

## ANALISIS SENTIMEN TERHADAP PENGGUNAAN DONASI KORBAN PENYIRAMAN AIR KERAS PADA MEDIA SOSIAL X.COM MENGUNAKAN METODE BERT

Helna Berliana<sup>1</sup>, Ridwan Yusuf<sup>2</sup>

Universitas Dharma Wacana, Lampung

e-mail: <sup>1</sup>helnaberliana2312@gmail.com, <sup>2</sup>ryusuf87@gmail.com

**Abstract:** *Donations, as a form of social solidarity, often face issues of transparency and misuse, particularly in humanitarian cases such as acid attacks. These issues lead to public distrust and become widely discussed on social media platforms like X.com. This study aims to analyze public sentiment toward the use of donations in such a case by collecting 1,077 tweets using specific keywords. The data underwent preprocessing, sentiment labeling using the VADER method, and classification with IndoBERT, a BERT-based model for the Indonesian language. The results show a dominance of neutral sentiment, causing class imbalance and bias toward the neutral class. As a result, the model showed low accuracy in detecting positive and negative sentiments, and signs of overfitting were observed. This study recommends the application of data balancing techniques and regularization methods to improve classification performance, particularly in handling minority sentiment classes.*

**Keywords:** *Sentiment Analysis, BERT, Donation, Social Media, VADER, IndoBERT, Acid Attack*

**Abstrak:** Donasi sebagai bentuk solidaritas sosial sering kali menghadapi isu transparansi dan penyalahgunaan, terutama dalam kasus kemanusiaan seperti penyiraman air keras. Isu ini menimbulkan ketidakpercayaan publik dan menjadi topik hangat di media sosial X.com. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap penggunaan donasi pada kasus tersebut dengan menggunakan 1.077 tweet yang dikumpulkan berdasarkan kata kunci tertentu. Data dianalisis melalui proses praproses, pelabelan sentimen menggunakan metode VADER, dan klasifikasi menggunakan IndoBERT, model BERT untuk Bahasa Indonesia. Hasil penelitian menunjukkan dominasi sentimen netral yang menyebabkan ketidakseimbangan kelas dan bias model terhadap kelas tersebut. Akurasi dalam mengidentifikasi sentimen positif dan negatif pun rendah, serta terdeteksi adanya overfitting. Penelitian ini merekomendasikan penggunaan teknik penyeimbangan data dan metode regularisasi untuk meningkatkan performa klasifikasi, khususnya pada sentimen minoritas.

**Kata kunci:** Analisis Sentimen, BERT, Donasi, Media Sosial, VADER, IndoBERT, Penyiraman Air Keras

### PENDAHULUAN

Donasi merupakan bentuk solidaritas sosial yang penting, namun sering kali dihadapkan pada isu transparansi dan penyalahgunaan dana, terutama dalam kasus-kasus kemanusiaan seperti bencana alam atau penyiraman air keras. Isu ini sering memicu ketidakpercayaan publik terhadap

pengelolaan donasi dan menjadi bahan diskusi hangat di media sosial, seperti X.com, yang memungkinkan analisis sentimen untuk mengidentifikasi respons masyarakat terhadap penyalahgunaan donasi.

X.com, sebagai platform media sosial populer, menyebarkan informasi secara cepat dan memungkinkan interaksi luas mengenai berbagai isu, termasuk isu

sosial sensitif. Namun, dengan volume interaksi yang tinggi, analisis sentimen menjadi tantangan kompleks. Pendekatan ini bertujuan untuk mengkategorikan opini dalam teks menjadi positif, negatif, atau netral, memberikan wawasan tentang persepsi publik terkait isu-isu tertentu.

Metode BERT (Bidirectional Encoder Representations from Transformers) kini digunakan dalam analisis sentimen di media sosial, menggantikan model leksikon seperti VADER. BERT memahami konteks kalimat secara mendalam dan lebih efektif dalam menangani bahasa gaul dan emotikon. Meskipun BERT memiliki akurasi yang tinggi dalam memahami sentimen kompleks, ia memerlukan sumber daya komputasi lebih besar dan waktu pemrosesan yang lebih lama dibandingkan metode lain, namun tetap memberikan hasil yang lebih akurat dalam konteks percakapan media sosial.

