Jurnal Kecerdasan Buatan dan Teknologi Informasi
Vol. 4, No. 2, Mei 2025, hlm. 152-163
e-ISSN: 2964-2922, p-ISSN: 2963-6191
DOI : https://doi.org/10.69916/jkbti.v4i2.322

ANALISIS SENTIMEN PUBLIK DI MEDIA SOSIAL TERHADAP KENAIKAN PPN 12% DI INDONESIA MENGGUNAKAN *INDOBERT*

Michael Reynald Manoppo[*1], Indri Claudia Kolang[2], Daffa Nur Fiat[3], Reza Michelly Cantika Mawara[4], Anggraini Dwi Putri Sumarno[5] , Ade Yusupa[6], Victor Tarigan[7]

[1,2,3,4,5,6,7]Teknik Informatika, Fakultas Teknik, Universitas Sam Ratulangi, Manado, Indonesia
Email: [1]mikelnoppo03@gmail.com, [2]claudiakolang@gmail.com, [3]nurfiatdaffa@gmail.com , [4]michellymawara@gamil.com, [5]anggrainipsumarno18@gmail.com, [6]ade@unsrat.ac.id, [7]victortarigan@unsrat.ac.id

(Diterima : 24 April 2025, Direvisi : 10 Mei 2025, Disetujui : 18 Mei 2025)

Abstrak

Penelitian ini menganalisis sentimen publik terkait rencana kenaikan Pajak Pertambahan Nilai (PPN) 12% di Indonesia menggunakan model *transformer* berbasis Bahasa Indonesia, *IndoBERT*. Dengan mengumpulkan 2.581 sampel data dari platform media sosial X, Instagram, dan TikTok, penelitian ini *BERT*ujuan untuk memahami respons publik secara mendalam. Data melalui tahapan pra-pemrosesan, *tokenisasi*, dan *label mapping* sebelum dibagi 80/10/10 menjadi set pelatihan, validasi, dan pengujian. Model *IndoBERT* dasar yang di-*fine-tuned* selama tiga *epoch* menunjukkan kinerja yang signifikan pada set pengujian. Secara kuantitatif, model mencapai *accuracy* 84,94%, *precision* 85,60%, *recall* 84,94%, dan *F1-score (weighted)* 84,37%. Analisis distribusi sentimen lebih lanjut menunjukkan bahwa sentimen publik yang dominan adalah negatif. Tingginya nilai metrik evaluasi ini menegaskan efektivitas *IndoBERT* untuk tugas klasifikasi sentimen berbahasa Indonesia pada data media sosial. Kesimpulannya, temuan ini tidak hanya menunjukkan kapabilitas model, tetapi juga memberikan dasar analisis yang kuat mengenai tingkat penerimaan atau penolakan publik terhadap kebijakan kenaikan PPN 12%, menawarkan nilai tambah bagi pemahaman dampak kebijakan publik.

Kata kunci: analisis sentimen, *IndoBert*, kenaikan PPN, media sosial, *natural language processing*

ANALYSIS OF PUBLIC SENTIMENT ON SOCIAL MEDIA TOWARDS THE 12% PPN INCREASE IN INDONESIA USING INDOBERT

Abstract

This research analyzes public sentiment regarding the planned 12% Value Added Tax (VAT) increase in Indonesia using an Indonesian-based transformer model, IndoBERT. By collecting 2581 data samples from social media platforms X, Instagram, and TikTok, this research aims to understand public responses in depth. The data went through pre-processing, tokenization, and label mapping stages before being divided 80/10/10 into training, validation, and testing sets. The basic IndoBERT model fine-tuned over three epochs showed significant performance on the testing set. Quantitatively, the model achieved 84.94% accuracy, 85.60% precision, 84.94% recall, and 84.37% F1-score (weighted). Further sentiment distribution analysis shows that the dominant public sentiment is negative. The high value of these evaluation metrics confirms the effectiveness of IndoBERT for the task of Indonesian sentiment classification on social media data. In conclusion, these findings not only demonstrate the capabilities of the model, but also provide a strong analytical basis regarding the level of public acceptance or rejection of the 12% VAT increase policy, offering added value to the understanding of public policy impact.

Keywords: *sentiment analysis, IndoBert, tax increase vat, social media, natural language processing.*

1. PENDAHULUAN

Pajak berfungsi sebagai salah satu cara untuk memperoleh pendapatan bagi suatu negara, termasuk Indonesia. Pajak juga menn dan kritik publik dalam politik. Informasi ini penting bagi pemerintah dalam menyusun kebijakan untuk meminimalkan dampak negatif dan memaksimalkan manfaat positif [1], [2]. Indonesia mengalami kemajuan dengan hadirnya turunan dari *BERT*, *IndoBERT*. Penelitian yang dilakukan oleh membuktikan bahwa *IndoBERT*

mampu mencapai tingkat akurasi yang lebih tinggi dalam mengklasifikasikan sentimen tweet berbahasa Indonesia dibandingkan metode machine learning konvensional.

BERT (*Bidirectional Encoder Representations from Transformers*) adalah model bahasa yang dikembangkan oleh Google pada tahun 2018. Model ini menggunakan arsitektur *Transformer* yang memungkinkannya memahami konteks kata secara dua arah (*bidirectional*), yakni dengan mempertimbangkan kata-kata sebelum dan sesudah dalam sebuah kalimat. Hal ini menjadi keunggulan dibandingkan model-model sebelumnya yang umumnya hanya memproses teks secara satu arah (*unidirectional*). Versi *BERT Base* terdiri atas 110 juta parameter, 12 lapisan (*layers*), ukuran tersembunyi (*hidden size*) sebesar 768, serta 12 kepala perhatian diri (*self-attention heads*). Sementara itu, *BERT Large* memiliki 340 juta parameter, 24 lapisan, *hidden size* sebesar 1.024, dan 16 *self-attention heads*. Konfigurasi ini memungkinkan *BERT* memahami kosakata, tata bahasa, serta hubungan antar-teks secara lebih mendalam. Dalam versi yang disesuaikan untuk bahasa Indonesia, yaitu *IndoBERT*, terdapat dua tugas utama yang digunakan selama proses *pre-training*: *Masked Language Modeling* (MLM) dan *Next Sentence Prediction* (NSP). Tugas MLM melatih model untuk memprediksi kata-kata yang sengaja dihilangkan dari sebuah kalimat, sementara NSP membantu model memahami hubungan logis antara dua kalimat. Dengan pendekatan tersebut, *IndoBERT* mampu memperoleh pemahaman yang kuat terhadap struktur dan konteks bahasa Indonesia bahkan sebelum proses pelatihan lebih lanjut (*fine-tuning*) [3].

