

PENERAPAN *LARGE LANGUAGE MODEL* DALAM ANALISIS SENTIMEN PADA PEMILIHAN CALON PRESIDEN

A.Ali Akbar Khaerun¹, Nur Afif², Mustikasari³

Jurusan Teknik Informatika, Universitas Islam Negeri Alauddin Makassar
e-mail: aaliakbar214@gmail.com¹, nur.afif@uin-alauddin.ac.id², mustikasari@uin-alauddin.ac.id³

***Koresponden Author:** A.Ali Akbar Khaerun, aaliakbar214@gmail.com
Accepted: 08 30, 2024 ; Revised: 08 28, 2024; Published: 08 31, 2024

Abstrak

Pemilihan calon presiden di Indonesia merupakan topik hangat di media sosial, terutama Twitter. Penelitian ini menganalisis sentimen masyarakat terhadap pemilihan calon presiden 2024 dengan menggunakan model IndoBERT, yang dirancang khusus untuk Bahasa Indonesia, pada data berjumlah 8442 tweet. Penelitian ini mengikuti metode CRISP-DM, mencakup pemahaman bisnis, data, persiapan, pemodelan, evaluasi, dan penerapan. Data dikumpulkan melalui crawling dengan kata kunci terkait pemilihan, kemudian dilakukan preprocessing dan pelabelan manual sebelum diproses oleh model. Hasil menunjukkan bahwa IndoBERT mencapai akurasi 98%, dengan presisi, recall, dan F1-score juga 98% pada epoch ke-10. Evaluasi ukuran batch menunjukkan bahwa batch 4 memberikan performa terbaik. Model ini efektif dalam mengklasifikasikan sentimen terkait pemilihan calon presiden 2024 dan menjadi alat yang berguna untuk memahami opini publik.

Kata kunci: IndoBERT, Large Language Model, analisis sentimen, pemilihan calon presiden, Twitter.

Abstract

[IMPLEMENTATION OF LARGE LANGUAGE MODELS IN SENTIMENT ANALYSIS FOR PRESIDENTIAL CANDIDATE ELECTIONS] The presidential candidate election in Indonesia is a hot topic on social media, especially Twitter. This study analyzes public sentiment regarding the 2024 presidential candidate election using the IndoBERT model, which is specifically designed for the Indonesian language, on a dataset of 8,442 tweets. This research follows the CRISP-DM methodology, which includes business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The data was collected through crawling with keywords related to the election, followed by preprocessing and manual labeling before being processed by the model. The results show that IndoBERT achieved an accuracy of 98%, with precision, recall, and F1-score also at 98% at the 10th epoch. Batch size evaluation indicated that a batch size of 4 yielded the best performance. This model is effective in classifying sentiment related to the 2024 presidential candidate election and serves as a useful tool for understanding public opinion.

Keywords: IndoBERT, Large Language Model, sentiment analysis, presidential candidate election, Twitter.

1. PENDAHULUAN

Pemilihan Presiden di Indonesia merupakan salah satu proses demokrasi terbesar dan paling penting yang melibatkan seluruh lapisan Masyarakat [1]. Pemilihan ini dilakukan setiap

lima tahun sekali dan bertujuan untuk memilih pemimpin tertinggi yang akan memimpin negara, sesuai dengan prinsip-prinsip demokrasi yang termaktub dalam UUD 1945. Keberlangsungan pemilihan presiden ini memunculkan berbagai opini dan



perbincangan publik, terutama di platform media sosial seperti Twitter, yang memungkinkan masyarakat untuk menyuarakan pendapat mereka secara terbuka dan luas. [2]

