Indonesian Journal of Computer Science
ISSN 2302-4364 (*print*) dan 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam, No. 1, Padang, Indonesia, Telp. (0751) 7056199, 7058325
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Analisis Sentimen Putusan Mahkamah Konstitusi Terhadap Batas Usia Capres Dan Cawapres Menggunakan IndoBERT

Luffi Septian[1], Teguh Aljauza[2], Christina Juliane[3]
luffi.septian@unigal.ac.id[1], teguhaljauza@gmail.com[2], christina.juliane@likmi.ac.id[3]
[1] Universitas Galuh
[2,3] STMIK LIKMI Bandung

Informasi Artikel

Diterima : 18 Dec 2024
Direview : 23 Dec 2024
Disetujui : 30 Dec 2024

Kata Kunci

Analisis Sentimen; IndoBERT; Putusan Mahkamah Konstitusi; Mahkamah Konstitusi; Capres Cawapres.

Abstrak

Putusan Mahkamah Konstitusi nomor 90/PUU-XXI/2023 tentang batas usia calon presiden dan wakil presiden telah memicu perbincangan masyarakat. Hal ini ditandai dengan kata kunci 'Putusan MK" pada media sosial Twitter/X menduduki peringkat tiga trending topik nasional selama pertengahan bulan Oktober. Putusan tersebut dinilai kontroversial karena berkaitan dengan momentum Pemilihan Presiden 2024. Penelitian ini menggunakan data dari media sosial Twitter/X untuk menganalisis sentimen masyarakat terhadap putusan Mahkamah Konstitusi. Model yang digunakan dalam penelitian ini adalah IndoBERT, sebuah arsitektur transformer BERT yang dikembangkan oleh tim IndoNLU. Metode ini dipilih berdasarkan efektivitasnya dalam memproses teks berbahasa Indonesia untuk mengidentifikasi dan mengategorikan opini publik menjadi positif, negatif, atau netral terkait dengan keputusan Mahkamah Konstitusi. Hasil awal menunjukkan model IndoBERT tanpa augmentasi data mencapai akurasi 0.81 dan F1 skor 0.58. Selanjutnya, penggunaan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) meningkatkan F1 skor namun tidak berdampak signifikan pada akurasi. Eksperimen selanjutnya dengan augmentasi random swap, menghasilkan peningkatan performa yang substansial, dimana model IndoBERT mencapai akurasi dan F1 skor sama-sama pada angka 0.90.

Keywords

Sentiment Analysis; IndoBERT; Constitutional Court Decision; Constitutional Court; Presidential Candidate

Abstract

The Constitutional Court decision number 90/PUU-XXI/2023 regarding the age limit for presidential and vice-presidential candidates has sparked public discussions. This is indicated by the keyword 'Putusan MK' (Constitutional Court Decision) on social media Twitter/X, which ranked third in the national trending topics during mid-October. The decision is considered controversial due to its connection with the 2024 Presidential Election momentum. This study utilizes data from social media Twitter/X to analyze the public response to the Constitutional Court decision. The model used in this research is IndoBERT, a BERT transformer architecture developed by the IndoNLU team. This method was selected based on its effectiveness in processing Indonesian language texts to identify and categorize public opinion as positive, negative, or neutral regarding the Constitutional Court's decision. Initial results show that the IndoBERT model without data augmentation achieved an accuracy of 0.81 and an F1 score of 0.58. Subsequently, the use of the Synthetic Minority Over-sampling Technique (SMOTE) improved the F1 score but did not significantly impact accuracy. Further experiments with random swap augmentation resulted in a substantial performance improvement, where the IndoBERT model achieved both an accuracy and F1 score of 0.90.

A. Pendahuluan

Keputusan Mahkamah Konstitusi mengenai batas usia calon presiden dan calon wakil presiden di Indonesia terhadap permohonan yang diajukan oleh Mahasiswa Universitas Surakarta, Almas Tsaqibbirru telah menemui tahap final, dalam Perkara Nomor 90/PUU-XXI/2023 Mahkamah Konstitusi menyatakan bahwa Pasal 169 huruf q dari Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum menetapkan usia minimal 40 tahun untuk capres dan cawapres bertentangan dengan UUD 1945 dan tidak memiliki kekuatan hukum mengikat, kecuali jika diinterpretasikan bahwa mereka yang berusia di bawah 40 tahun tetapi telah/sedang menduduki jabatan publik melalui pemilu, termasuk kepala daerah, dapat menjadi capres atau cawapres[1]. Dalam pertimbangan hukumnya, Mahkamah Konstitusi menekankan pentingnya partisipasi dan kualifikasi calon-calon yang berkualitas dan berpengalaman dalam pemilihan umum. Batasan usia semata bisa menghambat partisipasi generasi muda dalam kepemimpinan nasional dan potensial mengurangi peluang bagi tokoh-tokoh generasi muda. Putusan ini juga menegaskan bahwa seseorang yang berusia di bawah 40 tahun tetap bisa menjadi calon presiden atau wakil presiden, selama telah memiliki pengalaman memegang jabatan publik yang dipilih melalui pemilu, seperti anggota DPR, DPD, DPRD, Gubernur, Bupati, atau Walikota.

