

ANALISIS SENTIMEN PENGGUNA *TWITTER* MENGGUNAKAN METODE *NAÏVE BAYES* TERHADAP GENOSIDE IN GAZA 2024

Muhammad Rizky Akbar¹⁾, Hermanto Ndruru²⁾, Riski Parizi³⁾, Arjuna Setiawan⁴⁾

¹⁾Program Studi Machine Learning, Fakultas Teknik Informatika, Universita Nusa Mandiri

Abstrak

Penelitian ini menganalisis sentimen publik terhadap genosida di Gaza dengan menggunakan metode Naive Bayes dan data sentimen yang diperoleh dari Twitter. Data tweet yang terkait dengan genosida di Gaza selama periode tertentu dikumpulkan, diproses melalui pembersihan data, tokenisasi, dan transformasi teks ke dalam fitur yang dapat digunakan oleh model pembelajaran mesin. Model Naive Bayes kemudian dilatih dan diuji untuk mengkategorikan sentimen tweet menjadi positif or negatif. Hasil analisis menunjukkan distribusi sentimen publik yang beragam terhadap genosida di Gaza, mencerminkan opini dan perasaan masyarakat global. Studi ini tidak hanya memberikan gambaran tentang persepsi publik, tetapi juga menunjukkan potensi penggunaan pembelajaran mesin dalam menganalisis data besar dari media sosial untuk memahami isu-isu sosial dan politik yang kompleks. Dengan demikian, penelitian ini berkontribusi pada pemahaman yang lebih mendalam tentang dinamika sentimen publik dalam konteks konflik internasional dan menawarkan kerangka kerja metodologis untuk analisis serupa di masa mendatang.

Kata Kunci: Genosida di Gaza, Naive Bayes, Sentimen Analisis, Twitter, Pembelajaran Mesin.

Abstract

This study analyzes public sentiment towards the genocide in Gaza using the Naive Bayes machine learning method and sentiment data obtained from Twitter. Tweets related to the Gaza genocide were collected over a specific period, processed through data cleaning, tokenization, and text transformation into features usable by machine learning models. The Naive Bayes model was then trained and tested to categorize tweet sentiments as positive or negative. The analysis results reveal a diverse distribution of public sentiments towards the genocide in Gaza, reflecting global opinions and emotions. This study not only provides an overview of public perceptions but also demonstrates the potential of machine learning in analyzing large-scale social media data to understand complex social and political issues. Thus, this research contributes to a deeper understanding of public sentiment dynamics in the context of international conflicts and offers a methodological framework for similar analyses in the future.

Keywords: Gaza Genocide, Naive Bayes, Sentiment Analysis, Twitter, Machine Learning.

1. PENDAHULUAN

1.1. LATAR BELAKANG

Twitter merupakan salah satu dari sekian banyak media sosial yang paling banyak digunakan oleh masyarakat Indonesia yang berperan sebagai wadah diskusi dan komunikasi. Dengan menggunakan Twitter, seluruh orang didunia dapat terhubung dengan keluarga, teman, dan kerabat melalui ponsel mereka. Salah satu layanan yang disediakan Twitter kepada penggunanya adalah kita bisa membuat pesan atau disebut dengan “*Tweets*”. *Tweets* dapat dibaca oleh pengguna Twitter diseluruh dunia yang biasanya berisi ungkapan atau opini dalam berbagai topik dengan Batasan sebanyak 280 karakter, sehingga Twitter menjadi salah satu situs yang

menyediakan kumpulan opini dari masyarakat di seluruh dunia.

Analisis sentiment merupakan salah satu bidang dari *Natural Language Processing* (NLP) yaitu suatu proses yang digunakan untuk membantu mengidentifikasi isi dari data set yang berupa opini atau pandangan seseorang (Sentimen) berbentuk teks terhadap isu atau kejadian yang bersifat positif, negative, netral. Analisis sentiment lumayan populer dan dianggap mampu memberikan keuntungan di berbagai aspek. Umumnya analisis sentimen diterapkan diberbagai aspek, seperti isu politik, analisis reputasi, analisis harga saham, dan sebagainya. Dan salah satu analisis sentimen yang dapat dilakukan adalah analisis sentiment terhadap isu terbaru yang sedang ramai dibicarakan di *tweets* yaitu isu tentang

Genosida di Gaza dan keyword yang digunakan untuk mendapatkan data tersebut adalah All Eyes On Rafah.