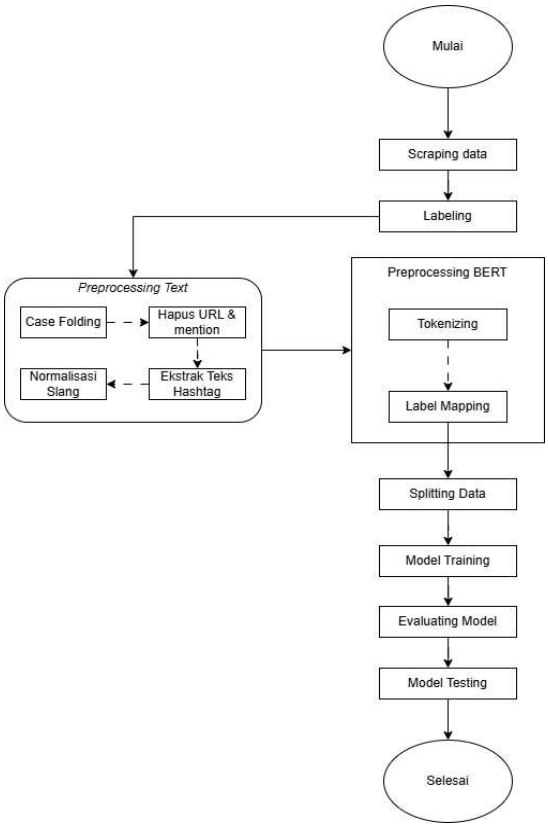
Berbagai ulasan dan tanggapan dari pengguna di platform seperti X, TikTok, dan Instagram mencerminkan penolakan mereka terhadap kenaikan PPN 12%, yang sering kali berisi kritik dan masukan. Beberapa masyarakat bahkan menyatakan enggan untuk membayar pajak di tahun berikutnya karena dianggap memberatkan. Hal ini menunjukkan adanya kebutuhan untuk memahami sentimen masyarakat terhadap kebijakan pemerintahan yang baru mengenai kenaikan PPN hingga 12%. Penelitian terbaru mengenai *IndoBERT* dalam *sentiment analysis* dan klasifikasi teks dalam bahasa Indonesia menunjukkan bahwa mengintegrasikan *Confident Learning* dengan *IndoBERT* meningkatkan akurasi analisis sentimen dari 85,5% menjadi 86,03% [4]. Dengan memanfaatkan teknologi *Natural Language Processing* (NLP), khususnya model *IndoBERT* yang dirancang untuk bahasa Indonesia, analisis ini dapat membantu dalam mengklasifikasikan ulasan menjadi kategori sentimen positif, negatif, dan netral. Model *IndoBERT* memiliki kemampuan untuk memahami bahasa Indonesia dengan baik, yang memungkinkan dihasilkannya analisis yang lebih akurat [5]. Penelitian [6] menggunakan *Naive Bayes* untuk melakukan *sentiment analysis* terhadap tanggapan masyarakat Indonesia mengenai rencana kenaikan PPN menjadi 12% di media sosial *X*, dan hasil pengujian memperoleh akurasi 83% dengan presisi sebesar 68,8% dan *recall* sebesar 78,6%. Penelitian [7] menerapkan *BERT* (*Bidirectional Encoder Representations from Transformers*) untuk melakukan *sentiment analysis* terhadap komentar YouTube terkait kenaikan PPN 12%. Penelitian ini juga menyoroti pentingnya memahami dinamika sentimen publik melalui platform digital seperti *YouTube* untuk menilai persepsi masyarakat terhadap kebijakan ekonomi. Temuan ini diharapkan dapat memberikan wawasan kepada pembuat kebijakan dalam merumuskan strategi komunikasi yang lebih efektif dan responsif terhadap kritik publik. Sementara itu, penelitian [8] membahas analisis sentimen publik di platform *X* terhadap rencana kenaikan PPN menjadi 12% menggunakan model *BERT*. Hasil penelitian menunjukkan dominasi sentimen negatif (48,58%), dengan kekhawatiran masyarakat terhadap kenaikan biaya hidup, disusul oleh sentimen netral (42,39%) dan positif (9,03%). Model *BERT* berhasil mencapai akurasi sebesar 83%, dan visualisasi *word cloud* mengungkap kata-kata seperti "harga", "rakyat", dan "beban" dalam kategori sentimen negatif. Penelitian ini memberikan wawasan berbasis data untuk mitigasi dampak kebijakan.

Penelitian ini *BERT*ujuan untuk menganalisis persepsi publik di Indonesia terhadap kebijakan kenaikan Pajak Pertambahan Nilai (PPN) menjadi 12%, dengan fokus pada pengguna platform X (sebelumnya Twitter). Analisis dilakukan melalui pendekatan *sentiment analysis* menggunakan model bahasa *IndoBERT*, yang dirancang khusus untuk memahami konteks dan nuansa bahasa Indonesia secara lebih akurat. Dengan pendekatan ini, penelitian tidak hanya mengungkap distribusi sentimen positif, negatif, dan netral dari masyarakat, tetapi juga memberikan pemahaman yang lebih dalam tentang ekspresi dukungan maupun penolakan terhadap kebijakan tersebut. Selain itu, hasil penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pemerintah dalam merumuskan strategi komunikasi kebijakan publik yang lebih efektif dan responsif terhadap opini masyarakat, serta menjadi landasan dalam pengembangan studi lanjutan terkait kebijakan yang fiskal.

2. METODE PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif dengan metode analisis sentimen berbasis *Natural Language Processing* (NLP). Proses penelitian ini mencakup beberapa tahapan utama, yaitu pengumpulan data, pra-pemrosesan data, implementasi model, serta evaluasi kinerja model. Model yang digunakan dalam penelitian ini adalah *IndoBERT* indobenchmark/*IndoBERT*-base-p1, yang merupakan model berbasis *Transformer* yang telah dioptimalkan untuk pemrosesan bahasa Indonesia. *IndoBERT* digunakan untuk mengklasifikasikan sentimen menjadi tiga kategori: positif, negatif, dan netral. Model ini dipilih karena memiliki keunggulan dalam memahami konteks bahasa Indonesia dibandingkan metode klasik seperti *Naïve Bayes* atau *Support Vector Machine* (SVM) [9].

Metode berbasis *pemrosesan bahasa alami* (*Natural Language Processing* atau NLP) ini menggunakan teknik *pemrosesan bahasa alami* untuk dapat lebih meningkatkan kinerja kedua metode yang digunakan. *BERT* dapat memahami makna kata dari berbagai sudut pandang karena kemampuan uniknya untuk menangkap konteks kata secara dua arah. Dalam hal akurasi dan skor *F1*, model ini meningkatkan sistem *baseline* sebesar 48,5% dan 46,49% untuk *TextBlob*, 2,5% dan 14,49% untuk *Multinomial Naïve Bayes*, dan 3,5% dan 13,49% untuk *Support Vector Machine*. Dengan metode *Bidirectional Long Short-Term Memory* dibandingkan dengan *Long Short-Term Memory*, hasil penelitian menunjukkan bahwa metode *Bidirectional Long Short-Term Memory* lebih baik daripada *Long Short-Term Memory*, dengan akurasi terbaik sebesar 91% dan kehilangan instruksi sebesar 28%. Penelitian yang dilakukan oleh Dloifur Rohman Alghifari akan dilakukan pada tahun 2022 [10].



Gambar 1 . Metodologi Penelitian

Dalam penelitian ini terdapat metodologi penelitian yang dapat dilihat pada Gambar 1, di mana akan dilakukan sebuah pengumpulan data dengan cara *scraping* dan juga menggunakan AI yang selanjutnya masuk ke tahap *labeling* untuk menentukan data tersebut apakah bersifat netral, negatif, atau positif. Setelah itu, masuk ke dalam pra-pemrosesan atau memproses teks dari non-informatif menjadi terstandarisasi dengan melakukan *case folding*, menghapus URL dan *mention*, normalisasi slang, dan ekstrak teks *hashtag*.

