

Analisis Sentimen Masyarakat terhadap Isu Redenominasi Rupiah pada *Platform YouTube* Menggunakan Algoritma *Naïve Bayes*

Alfieny Putri Heriyanto¹, Nono Heryana², Apriade Voutama³

Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

E-mail Korespondensi : 1alfienyputrih@gmail.com, 2nono@unsika.ac.id 3apriade.voutama@staff.unsika.ac.id

History Artikel

Diterima : 6 November 2025

Disetujui : 19 Februari 2026

Dipublikasikan : 26 April 2026

Abstract

Abstract. The issue of rupiah redenomination is one of the economic topics widely discussed by the public and has generated various opinions in society. The growth of social media, especially YouTube, allows users to express their views openly through the comment section. This study aims to analyze public sentiment toward the rupiah redenomination issue on YouTube using the Naïve Bayes Classifier algorithm. The research applies the Knowledge Discovery in Database (KDD) method, consisting of data selection, preprocessing, transformation, data mining, and Evaluation. The Inset Lexicon approach was employed to assign sentiment labels using positive and negative dictionaries, alongside manual labeling validated by a linguistics expert. The findings show that both labeling methods are dominated by positive sentiment. Inset Lexicon labeling identified 2,157 positive, 1,562 negative, and 1,600 neutral data points. Meanwhile, manual labeling resulted in 2,784 positive, 1,214 negative, and 1,214 neutral data points. Evaluation results reveal that Inset Lexicon labeling achieved better performance than manual labeling. Using SMOTE and a 90:10 data split ratio, the model reached 69% accuracy, 68% precision, and 68% recall, while manual labeling achieved 52% accuracy, 33% precision, and 33% recall.

Keywords: *Sentiment Analysis, Rupiah Redenomination, YouTube, Naïve Bayes Classifier*

Abstrak

Abstrak. Isu redenominasi rupiah merupakan salah satu topik ekonomi yang banyak dibahas oleh masyarakat dan menimbulkan berbagai opini di ruang publik. Perkembangan media sosial, khususnya *YouTube*, memungkinkan pengguna menyampaikan pandangan secara terbuka melalui kolom komentar. Penelitian ini bertujuan menganalisis sentimen publik terhadap isu redenominasi rupiah di *YouTube* menggunakan algoritma *Naïve Bayes Classifier*. Metode yang digunakan adalah *Knowledge Discovery in Database (KDD)* yang meliputi tahap *data selection*, *preprocessing*, *transformasi*, *data mining*, dan evaluasi. Pendekatan *Inset Lexicon* diterapkan untuk memberikan label sentimen berdasarkan kamus positif dan negatif, serta pelabelan manual yang divalidasi oleh ahli linguistik. Hasil penelitian menunjukkan bahwa kedua metode pelabelan didominasi sentimen positif. Pelabelan *Inset Lexicon* mengidentifikasi 2.157 data positif, 1.562 data negatif, dan 1.600 data netral. Sementara itu, pelabelan manual menghasilkan 2.784 data positif, 1.214 data negatif, dan 1.214 data netral. Hasil evaluasi menunjukkan bahwa pelabelan *Inset Lexicon* memiliki kinerja lebih baik dibandingkan pelabelan manual. Dengan penerapan SMOTE dan rasio pembagian data 90:10, model mencapai akurasi 69%, presisi 68%, dan *recall* 68%. Sebaliknya, pelabelan manual hanya mencapai akurasi 52%, presisi 33%, dan *recall* 33% dalam pengujian model.

Kata Kunci: *Analisis Sentimen, Redenominasi Rupiah, YouTube, Naïve Bayes Classifier*

How to Cite: Alfieny Putri Heriyanto (2026) Analisis Sentimen Masyarakat terhadap Isu Redenominasi Rupiah pada Platform *YouTube* Menggunakan Algoritma *Naïve Bayes*. *KOMPUTEK : Jurnal Teknik Universitas Muhammadiyah Ponorogo*, Vol 10 (1): Halaman 105-113

© 2026 Universitas Muhammadiyah Ponorogo. All rights reserved

ISSN 2614-0985 (Print)
ISSN 2614-0977 (Online)

PENDAHULUAN

Perkembangan teknologi informasi di era globalisasi mendorong masyarakat semakin aktif menyampaikan opini melalui media sosial (Sirait et al., 2024). Platform seperti *YouTube*, *Instagram*, *Facebook*, *WhatsApp*, *TikTok*, dan *X* menjadi ruang publik digital yang memungkinkan individu menyampaikan pandangan terhadap isu politik, sosial, ekonomi, maupun budaya (Alfifa et al., 2025). Berdasarkan data Datareportal (2025), *YouTube* termasuk platform dengan jumlah kunjungan bulanan tertinggi di Indonesia, sehingga menjadi sumber data yang relevan untuk menganalisis opini publik (datareportal, 2025). Kolom komentar pada *YouTube* merepresentasikan respons masyarakat secara langsung terhadap suatu isu yang sedang berkembang (Damayanti et al., 2023).