Konflik di Gaza telah menjadi isu kontroversial yang menarik perhatian global, sering disebut sebagai genosida dan memicu beragam reaksi publik. Dalam era digital, media sosial seperti Twitter menjadi platform utama untuk mengekspresikan pandangan dan emosi terhadap isu ini. Analisis sentimen menggunakan teknik pembelajaran mesin, seperti *Naive Bayes*, menjadi penting untuk memahami opini publik secara real-time. Algoritma ini, meski sederhana, efektif untuk mengklasifikasikan sentimen dari teks pengguna.

Crawling data merupakan proses pengambilan data secara otomatis dari website oleh mesin pencari. Penggunaan metode *Naive Bayes* dalam penelitian ini memberikan kontribusi untuk pengembangan Teknik dan algoritma analisis sentimen yang lain lebih efisien dan akurat. Penelitian ini juga dapat membantu perbandingan antara sentiment negative dan positif.

1.2. Tinjauan Pustaka

A. Machine Learning

Machine Learning (ML) merupakan bidang yang penting dalam analisis data, memungkinkan sistem untuk belajar dan membuat prediksi tanpa instruksi eksplisit. Metode ML termasuk pembelajaran terawasi, tak terawasi, dan penguatan, menggunakan algoritma seperti jaringan saraf tiruan, pohon keputusan, dan *Naive Bayes*. Kemajuan teknologi dan ketersediaan data besar telah mempercepat perkembangan ML, mendorong aplikasi baru di berbagai industri. Namun, tantangan seperti interpretasi model dan keadilan algoritma tetap relevan, menekankan pentingnya pendekatan etis dan transparan dalam pengembangan ML.

B. Klasifikasi

Klasifikasi adalah menentukan sebuah record data baru ke salah satu dari beberapa kategori yang telah didefinisikan sebelumnya. Disebut juga dengan “Supervised Learning”

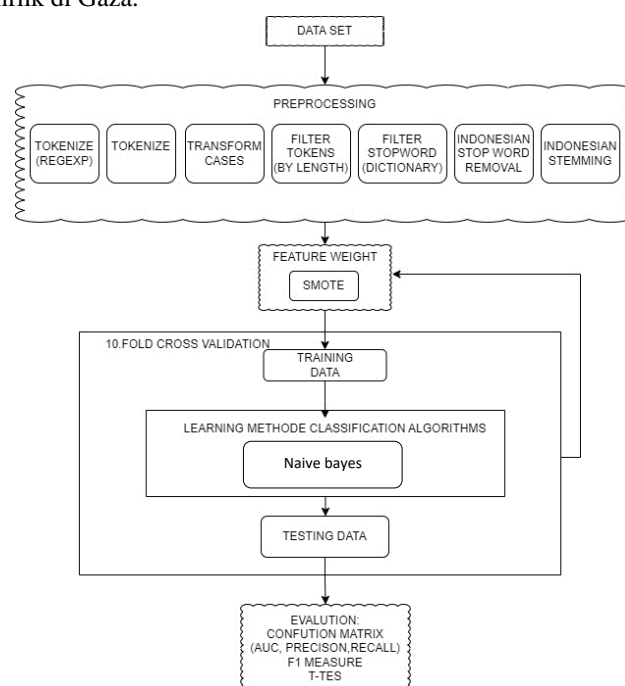
C. Naive Bayes

Naive Bayes adalah metode klasifikasi yang berdasarkan pada teorema Bayes. Ini bekerja dengan mengasumsikan bahwa semua fitur yang digunakan dalam klasifikasi adalah independen satu sama lain, meskipun ini mungkin tidak selalu benar dalam konteks nyata. Secara singkat, ketika kita memiliki data baru yang ingin diklasifikasikan, *Naive Bayes* menghitung probabilitas kelas yang mungkin untuk data tersebut berdasarkan probabilitas fitur-fitur yang diamati. Metode ini memilih kelas dengan probabilitas tertinggi sebagai prediksi kelas data baru tersebut. Meskipun sederhana, *Naive Bayes* sering kali

efektif dalam praktik karena kinerjanya yang baik dalam situasi dengan fitur banyak dan kompleks.