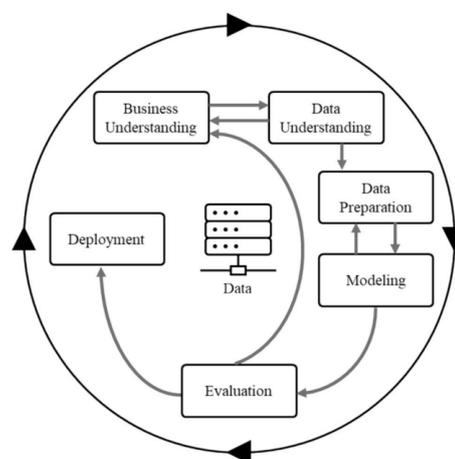
Media sosial Twitter sangat populer di kalangan masyarakat Indonesia dan menjadi salah satu *platform* utama untuk diskusi publik [3]. Dengan jumlah pengguna yang terus meningkat, Twitter menjadi sarana penting untuk menggali opini masyarakat terkait isu-isu politik. Pemilihan presiden menjadi topik hangat, dan sentimen publik terkait calon presiden sering kali tercermin dalam unggahan dan komentar pengguna. Dalam konteks ini, analisis sentimen dapat memberikan wawasan mengenai persepsi masyarakat, dengan mengelompokkan opini publik ke dalam kategori positif, negatif, atau netral. Informasi ini sangat penting bagi pengambil kebijakan dan pengamat politik dalam memahami aspirasi masyarakat secara *real-time*. [4]

Mengingat besarnya *volume* data yang dihasilkan dari platform media sosial, diperlukan pendekatan yang tepat untuk menganalisis data tersebut. Model bahasa besar (*Large Language Model*), seperti *Bidirectional Encoder Representations from Transformers* (BERT), yang dikembangkan oleh Google, telah terbukti efektif dalam menangkap konteks dan memahami hubungan antar kalimat dalam teks berbahasa alami [5]. IndoBERT, salah satu varian BERT yang dirancang khusus untuk bahasa Indonesia, merupakan model yang ideal untuk memahami sentimen masyarakat di Indonesia. IndoBERT dilatih menggunakan korpus berbahasa Indonesia yang sangat besar, sehingga dapat memahami konteks bahasa Indonesia dengan baik. [6][7]

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terkait pemilihan calon presiden di Indonesia dengan menggunakan model IndoBERT. Penelitian ini mengumpulkan data dari Twitter dan mengklasifikasikannya ke dalam sentimen positif, negatif, atau netral. Hasil penelitian ini diharapkan dapat memberikan pandangan yang lebih mendalam tentang opini publik terhadap calon presiden 2024, serta menunjukkan efektivitas penggunaan model IndoBERT dalam analisis sentimen untuk bahasa Indonesia.

2. METODE

Penelitian ini menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM), yang mencakup enam tahap utama: pemahaman masalah, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi. Setiap tahap memiliki peran penting dalam mengembangkan model yang akurat dan andal untuk analisis sentimen pada pemilihan calon presiden di Indonesia. Berikut ini adalah rincian dari tiap tahapan dalam metode penelitian ini:



Gambar 1.1. Metode CRISP-DM [8]

1) Pemahaman Masalah (*Business Understanding*)

Tahap ini bertujuan untuk memahami tujuan dan kebutuhan utama dari penelitian ini, yaitu menganalisis sentimen publik terkait pemilihan calon presiden 2024 dengan menggunakan model IndoBERT. Melalui tahap ini, dirumuskan langkah-langkah yang perlu diambil untuk mencapai hasil analisis yang akurat dan relevan terhadap kebutuhan pengambilan keputusan dalam konteks politik Indonesia.

2) Pemahaman Data (*Data Understanding*)

Data dalam penelitian ini dikumpulkan melalui teknik *data crawling* dari platform Twitter. Kata kunci seperti "pemilihan calon presiden 2024", "pilpres 2024", "pemilu 2024", serta nama calon-calon presiden yang

populer digunakan untuk menyaring data terkait opini publik. Data yang berhasil dikumpulkan kemudian dipelajari untuk memastikan relevansinya dengan topik penelitian serta keakuratan dalam konteks analisis sentimen.

3) Persiapan Data (*Data Preparation*)

Pada tahap ini, data yang dikumpulkan diolah melalui beberapa tahap *preprocessing*, yang meliputi:

- Case Folding*: Mengubah semua teks menjadi huruf kecil.
- Text Cleaning*: Menghapus elemen yang tidak relevan seperti URL, mention, hashtag, tanda baca, angka, simbol, serta spasi berlebihan.
- Text Normalization*: Mengubah kata-kata slang atau singkatan menjadi kata baku dan mengonversi kata berbahasa asing ke bahasa Indonesia.
- Anotasi Data*: Memberikan label secara manual ke dalam kategori sentimen positif, negatif, atau netral untuk setiap teks.

