan menjadi penting untuk menilai tingkat penerimaan masyarakat terhadap arah kebijakan pemerintah [5]. kebijakan pendidikan bukan hanya instrumen administratif, tetapi juga strategi sosial untuk menghadapi tantangan global seperti digitalisasi, pemerataan akses pendidikan, dan relevansi kurikulum. Ia menyatakan bahwa kebijakan pendidikan harus adaptif terhadap perubahan zaman, terutama dalam menghadapi era Society 5.0[6].

Selain itu, kebijakan pendidikan harus dipahami sebagai strategi sosial yang responsif terhadap dinamika global, seperti digitalisasi, pemerataan akses, dan penyesuaian kurikulum dengan kebutuhan zaman. Dalam konteks era Society 5.0, kebijakan pendidikan dituntut adaptif, inovatif, dan berbasis teknologi agar mampu menciptakan sumber daya manusia yang kompetitif dan berdaya saing tinggi. Dengan demikian, integrasi antara evaluasi berbasis data melalui analisis sentimen publik perumusan kebijakan yang adaptif menjadi kunci dalam mewujudkan sistem pendidikan yang inklusif, relevan, dan berkelanjutan.

3. METODE PENELITIAN

Metode penelitian ini merupakan komponen pendukung utama yang berperan dalam memastikan proses penelitian berjalan secara sistematis, terarah, dan ilmiah. Metode tidak hanya berfungsi sebagai pedoman dalam pengumpulan data, tetapi juga menjadi landasan dalam menentukan sumber penelitian, teknik pengolahan data, tahapan penelitian, hingga metode pengembangan model yang digunakan. Dengan adanya metode yang jelas dan terstruktur, penelitian dapat menghasilkan temuan yang valid, reliabel, dan dapat dipertanggungjawabkan secara akademi.

3.1 Alur Penelitian



Gambar 1. *Flowchart* Penelitian

Penelitian ini diawali dengan identifikasi masalah mengenai pandangan masyarakat terhadap arah kebijakan pendidikan di Indonesia yang banyak disampaikan melalui

media sosial. Selanjutnya dilakukan studi literatur untuk mempelajari teori terkait *Natural Language Processing* (NLP), sentiment analysis, dan model multilingual sebagai dasar dalam penelitian, lalu penentuan sumber data yang berasal dari media sosial seperti YouTube. Data kemudian dikumpulkan melalui proses *crawling* atau *scraping* menggunakan API dan bahasa pemrograman *Python*, lalu disimpan dalam bentuk *dataset* seperti DataFrame atau file CSV. dilakukan pra-pemrosesan data preprocessing hingga implementasi metode hinhha didapat hasil evaluasi model menggunakan metrik seperti accuracy, precision, recall, dan F1-score. Selanjutnya dilakukan analisis dan visualisasi data. Dan Tahap akhir penelitian adalah penarikan kesimpulan dan rekomendasi.

3.2 Sumber Data

Sumber data dalam penelitian ini merupakan komentar publik yang membahas tentang arah kebijakan Pendidikan yang diambil dari media sosial youtube <https://www.youtube.com/watch?v=1dK0WuqXZ4A>, Teknik pengumpulan data dilakukan dengan *web scripping Google Colaboratory*.

3.3 Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan *multilingual sentiment method* machine learning *Software Orange* data mining dalam visualisasi data, Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data numerik hasil transformasi teks dan pengukuran kinerja algoritma klasifikasi sentimen secara statistik. Metode analisis sentimen digunakan untuk mengidentifikasi opini publik terhadap arah kebijakan Pendidikan.