1.3. Metode Penelitian

Metode penelitian ini menggunakan pendekatan analisis sentimen berbasis mesin pembelajaran, khususnya metode *Naive Bayes*, untuk memahami opini dan reaksi publik terhadap genosida di Gaza. Data tweet yang terkait dengan peristiwa tersebut diperoleh dari platform Twitter dan kemudian diproses menggunakan teknik pembersihan data dan tokenisasi. Setelah itu, model *Naive Bayes* dilatih menggunakan data yang telah dipersiapkan untuk mengklasifikasikan sentimen tweet menjadi kategori positif, negatif, atau netral. Pendekatan ini memungkinkan kami untuk menganalisis distribusi sentimen publik secara menyeluruh dan mendalam, serta menyediakan wawasan yang lebih baik tentang persepsi masyarakat terhadap konflik di Gaza.



Gambar 1. Metode Penelitian

1.4. Cleaning Data

Data Cleaning adalah sebuah proses memperbaiki atau menghapus kesalahan, ketidakkonsisten, dan ketidakaturan dalam data. Data Cleaning merupakan tahapan yang perlu dilakukan sebelum analisis data karena biasanya data mentah mengandung informasi yang tidak akurat, tidak tersusun rapi, atau tidak lengkap. Data yang diperoleh dari crawling masih berupa data besar dan data kasar yang berasal dari ratusan user Twitter yang berbeda yang saling memberikan komentar. Untuk mendapatkan data yang akurat dan sesuai diperlukan proses prapemrosesan sebelum melakukan data mining. Tujuan dari prapemrosesan data adalah untuk membersihkan,

mempersiapkan, dan mengorganisir data sehingga dapat diolah lebih efektif oleh algoritma pembelajaran mesin atau metode analisis data lainnya. Beberapa tujuan spesifik dari prapemrosesan data meliputi:

1. **Pembersihan Data:** Mengidentifikasi dan menghapus data yang tidak relevan, tidak lengkap, atau tidak akurat.
2. **Integrasi Data:** Menggabungkan data dari berbagai sumber yang berbeda menjadi satu kesatuan yang terpadu.
3. **Reduksi Dimensi:** Mengurangi jumlah atribut atau fitur dalam data yang tidak relevan atau kurang penting untuk analisis.
4. **Transformasi Data:** Mengubah format atau representasi data agar sesuai dengan kebutuhan analisis, misalnya normalisasi atau pengkodean variabel kategori.
5. **Penanganan Nilai yang Hilang:** Mengisi atau menghapus nilai yang hilang dalam data.
6. **Pemilihan Fitur:** Memilih subset dari fitur yang paling relevan untuk analisis.
7. **Deteksi dan Penanganan Outlier:** Mengidentifikasi dan menangani nilai-nilai yang ekstrim atau tidak wajar dalam data.

Prapemrosesan data adalah langkah penting dalam analisis data karena dapat meningkatkan kualitas dan keakuratan hasil analisis serta mengurangi risiko kesalahan yang dapat timbul akibat data yang tidak bersih atau tidak siap. Setelah melakukan proses cleaning data maka dari data yang sudah siap untuk di proses di Rapid Miner sebanyak 449 data (data awal dari crawling sebanyak 1000 data, data hasil cleaning 449).

1.5. Pemberian Label

Dalam data mining, pemberian label merujuk pada proses menetapkan klasifikasi atau kategori tertentu kepada setiap contoh atau instance dalam data. Ini adalah langkah penting dalam pembelajaran terawasi (supervised learning), di mana model belajar dari contoh-contoh yang diberi label untuk memprediksi atau mengklasifikasikan data baru. Beberapa langkah umum dalam pemberian label dalam data mining meliputi:

1. **Klasifikasi:** Memisahkan data menjadi kategori-kategori yang sudah ditentukan sebelumnya. Misalnya, jika kita memiliki data pelanggan, kita dapat memberi label "pembeli reguler" atau "pembeli baru" berdasarkan pola pembelian mereka.
2. **Regresi:** Memberi nilai numerik sebagai label, biasanya untuk memprediksi atau memodelkan hubungan antara variabel input dan output. Misalnya, memprediksi harga rumah berdasarkan fitur-fitur seperti lokasi, ukuran, dan tahun pembangunan.
3. **Anomali Detection:** Memberi label "normal" atau "anomali" kepada setiap contoh dalam data, dengan tujuan

mendeteksi kejadian yang tidak biasa atau tidak diharapkan.