Setelah pembersihan, data dibagi menjadi tiga bagian: data pelatihan, data validasi, dan data pengujian dengan proporsi 70%, 18%, dan 12% secara berurutan.

4) Pemodelan (*Modeling*)

Proses pemodelan dilakukan dengan menggunakan IndoBERT, model bahasa yang dikembangkan khusus untuk bahasa Indonesia. IndoBERT digunakan untuk klasifikasi sentimen dengan menetapkan parameter-parameter penting, seperti panjang urutan maksimal (512 token), ukuran *batch* (32), dan jumlah *epoch* yang dioptimalkan untuk pelatihan model. Tokenisasi dilakukan dengan metode *WordPiece*, di mana setiap token diubah menjadi representasi numerik sebelum dimasukkan ke model.

5) Evaluasi (*Evaluation*)

Kinerja model dievaluasi menggunakan metrik evaluasi, yaitu akurasi, presisi, recall, dan F1-score. *Confusion matrix* digunakan untuk menilai hasil prediksi dari setiap kelas sentimen,

yang memberikan gambaran mengenai performa model dalam mengidentifikasi sentimen positif, negatif, dan netral. Jika hasil evaluasi menunjukkan kinerja yang belum memadai, model akan diperbaiki dan dievaluasi kembali hingga mencapai hasil yang optimal.

6) Implementasi (*Deployment*)

Setelah melalui proses pelatihan dan evaluasi, model yang telah dilatih kemudian diimplementasikan pada sebuah aplikasi berbasis web. *Framework* Laravel digunakan untuk membangun antarmuka yang memungkinkan pengguna memasukkan teks atau tweet dan menerima hasil klasifikasi sentimen secara langsung. Aplikasi ini dirancang agar mudah diakses oleh pengguna yang ingin mengetahui sentimen publik terhadap calon presiden tertentu dalam pemilihan 2024.

Metode CRISP-DM yang diterapkan dalam penelitian ini memberikan struktur sistematis yang membantu dalam pengembangan model yang akurat dan relevan, serta memastikan bahwa model dapat diimplementasikan dalam konteks dunia nyata.

3. HASIL DAN PEMBAHASAN

Pengembangan model IndoBERT yang digunakan dalam penelitian ini adalah varian *Large Language Model* yang telah dipra-latih dengan data berbahasa Indonesia, yaitu *Indobert-large-p1* dan *Indobert-large-p2*. Model ini diimplementasikan menggunakan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Tahapan penting dalam pengembangan model ini meliputi *data crawling*, *preprocessing*, pelabelan data, dan tokenisasi untuk persiapan input model. Dataset terdiri dari 8,442 data tweet yang telah dibersihkan dan diproses.

Hasil evaluasi model dilakukan untuk menilai akurasi prediksi dalam mengklasifikasikan data sentimen. Untuk mendapatkan hasil yang optimal dilakukan

perbandingan penggunaan *stopword*, ukuran *batch* serta nilai *epoch*.

Tabel 1. Perbandingan penggunaan *stopword*

<i>Stopword</i>	Akurasi
Tidak pakai	98%
Pakai	67%

Hasil dan visualisasi menunjukkan bahwa penghapusan *stopword* sering dilakukan untuk menyederhanakan data dan meningkatkan efisiensi pemrosesan. Namun, hal ini dapat menurunkan akurasi model jika informasi penting hilang. Pada IndoBERT, yang memahami struktur dan makna kalimat secara keseluruhan, penghilangan terlalu banyak kata dapat mengurangi performa model dalam analisis sentimen. Berdasarkan Tabel 1, terlihat bahwa penggunaan *stopword* yang tidak tepat dapat mengurangi akurasi model.