Keputusan mengenai batas usia calon presiden dan calon wakil presiden memunculkan berbagai pandangan dari kalangan masyarakat, salah satunya pada media sosial X atau dikenal sebelumnya sebagai twitter. Oleh karenanya kata kunci 'Putusan MK' menjadi trending topik di X/twitter selama periode pertengahan bulan september hingga akhir bulan oktober. twitter merupakan salah satu media sosial yang cukup populer di Indonesia dan menjadi platform untuk mengungkapkan pendapat masyarakat terhadap kebijakan publik yang dilakukan oleh pemerintah [2] [3]. Hal ini menjadikan peneliti menggunakan twitter sebagai media sosial untuk menganalisis sentimen terhadap putusan Mahkamah Konstitusi mengenai batas usia Capres dan Cawapres.

Analisis sentimen merupakan metodologi yang digunakan untuk mengevaluasi dan mengklasifikasikan pandangan yang dinyatakan dalam sebuah teks. Tujuannya adalah untuk menilai apakah reaksi terhadap suatu subjek spesifik bersifat positif, negatif, atau netral. Ini dilakukan melalui proses pembandingan yang memungkinkan identifikasi dan kategorisasi opini yang terkandung dalam teks tersebut [4] [5]. Analisis sentimen juga disebut sebagai bidang penelitian yang terus berkembang dan berada diantara beberapa disiplin ilmu, termasuk Data Mining, Pengolahan Bahasa Alami (*Natural Language Processing*, NLP), dan Pembelajaran Mesin (*Machine Learning*) yang berfokus pada proses penggalian sentimen dari sebuah kalimat, dengan mengacu pada isi atau konten dari kalimat tersebut [6]. Menganalisis sentimen dari teks yang berasal dari postingan dan komentar di media online merupakan tugas yang kompleks, hal ini disebabkan oleh ketidakteraturan yang ada dalam teks postingan maupun komentar, terutama pada teks berbahasa Indonesia yang memiliki karakteristik unik berbeda dari bahasa lain [7]. Atas dasar tersebut, penelitian ini menggabungkan pengolahan bahasa alami dengan teknologi pembelajaran mesin dalam menganalisis sentimen putusan Mahkamah Konstitusi mengenai batas usia Capres dan Cawapres.

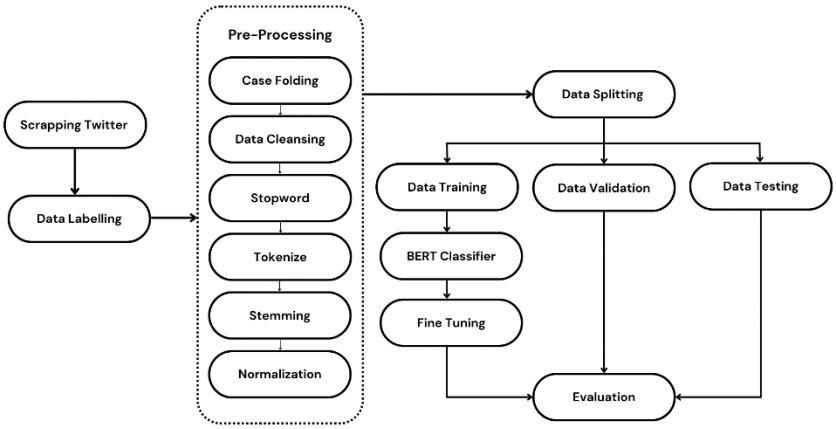
Analisis Sentimen semakin berkembang pesat dan mendapatkan perhatian yang signifikan dalam literatur penelitian [8]. Salah satu penelitian yang membahas analisis sentimen dilakukan oleh Lestandy, dkk pada tahun 2021 dengan judul "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes" dengan tujuan mendeteksi dan mencegah penyebaran informasi yang mungkin tidak benar atau membingungkan terkait vaksin, sehingga masyarakat dapat mengakses informasi yang akurat dan dapat dipercaya. Hasil dari pengujian dalam penelitian tersebut menunjukkan RNN (TF-IDF) memiliki akurasi lebih besar yaitu 97,77% dibandingkan Naïve Bayes (TF-IDF) sebesar 80% [9]. Penelitian serupa lainnya dilakukan pada tahun 2017 oleh Lestari, dkk dengan judul "Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji" Untuk memecahkan permasalahan diajukan solusi berupa sistem Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia dengan harapan dapat dijadikan sebagai perhitungan sekunder setelah lembaga survei dan quick count, serta mampu memperbaiki strategi kampanye para cagub dan cawagub. Dengan hasil pengujian akurasi, diperoleh 68,52% untuk kondisi pembobotan tekstual, 75,93% untuk pembobotan non-tesktual, dan 74,81% untuk kondisi penggabungan dengan nilai konstanta 0,5 untuk tekstual dan 0,5 untuk non-tekstual [10]. Adapun penelitian yang dilakukan Marwanta, dkk tentang Analisis Sentimen Pencitraan Perguruan Tinggi di Yogyakarta Menggunakan Metode Naïve Bayes Classifier pada tahun 2023, dengan tujuan untuk mengetahui sentimen masyarakat tentang Pendidikan tinggi di Yogyakarta, hasil penelitian menunjukan klasifikasi analisis sentimen dari data uji menunjukkan 82,1% data berkategori netral, 14,8% positif, dan 3,1% negatif, dengan nilai akurasi 73% [11].