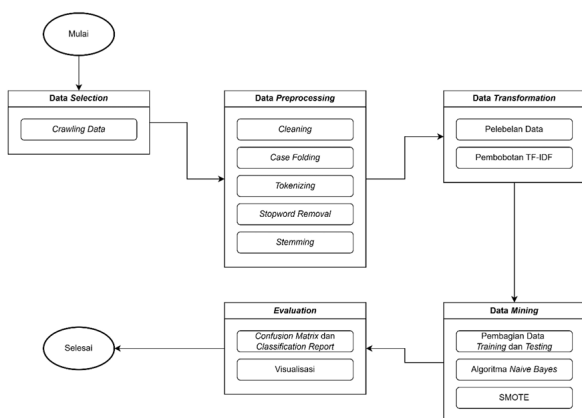
Salah satu isu yang ramai diperbincangkan adalah kebijakan redenominasi rupiah. Redenominasi merupakan penyederhanaan nilai nominal mata uang dengan mengurangi digit nol tanpa mengubah nilai riil atau daya beli. Kebijakan ini berbeda dengan sanering karena tidak mengurangi nilai uang, melainkan bertujuan meningkatkan efisiensi transaksi, mempermudah pencatatan keuangan, serta memperkuat citra rupiah di tingkat global (Dimas Waraditya Nugraha et al., 2025). Wacana ini telah muncul sejak 2010 dan kembali digulirkan pada 2025 melalui Rancangan Undang-Undang tentang Perubahan Harga Rupiah dengan target implementasi bertahap hingga 2029. Namun, keberhasilannya bergantung pada stabilitas ekonomi, kesiapan infrastruktur, literasi masyarakat, serta strategi komunikasi untuk mencegah kesalahpahaman seperti ilusi uang dan *praktik opportunistic rounding* (Diploma et al., 2023). Untuk memahami respons masyarakat terhadap isu tersebut, digunakan analisis sentimen sebagai bagian dari *Natural Language Processing* (NLP) (Fatah & Syarifah, 2025). Analisis sentimen bertujuan mengklasifikasikan opini teks ke dalam kategori positif, negatif, atau netral (Fikri et al., 2020).

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *Naïve Bayes* efektif dalam klasifikasi sentimen. Penelitian Rizal dan Hayati menganalisis sentimen ulasan film *anime One Piece* di platform *X* menggunakan *Multinomial Naïve Bayes* dan memperoleh akurasi 81,52%, *precision* 82,49%, *recall* 81,52%, dan *F1-score* 81,18% (Rizal & Hayati, 2024). Selain itu, penelitian Arita meneliti komentar *YouTube* terhadap *anime Spy x Family* di channel *Muse* Indonesia dan memperoleh hasil terbaik pada rasio data 80:20 dengan akurasi 86,03%, *precision* 81,08%, *recall* 98,90%, dan *F1-score* 89,11% (Arita, 2025). Sementara itu, Widodo, Fajri, dan Sari membandingkan *Naïve Bayes* dan *Support Vector Machine* (SVM) dalam analisis sentimen film animasi “JUMBO”, dan menunjukkan bahwa *Naïve Bayes* lebih unggul dengan akurasi 80,60%, *precision* 73,83%, *recall* 73,50%, dan *F1-score* 69,98% (Robi Widodo et al., 2025).

Meskipun berbagai penelitian telah membuktikan efektivitas *Naïve Bayes* dalam analisis sentimen, kajian yang secara khusus membahas sentimen publik terhadap isu redenominasi rupiah masih terbatas. Oleh karena itu, penelitian ini bertujuan menganalisis sentimen masyarakat terhadap isu redenominasi rupiah pada platform *YouTube* menggunakan algoritma *Naïve Bayes*. Untuk mengatasi ketidakseimbangan data sentimen, diterapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) agar distribusi kelas lebih seimbang dan performa klasifikasi meningkat (Sarumpaet & Suryono, 2025). Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian diharapkan dapat menggambarkan distribusi sentimen publik terhadap kebijakan redenominasi rupiah serta menjadi bahan pertimbangan bagi pemerintah dalam merancang strategi komunikasi publik yang lebih efektif.