Tabel 2. Perbandingan Ukuran *Batch*

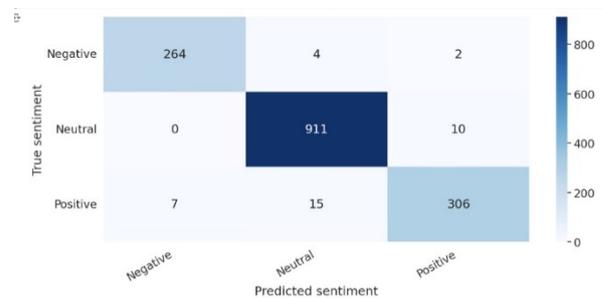
<i>Batch</i>	Akurasi p1	Akurasi p2
4	95%	98%
16	84%	85%
32	89%	90%

Hasil perbandingan ukuran *batch* menunjukkan bahwa *batch* 4, 16, dan 32 memiliki perbedaan akurasi dan durasi pelatihan yang tidak terlalu signifikan. *Batch* 16 kurang optimal dalam klasifikasi data. Semakin kecil ukuran *batch*, akurasi cenderung lebih baik, meskipun eksperimen tetap diperlukan untuk menemukan *batch* yang optimal. Model *indobert-large-p2* mencapai akurasi tertinggi 98%, sedangkan *indobert-large-p1* mencapai akurasi terbaik 95% pada *batch* 4. Oleh karena itu, *batch* 4 pada *indobert-large-p2* dinilai paling seimbang.

Tabel 3. Perbandingan Nilai *Epoch*

<i>Epoch</i>	Akurasi	Presisi	<i>Recall</i>	<i>F1</i>
5	98%	97%	97%	97%
10	98%	98%	98%	98%
15	98%	98%	97%	97%

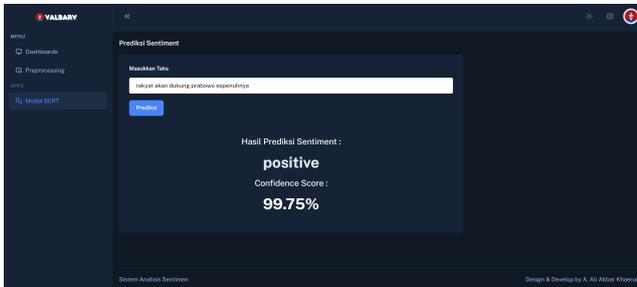
Hasil evaluasi model menunjukkan bahwa pada epoch ke-10, model mencapai akurasi, presisi, recall, dan *F1-score* sebesar 98%. Peningkatan jumlah epoch dari 5 ke 15 hanya memberikan perbedaan 1% pada presisi, recall, dan *F1-score*. Oleh karena itu, eksperimen dihentikan pada epoch ke-15 karena akurasi tetap stabil. Epoch ke-10 dipilih sebagai titik optimal karena model menunjukkan performa terbaik di semua metrik evaluasi. Hal ini menunjukkan bahwa tambahan epoch tidak selalu meningkatkan performa model secara konsisten.



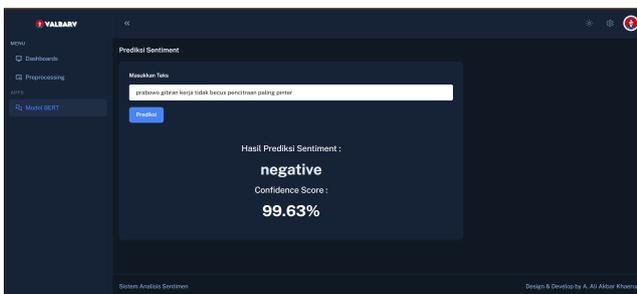
Gambar 3.1. *Confusion Matrix* Validasi data

Matriks menunjukkan performa model BERT dalam mengklasifikasikan sentimen negatif, netral, dan positif. Angka pada diagonal utama (264 untuk negatif, 911 untuk netral, 306 untuk positif) menandakan prediksi yang benar, sedangkan angka di luar diagonal menunjukkan kesalahan. Warna gelap pada diagonal utama menunjukkan akurasi tinggi, terutama pada sentimen netral. Namun, terdapat kesalahan prediksi, terutama pada sentimen positif yang terkadang diprediksi sebagai netral atau negatif, akibat data yang tidak seimbang dan kemiripan antar label.