3.4 Jenis dan Pendekatan Penelitian

Tahapan ini bertujuan untuk menghasilkan data teks yang lebih terstruktur dan siap digunakan dalam proses klasifikasi. Data teks dilakukan melalui beberapa tahapan pra-pemrosesan untuk meningkatkan kualitas data, meliputi:

- *Case folding* digunakan mengubah seluruh huruf menjadi huruf kecil.
- *Cleaning* digunakan menghapus tanda baca, angka, URL, emoji, dan karakter khusus.
- *Tokenizing* digunakan memecah teks menjadi unit-unit kata (token) agar dapat dianalisis lebih lanjut
- *Stopword removal*, digunakan menghapus kata-kata umum yang tidak memiliki makna sentimen.
- *Stemming*, yaitu mengubah kata ke bentuk dasar untuk menyederhanakan variasi kata dan meningkatkan konsistensi fitur

Setelah tahap pra-pemrosesan selesai, proses dilanjutkan ke tahap analisis sentimen. Pada tahap ini, teks yang telah dibersihkan direpresentasikan ke dalam bentuk numerik menggunakan teknik ekstraksi fitur seperti TF-IDF atau *word embedding*. Selanjutnya, model klasifikasi berbasis *machine learning* atau *deep learning* diterapkan untuk mengelompokkan teks ke dalam kategori sentimen seperti positif, negatif, atau netral. Proses ini juga mencakup pembagian data menjadi data latih dan data uji, pelatihan model, serta evaluasi performa menggunakan metrik seperti akurasi, presisi, distribusi. Dengan tahapan tersebut, analisis sentimen dapat dilakukan secara sistematis untuk menghasilkan klasifikasi yang lebih akurat dan representatif terhadap opini publik.

3.5 Jenis dan Pendekatan Penelitian

Hasil analisis untuk mengetahui pola sentimen publik terhadap arah kebijakan pendidikan dengan *software orange data mining*. Analisis dilakukan dengan melihat distribusi sentimen serta mengaitkannya dengan isu-isu kebijakan yang sedang berkembang. Hasil analisis ini diharapkan dapat menjadi masukan bagi pemangku kepentingan dalam merumuskan kebijakan pendidikan yang lebih responsif terhadap opini publik.

4. HASIL DAN PEMBAHASAN

4.1 *Scripting Data*

1. *Import Library*

Import library merupakan langkah awal yang digunakan untuk pengolahan dan analisis data, khususnya data berbentuk tabel (DataFrame).

2. *Crawling Data*

Fungsi *crawling* komentar digunakan untuk mengumpulkan data opini publik dari *platform* YouTube dengan memanfaatkan YouTube Data API v3. Proses ini dilakukan secara otomatis berdasarkan ID video tertentu, sehingga diperoleh data komentar yang relevan sebagai sumber data utama dalam analisis sentimen terhadap arah kebijakan pendidikan.

Pada proses *crawling* komentar dilakukan dengan memanfaatkan YouTube Data API menggunakan API Key yang valid. Data komentar diambil berdasarkan ID video tertentu yang relevan dengan topik kebijakan pendidikan. Komentar yang berhasil dikumpulkan selanjutnya digunakan sebagai sumber data utama dalam analisis sentimen publik.

Hasil *crawling* komentar YouTube kemudian diubah ke dalam bentuk DataFrame menggunakan library *Pandas* dengan menampilkan 10 sampel data awal. DataFrame ini menyajikan informasi komentar secara terstruktur yang terdiri dari waktu publikasi, nama penulis, isi komentar, dan jumlah *like*. Struktur tabular tersebut memungkinkan setiap atribut data tersusun dalam kolom yang sistematis sehingga memudahkan proses eksplorasi dan analisis awal (*exploratory data analysis*). Selanjutnya, data tersebut diekspor ke dalam format CSV agar dapat disimpan, dibagikan, dan diproses lebih lanjut pada tahapan penelitian berikutnya.

Secara interpretatif, transformasi ke dalam bentuk DataFrame memiliki peran penting dalam menjamin kualitas dan keterbacaan data. Dengan struktur yang jelas, peneliti dapat mengidentifikasi pola awal seperti distribusi waktu komentar, intensitas interaksi berdasarkan jumlah *like*, serta kecenderungan isi opini yang berkembang. Selain itu, format CSV memberikan fleksibilitas untuk dilakukan proses pra-pemrosesan seperti pembersihan data duplikat, penanganan nilai kosong (*missing values*), serta normalisasi teks sebelum masuk ke tahap analisis sentimen.