### Analisis Sentimen

Analisis sentimen adalah proses untuk menentukan sentimen atau perasaan yang terkandung dalam sebuah teks, baik berupa opini positif, negatif, maupun netral. Di media sosial, analisis sentimen sering digunakan untuk memahami pandangan publik terhadap isu-isu tertentu. Penelitian sebelumnya menunjukkan bahwa analisis sentimen bermanfaat untuk mengetahui opini masyarakat, terutama dalam konteks isu-isu yang sedang viral. (Hidayatullah & Susanto, 2022).

### Donasi

Donasi adalah pemberian sumbangan, baik uang, barang, atau jasa, untuk membantu individu atau organisasi yang membutuhkan tanpa mengharapkan imbalan. Tujuan utamanya adalah untuk mendukung mereka yang terkena bencana, kesulitan ekonomi, atau proyek sosial. Di era digital, donasi semakin mudah dilakukan melalui platform online atau media sosial, yang memungkinkan kontribusi secara cepat dan transparan. Partisipasi publik dalam donasi

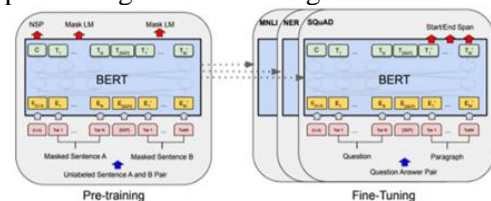
dipengaruhi oleh transparansi dan kemudahan akses, sementara interaksi di media sosial seperti komentar dan berbagi membantu meningkatkan kesadaran serta keterlibatan masyarakat (Pratama & Yulianto, 2022; Santika, 2021).

### Valence Aware Dictionary for sEntiment Reasoning (VADER)

Salah satu metode berbasis leksikon yang populer dan sering digunakan dalam penelitian adalah VADER. VADER merupakan alat untuk analisis sentimen yang berfungsi mengevaluasi teks serta menentukan sentimen positif, negatif, atau netral (Hutto dan Gilbert, 2020). Algoritma VADER mengombinasikan pendekatan berbasis leksikon dengan berbasis aturan untuk menganalisis sentimen teks. Metode ini menggunakan kamus berisi kata-kata yang dikategorikan sebagai positif, negatif, atau netral, serta aturan sintaksis untuk memahami konteks kata-kata tersebut dalam teks. Kamus leksikon VADER dilengkapi dengan bobot untuk masing-masing kata, yang berada dalam rentang -4 hingga +4. Setelah melakukan analisis terhadap teks, VADER menghasilkan skor gabungan (compound). Skor compound diperoleh melalui penjumlahan skor valensi setiap kata dalam leksikon yang kemudian dinormalisasi dalam rentang -1 (sangat negatif) hingga +1 (sangat positif) (Hutto dan Gilbert, 2020).

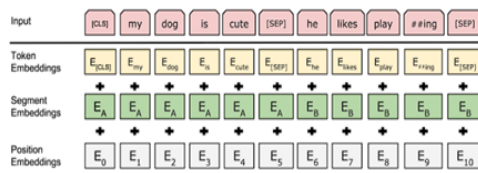
### BERT

Bidirectional Encoder Representations from Transformers (BERT) adalah model bahasa yang diperkenalkan oleh Google tahun 2018. BERT mempunyai dua tahap kerja yaitu pretraining dan fine-tuning.



**Gambar 1** Prosedur Pre-training dan Finetuning BERT

Gambar 1 menjelaskan tahapan dalam pretraining dan fine-tuning. Pre-training melibatkan model dalam memahami bahasa dan konteksnya melalui Masked Language Modeling (MLM) dan Next Sentence Prediction (NSP) bersamaan. MLM memungkinkan model menggabungkan konteks dari kedua arah, yaitu konteks kiri dan kanan. NSP menggabungkan representasi pasangan teks selama pre-train

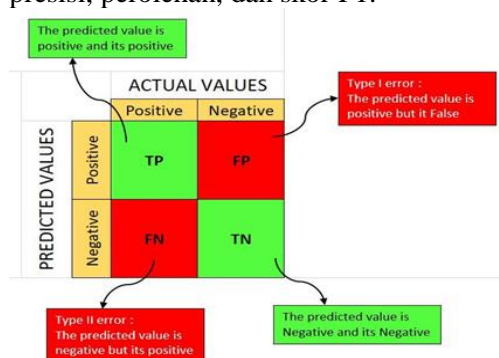


**Gambar 2 Representasi Input BERT**

Gambar 2 menunjukkan representasi input dari BERT. Setiap urutan dimulai dengan token khusus [CLS]. Token khusus [SEP] ditambahkan untuk memisahkan setiap kalimat. Setiap token memiliki segment embedding yang membedakan antara kalimat A dan B. Positional embedding untuk memberikan tanda posisi setiap token dalam urutan. Input encoder BERT dihasilkan dengan menggabungkan Token Embeddings, Segment Embeddings, dan Positional Embeddings.