Pada bagian selanjutnya akan dilakukan *preprocessing IndoBERT* dengan membuat *tokenizing* dan juga *label mapping* dengan mengganti netral menjadi nilai '0', negatif menjadi '1', dan positif menjadi nilai '3'. Setelah dilakukan sebuah *preprocessing IndoBERT*, data akan dibagi menjadi tiga bagian yaitu *x*, *y*, dan *z* di mana *x* merupakan data training, *y* data validasi, dan *z* sebagai data testing [11].

Pada data training akan digunakan untuk masuk ke dalam tahapan model training menggunakan *IndoBERT* yang nantinya hasil dari training model akan dilakukan evaluasi model untuk dapat mengetahui apakah model yang

digunakan bagus atau tidak dengan melihat dari akurasi, presisi, *recall*, dan *F1-score*. Jika sudah melewati evaluasi model, maka tahap selanjutnya merupakan model testing di mana hasil training akan dilakukan pengujian untuk dapat mengetahui seberapa baik model dalam melakukan training dari dataset yang telah diberikan. Kosakata bahasa Indonesia yang luas (lebih dari 220 juta kata) digunakan untuk membangun *IndoBERT*. Kosakata ini diperoleh dari sumber bahasa yang menggunakan bahasa Indonesia yang baik dan benar, seperti koran daring, korpus web Indonesia, dan sumber lainnya [12].

2.1. Pengumpulan Data

Pada bagian ini data akan diambil melalui ketiga platform yaitu Twitter, Instagram, dan TikTok yang dikumpulkan melalui API dan *scraping* menggunakan *tweet harvest*. Data yang sudah terkumpul akan digabungkan menjadi satu dataset utuh yang disimpan dalam bentuk *csv*. Pengumpulan data difokuskan dengan melihat sebuah postingan dan komentar yang mengandung unsur kata kunci terkait kenaikan PPN 12%, seperti, “PPN 12%”, “Pajak Pertambahan nilai naik”, “Dampak PPN 12%”, ”Kenaikan PPN”. Pada hasil pengumpulan dataset yang berhasil dikumpulkan untuk dilakukan sebuah klasifikasi sentimen mencakup rentang waktu dari 10 September 2024 (sebagai data terlama) hingga 17 April 2025 (sebagai data terbaru).

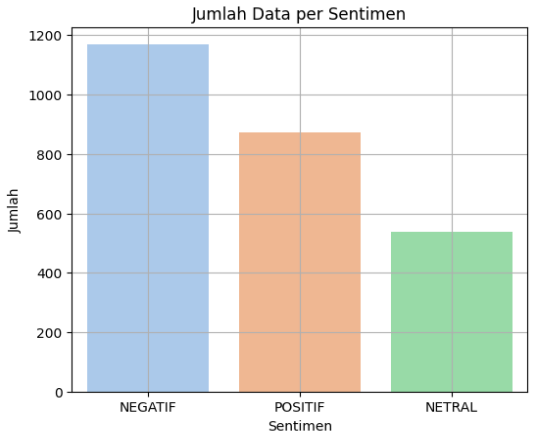
2.2. Labeling

Setelah, pengumpulan data telah dilakukan, maka berikutnya akan dilakukan fase labeling data secara manual, dimana awalan dari melakukan labeling dengan cara membagi menjadi 5 bagian yang dimana setiap bagian terdapat 500 data yang nantinya akan dilakukan labeling secara manual untuk menentukan apakah data tersebut bersifat Netral, Positif atau Negatif.

Tabel 1. Contoh hasil Labeling Data

Data	Kategori Sentiment atau Labeling
Gw liat dari berita sih PPN 12% ini memang udah direncanakan sejak lama.	**Netral**
Refund PPN 12% non produk mewah ini lumayan membantu konsumen. Semoga makin efisien prosesnya!	**Positif**
Gara-gara PPN naik 12% jadi harus mikirin harga barang sebelum belanja. Makin mahal aja hidup sekarang!	**Negatif**

Dapat dilihat pada tabel 1 menampilkan sebuah contoh pelabelan data secara manual berdasarkan analisis sentimen terhadap kenaikan Pajak Pertambahan Nilai (PPN) sebesar 12%. Proses pelabelan ini mengklasifikasikan data ke dalam tiga kategori: negatif, netral, dan positif. Kategori negatif diberikan pada data yang mengandung kata-kata atau frasa bernada negatif, sarkastik, atau kritik yang secara implisit maupun eksplisit tidak mendukung implementasi kenaikan PPN 12%. Kategori netral diterapkan pada data yang tidak menunjukkan dukungan maupun penolakan terhadap kenaikan PPN 12%, atau bersifat ambigu terkait isu tersebut. Terakhir, kategori positif diberikan pada data yang mengandung pernyataan dukungan terhadap pemberlakuan PPN 12%. Beberapa penelitian sebelumnya digunakan untuk melabelkan tweet sebagai berita dalam pedoman pelabelan yang kami buat [13].



Gambar 2 . Persentase data untuk setiap kelas sentimen

Gambar 2 menyajikan distribusi jumlah sampel data untuk setiap kelas sentimen (Negatif, Positif, dan Netral) dalam dataset yang digunakan setelah proses pelabelan awal. Secara visual, dapat diamati bahwa terdapat ketidakseimbangan dalam jumlah data antar kelas.Distribusi ini menunjukkan bahwa sentimen Negatif merupakan kelas mayoritas, diikuti oleh sentimen Positif, sementara sentimen Netral menjadi kelas minoritas. Ketidakseimbangan kelas (class imbalance) ini merupakan karakteristik penting dari dataset yang dapat mempengaruhi proses pelatihan dan evaluasi model. Model machine learning cenderung lebih bias atau memiliki performa yang lebih baik pada kelas mayoritas jika ketidakseimbangan ini tidak ditangani dengan tepat.

Untuk mengatasi potensi bias akibat ketidakseimbangan kelas selama pelatihan model *IndoBERT*, penelitian ini menerapkan teknik pembobotan kelas (class weighting). Secara spesifik, pada fungsi loss selama proses fine-tuning, bobot yang lebih besar diberikan kepada kelas minoritas (Netral dan Positif) dan bobot yang lebih kecil kepada kelas mayoritas (Negatif). Tujuannya adalah untuk memberikan "penalti" yang lebih besar kepada model jika salah mengklasifikasikan sampel dari kelas minoritas, sehingga mendorong model untuk belajar mengenali pola dari semua kelas secara lebih seimbang, meskipun jumlah sampelnya berbeda. Penggunaan weighted average untuk metrik evaluasi seperti F1-score juga dilakukan untuk memberikan gambaran performa yang lebih adil dalam situasi ketidakseimbangan kelas ini.