Keberadaan atribut jumlah *like* juga dapat dimanfaatkan sebagai indikator tingkat keterlibatan (*engagement*) atau dukungan terhadap suatu opini, sehingga analisis tidak hanya berfokus pada polaritas sentimen, tetapi juga pada bobot pengaruh komentar tersebut. Dengan demikian, proses konversi data ke dalam DataFrame dan CSV bukan sekadar langkah teknis, melainkan tahap fundamental yang mendukung akurasi, efisiensi, dan kedalaman interpretasi dalam analisis sentimen publik.

Tabel 1. Hasil *Crawling*

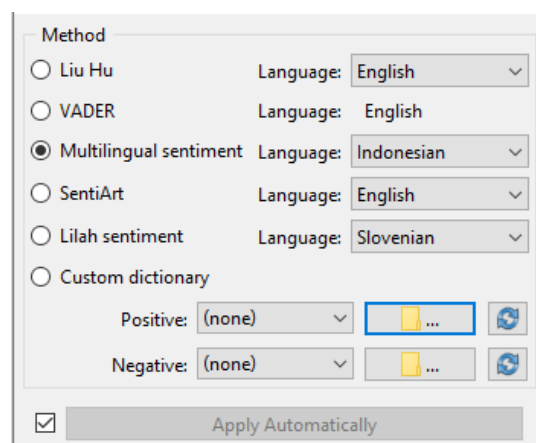
index	publishedAt	authorDisplayName	textDisplay	likeCount
0	2026-02-03T08:35:29Z	@FARUK-f1v	Tidak ada kemajuan pendidikan selagi kedaulatan ekonomi bagi tenaga pendidik yaitu guru di Indonesia allahhu akbar	0
1	2026-01-07T06:26:04Z	@KusriniKarangrejo	Semoga para mahasiswa yg masih tunda wisuda akan segera mendapatkan kebijakan baru terutama di UIN UIN d indonesia	0
2	2025-10-09T12:46:45Z	@KorneliaRedamata-p8h	Sudah hal yg. lumrah bg. Negeri kaya ini, Isi & tujuan Pendidikan tak pernah pudar yaitu memanusiakan manusia hingga berkualitas. Dan yang paling aneh adalah setiap ganti Menteri kerjanya hanya ganti label layaknya mi	0

			instan yg. terus diganti labrel kemasannya. Sistim yg. dibangun hanyalah industri para koruptor. Tak akan ada kelebihan sejak dari doeloe.	
3	2025-09-09T01:21:52Z	@haritzaryandra	semangat pak prof, semoga dengan bapak memimpin para pendidik, bangsa ini bisa menjadi bangsa yang tidak hanya terdepan dalam nilai negatof, tetapi bisa bersaing dalam kans global dengan bangsa lain, saya yakin anak2 kita memiliki otak yang ngga kalah canggih dengan bangsa lain, sampai eropa juga, tinggal pembimbing dan pendidiknya yang memiliki peranan sebagai ibarat pengganti orang tua diluar rumah dan pembekal dimasyarakat...	0
4	2025-08-29T07:56:42Z	@gatiday5730	BPK Mentri baru ini hanya membuat guru resa dn sibuk 🙄	0
5	2025-08-11T03:18:08Z	@IwanJawahir	Luar biasa semua kebijakan p.menteri,bagaimana dgn guru2x yg belum d ⁴ /s1,apa ada perhatian kusus dari pemerintah.tolong hal ini dipertimbangan dari p.menteri dan jajaranya..	0
6	2025-07-29T13:29:46Z	@utamaariwibawa3012	Kembalilan un, enak bener skrg otak moral gak nyampe nilai nya 9 njir 😄	0
7	2025-06-23T08:52:56Z	@justanjus5074	In syaa Allah semoga dunia Pendidikan lebih maju dengan berbagai macam program unggulannya	0
8	2025-06-20T12:54:12Z	@kasiammurtini8575	Emang btul pak jdi gru paud KLO lulusan SMA dengn lulusan KPGTk,dh beda jauh cara menghadapi murid yg masih 0,kbanyakn KLO lulusan SMA keminter....pak.	0
9	2025-06-20T12:54:10Z	@kasiammurtini8575	Emang btul pak jdi gru paud KLO lulusan SMA dengn lulusan KPGTk,dh beda jauh cara menghadapi murid yg masih 0,kbanyakn KLO lulusan SMA keminter....pak.	0