### CONFUSION MATRIK

Matriks konfusi biasanya digunakan untuk mengukur kinerja model klasifikasi, dan juga dapat digunakan untuk mengukur kinerja model melalui perhitungan metrik kinerja seperti akurasi, presisi, perolehan, dan skor F1.



**Gambar 3 Confusion Matrix**

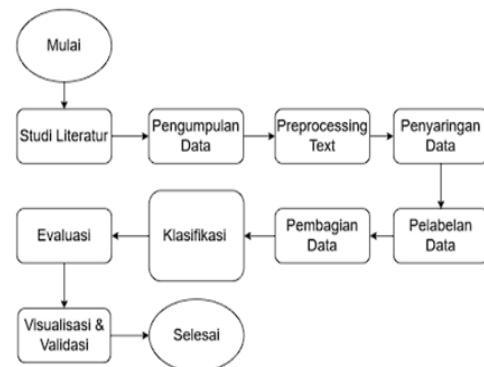
yang mana:

1. True Positive (TP) memprediksi ramalan positif dan itu benar.
2. True Negative (TN) memprediksi ramalan negatif dan itu benar.
3. False Positive (FP) memprediksi ramalan positif dan itu salah
4. False Negative (FN) memprediksi ramalan negatif dan itu salah.

### Kesalahan Tipe 2 ini sangat berbahaya.

Nilai False Positive (FP), False Negative (FN), True Positive (TP), dan True Negative (TN) digunakan untuk menentukan akurasi. Nilai akurasi menunjukkan kemampuan sistem untuk mengklasifikasikan data dengan benar. Dengan kata lain, nilai presisi adalah prediksi antara data yang diklasifikasikan dengan benar dan kumpulan data

### METODE



**Gambar 4 Metodologi penelitian**

Gambar 4 merupakan metologi yang digunakan dalam penelitian. Berikut merupakan penjabaran dari setiap tahapan.

1. Studi Literatur  
Tahapan ini melakukan pencarian rujukan teori terkait penelitian dari jurnal, buku, artikel ilmiah, penelitian terdahulu, dan sumber lain dari internet.
2. Pengumpulan Data  
Tahapan ini melakukan pengumpulan data dari X.com dari tanggal Data terdiri dari 1077 tweets dari tanggal 1 Oktober-1 November

- 2024 berdasarkan 3 kata kunci pencarian yaitu “Penyiraman Agus”, “Donasi Agus”, dan “Uang Donasi Agus Air Keras.” Pengumpulan data dilakukan dengan menggunakan tweet-harvest, kemudian disimpan dalam bentuk csv.
3. Preprocessing Text
 

Pada tahapan ini melakukan pemrosesan teks agar bisa memenuhi standar dari model dan siap untuk dilakukan pelatihan. Tahapan praproses teks sebagai berikut.

    - a. Case Folding
 

Tahap ini membuat seluruh karakter dalam teks menjadi format huruf kecil
    - b. Data Cleaning
 

Tahap ini melakukan pembersihan karakter berulang, tautan, nama pengguna, angka, tagar, simbol, spasi berlebih, dan tanda baca.
    - c. Normalization
 

Tahap ini melakukan pengubahan katakata yang tidak standar menjadi standar.
    - d. Stopwords Removal
 

Tahap ini melakukan penghapusan kata-kata tidak bermakna signifikan seperti 'dari', 'dan', 'di', 'ke', 'yang'.
    - e. Stemming
 

Tahap ini menghapus afiks seperti sufiks, prefiks, dan konfiks.
  4. Penyaringan Data
 

Pada tahap ini melakukan penyaringan tweet yaitu menghilangkan tweet selain Bahasa Indonesia. Tweet akan disaring untuk menghindari bahasa asing, bahasa daerah, serta bahasa melayu. Penyaringan data juga dilakukan untuk menghapus data duplikat.
  5. Pelabelan Data
 