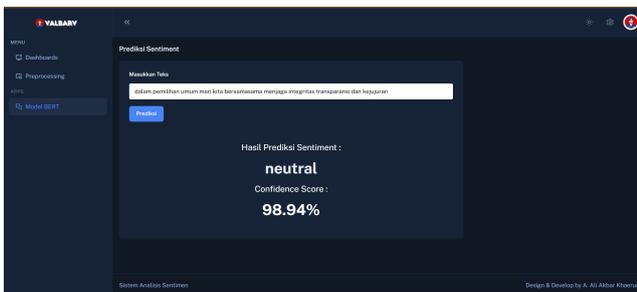
Setelah evaluasi dilakukan dan mendapatkan model yang optimal, model akan diimplementasikan kedalam sebuah website yang digunakan untuk menginput teks komentar yang ingin di prediksi.



Gambar 3.2. Tampilan Halaman Prediksi Positif



Gambar 3.3. Tampilan Halaman Prediksi Negatif



Gambar 3.4. Tampilan Halaman Prediksi Netral

4. KESIMPULAN

Penelitian ini berhasil dalam mengembangkan dan menerapkan model *Large Language Model*, khususnya IndoBERT-Large-P2, untuk mengenali teks sentimen berbahasa Indonesia dengan baik, memperhatikan variasi seperti gaya bahasa dan jenis teks. Model BERT yang digunakan menunjukkan performa tinggi dengan akurasi, presisi, recall, dan F1-

score sebesar 98%. Pelatihan dilakukan menggunakan Google Colab, yang menyediakan komputasi dan penyimpanan memadai. Antarmuka berbasis web yang dikembangkan dengan Laravel memudahkan pengguna dalam analisis sentimen.

5. DAFTAR PUSTAKA

- [1] F. Cindy Ambarsari, A. Fadillah Gizka, N. Zerlinda, N. Sasqya Mega Putri, M. Eka Putri Giyanti, dan D. Puspa Arum, "Bahasa Indonesia dalam Komunikasi Politik: Variasi Gaya Bahasa Calon Presiden dalam Debat Pertama Pilpres 2024," *Innov. J. Soc. Sci. Res.*, vol. 4, hal. 5702–5720, 2024, [Daring]. Tersedia pada: <https://j-innovative.org/index.php/Innovative/article/view/9989/6928>
- [2] P. A. M. Prasetyo dan A. Hermawan, "Analisis sentimen twitter terhadap pemilihan presiden menggunakan algoritma naïve bayes," *J. Inform. Teknol.*, vol. 4, no. 2, hal. 224–233, 2023, doi: 10.37373/infotech.v4i2.863.
- [3] M. W. Ward, "Twitter User Statistics 2024," *Search Logistics*, 2024. <https://www.searchlogistics.com/learn/statistics/twitter-user-statistics/>
- [4] D. Darwis, N. Siskawati, dan Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. Tekno Kompak*, vol. 15, no. 1, hal. 131–145, 2021, doi: <https://doi.org/10.33365/jtk.v15i1.744>.
- [5] F. Halim, Liliana, dan K. Gunadi, "Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT," *J. Infra*, vol. 10, no. 1, 2022.
- [6] A. Simanjuntak, R. Lumbantoruan, K. Sianipar, R. Gultom, M. Simaremare, dan S. Situmeang, "Studi dan Analisis Hyperparameter Tuning IndoBERT Dalam Pendeteksian Berita Palsu," *J. Nas. Tek. ELEKTRO DAN Teknol. Inf.*,

- vol. 13, hal. 60–67, 2024.
- [7] L. F. Simanjuntak, R. Mahendra, dan E. Yulianti, “We Know You Are Living in Bali: Location Prediction of Twitter Users Using BERT Language Model,” *Big Data Cogn. Comput.*, vol. 6, no. 3, 2022, doi: 10.3390/bdcc6030077.
- [8] N. Doede, P. Merkel, M. Kriwall, M. Stonis, dan B. A. Behrens, “Implementation of an intelligent process monitoring system for screw presses using the CRISP-DM standard,” *Prod. Eng.*, no. 0123456789, 2024, doi: 10.1007/s11740-024-01298-8.