Penelitian kami memperluas cakupan dan mendalami gap yang ditemukan dalam studi-studi sebelumnya di bidang analisis sentimen menggunakan model IndoBERT. Penelitian ini berfokus pada evaluasi pengaruh konfigurasi berbeda dari model IndoBERT, termasuk variasi learning rate dan batch size, terhadap efektivitas dalam analisis sentimen. Selanjutnya, penelitian ini menerapkan *Synthetic Minority Over-sampling Technique (*SMOTE) dan augmentasi random swap yang belum sepenuhnya dieksplorasi dalam penelitian-penelitian sebelumnya. Tujuan dari penelitian ini adalah untuk mengidentifikasi dan menganalisis sentimen publik terhadap keputusan politik yang kontroversial, menggunakan teknologi NLP yang disesuaikan untuk bahasa Indonesia. Harapan dari penelitian ini adalah memberikan kontribusi pada pemahaman yang lebih mendalam terhadap opini publik, serta membantu pembuat kebijakan terkait dalam merespon dinamika opini publik. Selain itu, penelitian ini bertujuan untuk memperkaya literatur dalam bidang NLP dan analisis sentimen dengan studi kasus yang spesifik pada konteks bahasa Indonesia.

B. Metode Penelitian

Dalam penelitian ini proses penelitian meliputi beberapa tahap, pertama pengambilan data Twitter (scrapping), dilanjutkan tahap pemberian label sentimen pada dataset secara manual. Dataset tersebut kemudian dipecah menjadi tiga bagian: 70% data pelatihan, 20% data validasi, dan 10% data pengujian.

Proses ini untuk memastikan bahwa model memiliki jumlah data yang cukup untuk belajar, serta data terpisah untuk menilai kinerja dan mencegah overfitting. Tahap berikutnya melibatkan pengembangan model klasifikasi sentimen yang memanfaatkan arsitektur IndoBERT. Selama tahap ini, fine-tuning dari hyperparameter dilakukan untuk meningkatkan efektivitas model dalam membedakan antara sentimen positif, negatif, atau netral dari teks yang dianalisis, diikuti dengan evaluasi untuk menguji keakuratan model. Gambaran umum tahapan penelitian terdapat pada gambar 1.



Gambar 1. Tahapan Penelitian IndoBERT

a. Pengumpulan Data

Proses pengambilan data untuk penelitian ini menggunakan library twitter scraper yang tersedia melalui PIP pada google colab. Dalam proses ini, library tersebut dioperasikan untuk mencari dan mengumpulkan tweet yang secara spesifik mencakup frasa "Putusan MK" sebagai kunci utama dalam pencarian. Fokus pencarian pada frasa ini bertujuan untuk mendapatkan data yang relevan dengan topik yang diteliti. Dalam kurun waktu mulai dari tanggal 1 September 2023 sampai dengan 25 Oktober 2023, Proses pengumpulan data (scrapping) berhasil mengumpulkan total 2819 tweet. Setelah melalui tahap penyaringan, jumlah data berkurang menjadi 2236 tweet yang layak analisis. Dari jumlah tersebut, teridentifikasi sebanyak 1698 tweet yang mengungkapkan sentimen negatif, 212 tweet dengan sentimen positif, dan 326 tweet yang bersifat netral. Kumpulan data ini kemudian akan dijadikan dasar untuk analisis lebih lanjut dalam memahami beragam sentimen dan pandangan yang diungkapkan oleh pengguna Twitter terkait dengan "Putusan MK".

b. Labelisasi Dataset

Labelling atau pemberian label pada dataset adalah proses penentuan kategori atau kelas untuk setiap elemen dalam dataset. Dalam penelitian ini labeling dilakukan secara manual untuk mengkategorikan sentimen menjadi "positif", "negatif", atau "netral" pada setiap data. Data yang diberi label kemudian diproses terlebih dahulu untuk memastikan kebersihan dan kesesuaian untuk klasifikasi [12].