10	2025-05-28T00:43:19Z	@SwisNdua	Bapak mau tanya,bisa minta tolong kah,anak ku pintar di sekolah,SMA kelas 2 sekarang,bagaimana cara bisa dapat beasiswa,bisa sambung kuliah gratis pak 👉👉👉👉👉	0
----	----------------------	-----------	---	---

4.2 Implementasi *Software Orange*

Secara interpretatif, pada bagian metode ini terdapat beberapa pilihan metode analisis sentimen yang dapat digunakan, yaitu Liu Hu, VADER, Multilingual Sentiment, SentiArt, dan Lillah Sentiment. Masing-masing metode memiliki karakteristik dan kemampuan yang berbeda dalam menganalisis sentimen berdasarkan bahasa tertentu.



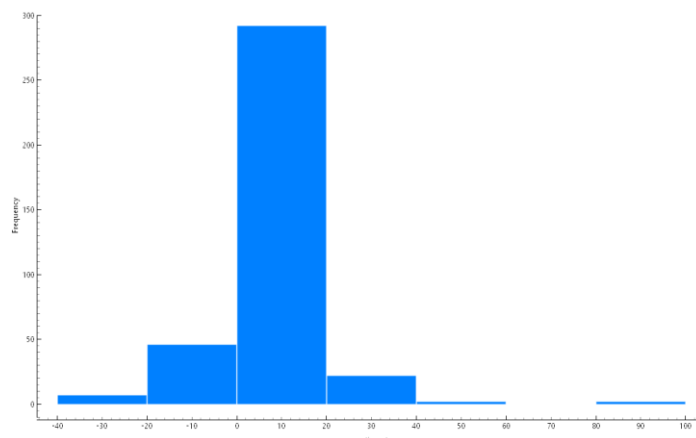
Gambar 2. Konfigurasi Analisis *Multilingual sentiment*

Pada gambar 2 terlihat bahwa metode yang dipilih adalah *Multilingual Sentiment* dengan bahasa Indonesian. Hal ini menunjukkan bahwa sistem akan menganalisis teks menggunakan model yang mampu memahami bahasa Indonesia serta kemungkinan variasi bahasa lain. Pemilihan metode ini sangat relevan ketika data yang dianalisis berasal dari komentar masyarakat Indonesia di media sosial. Metode lain seperti VADER dan Liu Hu umumnya digunakan untuk teks berbahasa Inggris, sedangkan SentiArt dan Lillah Sentiment mendukung bahasa tertentu lainnya. Oleh karena itu, pemilihan metode *Multilingual Sentiment* menunjukkan bahwa penelitian atau analisis difokuskan pada analisis sentimen berbasis multibahasa dengan penekanan pada bahasa Indonesia.

Gambar 3 menjelaskan bahwa visualisasi *word cloud* yang menampilkan kumpulan kata paling sering muncul dalam komentar terkait kebijakan pendidikan hasil komentar YouTube. Pada *word cloud*, ukuran kata menunjukkan tingkat frekuensi kemunculan semakin besar ukuran kata, semakin sering kata tersebut muncul dalam data *word cloud* ini. Visualisasi ini belum menunjukkan polaritas (positif/negatif), tetapi memberikan gambaran isu dominan yang menjadi perhatian masyarakat. Untuk mengetahui kecenderungan sentimen (positif, negatif, netral), diperlukan analisis klasifikasi lanjutan.