2.3. Pra-pemrosesan Data

Pra-pemrosesan dilakukan setelah melakukan labeling data sesuai dengan kategori yang sudah ditentukan yaitu netral, positif, negatif [14]. Pada tahap ini data yang sudah dilakukan label, selanjutnya akan dilakukan *Case folding* yaitu melakukan perubahan teks menjadi huruf kecil untuk konsistensi, Jika sudah maka akan masuk ke tahap berikutnya dengan menghilangkan elemen yang tidak relevan atau tidak bisa digunakan dalam klasifikasi sentimen seperti username dan URL. Selanjutnya, akan masuk kedalam normalisasi slang yang dimana akan mengubah kata-kata seperti siang atau tidak baku dengan bentuk standar menggunakan kamus slang Bahasa Indonesia. Menghilangkan tanda baca standar juga diperlukan dalam melakukan *Fine-tuning* dengan menghapus kalimat yang tidak diperlukan. Pembersihan ini sangat diperlukan untuk analisis yang dapat dilakukan untuk lebih akurat. Dengan melakukan *praprocessing* data yang baik, kita dapat menghasilkan dataset yang berkualitas dan dapat diandalkan untuk analisis lebih lanjut serta memastikan bahwa model yang akan dibangun berhasil [15].

2.4. *Preprocessing BERT*

Data yang telah melalui pembersihan dan normalisasi di bagian pra-pemrosesan data. Selanjutnya akan dilakukan sebuah tahapan *preprocessing BERT* untuk menyiapkan data agar dapat kompatibel dan arsitektur model *IndoBERT*. Dalam bagian ini dilakukan dua tahapan *preprocessing BERT*, yaitu tokenisasi dan label mapping. Metode ini digunakan dalam Python, bahasa pemrograman, untuk membagi komentar pengguna ke dalam tiga kategori sentimen: positif, netral, dan negatif [16].

a. *Tokenizing*

   Proses ini merupakan bagian dasar yang *BERT*ujuan untuk mengkonversi data teks hasil pra-pemrosesan (cleaned_text) menjadi representasi numerik berupa urutan token ID yang dapat dipahami dan diproses oleh model *IndoBERT*. Tokenisasi dilakukan menggunakan tokenizer yang sesuai dengan model *IndoBERT* yang dipilih, yaitu AutoTokenizer dari checkpoint *indobenchmark/IndoBERT-base-p1* yang diperoleh dari library Transformers Hugging Face.

b. *Label Mapping*

   Label sentimen kategorikal (Positif, Negatif, Netral) yang telah diatribusikan pada setiap data perlu dikonversi menjadi format numerik agar dapat diproses oleh model klasifikasi selama pelatihan dan evaluasi. Tahap ini, yang dikenal sebagai *Label Mapping*, dilakukan dengan menetapkan nilai integer unik untuk setiap kategori sentimen. Secara spesifik dalam penelitian ini, label 'NETRAL' dipetakan ke nilai 0, label 'NEGATIF' dipetakan ke nilai 1, dan label 'POSITIF' dipetakan ke nilai 2. Pemetaan ini diterapkan secara konsisten pada seluruh sampel data di set pelatihan, validasi, dan pengujian, di mana nilai numerik ini berfungsi sebagai target atau ground truth yang akan dipelajari dan diprediksi oleh model *IndoBERT*.

2.4. Pemisahan Data (*Data Splitting*)

Setelah proses *preprocessing* selesai, maka langkah berikutnya adalah membagi data menjadi tiga subset: 80% untuk pelatihan (2064 sampel), 10% untuk validasi (258 sampel), dan 10% untuk pengujian (259 sampel). Pembagian dilakukan untuk dapat memastikan model yang dilatih secara optimal dan pada pembagian data menggunakan metode *stratified splitting* untuk mempertahankan distribusi kelas sentimen asli (Positif, Negatif, Netral) yang tidak seimbang di setiap subset. Penggunaan random_state (nilai 42) memastikan reprodusibilitas pembagian data. Set pelatihan digunakan untuk *fine-tuning model*, set validasi untuk pemantauan kinerja selama pelatihan dan pemilihan checkpoint terbaik, dan set pengujian untuk evaluasi final kemampuan generalisasi model.

2.5. Model Training

Pada tahap ini, dilakukan proses melatih model klasifikasi teks menggunakan data untuk pelatihan model yang telah dilakukan pembagian dan untuk arsitektur dalam pemodelan ini menggunakan *IndoBERT*, dimana merupakan model pre-trained *BERT* yang dilatih menggunakan sebuah korpus bahasa Indonesia. Di pemodelan ini menggunakan *IndoBERT*-base-p1 yang telah dilakukan pre-train selanjutnya masuk ke bagian fine-tuning untuk tugas klasifikasi sentimen tiga kelas yaitu Netral, Positif, dan Negatif. Proses ini dilakukan pada set pelatihan 2064 sampel menggunakan framework Hugging Face Trainer.

Tabel 2. Parameter Inti Pelatihan Model

Parameter	Nilai	Keterangan
num_train_*epoch*s	3	Jumlah iterasi kepada pelatihan dataset
per_devixe_train_batch_size	16	Jumlah sampel per batch di setiap GPU saat pelatihan
warmup_steps	500	Menstabilkan model dengan learning rate scheduler
weight_decay	0.01	Regulasi dalam pencegahan overfitting

Tabel 3. Parameter Manajemen Pelatihan dan Evaluasi Model

Parameter	Nilai	Keterangan
eval_strategy	*'epoch'*	Evaluasi terjadi setiap akhir *epoch*
save_strategy	*'epoch'*	Model akan tersimpan setiap akhir *epoch*
load_best_model_at_end	*True*	Model terbaik akan disimpan secara otomatis berdasarkan evaluasi
metric_for_best_model	*'f1'*	Penentuan model terbaik sesuai dengan Metrik 'f1'
greater_is_better	*True*	Nilai metrik yang tinggi akan dianggap lebih baik
logging_steps	*10*	Frekuensi log dalam satuan langkah pelatihan
per_device_eval_batch_size	*16*	Ukuran batch per perangkat ketika evaluasi
report_to	*'none'*	Tidak melaporkan ke platform eksternal

Pada pelatihan model dibagi menjadi dua kelompok utama yang bisa dilihat pada Tabel 2 dan Tabel 3. Parameter Inti Pelatihan yang dapat dilihat pada Tabel 2 merupakan parameter yang langsung mempengaruhi proses pembelajaran model yang dilakukan pengaturan dari jumlah *epoch*, ukuran *epoch*, langkah warm-up, dan regularisasi. Parameter ini akan menentukan seperti apa model dapat mengerti pola dari data pelatihan yang telah diberikan. Selanjutnya untuk Tabel 3 yaitu Parameter Manajemen dan Evaluasi Model digunakan untuk melakukan manajemen dalam evaluasi model penyimpanan model, serta melacak model yang paling terbaik pada pelatihan model dilakukan. Parameter tersebut digunakan agar pelatihan dapat berjalan dengan akurat, efisien dan mendapatkan hasil model terbaik dan dapat memonitoring secara berkala yang berguna untuk terjadinya sebuah overfitting.

2.6. Evaluating Model

Setelah proses pelatihan model selesai, tahap berikutnya adalah evaluasi untuk mengukur performa dan kemampuan generalisasi model pada data yang belum pernah digunakan selama pelatihan. Evaluasi ini dilakukan pada set pengujian, yang merupakan 10% dari total dataset dan telah dipisahkan secara stratifikasi untuk merepresentasikan distribusi kelas sentimen pada data aslinya. Evaluasi model memanfaatkan metode *evaluate()* yang tersedia dalam pustaka *Hugging Face Trainer*. Metode ini secara otomatis menghitung berbagai metrik performa dengan membandingkan prediksi model terhadap label sebenarnya pada seluruh sampel di set pengujian. Metrik evaluasi yang dihitung dalam penelitian ini meliputi:

a. *Accuracy* (Akurasi): Proporsi prediksi yang benar dari total sampel.
b. *Precision* (Presisi): Proporsi positif sejati (*True Positive*) dari total prediksi positif oleh model.
c. *Recall* (*Recall*): Proporsi positif sejati (*True Positive*) dari total sampel yang seharusnya positif.
d. *F1-score* (F1): Rata-rata harmonis dari presisi dan *recall*.