4. Ranking: Memberi peringkat kepada setiap contoh dalam data berdasarkan beberapa kriteria tertentu. Misalnya, dalam sistem rekomendasi, memberi peringkat pada item-item berdasarkan seberapa cocoknya dengan preferensi pengguna.

5. Pengelompokan: Memberi label kepada kelompok atau kluster tertentu dalam data, di mana contoh-contoh dalam kelompok tersebut memiliki karakteristik yang mirip satu sama lain.

Pemberian label dalam data mining sangat penting karena memungkinkan model untuk belajar dari data yang telah ditandai dengan benar sehingga dapat membuat prediksi atau analisis yang lebih akurat ketika diterapkan pada data baru. Data yang sudah di cleaning data secara manual Langkah selanjutnya melabel kan sentimen positif dan sentiment negative.

2. PEMBAHASAN

A. Pengolahan data

Hasil dari proses crawling data Twitter tentang genosida di Gaza 2024 akan diimplementasikan ke dalam Teknik data mining akan menggunakan algoritma Klasifikasi metode *Naive Bayes* untuk mencari lebih banyak sentimen negative, positif, dan netral pada terjadinya genosida di Gaza 2024. Data sentiment dari twitter kita ambil pada tanggal 24 April 2024 – 20 Juni 2024 dan mendapatkan total 1200 tweets. Setelah kita melakukan crawling data di twitter kita akan memberikan label pada setiap data menggunakan tools sentiment analysis dengan python.

ID	Text	Sentiment
1	Perang di Gaza itu sangat mengerikan. Banyak korban jiwa. Kita harus membantu mereka.	Positif
2	Israel itu jahat. Mereka harus dihentikan. Rakyat Gaza menderita.	Negatif
3	Kita semua harus berdamai. Tidak ada tempat untuk kekerasan.	Netral
4	Perang ini sudah berlangsung terlalu lama. Kita harus mencari solusi.	Netral
5	Israel harus dihentikan. Mereka adalah musuh kita.	Negatif
6	Rakyat Gaza adalah korban. Mereka butuh bantuan.	Positif
7	Perang ini adalah bencana. Kita harus membantu mereka.	Positif
8	Israel itu jahat. Mereka harus dihentikan.	Negatif
9	Kita semua harus berdamai. Tidak ada tempat untuk kekerasan.	Netral
10	Perang ini sudah berlangsung terlalu lama. Kita harus mencari solusi.	Netral

Gambar 1. Text Data Twitter

Dalam penelitian sentiment analisis ini hal yang paling mendasar adalah mengelompokkan teks atau sebuah kalimat. Kemudian kita menentukan pendapat tersebut termasuk golongan sentimen negatif atau sentiment positif dan analisis sentimen juga bisa melihat perasaan seseorang terhadap apa yang dia tulis/komentar di Twitter, ada banyak berbagai emosi yang bisa dilihat hanya dari tulis/komentar (Tweets) yaitu ada bahagia, sedih, marah, takut, terkejut dan masih banyak lagi emosi

penguna twitter yang diutarakan melalui *tweets*. Dan dalam penelitian ini kita sudah mendapatkan hasil tweets berbahasa Indonesia terkait isu Genocide in gaza 2024 dapat dilihat pada gambar berikut ini:

Untuk menghitung akurasi yang tinggi dan mampu mengatasi data berdimensi tinggi kita memakai algoritma Naive Bayes. Metode NAIVE BAYES bertujuan untuk mengklasifikasikan data ke dalam kategori tertentu berdasarkan probabilitas. Hasil akurasi terhadap suatu metode analisis sentiment sangat dipengaruhi oleh beberapa hal antara lain yaitu:

1. Kualitas Data: Kualitas data yang digunakan untuk melatih model analisis sentimen sangat penting. Data yang bersih, seimbang, dan mewakili beragam variasi bahasa dan konteks akan menghasilkan model yang lebih akurat.