METODE PENELITIAN

Penelitian ini menerapkan pendekatan eksperimen kuantitatif dengan menggunakan algoritma *Naïve Bayes Classifier* dengan pendekatan *Multinomial Naïve Bayes* untuk melakukan analisis sentimen. Pemilihan metode menggunakan *Knowledge Discovery of Data* (KDD) ini didasarkan pada kemampuannya dalam mengklasifikasikan data teks secara efektif dan efisien, khususnya pada data media sosial yang memiliki karakteristik tidak terstruktur.



Gambar 1. Metode Penelitian

a. Data Selection

Pada tahap ini, penelitian akan melakukan pencarian dan pengambilan data dari media sosial *YouTube*. Data akan dikumpulkan menggunakan teknik *web crawling* dengan *Python* untuk rentang waktu November hingga Desember 2025. Pengambilan data dilakukan dari enam video *YouTube* yang membahas topik redenominasi rupiah, dengan setiap video diambil sebanyak 1.000 komentar. Data yang digunakan berupa komentar berbahasa Indonesia yang memuat opini pengguna *YouTube* terkait isu tersebut. Selanjutnya, komentar-komentar yang ada akan diklasifikasikan ke dalam tiga kelas sentimen, yaitu sentimen positif, sentimen negatif, dan sentimen netral.

b. Data Preprocessing

Pada tahap ini akan melakukan *Preprocessing*, berupa data cleansing agar data mentah dan kotor yang diambil melalui komentar menjadi konsisten dan tidak ada *noise* (Anuar et al., 2023). Selain itu, ada pula beberapa tahap *preprocessing* yang lain, seperti berikut:

1. Cleaning

Cleaning akan menghilangkan beberapa data dalam dataset yang memerlukan penanganan melalui berbagai metode seperti menghapus bahasa selain Indonesia, data yang terduplikat, nilai yang hilang, dan penyesuaian tipe data (Aryanti et al., 2024).

2. Case Folding

Case Folding untuk mengubah seluruh huruf yang ada dalam dokumen menjadi huruf kecil. Tujuannya agar semua huruf sama dan mudah untuk diolah. Hanya karakter huruf dari 'a' hingga 'z' yang akan diterima. Karakter selain itu akan dibuang (Suhendra et al., 2024).

3. Tokenizing

Tokenizing akan membagi teks menjadi token atau kata-kata individual agar sesuai format yang diperlukan oleh model. Untuk memisahkan antar kata akan digunakan spasi (Aryanti et al., 2024).

4. Normalization

Normalization untuk mengubah kata atau frasa yang tidak baku atau tidak standar menjadi bentuk yang lebih konsisten dan sesuai dengan kaidah bahasa yang baku (Ramdhanian et al., 2025).

5. Stopword Removal

Stopword Removal akan menghilangkan kata-kata penghubung, tanda baca, URL, atau karakter khusus yang tidak diperlukan dalam analisis sentimen. Untuk dokumen berbahasa Indonesia, contoh dari kata-kata penghubung yaitu “di”, “dari”, “yang”, “dan” dan lain-lain (Ramdhanian et al., 2025).

6. Stemming

Stemming akan dilakukan untuk mengubah kata-kata dalam sebuah dokumen menjadi bentuk dasarnya dengan cara menghapus semua bagian tambahan seperti awalan, akhiran, sisipan, serta kombinasi dari awalan akhiran (Thaariq Razaq et al., 2023).

c. **Data Transformation**

Pada tahap ini, data komentar *YouTube* yang telah melalui proses *preprocessing* akan diubah menjadi representasi numerik agar dapat diolah oleh model. Proses ini diawali dengan pelabelan sentimen menggunakan *Inset Lexicon* based dalam format TSV dan pelabelan secara manual dengan validasi oleh Ahli Bahasa, di mana pada pelabelan dengan *Inset Lexicon* setiap kata pada komentar dicocokkan dengan daftar kata berkategori positif atau negatif untuk memberikan label sentimen awal. Setelah proses *Inset Lexicon* based selesai, komentar tersebut kemudian dibobot menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengukur pentingnya setiap kata dalam keseluruhan dokumen.

d. **Data Mining**

Pada tahap ini diterapkan teknik data *mining* menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan sentimen komentar *YouTube* terkait isu redenominasi rupiah. Untuk mengatasi ketidakseimbangan kelas, digunakan metode *Synthetic Minority Over-sampling Technique* (SMOTE) guna menambah data pada kelas minoritas secara sintesis sehingga model tidak bias dan mampu menghasilkan prediksi yang lebih seimbang. Penelitian ini termasuk dalam kategori *supervised learning* karena data telah diberi label sentimen menggunakan pendekatan *Inset Lexicon* dan pelabelan manual. Evaluasi model dilakukan menggunakan skema *train-test split* dengan empat skenario pembagian data untuk menganalisis pengaruh proporsi data pelatihan terhadap performa model. Skenario yang digunakan meliputi rasio 90:10, 80:20, 70:30, dan 60:40 antara *data training* dan *data testing* guna menentukan konfigurasi terbaik berdasarkan hasil evaluasi.