Karena dataset ini memiliki distribusi kelas sentimen yang tidak seimbang, penggunaan metrik F1 dengan average 'weighted' dipilih. Metode weighted average menghitung metrik untuk setiap kelas secara independen, lalu mengambil rata-rata dengan bobot berdasarkan jumlah sampel di setiap kelas. Pendekatan ini memberikan gambaran performa model secara keseluruhan yang lebih akurat, terutama pada kelas minoritas.

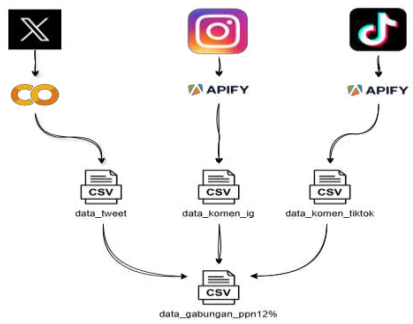
2.7. Pengujian Model

Dilakukan pengujian terhadap model yang telah dilakukan pelatihan dengan beberapa contoh teks yang berikan sesuai dari opini publik terhadap isu kenaikan PPN 12%. Pengujian ini dilakukan untuk mengevaluasi seberapa bagus model dapat melakukan klasifikasi, memahami dan menerjemahkan berbagai bentuk opini masyarakat terhadap kenaikan PPN 12%, baik dari sebuah bentuk dukungan, penolakan atau tidak paham tentang kebijakan kenaikan PPN 12%. Model yang telah dilatih ini *BERT*ujuan untuk melihat kemampuan model dalam mencerna makna teks yang diberikan beberapa kalimat yang mewakili setiap jenis reaksi publik yaitu Netral, Positif, dan Negatif seperti kalimat yang memberikan sebuah penolakan terhadap kebijakan kenaikan PPN 12%, Pernyataan yang ketidakpahaman tentang kebijakan kenaikan PPN 12%, dan Kalimat setuju terhadap kebijakan kenaikan PPN 12%.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil eksperimen klasifikasi sentimen menggunakan model *IndoBERT* yang telah disetel (*fine-tuned*) pada dataset media sosial terkait kenaikan PPN 12% di Indonesia. Pembahasan mencakup performa model yang diukur menggunakan beberapa metrik evaluasi, perbandingan hasil dengan penelitian terkait, dan interpretasi temuan mengenai distribusi sentimen publik yang terekam dalam dataset. Pada tahap pengambilan data untuk masing-masing platform dilakukan secara berbeda-beda, untuk data tweet dari X kita menggunakan teknik *web scraping*. Proses ekstrak data dimulai dengan mendapatkan *auth token* dari akun X, dan diberikan ke library *tweet harvest* agar bisa memulai *scraping* dengan *keyword* PPN12%. Untuk data dari Instagram dan TikTok, kita menggunakan API dari Apify, sehingga hanya perlu mencari sebuah postingan yang membahas tentang PPN12%, setelah itu menyalin tautan (*link*) postingan tersebut dan memberikannya ke Apify, yang nanti akan melakukan proses *scraping* data, kemudian hasilnya dapat didownload sebagai file CSV. Dataset yang digunakan dalam penelitian ini berjumlah total 2581 sampel data teks dari media sosial (X, Instagram, dan TikTok) yang relevan dengan topik kenaikan PPN 12% di Indonesia. Tahapan selanjutnya masuk ke bagian pelabelan data, dimana 2581 data dibagi menjadi 5 batch, masing-masing berisi kurang lebih 500 data untuk melakukan pelabelan secara manual. Dataset yang telah melalui proses pelabelan manual kemudian dilanjutkan ke tahap pra-pemrosesan awal. Tahap ini *BERT*ujuan untuk membersihkan noise dan artefak yang tidak relevan untuk pelatihan model. Proses pra-pemrosesan mencakup *case folding*, penghapusan *Uniform Resource Locators* (*URL*) dan *mention* pengguna, serta normalisasi kata-kata *slang* menggunakan kamus *slang* bahasa Indonesia. Hasil dari pra-pemrosesan ini akan menghasilkan teks yang bersih dan siap untuk masuk pada tahap pra-pemrosesan spesifik untuk model *IndoBERT*. Pada tahap pra-pemrosesan *IndoBERT*, data akan diubah dari bentuk teks ke dalam format numerik yang dapat dipahami dan diproses oleh arsitektur model berbasis *transformer*. Pra-pemrosesan spesifik *IndoBERT* dalam penelitian ini meliputi dua komponen utama: *tokenization* dan *label mapping*.

Dalam penelitian ini, proses *tokenization* dilakukan menggunakan *AutoTokenizer* dari library Hugging Face *transformers*. *Tokenizer* yang spesifik diambil dari *checkpoint* model *indobenchmark/IndoBERT-base-p1* yang dipilih. *Tokenizer* ini menggunakan skema *subword tokenization* (seperti WordPiece) yang efisien dalam menangani kosakata yang besar dan kata-kata yang belum pernah ditemui sebelumnya (*out-of-vocabulary*). Setiap sampel teks dari kolom *cleaned_text* diproses oleh *tokenizer* ini. Untuk memastikan semua urutan input memiliki panjang yang seragam, yang merupakan syarat untuk pemrosesan batch oleh model, parameter *padding=True* diaktifkan. Selain itu, untuk menangani teks yang panjangnya melebihi batas maksimum input model (pada *IndoBERT* dasar biasanya 512 token), parameter *truncation=True* diaktifkan. Dalam implementasi ini, panjang maksimum urutan (*max_length*) secara eksplisit diatur menjadi 512 token.



Gambar 3 . Prosess pengambilan data

Model klasifikasi sentimen, terutama pada lapisan outputnya, memerlukan target berupa nilai numerik, bukan label kategorikal berupa string (seperti 'NETRAL', 'NEGATIF', 'POSITIF'). Oleh karena itu, tahap label mapping dilakukan untuk mengkonversi label sentimen asli menjadi representasi integer yang unik untuk setiap kategori sentimen. Secara spesifik dalam penelitian ini, label 'NETRAL' dipetakan ke nilai 0, label 'NEGATIF' dipetakan ke nilai 1, dan label 'POSITIF' dipetakan ke nilai 2. Pemetaan ini diterapkan secara konsisten pada seluruh sampel data di set pelatihan, validasi, dan pengujian, di mana nilai numerik ini berfungsi sebagai target atau ground truth yang akan dipelajari dan diprediksi oleh model *IndoBERT*.