Tabel 2. Distribusi Sentimen

Kategori Sentimen	Jumlah	Persentase
Positif	290	80.6 %
Netral	45	12.5 %
Negatif	25	6.9 %
Total	360	100 %

**Gambar 4.** Frekuensi Distribusi Sentimen

Tabel 2 menunjukkan bahwa sebagian besar komentar atau opini yang dianalisis memiliki sentimen positif, yaitu sebanyak 290 data atau sekitar 80,6% dari total dataset. Hal ini menunjukkan bahwa mayoritas respon atau pandangan yang muncul dalam data penelitian cenderung memberikan tanggapan yang positif terhadap topik yang dianalisis. Sementara itu, sentimen netral berjumlah 45 data atau sekitar 12,5%, yang menunjukkan adanya komentar yang bersifat informatif atau tidak secara jelas menunjukkan sikap mendukung maupun menolak.

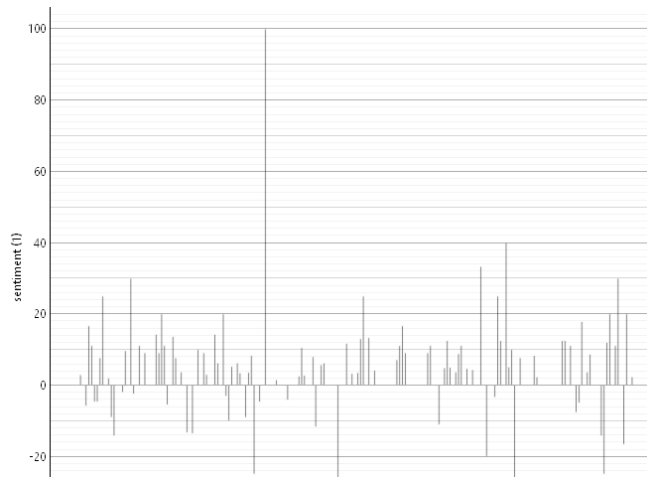
Adapun sentimen negatif berjumlah 25 data atau sekitar 6,9%, yang mencerminkan adanya sebagian kecil opini publik yang memberikan kritik atau pandangan kurang mendukung terhadap isu yang dibahas.

Berdasarkan distribusi tersebut dapat diinterpretasikan bahwa kecenderungan sentimen dalam dataset penelitian didominasi oleh sentimen positif, sedangkan sentimen negatif relatif lebih sedikit

Beberapa hal yang dapat dijelaskan dari gambar 4 adalah sebagai berikut:

- Sebagian besar nilai sentiment publik berada pada rentang -10 hingga 20, menunjukkan dominasi sentimen netral hingga positif ringan.
- Terdapat beberapa lonjakan (spike) positif yang sangat tinggi, bahkan mendekati 100. Ini menunjukkan adanya data dengan sentimen sangat positif (outlier).

- Ada pula beberapa nilai negatif ekstrem hingga sekitar -25, namun jumlahnya lebih sedikit dibanding lonjakan positif.
- Pola distribusi terlihat fluktuatif, artinya sentimen berubah-ubah antar data, tetapi tidak menunjukkan tren kenaikan atau penurunan yang konsisten



Gambar 5. Distribusi Sentimen

Dapat disimpulkan berdasarkan hasil evaluasi model, diperoleh nilai accuracy sebesar 80.6%, yang menunjukkan bahwa model memiliki kinerja yang cukup baik dalam mengklasifikasikan sentimen komentar. Model memiliki *precision* dan *recall* yang tinggi pada kategori sentimen positif, sehingga mampu mengidentifikasi komentar positif dengan bahwa opini publik terhadap kebijakan pendidikan di Indonesia sangat beragam, dengan dominasi sentimen positif, negatif, dan netral yang tersebar pada berbagai komentar. Hal ini menunjukkan bahwa kebijakan pendidikan memperoleh respon yang berbeda-beda dari masyarakat, sehingga analisis sentimen multibahasa.