Tabel 1. Contoh data yang sudah diberi label

Tweet	Label
Putusan mk justru ngasih kesempatan lebih banyak org jadi pemimpin	Positif
Rame banget komentar soal putusan mk ya	Netral
Putusan mk ini sungguh tidak bisa diterima dikalangan masyarakat	Negatif

Berdasarkan Tabel 1, tweet "Putusan mk justru ngasih kesempatan lebih banyak org jadi pemimpin" diberi label sebagai sentimen positif. Kalimat ini mengandung dukungan atau pandangan positif terhadap keputusan yang memberikan kesempatan lebih banyak kepada orang untuk menjadi pemimpin. Fokus kalimat ini adalah pada harapan positif atas adanya peluang yang diberikan oleh keputusan tersebut. Selanjutnya, tweet "Rame banget komentar soal putusan mk ya" diberi label netral. Kalimat ini hanya berfokus pada kenyataan bahwa ada banyak komentar mengenai keputusan mk, tanpa menunjukkan dukungan atau penolakan terhadap isi keputusan tersebut.

Terakhir, tweet "Putusan mk ini sungguh tidak bisa diterima dikalangan masyarakat" diberi label sebagai sentimen negatif. Kalimat ini menunjukkan adanya sentimen negatif atau kritikan terhadap keputusan MK yang tidak diterima oleh masyarakat. Ini menandakan adanya ketidakpuasan atau penolakan terhadap keputusan tersebut.

c. Pre-Proccessing Data

Preprocessing data dilakukan sebelum tweet dianalisis. Pemrosesan data awal menggabungkan beberapa prosedur untuk membersihkan dan mengubah data mentah menjadi format yang dapat dimengerti [13], [14]. Persiapan data diperlukan karena sebagian besar data masih menggunakan ejaan yang berbeda untuk bahasa yang berbeda atau karena beberapa fitur data tidak diperlukan untuk prosedur penelitian selanjutnya [15]. Tahapan pra-pemrosesan dalam penelitian ini mencakup beberapa langkah, meliputi :

1. Data cleansing
   Pembersihan data merupakan langkah penting untuk mengurangi noise tweet di dataset. Proses pembersihan data kami meliputi penghapusan data duplikat, tautan, nama pengguna (@namapengguna), tagar (#), angka, simbol, spasi tambahan, dan tanda baca [16] [17].
2. Case folding
   Selanjutnya pada tahap case folding, semua huruf akan diubah menjadi huruf kecil.
3. Stopword
   Stopword atau stopword removal dilakukan dimana kata-kata yang tidak memiliki signifikansi semantik dalam konteks analisis, seperti "dan", "atau", "tapi", "adalah", dan lain-lain, dihilangkan dari teks.
4. Stemming
   Dalam proses stemming kata-kata direduksi ke bentuk dasar atau akar kata, sehingga variasi dari kata yang sama, seperti "bermain", "bermain-main", dan "main" akan direduksi menjadi "main".

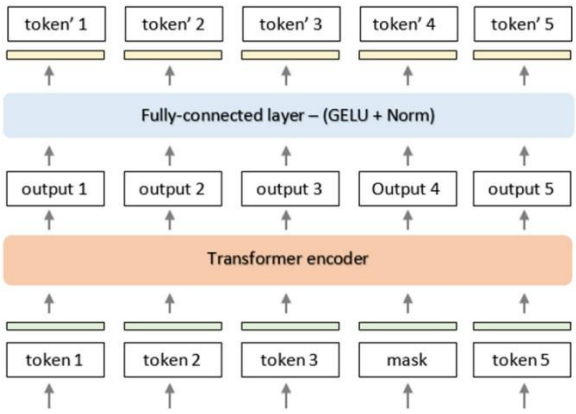
5. Tokenization
   Proses selanjutnya adalah pemisahan teks menjadi unit-unit yang lebih kecil.
6. Normalization
   Tahap normalization bertujuan untuk mengubah varian kata menjadi bentuk standarnya, sehingga kata-kata dengan makna yang sama namun bentuk yang berbeda dapat dianalisis secara konsisten. Proses-proses ini membantu dalam meningkatkan akurasi model NLP yang akan digunakan untuk menganalisis sentimen dari data teks yang dikumpulkan.

d. Data Splitting

Setelah melalui tahapan pre-processing, maka akan dibagi menjadi tiga bagian: data pelatihan, data validasi, dan data pengujian [16] Data pelatihan digunakan untuk mengembangkan model, data validasi berfungsi mengurangi overfitting dan meningkatkan kemampuan generalisasi model, sementara data pengujian digunakan untuk mengevaluasi kinerja model.

e. BERT Classifier

Penggunaan IndoBERT sebagai metode analisis sentimen bertujuan untuk menghasilkan representasi teks bahasa Indonesia yang akurat dan kaya secara linguistik. IndoBERT adalah model berbasis BERT yang telah pra-latih menggunakan korpus besar Indonesia. Arsitektur IndoBERT sama dengan BERT, yang terdiri dari tumpukan enkoder Transformer. IndoBERT memiliki dua varian arsitektur: IndoBERT-base dan IndoBERT-large. IndoBERT-base memiliki 12 lapisan transformer, sedangkan IndoBERT-large memiliki 24 lapisan. Arsitektur indoBERT dapat dilihat digambar 2.