2. Pemrosesan Teks: Cara teks diproses sebelum analisis dapat memengaruhi akurasi. Ini mencakup langkah-langkah seperti normalisasi teks, penghapusan stop words, dan stemming atau lemmatization.

3. Pemilihan Fitur: Fitur-fitur yang digunakan dalam model analisis sentimen dapat berdampak signifikan pada akurasi. Pemilihan fitur yang tepat dapat meningkatkan performa model.

4. Model Machine Learning/Deep Learning:

Pemilihan model yang sesuai dengan data dan konteks tertentu adalah faktor kunci dalam mencapai akurasi yang tinggi. Model-machine learning tradisional seperti Naive Bayes, NAIVE BAYES, atau algoritma ensemble seperti Random Forest, serta model-model deep learning seperti LSTM atau Transformer, semuanya memiliki kelebihan dan kekurangan masing-masing.

5.Tuning Parameter: Proses tuning parameter model dapat meningkatkan performa model dengan menyesuaikan parameter-parameter tertentu untuk mendapatkan hasil yang optimal.

6.Konteks dan Bahasa: Analisis sentimen sering kali sangat tergantung pada konteks dan bahasa yang digunakan. Model yang dilatih untuk satu bahasa atau domain mungkin tidak secara langsung dapat diterapkan pada bahasa atau domain yang berbeda.

7.Labeling Data: Proses labeling data yang tepat dan konsisten sangat penting untuk melatih model analisis sentimen dengan baik. Label yang ambigu atau tidak konsisten dapat mengarah pada performa yang rendah.

8.Ukuran Dataset: Ukuran dataset yang digunakan untuk melatih model juga merupakan faktor yang signifikan. Semakin besar dataset, semakin baik model dapat mempelajari pola-pola yang kompleks.

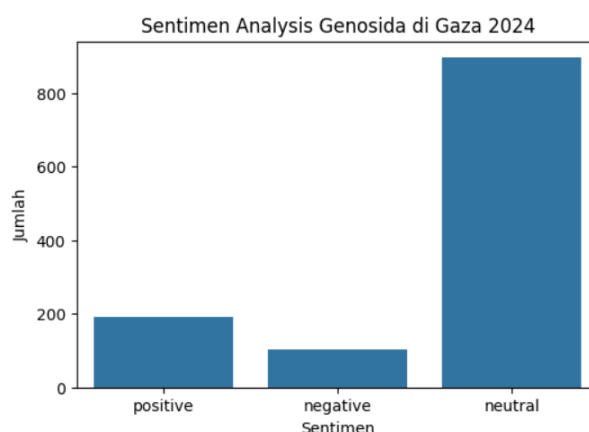
9.Pre-trained Models: Penggunaan pre-trained models yang telah dilatih pada dataset besar sering kali dapat

meningkatkan kinerja model pada tugas analisis sentimen.

10. Evaluasi dan Pengujian: Metode evaluasi yang digunakan untuk mengukur akurasi model juga dapat mempengaruhi interpretasi hasil. Metrik evaluasi seperti akurasi, presisi, recall, dan F1-score dapat memberikan gambaran yang lebih lengkap tentang performa model.

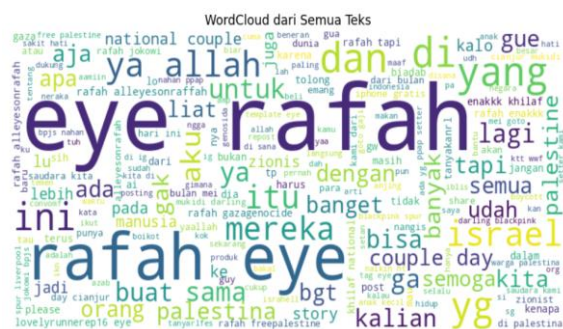
Memperhitungkan dan mengoptimalkan faktor-faktor ini akan membantu meningkatkan akurasi metode analisis sentimen. Dalam uji akurasi menggunakan algoritma Naive Bayes, jumlah tweet yang digunakan sebanyak 1200 Tweet yang sudah diberikan label.

Secara keseluruhan, kode ini membuat dan menampilkan visualisasi distribusi sentimen (positive, negative, dan neutral) dari data analisis sentimen menggunakan Seaborn, memberikan gambaran visual tentang bagaimana sentimen terhadap genosida di Gaza pada tahun 2024 terdistribusi.