Pada tahap ini dilakukan pelabelan otomatis menggunakan Vader Lexicon. Pelabelan otomatis dilakukan untuk menentukan label teks dalam kategori positif atau negatif.
  6. Pembagian Data
 

Data akan dibagi menjadi dua yang terdiri dari data latih (train) dan data uji (test). Data akan dibagi dengan pembagian 80:20 yaitu 80% untuk data latih dan 20% untuk data uji.
  7. Klasifikasi
 

Tahap klasifikasi akan dilakukan dengan menggunakan IndoBERT ('indobenchmark/indobert-base-p1') yang merupakan model BERT untuk bahasa Indonesia. Sebelum dilakukan pelatihan, dilakukan word embedding untuk mengonversi kata-kata atau token dalam sebuah kalimat menjadi vektor numerik yang dapat dimengerti oleh BERT. Tahapan word embedding yaitu sebagai berikut.

    - a. Tokenisasi
 

Tokenisasi dilakukan untuk memisahkan teks menjadi unit kecil yaitu token.
    - b. Menambahkan Token [CLS] dan [SEP]
 

Token [CLS] digunakan sebagai token klasifikasi pada awal teks, sedangkan token [SEP] digunakan sebagai token penanda akhir teks atau pemisah antar teks.
    - c. Mengonversi Token ke ID Token
 

akan diubah menjadi ID sesuai dengan kamus model yang digunakan yaitu "indobenchmark/indobert-base-p1"
    - d. Menambahkan Padding
 

Penambahan token padding dilakukan agar semua input teks memiliki panjang yang sama sesuai panjang maksimum yang ditentukan.
    - e. Menambahkan Attention Mask
 

Attention mask digunakan untuk membedakan token yang sebenarnya dari token padding.
    - f. Menambahkan Positional Embedding
 

Penambahan positional embedding untuk memberikan informasi posisi absolut dalam teks.
  8. Evaluasi
 

Tahap evaluasi dilakukan untuk menilai seberapa baik model melakukan klasifikasi sentimen.

- Performa model akan dievaluasi menggunakan confusion matrix. Selanjutnya dilakukan perhitungan perhitungan accuracy, precision, recall, dan F1-Score.
9. Visualisasi dan Validasi Tahap visualisasi melakukan penulisan kode untuk tampilan antarmuka agar mempermudah pengguna memahami hasil klasifikasi. Visualisasi dilakukan menggunakan framework streamlit dengan menampilkan grafik hasil sentimen, visualisasi klasifikasi teks dan file csv. Tahap validasi yaitu mencocokkan hasil prediksi dengan analisis sebenarnya.

## HASIL DAN PEMBAHASAN

### Pengumpulan Data

Scraping tweet dilakukan dengan menggunakan tweet-harvest. Hasil scrapping yang telah dilakukan dengan 3 kata kunci mendapatkan sebanyak 1.007 tweet. Contoh hasil dari pengumpulan data pada tabel 1.

create_d_at	username	full_text
Sat Oct 26 04:53:59 +0000 2024	panggilakui koo	Jujur dari lubuk hati saya tuh kasian sama si Agus Salim. Udah ga bisa liat tapi pake kacamata. Kerennnn nomor 1. Dari awal kalau mas Agus transparan mengenai uang donasi saya yakin ga sampai seperti ini. Pengaruh orang disekitar mas Agus juga membuat mentalnya kena bisikan. <a href="https://t.co/9SX Xqw9ePO">https://t.co/9SX Xqw9ePO</a>

Fri Oct 25 12:21:45 +0000 2024	KenKenMannhattan	Agus mempermasalah kan uang donasi 1.5 M masih wajar. (namanya juga manusia liat uang segitu banyak) Tp Agus melaporkan Teh Novi dgn pasal terkait Pencemaran Nama baik itu kurang ajari, • Mata tersiram air keras : Ke Rumah Sakit Ke Kantor Polisi
--------------------------------	------------------	---

### Preprocessing Text

Pada tahapan ini terdiri dari 5 proses yaitu case folding, data cleaning, normalization, stopwords removal, dan stemming. Berikut merupakan contoh hasil dari preprocessing text pada tabel 2.