Gambar 2. Arsitektur IndoBERT [17]

Secara umum, model BERT (termasuk IndoBERT) dilatih menggunakan tugas MLM dan NSP untuk menangkap representasi teks secara dua arah. Setelah pra-latih, IndoBERT dapat disesuaikan dengan menambahkan lapisan keluaran untuk tugas spesifik. Dalam proses penyesuaian, bobot yang dipelajari selama pra-latih akan diperbarui berdasarkan dataset spesifik. Kami bereksperimen dengan beberapa versi IndoBERT dengan data latih dan hyperparameter yang berbeda selama proses pra-latih.

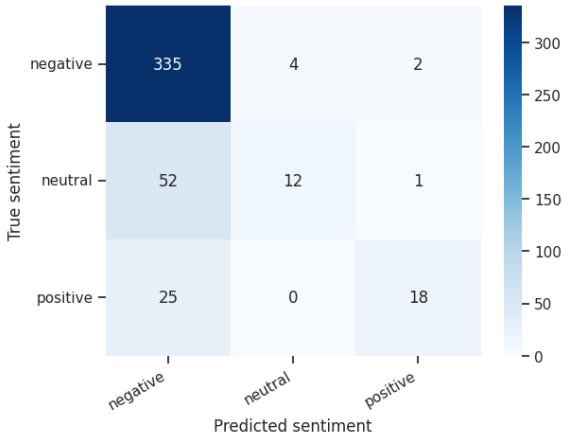
C. Hasil dan Pembahasan

Dalam penelitian ini, IndoBERT diuji dengan berbagai konfigurasi untuk mengevaluasi akurasi dan skor F1 dalam tugas analisis sentimen. Model IndoBERT base-p1 dengan learning rate 2e-5 dan batch size 16 menunjukkan akurasi sebesar 0.79 dan skor F1 0.59. Sementara itu, model large-p2 pada learning rate yang sama mencapai akurasi yang identik namun dengan skor F1 yang sedikit lebih tinggi, yaitu 0.60. Penggunaan learning rate 3e-6 dengan batch size 32 menunjukkan peningkatan akurasi menjadi 0.81 untuk base-p1, akan tetapi skor F1 menurun menjadi 0.58. Hasil ini menunjukkan bahwa peningkatan learning rate dan batch size dapat mempengaruhi akurasi dan skor F1 model, penggunaan base-p1 secara konsisten menunjukkan performa yang lebih baik dibandingkan dengan varian lain seperti yang disajikan pada tabel 2.

Tabel 2. Hasil Performa Model

	Learning Rate	Batch	Model	Accuracy	F1-Score
IndoBERT	2e 5	16	base-p1	0.79	0.59
			base-p2	0.77	0.59
			large-p1	0.78	0.57
			large-p2	0.79	0.60
	3e 6	32	**base-p1**	**0.81**	**0.58**
			base-p2	0.81	0.57
			large-p1	0.80	0.50
			large-p2	0.80	0.51

Matrix pada gambar 3 memberikan informasi tentang kemampuan model dalam mengklasifikasikan data dan jumlah kesalahan klasifikasi yang dihasilkan oleh model tersebut. Model ini dapat memprediksi sentimen negatif dengan sangat baik, dengan 335 tweet diklasifikasikan dengan benar. Namun, terdapat beberapa kesalahan klasifikasi, terutama dalam memprediksi sentimen netral, di mana 52 tweet netral salah diklasifikasikan sebagai negatif dan 25 tweet positif salah diklasifikasikan sebagai negatif. Ini menunjukkan bahwa model tersebut memiliki bias terhadap sentimen negatif.



Gambar 3. Confusion matrix IndoBERT menggunakan base-p1 dengan learning rate 3e 6

Kesalahan klasifikasi dapat disebabkan oleh ketidakseimbangan dalam dataset. Seperti yang dapat kita lihat dari Tabel 2, lebih dari 70% tweet dalam dataset kami memiliki sentimen negatif. Hal ini membuat model masih membutuhkan pengetahuan lebih untuk dapat lebih memahami karakteristik tweets positif dan netral.

a. Eksplorasi Dengan Metode SMOTE

Peneliti mengeksplorasi penggunaan *Synthetic Minority Over-sampling Technique* (SMOTE) untuk meningkatkan kinerja model klasifikasi pada dataset yang tidak seimbang. SMOTE merupakan salah satu metode resampling dengan tujuan untuk menyamakan distribusi kelas yang bekerja dengan cara mengambil sampel kelas minoritas dan melakukan penyisipan sampel sintesis untuk menambahkan jumlah data pada kelas minoritas [18]. Penelitian ini mengevaluasi efektivitas SMOTE pada empat model klasifikasi yang berbeda, yaitu Random Forest Classifier, Gradient Boosting, Naive Bayes, dan Support Vector Machine (SVM), metode validasi yang digunakan adalah K-Fold Cross-Validation dengan lima lipatan. Hasil evaluasi terdapat pada tabel 3.