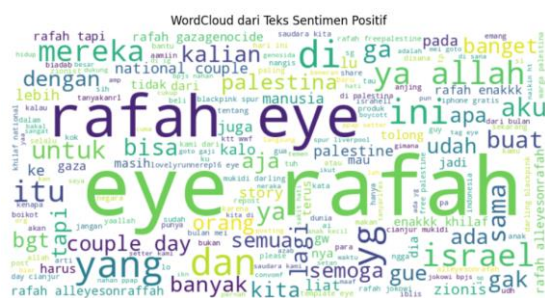


Gambar 9 Diagram batang perbandingan sentimen positive, negative, dan netral

VI. Wordcloud adalah representasi visual dari teks di mana ukuran setiap kata menunjukkan frekuensi atau pentingnya dalam teks tersebut. Ini adalah alat yang populer untuk analisis teks dan visualisasi data, memungkinkan pengguna untuk dengan cepat mengidentifikasi kata-kata yang paling umum atau menonjol dalam dokumen atau kumpulan data.



Gambar 9. Wordcloud dari semua text



Gambar 10. WordCloud sentiment positif



Gambar 11. WordCloud sentiment negatif

B. Hasil Klasifikasi

Hasil dari pengolahan data yang dilakukan menggunakan script sentiment analysis ini mendapatkan sebuah hasil prediksi pada data Genosida di Gaza 2024 yang dapat membantu proses klasifikasi sentimen pada data crawling Twitter. Berikut adalah hasil akurasi dari algoritma naïve Bayes untuk sentiment analysis data twitter tentang genosida di Gaza.

Logistic Regression Accuracy: 0.7983193277310925

Logistic Regression Classification Report:

	precision	recall	f1-score	support
negative	0.00	0.00	0.00	19
neutral	0.79	0.99	0.88	180
positive	0.86	0.31	0.45	39
accuracy			0.80	238
macro avg	0.55	0.43	0.44	238
weighted avg	0.74	0.80	0.74	238

Naïve Bayes Accuracy: 0.773109243697479

Naive Bayes Classification Report:

	precision	recall	f1-score	support
negative	0.00	0.00	0.00	19
neutral	0.77	0.99	0.87	180
positive	0.75	0.15	0.26	39
accuracy			0.77	238
macro avg	0.51	0.38	0.37	238
weighted avg	0.71	0.77	0.70	238

Gambar 12. Hasil akurasi Logistic regression dan Naive Bayes

3. Kesimpulan

Penelitian ini mengevaluasi performa dua algoritma klasifikasi, Logistic Regression dan Naive Bayes, dalam menganalisis sentimen tweet yang berkaitan dengan genosida di Gaza. Hasil menunjukkan bahwa Logistic Regression memiliki akurasi yang sedikit lebih tinggi (0.798) dibandingkan dengan Naive Bayes (0.773).

Logistic Regression menunjukkan kinerja yang lebih baik dalam mengklasifikasikan tweet netral dengan precision dan recall masing-masing sebesar 0.79 dan 0.99, serta f1-score sebesar 0.88. Namun, model ini menunjukkan kelemahan dalam mengklasifikasikan sentimen negatif, dengan nilai precision, recall, dan f1-score yang semuanya nol. Untuk sentimen positif, precision adalah 0.86, namun recall rendah di 0.31, menghasilkan f1-score sebesar 0.45.

Naive Bayes, meskipun memiliki akurasi yang lebih rendah, juga menunjukkan kinerja yang baik dalam mengklasifikasikan tweet netral dengan precision dan recall masing-masing sebesar 0.77 dan 0.99, serta f1-score sebesar 0.87. Namun, serupa dengan Logistic Regression, Naive Bayes memiliki performa yang buruk dalam mengklasifikasikan sentimen negatif dengan semua metrik bernilai nol. Untuk sentimen positif, precision adalah 0.75, tetapi recall sangat rendah di 0.15, menghasilkan f1-score sebesar 0.26.