PraProses	Teks
Data	@whoopziy Agus syedih Agus syedih banget uang donasi Agus mau di minta lagi
<i>Case Folding &amp; Data Cleaning</i>	agus syedih agus syedih banget uang donasi Agus mau di minta lagi
<i>Normalization</i>	agus sedih agus sedih banget uang donasi Agus mau di minta lagi
<i>Stopwords Removal</i>	agus sedih agus sedih uang donasi Agus mau minta
<i>Stemming</i>	agus sedih agus sedih uang donasi Agus mau minta

### Penyaringan Data

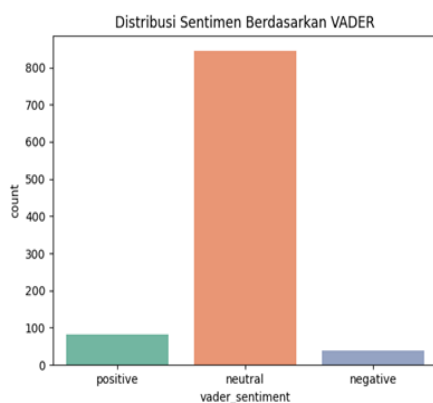
Penyaringan data dilakukan untuk menghapus data teks bahasa asing dan data duplikat. Dari hasil pengumpulan data, ditemukan bahasa asing seperti bahasa melayu, bahasa inggris, bahasa sunda, dan bahasa jawa.

Data teks dengan bahasa asing sebanyak 19. Kemudian, hasil pengecekan tweet yang duplikat ditemukan sebanyak 42 tweet. Jadi, data yang siap digunakan sebanyak 1.016 data. Dari data tersebut, 95% akan digunakan untuk skenario klasifikasi dan 5% dari dataset akan digunakan untuk validasi sistem.

### Pelabelan Data

Pada tahap ini dilakukan pelabelan sentimen positif dan negatif menggunakan Vader. Penentuan label ditentukan dari nilai compound score. Nilai compound score  $> 0.05$ , labelnya positif dan nilai compound score  $< -0.05$ , labelnya negatif. Dari hasil pelabelan, diperoleh data dengan sentimen positif sebanyak 265 dan sentimen negatif sebanyak 751. Berikut merupakan contoh data yang telah dilabeli pada tabel 3.

### Distribusi Sentimen Berdasarkan VADER



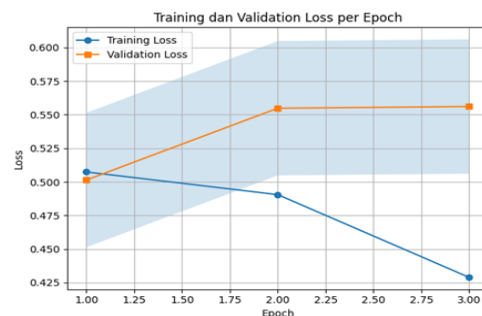
**Gambar 5 Distribusi Sentimen Berdasarkan VADER**

Hasil analisis sentimen menggunakan metode VADER (Valence Aware Dictionary and sEntiment Reasoner) ditampilkan pada Gambar 4. Berdasarkan visualisasi tersebut, distribusi sentimen pada data

menunjukkan dominasi yang sangat tinggi pada kategori netral. Jumlah data yang termasuk dalam sentimen netral tercatat lebih dari 800 data, yang menunjukkan bahwa sebagian besar teks atau kalimat dalam dataset tidak memiliki kecenderungan emosional yang kuat, baik secara positif maupun negatif. Sementara itu, jumlah data dengan sentimen positif berada di kisaran 80, dan data dengan sentimen negatif merupakan yang paling sedikit, yakni sekitar 40 data.

Ketidakseimbangan distribusi sentimen ini merupakan faktor penting yang perlu dicermati, khususnya dalam konteks pemodelan klasifikasi berbasis pembelajaran mesin atau deep learning. Ketimpangan jumlah data antar kategori dapat menyebabkan model menjadi bias terhadap kelas mayoritas, dalam hal ini kelas netral, sehingga model akan cenderung memberikan prediksi ke kelas tersebut meskipun data yang diuji berasal dari kelas minoritas. Kondisi ini sangat mungkin terjadi dalam proses pelatihan, karena model akan belajar lebih banyak dari data yang jumlahnya dominan. Akibatnya, akurasi terhadap kelas minoritas dapat menurun drastis atau bahkan tidak terdeteksi sama sekali. Oleh karena itu, distribusi ini menjadi dasar penting untuk mempertimbangkan metode penyeimbangan data (seperti oversampling, undersampling, atau penggunaan teknik augmentasi) pada tahapan preprocessing berikutnya guna meningkatkan performa klasifikasi.