e. **Evaluation**

Tahap *evaluation* menjadi tahapan terakhir dalam *Knowledge Discovery of Data* (KDD) yang nantinya menghasilkan kesimpulan dari pengolahan data *mining* (Utomo et al., 2023). Dalam penelitian ini digunakan *Confusion Matrix* untuk menunjukkan jumlah prediksi

benar dan salah pada setiap kelas, sehingga dapat terlihat bagaimana model membedakan sentimen positif, negatif, dan netral (Sarumpaet & Suryono, 2025). Selain itu, *Classification Report* digunakan untuk menampilkan metrik evaluasi seperti *precision*, *recall*, dan *f1-score* yang memberikan gambaran lebih komprehensif mengenai performa model dalam mengklasifikasikan setiap kelas (Harungguan et al., 2023).

HASIL DAN PEMBAHASAN

a. **Data Selection**

Tahap data *selection* dilakukan dengan mengumpulkan dataset melalui proses *crawling* komentar *YouTube* menggunakan *YouTube Data API v3* pada *Google Cloud Platform*. Data diambil dari enam video berbeda yang membahas isu redenominasi rupiah, baik dari kanal media arus utama maupun kreator independen, dengan mempertimbangkan tingkat interaksi seperti jumlah komentar dan keterlibatan *audiens*. Pengambilan komentar dilakukan pada rentang November hingga Desember 2025, dengan total 1.000 komentar dari setiap video sehingga diperoleh 6.000 data awal dalam format CSV. Seluruh data kemudian digabungkan menjadi satu dataset utama, dan dilakukan seleksi atribut dengan mempertahankan hanya kolom komentar sebagai fokus analisis. Selanjutnya, dilakukan penghapusan data duplikat untuk memastikan kualitas dataset, sehingga dari 6.000 data awal diperoleh 5.669 data bersih yang siap digunakan pada tahap pengolahan berikutnya. Ringkasan komentar *YouTube* terkait Isu redenominasi rupiah yang ditemukan dalam data hasil *crawling* dapat ditampilkan pada Gambar 2.

	author	comment	likes	published_at
0	@OptimisPrime-fbs	Mantep Menkeu pintar	0	2025-12-30T01:50:14Z
1	@PieterWenyi	Di percepat saja supaya uang yang sedang ada d...	0	2025-12-27T13:15:45Z
2	@georgetogas4877	Beritamu hoax alias bohong saja	0	2025-12-23T11:49:21Z
3	@AnangBor	Asal pertama kali bikin pemangkasan nol nya ta...	0	2025-12-16T06:15:36Z
4	@BustanuddinLubis	Kembali ke thn 1970	0	2025-12-11T14:39:28Z

Gambar 2. Hasil *Crawling* Data

b. **Data Preprocessing**

Pada tahap *preprocessing*, data komentar *YouTube* dibersihkan untuk meningkatkan kualitas dan kesiapan data sebelum proses klasifikasi. Tahapan yang dilakukan meliputi *cleansing* untuk menghapus karakter khusus,

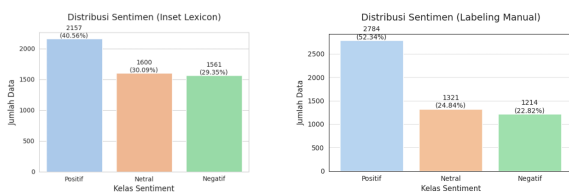
tanda baca, angka, emoji, dan tautan, case folding untuk mengubah seluruh teks menjadi huruf kecil, *Tokenizing* untuk memecah kalimat menjadi kata, *Stopword Removal* untuk menghapus kata-kata umum yang tidak memiliki makna sentimen, serta *Stemming* untuk mengubah kata ke bentuk dasarnya (Aldino et al., 2022). Setelah seluruh tahapan *preprocessing* dilakukan serta penghapusan data yang tidak relevan, jumlah data yang semula 5.669 menjadi 5.318 data yang siap digunakan pada tahap pelabelan dan klasifikasi. Proses ini menghasilkan data yang lebih terstruktur dan representatif sehingga dapat meningkatkan efektivitas ekstraksi fitur serta kinerja model klasifikasi.