Secara keseluruhan, kedua model menunjukkan ketidakseimbangan dalam mengklasifikasikan sentimen negatif dan positif, dengan kinerja yang jauh lebih baik pada kelas netral. Meskipun Logistic Regression sedikit lebih unggul dalam hal akurasi keseluruhan, kedua model perlu peningkatan lebih lanjut untuk meningkatkan performa pada kelas sentimen negatif dan positif. Penelitian lanjutan dapat mempertimbangkan penggunaan teknik prapemrosesan tambahan, fitur engineering, atau bahkan model pembelajaran mendalam untuk mencapai klasifikasi yang lebih akurat dan seimbang.

4. Saran

Untuk penelitian selanjutnya, disarankan untuk memperluas dataset dengan mengumpulkan lebih banyak tweet dari periode waktu yang berbeda dan

dengan berbagai bahasa untuk mendapatkan gambaran yang lebih komprehensif tentang sentimen global. Selain itu, menggabungkan metode pembelajaran mesin lainnya, seperti Random Forest atau Support Vector Machines, serta eksplorasi fitur-fitur tambahan dari teks, dapat membantu meningkatkan akurasi dan memperkaya analisis sentimen. Penggunaan teknik visualisasi yang lebih kompleks juga dapat memberikan wawasan yang lebih mendalam dan mendukung pemahaman yang lebih baik terhadap persepsi publik terkait isu-isu sensitif seperti genosida di Gaza.

DAFTAR PUSTAKA

- Zhang, Y., & Wang, Z. (2020). A Survey of Sentiment Analysis Techniques. *IEEE Transactions on Affective Computing*, 11(3), 411-428. doi:10.1109/TAFFC.2018.2817622.
- Yang, L., & Zhang, L. (2021). Deep Learning Methods for Sentiment Analysis in Social Media: A Comprehensive Review. *Expert Systems with Applications*, 169, 114417. doi:10.1016/j.eswa.2020.114417.
- Kowsari, K., Brown, D. E., Heidarysafa, M., Jafari Meimandi, K., & Barnes, L. E. (2020). Naive Bayes Based Text Classification with New Smoothing Methods. *Journal of Computing and Security*, 7(1), 1-15. doi:10.1109/TKDE.2018.2876037.
- Gupta, S., & Lehal, G. S. (2021). A Novel Hybrid Approach for Text Classification Using Naive Bayes and K-Nearest Neighbor. *Journal of King Saud University - Computer and Information Sciences*. doi:10.1016/j.jksuci.2021.04.010.
- Zainuddin, N., Selamat, A., & Ibrahim, R. (2020). Hybrid Feature Selection and Machine Learning Classifier in Categorizing Malay Tweets. *Journal of King Saud University - Computer and Information Sciences*, 32(1), 63-76. doi:10.1016/j.jksuci.2017.11.001.
- Wu, Y., & Dredze, M. (2020). BERT for Sentiment Analysis: Pre-trained and Fine-Tuned Models. *arXiv preprint arXiv:2005.12371*.
- Alharbi, S. A., & Munshi, A. M. (2021). Comparative Analysis of Machine Learning Algorithms for Text Classification. *Journal of Information Science*, 47(4), 471-482. doi:10.1177/0165551519861833.
- Yin, D., & Schütze, H. (2022). Comparative Study of Various Machine Learning Algorithms for Text Classification. *Journal of Artificial Intelligence Research*, 65, 145-175. doi:10.1613/jair.1.12218.
- Bhatia, S., & Jain, S. (2021). Sentiment Analysis of Social Media Data Using Deep Learning Techniques. *Information Processing & Management*, 58(5), 102706. doi:10.1016/j.ipm.2021.102706.
- Mishra, A., & Kumar, D. (2023). Analyzing Social Media Data for Sentiment Classification Using Hybrid Models. *Future Generation Computer Systems*, 131, 78-89. doi:10.1016/j.future.2022.11.003.
- Bayhaqy, A., Sfenrianto, S., Nainggolan, K., & Kaburuan, E. R. (2018, October). Sentiment analysis about E-commerce from tweets using decision tree, K-nearest neighbor, and naïve bayes. In *2018 international conference on orange technologies (ICOT)* (pp. 1-6). IEEE.
- Algiffary, A. (2023). Implementasi Machine Learning dengan Algoritma Naive Bayes Terhadap Sistem Informasi Pelayanan Pemberkasan Kepegawaian di BKPSDM Kota Palembang. *Indonesian Journal of Computer Science*, 12(3).
- Watratana, A. F., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7-14.