Tabel 3. Hasil Evaluasi Penggunaan SMOTE Dengan Empat Model Klasifikasi

Model	Accuracy	F1-Score
Random Forest Classifier	**0.80**	**0.77**
Gradient Boosting	0.75	0.74
Naive Bayes	0.64	0.67
Support Vector Machine	0.80	0.76

Hasil evaluasi menunjukkan peningkatan signifikan dalam F1 skor untuk semua model klasifikasi yang diuji, sementara akurasi juga meningkat atau tetap stabil. Model Random Forest mencapai rata-rata akurasi 0.80 dan F1 skor 0.77, sementara model SVM mencapai akurasi 0.80 dan F1 skor 0.76. Hasil ini menunjukkan bahwa penerapan SMOTE secara efektif meningkatkan kemampuan model dalam mengklasifikasikan sampel dari kelas minoritas, yang sebelumnya menjadi tantangan dengan dataset tidak seimbang. Penggunaan SMOTE dalam penelitian ini memberikan peningkatan dengan kinerja sebelumnya menggunakan IndoBERT, yang mencapai akurasi 0.81 dan F1 skor 0.58. Hal ini mengindikasikan bahwa meskipun akurasi tidak banyak berubah, SMOTE secara signifikan meningkatkan F1 skor, yang menunjukkan peningkatan dalam keseimbangan antara presisi dan recall, terutama dalam menangani kelas minoritas.

b. Eksperimen Menggunakan IndoBERT Dengan Metode Augmentasi Random Swap

Wei dan Zou pada penelitiannya yang berjudul EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks [19] Menjelaskan beberapa teknik augmentasi data yang mudah (EDA), yaitu diantaranya penggantian sinonim kata (namely synonym word), pengacakan urutan kata (random swap), penyisipan kata acak (random insert), dan penghapusan kata acak (random delete) untuk menghasilkan data baru. Penggunaan metode random swap pada penelitian ini selain karena

kemudahannya dalam implementasi, random swap juga dipercaya kemampuannya untuk mempertahankan makna semantik asli. random Swap akan menukar dua kata yang bukan stopword secara acak sebanyak n kali dalam kalimat [20]. Setiap kalimat sentimen positif dan netral dalam dataset asli penelitian ini dikalikan sebanyak tiga kali, menghasilkan variasi baru yang meningkatkan ukuran dataset tanpa kehilangan konteks asli, contoh hasil impementasi random swap terdapat pada tabel 4.

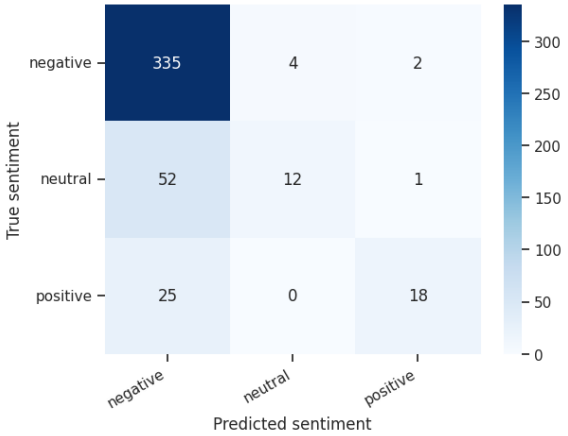
Tabel 4. Contoh Hasil Random Swap pada Kalimat

Teks	Keterangan
"Putusan mk sudah jelas sah mengikat"	Teks Asli
"Putusan sudah mk jelas sah mengikat"	*Swap* 1
"Sah mk sudah jelas putusan mengikat"	*Swap* 2
"Mk mengikat sudah jelas sah putusan"	*Swap* 3

Dataset awal terdiri dari 212 kalimat sentimen positif dan 326 kalimat sentimen netral, kemudian data di proses menggunakan metode augmentasi random swap. Proses ini menghasilkan tambahan 636 kalimat dengan sentimen positif dan 978 kalimat dengan sentimen netral. Langkah berikutnya adalah menggabungkan data yang telah di-augmentasi ini dengan data asli yang meliputi 1698 sentimen negatif, 212 sentimen positif dan 326 sentimen netral. Total keseluruhan dataset baru hasil augmentasi menggunakan random swap menjadi 3850, diantaranya 1698 sentimen negatif, 848 sentimen positif, 1304 sentimen netral. Eksperimen augmentasi ini bertujuan untuk memperkaya dataset, meningkatkan variabilitas data, dan mengatasi isu ketidakseimbangan kelas, yang diharapkan dapat meningkatkan kemampuan model klasifikasi sentimen dalam memahami dan memprediksi berbagai bentuk ekspresi sentimen.