comment	cleaning	case_folding	tokenize	normalisasi	stopword_removal	stening_data
0 Terima kasih semoga harkat dan \nMartabat Rupi...	Terima kasih semoga harkat dan \nMartabat Rupi...	terima kasih semoga harkat dan \nmartabat rupi...	[terima, kasih, semoga, harkat, dan, martabat, ...]	[terima, kasih, semoga, harkat, dan, martabat, ...]	[terima, kasih, semoga, harkat, martabat, rupi, ...]	terima kasih moga harkat martabat rupiah di ma...
1 Sangat bagus	Sangat bagus	sangat bagus	[sangat, bagus]	[sangat, bagus]	[bagus]	bagus
2 Tolong munculkan benggol kembali atau sejenis...	Tolong munculkan benggol kembali atau sejenis...	tolong munculkan benggol kembali atau sejenis...	[tolong, munculkan, benggol, kembali, atau, se...	[tolong, munculkan, logam, kembali, atau, se...	[tolong, munculkan, logam, jenis ganti, uang logam ala...	
3 Kalau bisa secepat nya\n2016\n2027 keburu saya...	Kalau bisa secepat nya\nn\nkeburu saya mati du...	kalau bisa secepat nya\nn\nkeburu saya mati du...	[kalau, bisa, secepat, nya, keburu, saya, mati, ...]	[kalau, bisa, secepat, nya, keburu, saya, mati, ...]	[secepat, keburu, mati, duluan, merasakan, uan...	cepat keburu mati duluan rasa uang umur
4 ngk sabar kapan dilaksanakan redenominasi labi...	ngk sabar kapan dilaksanakan redenominasi labi...	ngk sabar kapan dilaksanakan redenominasi labi...	[ngk, sabar, kapan, dilaksanakan, redenominasi, ...]	[tidak, sabar, kapan, dilaksanakan, redenomina...	[sabar, dilaksanakan, redenominasi, cepat, bagus]	sabar laksana redenominasi cepat bagus

Gambar 3. Hasil Data Preprocessing

c. *Data Transformation*

Pada tahap *Data Transformation*, komentar yang telah melalui *preprocessing* terlebih dahulu dilakukan pelabelan sentimen menggunakan pendekatan *Inset Lexicon* untuk mengklasifikasikan data ke dalam kategori positif, negatif, dan netral. Setelah proses pelabelan, teks ditransformasikan ke dalam bentuk numerik menggunakan metode TF-IDF untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam dokumen dan keseluruhan korpus (Widaningrum et al., 2022). Hasil transformasi berupa matriks fitur numerik yang merepresentasikan setiap komentar dan selanjutnya digunakan sebagai input dalam proses pelatihan serta pengujian model *Naïve Bayes*.



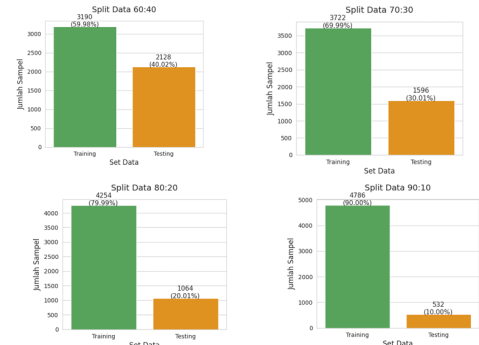
Gambar 4. Distribusi Sentimen (Inset Lexicon dan Manual)

aaaaaamiiiiinnnnnnnnn aaah aaaminnn aah aali aamiiin aamin aaminn aanjiing abadi	0	1	2	3	4
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0

Gambar 5. Proses TF-IDF

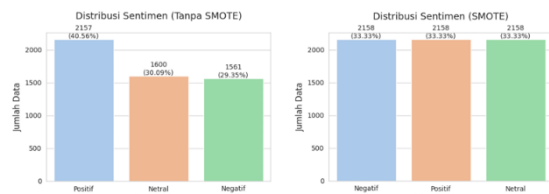
d. *Data Mining*

Pada tahap *data mining*, dilakukan proses pelatihan dan pengujian model menggunakan algoritma *Multinomial Naïve Bayes* untuk mengklasifikasikan sentimen komentar *YouTube* terkait isu redenominasi rupiah. Data kemudian dibagi menjadi data training dan data testing menggunakan empat skenario rasio pembagian, yaitu 90:10, 80:20, 70:30, dan 60:40.