Setelah proses *preprocessing* selesai, maka langkah berikutnya adalah membagi data menjadi tiga subset: 80% untuk pelatihan (2064 sampel), 10% untuk validasi (258 sampel), dan 10% untuk pengujian (259 sampel). Pembagian dilakukan untuk dapat memastikan model yang dilatih secara optimal dan pada pembagian data menggunakan metode *stratified splitting* untuk mempertahankan distribusi kelas sentimen asli (Positif, Negatif, Netral) yang tidak seimbang di setiap subset. Setelah dataset pra-pemrosesan dan label mapping selesai, model *IndoBERT* (indobenchmark/*IndoBERT*-base-p1) disetel (*fine-tuned*) untuk tugas klasifikasi sentimen tiga kelas (Netral, Negatif, Positif). Proses fine-tuning dilakukan selama 3 *epoch* menggunakan set pelatihan, dengan pemantauan performa pada set validasi untuk memilih model terbaik. Evaluasi akhir kinerja model dilakukan pada set pengujian yang sepenuhnya terpisah dan belum pernah digunakan selama pelatihan. Evaluasi ini *BERT*ujuan untuk mendapatkan estimasi objektif mengenai kemampuan generalisasi model pada data baru. Metrik yang diukur meliputi *Accuracy*, *Precision*, *Recall*, *dan F1-score*, dengan *F1-score* dihitung menggunakan weighted average untuk menangani ketidakseimbangan distribusi kelas.

Model *IndoBERT (indobenchmark/IndoBERT-base-p1)*, yang merupakan model bahasa berbasis Transformer yang telah dilatih pada korpus Bahasa Indonesia, kemudian dilatih lanjut (*fine-tuned*) pada set pelatihan selama 3 *epoch*. Konfigurasi parameter pelatihan, termasuk ukuran batch 16 per perangkat, langkah warmup 500, weight decay 0.01, dan penggunaan metrik F1 pada set validasi untuk menentukan model terbaik yang akan dimuat di akhir pelatihan, dijelaskan secara detail pada Tabel 2 dan Tabel 3 di bagian Metodologi. Setelah proses pelatihan selesai, model terbaik yang diperoleh dievaluasi secara komprehensif pada set pengujian (259 sampel) menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Mengingat adanya ketidakseimbangan distribusi kelas, metrik F1 dihitung menggunakan average 'weighted' untuk memberikan bobot performa pada setiap kelas sesuai dengan jumlah sampelnya, sehingga hasil evaluasi performa keseluruhan model lebih representatif. Hasil numerik dari evaluasi model pada set pengujian disajikan pada Tabel 5.

Tabel 4. Hasil Evaluasi Model Terhadap Data Uji

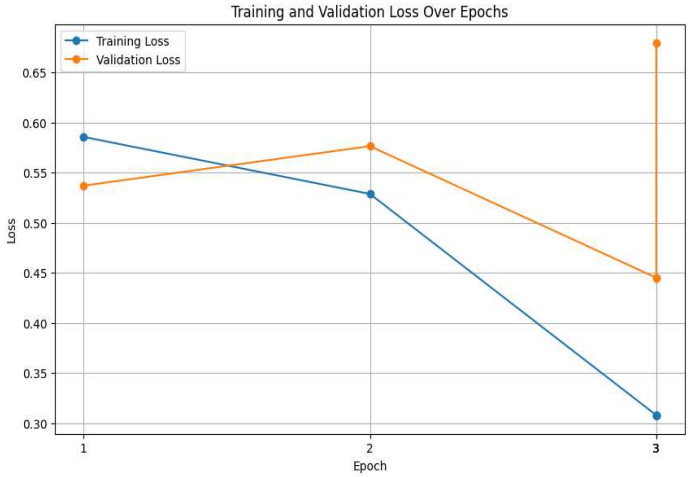
Parameter	Nilai
eval_loss	0.3584
eval_accuracy	0.8494
eval_f1	0.8437
eval_precision	0.856
eval_recall	0.8494
eval_runtime	0.9788
eval_samples_per_second	264.62
eval_steps_per_second	17.369
epoch	3

Tabel 4 menyajikan hasil evaluasi akhir model *IndoBERT* pada data uji. Berdasarkan tabel tersebut, kinerja model yang diperoleh dalam penelitian ini tergolong baik dan kompetitif dalam konteks analisis sentimen berbahasa Indonesia. Performa akurasi mencapai 84,94% dan *F1-score (weighted)* sebesar 84,37% menunjukkan bahwa model *IndoBERT* sangat efektif untuk tugas klasifikasi sentimen ini. Hasil ini sebanding atau bahkan sedikit lebih tinggi dibandingkan beberapa penelitian sebelumnya yang menggunakan model bahasa serupa atau metode tradisional untuk analisis sentimen berbahasa Indonesia. Sebagai contoh, studi oleh [6] yang menganalisis sentimen terkait kenaikan PPN 12% menggunakan metode *Decision Tree* pada data dari media sosial X melaporkan akurasi sebesar 81,34%. Sementara itu, penelitian oleh [17] yang juga menggunakan model *IndoBERT* untuk analisis sentimen calon presiden di Twitter melaporkan akurasi keseluruhan 80%. Perbandingan ini mengindikasikan bahwa model *IndoBERT* memiliki kapabilitas yang kuat dan dalam kasus penelitian ini, mampu menghasilkan kinerja yang solid, meskipun perbedaan hasil dapat dipengaruhi oleh karakteristik spesifik dataset dan isu yang dibahas. Parameter lain seperti presisi (0.856) dan *recall* (0.8494) pada data uji juga menunjukkan keseimbangan yang baik dalam kemampuan model untuk mengidentifikasi kelas sentimen secara benar.

Tabel 5. Ringkasan Metrik Validasi per *Epoch*

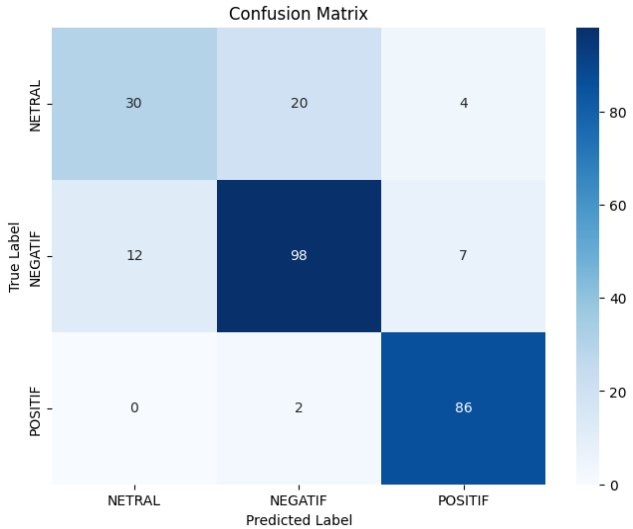
Epoch	*Training Loss*	*Validation Loss*	*Validation Accuracy*	*Validation F1*	*Learning rate*
1.0	*None*	0.537049	0.751938	0.741770	0.000012

2.0	*None*	0.576449	0.755814	0.706825	0.000025
3.0	*None*	0.679665	0.759690	0.726284	0.000038
3.0	*None*	0.445119	0.826255	0.819790	0.000038
3.0	*None*	0.445119	0.826255	0.819790	0.000038