c. Evaluasi

Pengujian kedua kembali dilakukan dengan model IndoBERT, namun kali menggunakan dataset yang telah di-augmentasi dengan random swap, hasil terbaik diraih menggunakan konfigurasi IndoBERT base-p1 dan learning rate sebesar 2e-5 dengan batch size 16. Hasil pengujian menunjukan akurasi dan F1 skor sama-sama mencapai angka 0.90, lebih tepatnya 0.9030 dan F1 Skor pada angka 0.9020. Confusion matrix pada gambar 5, dapat kita lihat bahwa model efektif dalam mengidentifikasi sentimen negatif, dengan 314 dari 340 kasus diklasifikasikan dengan benar. Namun, ada beberapa tantangan dalam klasifikasi sentimen netral dan positif, dimana ada 52 tweet netral yang salah diklasifikasikan sebagai negatif dan 25 tweet positif yang juga salah diklasifikasikan sebagai negatif. Meskipun demikian, model masih berhasil mengklasifikasikan 232 dari 262 kasus netral dan 152 dari 171 kasus positif dengan benar.



Gambar 5. Confusion matrix IndoBERT Menggunakan Metode Augmentasi Random Swap

Hasil evaluasi menunjukan adanya peningkatan signifikan dibandingkan dengan konfigurasi sebelumnya tanpa augmentasi, dimana hasil terbaik model IndoBERT sebelumnya tanpa augmentasi random swap hanya mencapai akurasi 0.81 dan skor F1 0.58. Lebih lanjut, hasil ini juga menunjukkan perbaikan dibandingkan dengan penggunaan SMOTE, yang meskipun meningkatkan skor F1, namun tidak memberikan peningkatan pada akurasi. Evaluasi ini menegaskan bahwa augmentasi data dapat mengatasi kelemahan inherent dari dataset tidak seimbang dan meningkatkan kemampuan model dalam mengidentifikasi dan memprediksi sentimen dengan lebih akurat.

D. Simpulan

Penelitian ini telah berhasil mengevaluasi efektivitas berbagai konfigurasi model IndoBERT dalam analisis sentimen, menyoroti bagaimana perubahan dalam learning rate dan batch size memengaruhi akurasi dan skor F1. Dari pengujian yang dilakukan, konfigurasi base-p1 dengan learning rate 2e-5 dan batch size 16 menunjukkan performa yang paling konsisten dan efektif, yang menandakan pentingnya pemilihan parameter yang tepat untuk tugas analisis sentimen dalam konteks bahasa Indonesia. Selanjutnya, penelitian ini menunjukkan bahwa penggunaan *Synthetic Minority Over-sampling Technique* (SMOTE) dapat meningkatkan F1 skor secara signifikan, terutama dalam mengatasi tantangan klasifikasi pada dataset yang tidak seimbang. Hal ini mengindikasikan bahwa meskipun akurasi tidak mengalami peningkatan yang signifikan, SMOTE memperbaiki keseimbangan antara presisi dan recall, khususnya dalam mengklasifikasikan kelas minoritas. Eksperimen dengan augmentasi random swap membawa peningkatan performa yang lebih lanjut, dengan kedua akurasi dan F1 skor mencapai sekitar 0.90, yang menegaskan efektivitas augmentasi data dalam meningkatkan kinerja model analisis sentimen. Namun, terdapat keterbatasan dalam klasifikasi sentimen netral dan positif, yang menunjukkan bahwa model masih memerlukan penyesuaian lebih lanjut untuk meningkatkan kemampuan dalam mengenali dan membedakan beragam jenis sentimen dengan lebih akurat. enggunaan model IndoBERT, terutama dengan konfigurasi yang tepat dan

penerapan teknik augmentasi data seperti SMOTE dan random swap, menawarkan peningkatan yang signifikan dalam analisis sentimen. Namun, penelitian ini juga menunjukkan bahwa masih ada ruang untuk perbaikan, terutama dalam meningkatkan akurasi klasifikasi sentimen netral dan positif. Hal ini membuka peluang untuk penelitian lebih lanjut dalam pengembangan dan penyesuaian model IndoBERT, serta eksplorasi teknik augmentasi data lainnya yang mungkin memberikan hasil yang lebih akurat dalam analisis sentimen. Diharapkan, penelitian ini dapat memberikan dasar yang kuat untuk studi-studi mendatang dalam bidang NLP dan analisis sentimen, khususnya dalam konteks bahasa Indonesia.