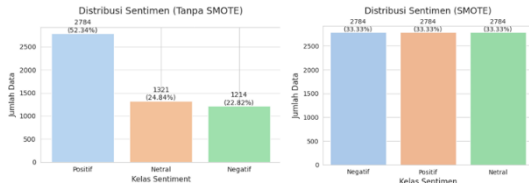


Gambar 6. Data Training dan Data Testing

Untuk mengatasi ketidakseimbangan distribusi kelas, diterapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE) pada data training guna menghasilkan distribusi yang lebih seimbang antar kelas sentimen (Sunata et al., 2024). Model kemudian dilatih menggunakan *data training* dan diuji pada *data testing* untuk mengevaluasi performanya berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Tahap ini bertujuan untuk mengetahui konfigurasi pembagian data dan penerapan SMOTE yang menghasilkan performa klasifikasi terbaik.



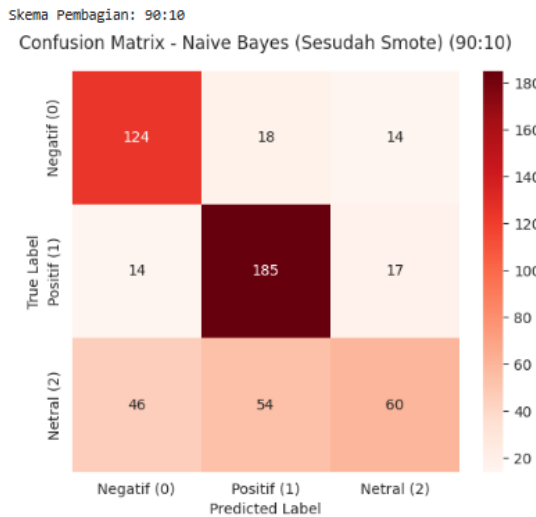
Gambar 7. Distribusi Sentimen Sebelum dan Sesudah SMOTE (Inset Lexicon)



Gambar 8. Distribusi Sentimen Sebelum dan Sesudah SMOTE (Manual)

e. Evaluation

Berdasarkan hasil evaluasi pada skenario pembagian data 90:10, konfigurasi dengan pelabelan *Inset Lexicon* dan penerapan SMOTE menunjukkan performa terbaik dalam penelitian ini. Model *Naïve Bayes* pada skenario tersebut menghasilkan *accuracy* sebesar 0,69, *precision* 0,68, *recall* 0,68, dan *F1-score* 0,66. Nilai ini menunjukkan bahwa model mampu mengklasifikasikan ketiga kelas sentimen (positif, negatif, dan netral) secara relatif seimbang. Penerapan SMOTE terbukti efektif dalam meningkatkan kemampuan model mendeteksi kelas minoritas, khususnya kelas netral yang sebelumnya memiliki *recall* rendah pada kondisi tanpa SMOTE. Selain itu, proporsi *data training* sebesar 90% memberikan ruang belajar yang lebih besar bagi model untuk mengenali pola distribusi kata secara lebih optimal.



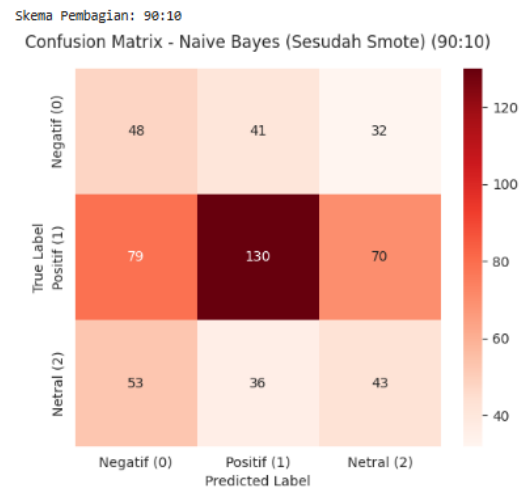
Gambar 9. Confusion Matrix Rasio 90:10 *Inset Lexicon* dengan SMOTE

Classification Report (Sesudah Smote):

	precision	recall	f1-score	support
Negatif (0)	0.67	0.79	0.73	156
Positif (1)	0.72	0.86	0.78	216
Netral (2)	0.66	0.38	0.48	160
accuracy			0.69	532
macro avg	0.68	0.68	0.66	532
weighted avg	0.69	0.69	0.68	532

Gambar 10. Classification Report Rasio 90:10 *Inset Lexicon* dengan SMOTE

Sementara itu, pada skenario 90:10 dengan pelabelan manual dan penerapan SMOTE, performa model menunjukkan peningkatan dibandingkan tanpa SMOTE, namun masih berada di bawah pendekatan *Inset Lexicon*. Pada konfigurasi ini, model memperoleh *accuracy* sebesar 0,42 dengan *macro average F1-score* sekitar 0,40. Meskipun SMOTE membantu mengurangi bias terhadap kelas mayoritas dan meningkatkan kemampuan deteksi kelas negatif dan netral, distribusi prediksi masih belum seimbang sepenuhnya. Hal ini mengindikasikan bahwa kualitas dan konsistensi pelabelan turut memengaruhi kinerja model.