### Evaluasi Performa Model: Training dan Validation Loss



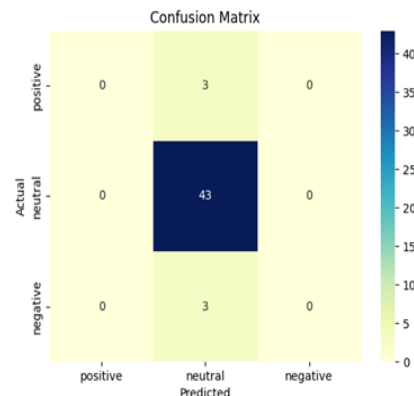
**Gambar 6 Evaluasi Performa Model: Training dan Validation Loss**

Evaluasi performa model dilihat dari grafik loss pada Gambar 5 yang menunjukkan perubahan nilai training loss dan validation loss selama tiga epoch proses pelatihan. Pada epoch pertama, baik nilai training loss maupun validation loss berada pada kisaran 0.50, menunjukkan bahwa model memulai proses pelatihan dengan performa awal yang cukup seimbang antara data latih dan validasi. Namun demikian, pada epoch kedua, training loss mengalami sedikit penurunan, sedangkan validation loss justru mengalami kenaikan menuju angka 0.55. Selanjutnya, pada epoch ketiga, training loss terus menurun secara konsisten hingga mencapai nilai sekitar 0.43, sedangkan validation loss cenderung stagnan pada nilai sebelumnya, tidak mengalami peningkatan maupun penurunan yang signifikan.

Polanya menunjukkan indikasi awal dari fenomena overfitting, yaitu kondisi di mana model terlalu menyesuaikan diri terhadap data pelatihan, sehingga kemampuan generalisasinya terhadap data baru atau data validasi menurun. Hal ini ditunjukkan oleh semakin jauhnya selisih antara training loss dan validation loss. Penurunan training loss tanpa diikuti penurunan validation loss menandakan bahwa model mulai kehilangan kemampuan untuk mengenali pola secara umum dan hanya berfokus pada data yang sudah pernah dilihat selama pelatihan.

Faktor utama yang kemungkinan besar mempengaruhi munculnya overfitting ini adalah distribusi kelas yang tidak seimbang dalam data pelatihan. Karena mayoritas data berlabel netral, model cenderung lebih mudah mencapai nilai loss yang rendah dengan cara memprediksi netral terhadap sebagian besar input. Oleh karena itu, strategi peningkatan performa model perlu difokuskan pada penyamaan distribusi kelas dan penerapan metode regularisasi, misalnya dengan menambahkan dropout, early stopping, atau menggunakan data augmentation berbasis teks.

### Evaluasi Akurasi Model: Confusion Matrix



**Gambar 7 Evaluasi Akurasi Model: Confusion Matrix**

Visualisasi confusion matrix pada Gambar 6 memberikan informasi lebih rinci mengenai performa model dalam mengklasifikasikan data uji ke dalam tiga kategori sentimen: positif, netral, dan negatif. Matriks tersebut memperlihatkan bahwa seluruh data yang sebenarnya berlabel positif dan negatif secara keliru diklasifikasikan sebagai netral. Secara spesifik, terdapat 3 data positif yang diprediksi sebagai netral, dan 3 data negatif juga diprediksi sebagai netral. Satu-satunya kategori yang terklasifikasi dengan benar oleh model adalah kelas netral, dengan total 43 data yang sesuai antara prediksi dan label aktual.

Kondisi ini mengindikasikan bahwa model benar-benar bias terhadap kelas mayoritas, yaitu netral. Akurasi model dalam mengidentifikasi data dari kelas minoritas (positif dan negatif) adalah 0%, yang berarti bahwa model gagal mengenali fitur atau pola linguistik khas dari teks berlabel positif maupun negatif. Akurasi keseluruhan yang mungkin tampak tinggi (karena banyaknya prediksi benar pada kelas netral) tidak mencerminkan kinerja sesungguhnya dari model terhadap ketiga kategori sentimen secara seimbang.