E. Referensi

[1] Mahkamah Konstitusi Republik Indonesia, "Batas Usia Capres-Cawapres 40 Tahun Atau Menduduki Jabatan yang Dipilih dari Pemilu/Pilkada," Mahkamah Konstitusi Republik Indonesia, "Batas Usia Capres-Cawapres 40 Tahun Atau Menduduki Jabatan yang Dipilih dari Pemilu/Pilkada." Accessed: Oct. 16, 2023. [Online].Available:www.mkri.id/index.php?page=web.Berita&id=19660&menu=2#

[2] B. W. Sari and F. F. Haranto, "Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet," Jurnal Pilar Nusa Mandiri, vol. 15, no. 2, pp. 171–176, Sep. 2019, doi: 10.33480/pilar.v15i2.699.

[3] H. B. Jatmiko, N. T. Kurniadi, and D. Maulana, "Optimasi Naïve Bayes Dengan Particle Swarm Optimization Untuk Analisis Sentimen Formula E-Jakarta Optimization of Naïve Bayes With Particle Swarm Optimization for Sentimen Analysis of Jakarta E-Prix," 2022.

[4] R. Jannati, R. Mahendra, C. W. Wardhana, and M. Adriani, "Stance Classification Towards Political Figures on Blog Writing," in 2018 International Conference on Asian Language Processing (IALP), IEEE, Nov. 2018, pp. 96–101. doi: 10.1109/IALP.2018.8629144.

[5] Oxford Disctionaries, "Sentiment Analysis." Accessed: Nov. 22, 2023. [Online]. Available: https://en.oxforddictionaries.com/definition/sentiment_analysis/

[6] M. Al-Ayyoub, A. A. Khamaiseh, Y. Jararweh, and M. N. Al-Kabi, "A comprehensive survey of arabic sentiment analysis," Inf Process Manag, vol. 56, no. 2, pp. 320–342, Mar. 2019, doi: 10.1016/j.ipm.2018.07.006.

[7] T. Apriyanti, H. Wulandari, M. Safitri, and N. Dewi, "Translating Theory of English into Indonesian and Vice-Versa," 2016.

[8] I. P. Windasari, F. N. Uzzi, and K. I. Satoto, "Sentiment analysis on Twitter posts: An analysis of positive or negative opinion on GoJek," in 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), IEEE, Oct. 2017, pp. 266–269. doi: 10.1109/ICITACEE.2017.8257715.

[9] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 5, no. 4, pp. 802–808, Aug. 2021, doi: 10.29207/resti.v5i4.3308.

[10] A. Rossi, T. Lestari, R. Setya Perdana, and M. A. Fauzi, "Analisis Sentimen Tentang Opini Pilkada Dki 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji," 2017. [Online]. Available: http://j-ptiik.ub.ac.id

[11] Yy. Marwanta, "Analisis Sentimen Pencitraan Perguruan Tinggi di Yogyakarta Menggunakan Metode Naïve Bayes Classifier," 2023. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[12] H. Imaduddin, F. Yusfida A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach." [Online]. Available: www.ijacsa.thesai.org

[13] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.

[14] A. Thakkar, D. Mungra, A. Agrawal, and K. Chaudhari, "Improving the Performance of Sentiment Analysis Using Enhanced Preprocessing Technique and Artificial Neural Network," IEEE Trans Affect Comput, vol. 13, no. 4, pp. 1771–1782, Oct. 2022, doi: 10.1109/TAFFC.2022.3206891.

[15] F. Putra Herlambang and D. Avianto, "Jurnal Media Informatika Budidarma Analisis Sentimen Opini Pengguna Twitter Terhadap Tragedi Kanjuruhan Malang dengan Metode Support Vector Machine," vol. 7, pp. 1727–1739, 2023, doi: 10.30865/mib.v7i4.6332.

[16] Merfat. M. Altawaier and S. Tiun, "Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis," Int J Adv Sci Eng Inf Technol, vol. 6, no. 6, p. 1067, Dec. 2016, doi: 10.18517/ijaseit.6.6.1456.

[17] C. H. Miranda, G. Sanchez-Torres, and D. Salcedo, "Exploring the Evolution of Sentiment in Spanish Pandemic Tweets: A Data Analysis Based on a Fine-Tuned BERT Architecture," Data (Basel), vol. 8, no. 6, p. 96, May 2023, doi: 10.3390/data8060096.

[18] H. Jurnal, A. Nurhopipah, and C. Magnolia, "Jurnal Publikasi Ilmu Komputer Dan Multimedia Perbandingan Metode Resampling Pada Imbalanced Dataset Untuk Klasifikasi Komentar Program MBKM," JUPIKOM, vol. 1, no. 2, 2022.

[19] J. Wei and K. Zou, "EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks," Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.11196

[20] H. T. Duong and T. A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis," Comput Soc Netw, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40649-020-00080-x.