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan bahwa *Multilingual Sentiment Analysis* terbukti efektif untuk diterapkan dalam menemukan dan menggambarkan pandangan sentimen publik terhadap reorientasi kebijakan. Pendekatan ini memungkinkan analisis teks berbahasa Indonesia secara lebih relevan karena memanfaatkan kamus sentimen yang sesuai dengan konteks bahasa dan kosakata lokal. Hasil akan analisis menunjukkan mayoritas pendapat publik cenderung memiliki sentimen yang netral ke positif. Namun demikian masih ditemukan beberapa pendapat yang memiliki sentimen negatif melalui kritisisme atau ketidakpuasan publik atas fasilitas tertentu yang dilakukan oleh kebijakan Pendidikan. Distribusi nilai sentimen yang terkonsentrasi di sekitar nilai netral hingga positif menunjukkan bahwa respons masyarakat tidak bersifat ekstrem, melainkan lebih rasional dan seimbang. Hal ini menegaskan bahwa kebijakan pendidikan yang dianalisis tidak menimbulkan penolakan publik secara luas, namun tetap memerlukan perhatian terhadap masukan dan evaluasi dari masyarakat.

DAFTAR PUSTAKA

- [1] A. Kumar and A. Jaiswal, "Sentiment Analysis and Opinion Mining: A Survey," *Journal of Information Systems and Technology Management*, vol. 18, no. 4, pp. 325–340, 2021
- [2] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, 2nd ed. Cambridge: Cambridge University Press, 2022.
- [3] Silitonga, P. D. P., Purba, L. I., Morina, I. S., & Rikki, A. (2023). *Multi-Language sentiment analysis using machine learning*. Jurnal Info Sains: Informatika dan Sains, 13(02), 167–174
- [4] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, 2nd ed. Cambridge: Cambridge University Press, 2021.
- [5] D. N. Putri and S. Rahmawati, "Analisis Sentimen terhadap Kebijakan Pemerintah Menggunakan Algoritma Naïve Bayes," *Jurnal Teknologi Informasi dan Sains*, vol. 12, no. 1, pp. 45–53, 2023
- [6] S. Nugroho, "Implementasi Kurikulum Merdeka dalam Kebijakan Pendidikan Nasional," *Jurnal Kebijakan dan Inovasi Pendidikan*, vol. 5, no. 3, pp. 85–94, 2023
- [7] R. Rahardjo, "Kebijakan Pendidikan di Era Society 5.0," *Jurnal Ilmu Pendidikan dan Kebijakan Publik*, vol. 8, no. 2, pp. 112–123, 2023
- [8] E. Cambria, B. Schuller, F. Xing, and C. Havasi, "Affective Computing and Sentiment Analysis: Trends and Challenges," *IEEE Transactions on Affective Computing*, vol. 14, no. 2, pp. 245–260, 2023.
- [9] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*, 4th ed. Burlington: Morgan Kaufmann, 2022.
- [10] Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek), *Laporan Kebijakan Pendidikan Nasional 2023*. Jakarta: Kemendikbudristek, 2023
- [11] L. Zhang and S. Wang, "A Comparative Study on Naïve Bayes Variants for Text Classification," *Expert Systems with Applications*, vol. 194, p. 116468, 2022.
- [12] N. Rana and J. H. Cheah, "Understanding Public Sentiment Toward Government Policies Through Social Media Analytics," *Government Information Quarterly*, vol. 39, no. 5, p. 101731, 2022.
- [13] R. Sari and D. Prasetyo, "Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen Bahasa Indonesia pada Media Sosial," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 10, no. 2, pp. 120–128, 2023.
- [14] UNESCO, *Global Education Monitoring Report 2024: Learning for Lasting Change*. Paris: UNESCO Publishing, 2024
- [15] Widiyanto, F. (2025). Analisis Sentimen Komentar Youtube tentang Konflik Iran-Israel Menggunakan Orange Data Mining. *Sains Data Jurnal Studi Matematika dan Teknologi*, 3(2), 81-88.
- [16] Y. Yin and H. Li, "Efficient Text Classification Using Improved Naïve Bayes for Social Media Data," *Applied Intelligence*, vol. 54, no. 3, pp. 2850–2863, 2024
- [17] F. Widiyanto, "Analisis Orange Data Mining Untuk Klasifikasi Penyakit Diabetes Menggunakan Model Decision Tree," vol. 19, no. 02, 2025.