Gambar 4 . *Diagram Training* dan *Validation Loss Over Epochs*

Tabel 4 menyajikan ringkasan metrik validasi yang diamati pada setiap akhir *epoch*. Tabel ini mencakup nilai *training loss*, *validation loss*, akurasi validasi, *F1-score* validasi, dan laju pembelajaran (*learning rate*) yang digunakan pada setiap *epoch*. Dari data pada Tabel 5 dan didukung oleh kurva loss pada Gambar 4, terlihat bahwa *training loss* menurun secara konsisten dari 0.5857 pada *epoch* pertama menjadi 0.3080 pada *epoch* ketiga. *Validation loss*, setelah mengalami sedikit kenaikan dari 0.5370 (*epoch* 1) menjadi 0.5764 (*epoch* 2), kemudian menurun secara signifikan ke nilai 0.4451 pada *epoch* ketiga. Fenomena kenaikan *validation loss* pada *epoch* kedua sementara *training loss* menurun mengindikasikan adanya potensi *overfitting*, namun penurunan kembali pada *epoch* ketiga menunjukkan bahwa model berhasil menemukan representasi fitur yang lebih *generalizable*. Laju pembelajaran yang tercatat menunjukkan penyesuaian yang dilakukan oleh *scheduler* selama proses pelatihan, dimulai dari 0.000012 pada *epoch* pertama dan berakhir pada 0.000038 di *epoch* ketiga. Kinerja *F1-score* tertinggi pada set validasi (0.819790) dicapai pada *epoch* ketiga, bersamaan dengan akurasi validasi tertinggi (0.826255), yang mengindikasikan bahwa model pada *epoch* ini memiliki keseimbangan terbaik antara presisi dan recall pada data validasi.



Gambar 5 . Confusion Matrix Model pada Data Uji

Gambar 5 menyajikan *confusion matrix* yang dihasilkan *model IndoBERT* pada data uji. Diagram ini memberikan visualisasi kinerja klasifikasi untuk masing-masing kelas sentimen: NETRAL, NEGATIF, dan POSITIF. Dari

diagonal utama, terlihat bahwa model berhasil mengklasifikasikan 98 sampel NEGATIF dengan benar, 86 sampel POSITIF dengan benar, dan 30 sampel NETRAL dengan benar. Hal ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali sentimen NEGATIF dan POSITIF. Namun, terdapat beberapa kesalahan klasifikasi yang perlu diperhatikan. Sebanyak 20 sampel NETRAL salah diklasifikasikan sebagai NEGATIF, yang merupakan jenis kesalahan paling umum untuk kelas NETRAL. Selain itu, 12 sampel NEGATIF juga salah diklasifikasikan sebagai NETRAL. Kesalahan klasifikasi antara kelas NETRAL dan NEGATIF ini kemungkinan disebabkan oleh ambiguitas dalam teks atau nuansa bahasa yang sulit dibedakan oleh model. Sementara itu, model menunjukkan kinerja yang baik dalam membedakan sentimen POSITIF, dengan hanya 2 sampel POSITIF yang salah diklasifikasikan sebagai NEGATIF dan tidak ada yang salah diklasifikasikan sebagai NETRAL.

Secara keseluruhan, *confusion matrix* ini mengkonfirmasi bahwa *model IndoBERT* yang dilatih cukup efektif dalam tugas klasifikasi sentimen terkait kenaikan PPN, dengan performa yang lebih kuat pada kelas NEGATIF dan POSITIF dibandingkan kelas NETRAL. Setelah tahap *training* dan *evaluasi* kita akan melakukan *pengujian* terhadap model yang dilatih. *Pengujian* ini dilakukan untuk mengevaluasi seberapa bagus model dapat melakukan klasifikasi, memahami dan menerjemahkan berbagai bentuk opini masyarakat terhadap kenaikan PPN 12%, baik dari sebuah bentuk dukungan, penolakan, atau ketidakpahaman tentang kebijakan kenaikan PPN 12%.

Tabel 5. Hasil Prediksi Model Terhadap Uji Teks

Teks	Hasil	Akurasi
'PPN naik = gaji harus naik'	NEGATIF	0.9045
'ppn itu apa bg ?'	NETRAL	0.7216
'saya suka kebijakan ppn12% naik ini'	POSITIF	0.9359
'Kebijakan PPN 12% sangat membebani rakyat.'	NEGATIF	0.8281

Dapat dilihat pada Tabel 5 yang merupakan hasil pengujian terhadap beberapa contoh teks mengenai kebijakan kenaikan PPN 12% bahwa model dapat mengklasifikasi teks dengan tingkat akurasi lumayan tinggi seperti pada klasifikasi pernyataan yang berunsur Negatif 'PPN naik = gaji harus naik' dan 'Kebijakan PPN 12% sangat membebani rakyat.' memiliki akurasi diatas 0.8 yang menunjukkan bahwa model dapat membaca teks yang mengandung unsur ketidakpuasan dan keluhan terhadap kebijakan kenaikan PPN 12%. Selain itu, pada kalimat 'ppn itu apa bg ?' yang masuk ke bagian teks bersifat netral berhasil di klasifikasikan oleh model dengan tingkat akurasi lumayan yaitu di atas 0.7. Sementara itu, pada kalimat yang berunsur positif 'saya suka kebijakan ppn12% naik ini' memiliki nilai akurasi yang sangat tinggi yaitu 0.9359. Secara keseluruhan, model yang telah dilakukan uji teks dapat membedakan kelas sentimen secara akurat. Selain evaluasi kuantitatif, dilakukan analisis kualitatif terhadap kasus-kasus misklasifikasi oleh model *IndoBERT* pada set pengujian untuk memahami batasan dan area potensial perbaikan. Beberapa contoh teks yang salah diklasifikasikan beserta analisis penyebabnya disajikan pada Tabel 6.

Tabel 6. Contoh Misklasifikasi Model *IndoBERT* beserta Analisis Penyebab

Teks	Prediksi Model	Label Sebenarnya	Analisis Kemungkinan Penyebab Misklasifikasi
Mantap betul PPN 12%! Dompet makin tebel nih gara-gara hemat pengeluaran. Makasih ya pemerintah!	POSITIF 0.9002	NEGATIF	Model cenderung melihat kata-kata "mantap betul", "dompet makin tebel", "hemat", "makasih" sebagai indikator positif secara literal, tanpa menangkap nada sarkastik yang menyindir dampak Tentu, berdasarkan sebaliknya.
Kenaikan tarif PPN menjadi 12 persen mulai 2025 akan berdampak pada pengeluaran masyarakat	POSITIF 0.4579	NEGATIF	Ambiguitas Semantik & Kurangnya Sinyal Negatif Kuat: Kata "berdampak" bersifat netral. Model mungkin gagal menangkap konotasi negatif yang sering diasosiasikan dengan "peningkatan pengeluaran masyarakat" tanpa adanya kata kunci negatif yang lebih eksplisit. Probabilitas yang tidak terlalu tinggi (0.4579) juga menunjukkan ketidakpastian model.

Ayo membayar pajak demi kesejahteraan para penjabat	POSITIF 0.9061	NEGATIF	Kegagalan Mendeteksi Sarkasme Kuat: Model menginterpretasikan frasa "Ayo membayar pajak" dan "demi kesejahteraan" secara literal sebagai ajakan positif. UnTentusur ironi yang sangat kuat ("kesejahteraan para penjabat") tidak berhasil mengubah klasifikasi model, yang terlihat dari probabilitas prediksi positif yang sangat tinggi.