Kinerja model seperti ini sangat tidak ideal dalam konteks analisis sentimen yang bertujuan untuk



memahami variasi opini, terutama dalam situasi di mana sentimen positif dan negatif memiliki dampak penting terhadap pengambilan keputusan. Oleh karena itu, diperlukan pendekatan yang lebih sensitif terhadap kelas minoritas, seperti penggunaan teknik class weighting, focal loss, atau pendekatan sampling yang dapat memperkuat sinyal dari kategori minoritas dalam proses pelatihan model.

## SIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap penggunaan donasi korban penyiraman air keras yang tersebar di media sosial X.com menggunakan model BERT (Bidirectional Encoder Representations from Transformers). Dari hasil pengumpulan data sebanyak 1.077 tweet, dilakukan tahapan praproses data, pelabelan dengan VADER, serta klasifikasi sentimen menggunakan IndoBERT.

Hasil analisis menunjukkan bahwa distribusi sentimen sangat tidak seimbang, dengan dominasi kelas netral yang signifikan. Hal ini menyebabkan model mengalami bias terhadap kelas mayoritas (netral), sehingga performa dalam mengklasifikasikan sentimen positif dan negatif sangat rendah. Evaluasi dengan confusion matrix menunjukkan bahwa model gagal mengenali sentimen minoritas, karena seluruh data positif dan negatif diprediksi sebagai netral.

Overfitting juga terdeteksi pada proses pelatihan, yang ditunjukkan oleh penurunan training loss tanpa diikuti penurunan validation loss. Hal ini memperkuat bahwa ketidakseimbangan data menjadi tantangan utama dalam pemodelan sentimen.

Oleh karena itu, disarankan untuk menggunakan teknik penyeimbangan data dan metode regularisasi pada penelitian lanjutan guna meningkatkan akurasi model terhadap seluruh kelas sentimen secara seimbang. Penggunaan pendekatan tambahan seperti class weighting atau

focal loss juga dapat menjadi solusi untuk meningkatkan sensitivitas model terhadap sentimen minoritas yang penting dalam pengambilan keputusan publik.

## DAFTAR PUSTAKA

- Firmansyah, D. A., & Damayanti, D. (2024). ANALISIS SENTIMEN MASYARAKAT TERHADAP THRIFTING PADA MEDIA SOSIAL TWITTER (X. COM) MENGGUNAKAN METODE K-NEAREST NEIGHBOR. JSR: Jaringan Sistem Informasi Robotik, 8(1), 10-14.
- Susilawati, L., & Christin, M. (2024). Analisis Tema Fantasi Komentar Donasi Palestina Pada Akun Youtube Nusantara Palestina Center. Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi), 8(4), 1023-1028.
- Wijaya, A. S. (2022). Analisis Exploratory Kata “donasi” Akibat Pandemi Covid-19 Pada Media Sosial Twitter. INTEK: Jurnal Informatika dan Teknologi Informasi, 5(1), 76-80.
- Blesyova, N., & Hasan, F. N. (2024). Analisis Sentimen Terhadap Bea Cukai Menggunakan Support Vector Machine Dan K-Fold Cross Validation. JATI (Jurnal Mahasiswa Teknik Informatika), 8(6), 12051-12056.
- Abimanyu, D. (2022). Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER. Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER, 5(03), 423-431.
- Fauziah, N. (2024). Analisis Sentimen Publik Terhadap Kenaikan Tarif PPN di Indonesia dengan Pendekatan VADER. Jurnal Akuntansi dan Keuangan, 12(2), 228-238.
- Ramadhan, M. F., & Siswoyo, B. (2024). Mengenal Model BERT dan Implementasinya untuk Analisis



- 
- Sentimen Ulasan Game. Prosiding SISFOTEK, 8(1), 395-398.
- Khadapi, M., & Pakpahan, V. M. (2024). Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi Twitter tentang Pemilu 2024. JUKI: Jurnal Komputer dan Informatika, 6(2), 130-137.
- Rininda, G., Santi, I. H., & Kirom, S. (2023). Penerapan Svm Dalam Analisis Sentimen Pada Edlink Menggunakan Pengujian Confusion Matrix. JATI (Jurnal Mahasiswa Teknik Informatika), 7(5), 3335-3342.
- Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis sentimen masyarakat terhadap program kartu prakerja pada twitter dengan metode support vector machine. Jurnal Informatika: Jurnal Pengembangan IT, 6(3), 150-155