Gambar 11. Confusion Matrix Rasio 90:10 Manual dengan SMOTE

Classification Report (Sesudah Smote):

	precision	recall	f1-score	support
Negatif (0)	0.27	0.40	0.32	121
Positif (1)	0.63	0.47	0.53	279
Netral (2)	0.30	0.33	0.31	132
accuracy			0.42	532
macro avg	0.40	0.40	0.39	532
weighted avg	0.46	0.42	0.43	532

Gambar 12. Classification Report Rasio 90:10 Manual dengan SMOTE

Dengan demikian, dapat disimpulkan bahwa pada rasio pembagian data 90:10, kombinasi pelabelan *Inset Lexicon* dan SMOTE memberikan performa klasifikasi yang lebih stabil dan optimal dibandingkan pelabelan manual dengan SMOTE. Sehingga menunjukkan bahwa metode pelabelan yang sistematis dan distribusi data yang seimbang memiliki peran penting dalam meningkatkan akurasi model *Naïve Bayes* dalam analisis sentimen.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, beberapa kesimpulan. *Algoritma Naïve Bayes* berhasil diterapkan untuk mengklasifikasikan sentimen opini masyarakat terkait isu redenominasi rupiah di *YouTube*. Penelitian menggunakan tahapan KDD. Representasi teks dilakukan menggunakan TF-IDF, serta SMOTE. Hasil menunjukkan bahwa *Naïve Bayes* efektif dan efisien untuk klasifikasi teks media sosial. Sementara itu, opini masyarakat didominasi sentimen positif, yang menunjukkan dukungan terhadap kebijakan redenominasi rupiah. Meski demikian, terdapat sentimen negatif yang mencerminkan kekhawatiran terhadap inflasi, kenaikan harga, dan kepercayaan publik. Sentimen netral umumnya bersifat informatif atau belum menunjukkan sikap yang jelas. Serta, performa model dipengaruhi oleh teknik pelabelan, rasio data latihan-uji, dan penerapan SMOTE. Hasil terbaik diperoleh pada pelabelan *Inset Lexicon* dengan SMOTE rasio 90:10 dengan akurasi 69%, precision 0,68, recall 0,68, dan F1-score 0,66. Penerapan SMOTE terbukti meningkatkan performa model

REFERENSI

- Aldino, A. A., Suryono, R. R., & Ambarwati, R. (2022). Analysis of Covid-19 Cash Direct Aid (BLT) Acceptance Using K-Nearest Neighbor Algorithm. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 16(2), 193. <https://doi.org/10.22146/ijccs.70801>
- Alfifa, A. N., Fahira, V., Batubara, R., Sakinah, K., Manik, J., & Yoes, R. A. (2025). Media Sosial dan

Pembentuka Opini Publik (Analisis Studi Kasus Echo Chamber Pada Interaksi Komentar di Akun Instagram @Turnbackhoaxid Dalam Konteks Post-Truth). *Jurnal Penelitian Ilmu-Ilmu Sosial*, 2(6). <https://doi.org/10.5281/zenodo.14709277>

- Anuar, F., Putra, R., Firman Fadilah, F., Enri, U., & Karawang, U. S. (2023). ANALISIS SENTIMEN ULASAN FILM OPPENHEIMER PADA SITUS IMDB MENGGUNAKAN METODE NAIVE BAYES.
- Arita, N. (2025). ANALISIS SENTIMEN PENGGUNA YOUTUBE TERHADAP ANIME SPY X FAMILY BAHASA INDONESIA MENGGUNAKAN NAÏVE BAYES. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3). <https://doi.org/10.23960/jitet.v13i3.7362>
- Aryanti, R., Fitriani, E., Royadi, R., Ardiansyah, D., & Saepudin, A. (2024). Sentiment Analysis of E-Grocery Application Reviews Using Lexicon-Based and Support Vector Machine. *Jurnal Riset Informatika*, 6(3), 149–158. <https://doi.org/10.34288/jri.v6i3.301>
- Damayanti, A., Diah Delima, I., & Suseno, A. (2023). *Pemanfaatan Media Sosial Sebagai Media Informasi dan Publikasi (Studi Deskriptif Kualitatif pada Akun Instagram @rumahkimkotatangerang)*.
- datareportal. (2025). *Digital 2025 Global Overview Report*. <https://datareportal.com/>
- Dimas Waraditya Nugraha, Agustinus Yoga Primantoro, Antonius Purwanto, Defri Werdiono, & Kris Mada. (2025, November 2). *Apakah Redenominasi Rupiah Perlu Dilakukan?* <https://www.kompas.id/artikel/apakah-redenominasi-rupiah-perlu-dilakukan>
- Diploma, M. A., Manajerial, A., Akuntansi, J., Negeri, P., & Pandang, U. (2023). Dampak Redenominasi Rupiah terhadap Penyajian Laporan Keuangan Audited. *AKUNSIKA: Jurnal Akuntansi Dan Keuangan*, 4(2). <http://jurnal.poliupg.ac.id/index.php/akunsika>