4. KESIMPULAN

Penelitian ini *BERT*ujuan membangun dan mengevaluasi model klasifikasi sentimen terhadap kebijakan kenaikan *PPN 12%* di Indonesia menggunakan pendekatan *Natural Language Processing (NLP)* berbasis model *Transformer*, yaitu *IndoBERT*. Data dikumpulkan dari platform media sosial *X*, Instagram, dan TikTok, kemudian melalui tahap pra-pemrosesan teks untuk menghasilkan data bersih, tokenisasi, dan pemetaan label. Model *IndoBERT* (checkpoint *indobenchmark/IndoBERT-base-p1*) di-*fine-tune* pada dataset yang dibagi secara stratifikasi (80% pelatihan, 10% validasi, 10% pengujian) selama tiga *epoch*. Evaluasi akhir pada data uji menunjukkan performa model yang kuat dengan *akurasi* 84,94%, *presisi* (weighted) 0,8460, *recall* (weighted) 0,8494, dan *F1-score* (weighted) 0,8437. Angka-angka ini mencerminkan kemampuan model dalam memahami konteks bahasa dan nuansa sentimen pada teks media sosial terkait isu kenaikan *PPN 12%*. Analisis distribusi sentimen mengungkap bahwa sentimen negatif mendominasi respons publik, memberikan wawasan penting bagi pemangku kepentingan mengenai potensi penolakan kebijakan. Penelitian ini menegaskan keunggulan *IndoBERT* dalam menangkap representasi linguistik yang kompleks pada teks berbahasa Indonesia dari media sosial.

Untuk pengembangan lebih lanjut, disarankan memperluas cakupan data ke sumber lain seperti forum diskusi atau komentar berita guna meningkatkan kemampuan generalisasi model terhadap variasi gaya bahasa. Optimalisasi tahap *fine-tuning* dan eksplorasi fitur linguistik tambahan juga dapat meningkatkan performa model. Meskipun model *Transformer* ini memiliki performa yang kuat, penelitian sebelumnya menunjukkan kerentanannya terhadap input yang dimanipulasi atau kondisi *stress test*. Oleh karena itu, menguji ketahanan model terhadap gangguan atau *adversarial examples* merupakan area penting untuk penelitian berikutnya demi memahami kelemahan model dalam kondisi dunia nyata yang tidak ideal.

DAFTAR PUSTAKA

[1] U. Faruq, S. Adipurno, A. Aziz, N. Faadhilah, and M. Ridwan, "Konsep Dasar Pajak dan Lembaga yang Dikenakan Pajak : Tinjauan Literatur dan Implikasi untuk Kebijakan Fiskal," *J. Ekon. dan Bisnis*, vol. 16, no. 2, pp. 65–70, 2024, doi: 10.55049/jeb.v16i2.306.

[2] N. Fauziah, M. Alkautsar, Y. Suryaman, and F. F. Roji, "Pelabelan VADER Dalam Menganalisis Persepsi Masyarakat Terhadap Kenaikan Tarif PPN di Indonesia," *J. Akunt. DAN Keuang.*, vol. 12, no. 2, pp. 228–238, 2024.

[3] M. G. Al-kadzim, "Analisis Perubahan Sentimen Publik di Media Sosial X terhadap Konflik Palestina-Israel Menggunakan Model *IndoBERT*," *Digit. Transform. Technol.*, vol. 4, no. 2, pp. 1167–1174, 2024.

[4] D. Al Akhdaan, T. E. Sutanto, and M. Liebenlito, "Confident Learning pada *IndoBERT*: Peningkatan Kinerja Klasifikasi Sentimen," *Indones. J. Comput. Sci.*, vol. 13, no. 5, pp. 2357–2386, 2024.

[5] Tarwoto, R. Nugroho, N. Azka, W. Sayudha, and R. Graha, "Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi) Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 9, no. June, pp. 495–505, 2025.

[6] J. K. Siagian and Painem, "ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP RENCANA KENAIKAN PPN MENJADI 12 % DI MEDIA SOSIAL X ANALYSIS OF INDONESIAN PUBLIC SENTIMENT TOWARDS THE PLAN TO INCREASE VAT TO 12 % ON X SOCIAL MEDIA USING THE NAÏVE BAYES METHOD," in *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2024, pp. 779–786.

[7] W. Irmayani, "PERSEPSI PUBLIK TERHADAP KENAIKAN PPN 12%: PENDEKATAN SENTIMEN PADA KOMENTAR YOUTUBE," *J. KHATULISTIWA Inform.*, vol. 12, no. 2, pp. 112–118, 2024.

[8] F. Imawan, D. F. Shiddieq, and F. F. Roji, "Analisis Sentimen Publik di X Terhadap Rencana Kenaikan PPN 12% Menggunakan *BERT*," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 10, no. 1, pp. 136–148, 2025.

[9] A. F. Setyawan, A. Devi, P. Ariyanto, F. K. Fikriah, and R. I. Nugraha, "Analisis Sentimen Ulasan iPhone di Amazon Menggunakan Model Deep Learning *BERT* Berbasis Transformer," *J. ELEKRONIKA DAN Komput.*, vol. 17, no. 2, pp. 447–452, 2024.

[10] P. Zahwa, S. Agustian, F. Yanto, and S. Baru, "IMPLEMENTASI BI-DIRECTIONAL LONG SHORT

TERM MEMORY TERHADAP KLASIFIKASI SENTIMEN DI," *Zo. J. Sist. Inf.*, vol. 7, no. 1, pp. 11–24, 2025.

[11] P. P. Wulan and H. Basri, "Analisis Sentimen Terhadap Layanan Nasabah Bank Menggunakan Teknik Klasifikasi Naive Bayes Sentiment Analysis of Banking Customer Service Using Naive Bayes Classification Technique," *J. Kecerdasan Buatan dan Teknol. Inf.*, vol. 3, no. 2, pp. 68–74, 2024.

[12] D. Nuryadi *et al.*, "FINE TUNING *INDOBERT* UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI TIKET . COM DI GOOGLE PLAY STORE," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 3577–3583, 2025.

[13] M. T. Uliniansyah *et al.*, "Twitter dataset on public sentiments towards biodiversity policy in Indonesia," *Data Br.*, vol. 52, p. 109890, 2024, doi: 10.1016/j.dib.2023.109890.

[14] B. Yanuargi, "Analisis Sentemen Terhadap Aplikasi Bukalapak Sebelum IPO dan Sesudah IPO Menggunakan Algoritma Naive Bayes," *Jnanaloka*, vol. 3, no. 1, pp. 17–25, 2022, doi: 10.36802/jnanaloka.2022.v3-no1-17-25.

[15] R. Merdiansah and A. Ali Ridha, "Sentiment Analysis of Indonesian X Users Regarding Electric Vehicles Using *IndoBERT*," *J. Ilmu Komput. dan Sist. Inf. (JIKOMSI*, vol. 7, no. 1, pp. 221–228, 2024.

[16] A. Wirayudha, P. S. Informasi, U. Gunadarma, D. Learning, and G. P. Store, "Analisis Sentimen Terhadap Ulasan Access By KAI Pada Google Play Store Menggunakan Metode *IndoBERT*," *PRINSIP Portal Ris. Inov. Sist. Perangkat Lunak*, vol. 3, pp. 9–20, 2025.

[17] P. Sayarizki and H. Nurrahmi, "Implementation of *IndoBERT* for Sentiment Analysis of Indonesian Presidential Candidates," *J. Comput.*, vol. 9, no. 2, pp. 61–72, 2024, doi: 10.34818/indojc.2024.9.2.934.