- Fatah, Z., & Syarifah, L. (2025). Analisis Sentimen Komentar YouTube pada Video Terkait Insiden Pengemudi Ojek Online dan Anggota Brimob Menggunakan Algoritma Naive Bayes. In *JAMASTIKA* (Vol. 4).
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naive Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *SMATIKA JURNAL*, 10(02), 71–76. <https://doi.org/10.32664/smatika.v10i02.455>
- Harungguan, A. R., Napitupulu, H., & Firdaniza, F. (2023). *Analisis Sentimen Dengan Metode Klasifikasi Naive Bayes dan Seleksi Fitur Chi-Square*.
- Pyrenia Iskandar, T., Ariza, R., & Nadhifa, F. (2021). Fenomena Penggunaan Youtube Channel Pada Anak Usia Dini Di Masa Pendemic Covid-19. In *Jurnal Ilmiah LISKI (Lingkar Studi Komunikasi)* (Vol. 7, Number 2). <http://journals.telkomuniversity.ac.id/liski106JurnalIlmiahLISKI>
- Ramdhania, N. A., Siregar, A. C., & Sucipto. (2025). Analisis sentimen pengguna X terhadap Istana Garuda IKN menggunakan algoritma Naive Bayes. *Jurnal Pendidikan Informatika Dan Sains*, 14(1), 36–53. <https://doi.org/10.31571/saintek.v14i1.8735>
- Rizal, M., & Hayati, U. (2024). ANALISIS SENTIMEN PENGGUNA TWITTER (X) TERKAIT FILM ONE PIECE MENGGUNAKAN METODE ALGORITMA NAIVE BAYES. *Jurnal Sistem Informasi Kaputama (JSIK)*, 8(1).
- Robi Widodo, T., Nur Fajri, I., & Wulan Sari, B. (2025). Sentiment Analysis of the Film ‘JUMBO’ on Twitter Using the Naive Bayes Method and Support Vector Machine (SVM) with a Text Mining Approach. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 9, Number 5). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Sarumpaet, L. H., & Suryono, R. R. (2025). Analisis Sentimen Publik Program PPPK di Media Sosial X menggunakan Naive Bayes dan SVM. *Edumatic: Jurnal Pendidikan Informatika*, 9(2), 362–371. <https://doi.org/10.29408/edumatic.v9i2.30065>
- Sirait, F., Irpan, D., Fadillah, R., & Syahputra Damanik, R. (2024). Sentiment Analysis on Twitter Social Media towards Najwa Shihab Using Naive Bayes Algorithm and Support Vector Machine (SVM). In *International Journal Of Health, Engineering And Technology* (Vol. 3, Number 1). <https://ijhet.com/index.php/ijhess/>
- Suhendra, T., Intan, B., & Martadinata, A. T. (2024). *ANALISIS SENTIMEN PENGGUNA APLIKASI NETFLIX PADA ULASAN GOOGLE PLAYSTORE MENGGUNAKAN METODE NAIVE BAYES*.
- Sunata, M. H. A., Irwiensyah, F., & Hasan, F. N. (2024). Analisis Sentimen Calon Presiden 2024 di Media Sosial X Menggunakan Naive Bayes dan SMOTE. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(3), 1313. <https://doi.org/10.30865/mib.v8i3.7708>
- Thaariq Razaq, M., Nurjanah, D., & Nurrahmi, H. (2023). *Analisis Sentimen Review Film Menggunakan Naive Bayes Classifier Dengan Fitur TF-IDF*.
- Utomo, Y. B., Kurniasari, I., & Yanuartanti, I. (2023). PENERAPAN KNOWLEDGE DISCOVERY IN DATABASE UNTUK ANALISA TINGKAT KECELAKAAN LALU LINTAS. *Jurnal Teknik Informatika Kaputama (JTIK)*, 7(1).
- Widaningrum, I., Mustikasari, D., Arifin, R., Lathifah Tsaqila, S., & Fatmawati, D. (2022). *Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan K-Means Clustering Untuk Menentukan Kategori Dokumen*.